JMIR Rehabilitation and Assistive Technologies

Development and Evaluation of Rehabilitation, Physiotherapy and Assistive Technologies, Robotics, Prosthetics and Implants, Mobility and Communication Tools, Home Automation and Telerehabilitation
Volume 13 (2026) ISSN 2369-2529 Editor in Chief: Sarah Munce, MS, PhD

Contents

Original Papers

Validity and Reliability of the Track-UL Algorithm Compared With Kinovea Software for Measuring Upper-Limb Functional Range of Motion in People After Stroke: Cross-Sectional Observational Study (e87128)
Hatem Lazem, David Harris, Abi Hall, Thomas Richards, Phaedra Leveridge, Maedeh Mansoubi, Xiaohan Xu, Paul Newell, Sarah Lamb, Helen Dawes. . . . 11

Intelligent Robot-Aided, Task-Oriented Physiotherapy Method for Upper Limb Rehabilitation: Development and Usability Study (e92898)
Piotr Falkowski, Kajetan Jeznach, Jan Oleksiuk, Tomasz Osiak, Maciej Pikuli ski, Andrzej Zakr cki, Krzysztof Zawalski, Piotr Kołodziejski, Daniel Ii . . . 28

Scientific and Technological Developments in Brain-Computer Interfaces: Dual Bibliometric Analysis of Articles and Patents (e95902)
Junhui Wang, Yakun Yuan, Hang Xu, Guannan Luan. . . . 47

Digitally Assisted Training System for Home-Based Scoliosis Therapy: Proof-of-Concept and Technical Verification Study (e93721)
Florian Günther, Sandra Hunger, Michael Werner, Nikolaus Landerer, Christoph-Eckhard Heyde, Welf-Guntram Drossel. . . . 64

A Benchmarking Framework for Real-Time 3D Pose Estimation in Motor Rehabilitation: Development and Proof-of-Concept Study (e91360)
Emanuel Lorenz, Beatrix Vereijken, Kerstin Bach, Xiaomeng Su. . . . 85

User Requirements of the Integrated Home-Based Rehabilitation Tool, Neurorehabilitation Ecosystem for Sustained Therapy: Multicenter Focus Group Study With Stroke Survivors, Caregivers, and Health Care Professionals (e79382)
Merle Matijsen, Cheriel Hofstad, Susana Rodriguez Gonzalez, Xavier Buxó, Vera Stara, Sara Leonzi, Margherita Rampioni, Anna Mura, Noël Keijsers. . . . 105

Development of Mobile Software “SRCardioCare” Prototype for Implementing Home-Based Exercise Program Among Patients After Adult Cardiac Surgical Revascularization: Qualitative Feasibility Study (e69197)
Ajith Kumar Pichai, A Sathya, Thillaigovindarajan SenthilKumar, Sadhanandham Shanmugasundaram, R Karthik. . . . 121

XSL•FO
RenderX

Leveraging Large Language Models for Early Detection of Anomaly Work Injury Cases: Data-Driven Approach to Rehabilitation Efficiency (e80607)
Peter Chen, Hayley Gu, Heidi Lo, Wing Chang, Cameron Lai, Sun Lai, Andy Cheng, Peter Ng. 137

Predicting Geriatric Rehabilitation Stays of ≤4 Weeks After Hip Fracture Surgery: Machine Learning Approach Using Physical Activity and Patient Data (e79331)
Sanne Krakers, Frank Wouda, Dieuwke van Dartel, Miriam Vollenbroek-Hutten, Johannes Hegeman, Up&Go After a Hip Fracture Group. . .
1 5 3

Enhancing Knee Joint Proprioception in Healthy Adults Through Exergame Training With Augmented Feedback: Randomized Controlled Pilot Trial (e78525)
Yiling Zhang, Luis García Arias, Hans Timmerman, Ming Cao, Elisabeth Wilhelm. 169

Evaluation of a Stretching Forearm Sleeve for Lateral Epicondylitis: Repeated Measures Study (e80400)
Adriana Ríos Rincón, Christine Guptill, Ann Tran, Rija Kamran, Salamah Alshammari, Antonio Miguel Cruz. 185

Usability, Acceptability, and Feasibility of a Personalized Adaptive Mirror Therapy for Upper-Limb Poststroke Rehabilitation Using Immersive Virtual Reality and Myoelectric Control: Single-Arm Pre-Post Study (e81894)
Daniela De Bartolo, Paolo De Pasquale, Marta Russo, Denise Berger, Antonella Maselli, Daniele Borzelli, Christian Nissler, Markus Nowak, Claudio Castellini, Giovanni Morone, Andrea d'Avella. 200

Understanding Education and Training Needs in Rehabilitation Technologies Among Those With Lived Experience, Rehabilitation Health and Care Professionals, and Technology Professionals: Qualitative Study (e97602)
Angela Thornton, Matthew Horrocks, Victoria Abbott-Fleming, Holly Blake. 220

Development, Usability, and Feasibility of a Pain-Stratified Smartphone-Based Home Intervention System for Patients With Adhesive Capsulitis: Exploratory Feasibility and Usability Trial (e107314)
Hsun-Han Hsu, Kai-Chia Cheng, Ya-Lan Chiu. 236

Comparing the Feasibility of S-Patch vs 24-Hour Holter in Frail Older People With Falls: Randomized Open-Label Crossover Pilot Study (e81761)
Barbara Rosario, Vern Tan, Vivian Barrera, Hnin Mon, Clara Seah, Colin Yeo. 253

Reliability and Discriminant Ability of an Instrumented Timed Up and Go Test in People With Postsurgical Orthopedic Conditions: Quantitative Study (e82632)
Marica Giardini, Ilaria Arcolin, Valerio Arcobelli, Michela Picardi, Sabato Mellone, Marco Godi. 270

Effects of an Exercise-Assisting Mobile App (Osteoarthritis-Rehabilitation Assistant [O-RA]) on Rehabilitation Outcomes in Older Adults: Randomized Controlled Parallel Clinical Trial (e80971)
Pajaree Sornmayura, Sintip Pattanakuhar, Napaschol Intapan, Krittipat Tragoolpua, Kaewkla Sroykabkaew, Rungkan Wangboon, Krittipong Wachirangkul, Sartita Wannarat, Jakkrit Klaphajone. 289

Smartphone-Supported Vestibular Rehabilitation in Individuals With Vestibular Dysfunction: Pilot Randomized Crossover Trial Assessing Functional Clinical Outcomes and Anxiety (e84207)
Azriel Kaplan, Liran Kalderon, Amit Wolfovitz, Yoav Gimmon, Shelly Levy-Tzedek. 302

Smartphone-Based Grading and Rehabilitation in Patients With Facial Palsy Using Computer Vision: Prospective Validation Study (e81963)
Franz-Tassilo Müller-Graff, Fabian Essig, Maximilian Friedrich, Kathrin Hoika, Stephan Hackenberg, Kristen Rak, Johannes Taeger. 318

Reducing Educational Bias in Cognitive Assessment via Dynamic Support Vector Machine Weighting: Validation Study on an Education-Stratified Dataset (e79841)
Qing Liu, Chi Ma, Mengyuan Liu, Suhui Chen, Mengting Yu, Lijuan Xia, Qi Zhang, Ming Wu. 336

XSL•FO
RenderX

Using a Co-Designed Digital Self-Management Program to Prepare Patients for Hip or Knee Replacement Surgery: Pragmatic Pilot Study (e68286)
Elizabeth Horton, Hayley Wright, Andy Turner, Louise Moody, Lucy Aphramor, Anna Carlson, Hesam Ghiasvand, Shea Palmer. 347

Stakeholder Perspectives on Facilitators of and Barriers to the Implementation of Internet of Things–Equipped Powered Wheelchairs in Sweden: Qualitative Case Study (e92414)
Johan Borg, Marie Elf, Anders Brantnell, Klas Palm. 365

Participant-Reported Reductions in Falls and Fall-Related Injuries Associated With Variable-Friction Footwear: Retrospective Pre-Post Survey Study (e104177)
Sara Nataletti, Jennifer Bartloff, Sara Prokup, Joseph Harris, Anushka Larson, Tyler Susko, Arun Jayaraman. 380

Indoor Navigation for People With Visual Impairment in Canada: Participatory Co-Design and Interdisciplinary Study of the Edge A-Eye Platform (e81347)
Soutongnoma Kaboré, Deborah Annan, Prajjol Puri, Sergio Mejia Romero, Nathalie Gingras-Royer, Justin Gauthier, H ng Nhung, Phuong Nguyen, Benoit Pelletier, Kim Nguyen, Hassane Alami, Joseph Nemargut. 406

Accuracy in the Estimation of Self-Reported Knee Brace Wear Time in Young Adults With a Symptomatic Knee Following ACL Reconstruction: Secondary Analysis of a Pilot Randomized Controlled Trial (e79725)
Matthew Savage, Benjamin Mentiplay, Harriette Slater, Fernanda Serighelli, David Carey, Jamon Couch, Andrea Bruder, Adam Culvenor. . .
4 2 9

Health Care Providers' Perspectives on a Hybrid Outpatient Stroke Telerehabilitation Program: Qualitative Implementation Study (e83081)
Marina Wasilewski, Shaghayegh Mirbaha, Karl Wong, Gary Siu, Siobhan Donaghy, Justin Huynh, Sarmiya Uruthiralingam, Michelle Nelson, Elizabeth Linkewich, Ron Lacombe, Matthew Godleski. 453

Clinical Outcomes of Individualized Electrostimulation Using a Wearable Electro Suit and Qualitative Feedback From a Mixed Cohort of Survivors of Stroke and Spinal Cord Injury With Spasticity: Case Series (e81522)
Jia Yen, Nur Kamsani, Hwa Lai, Ning Tang, Effie Chew. 468

Design Implications of Comfort and Usability of Manual Stairclimbing Wheelchair: Ergonomic Assessment and Pilot Study Using Surface Electromyography Inputs (e78965)
Abhishek Verma, Rohit Kumar, J Ramkumar. 487

A Social Justice Approach to Assistive Technology and Well-Being of People With Visual Disabilities in Low- and Middle-Income Countries: Qualitative Narrative Study (e72306)
Luisa Ortiz-Escobar, Mario Chavarria, Samia Hurst-Majno, Oscar Campo Salazar, Celia Escobar-Hurtado, Michael Stein, Minerva Rivas Velarde. 513

Comparison of Upper Body Joint and Hand Motions in Eating Solid Foods With Chopsticks and Semisolid Foods With a Spoon in Healthy Males and Females: Observational Study (e76239)
Jun Nakatake, Shogo Maeda, Shigeaki Miyazaki, Hideki Arakawa, Etsuo Chosa. 534

Exploring Barriers and Enablers for the Intention to Use Assistive Robotics Among People With Spinal Cord Injury and Those Involved in Their Care: Qualitative Study (e72080)
Susanne Frennert, Johanna Persson, Eva Díez-Rodríguez, Monica Alcobendas-Maestro, Fátima Villamayor Vega, Antonio Oliviero. 557

Blockchain-Based Mobile App for Digital Identification of Older Adults in Rural Peru: Design and Usability Evaluation Study (e79553)
Wilver Arana-Ramos, Aldo Pastrana-Leon, Juan Morales-Arevalo. 572

Role of Technology Acceptance in the Telerehabilitation of Patients With Metabolic Syndrome: Longitudinal Study (e82161)
Zsanett Tesch, Tamás Ujházi, István Kósa, Norbert Buzás. 589

XSL•FO
RenderX

Introduction of Nurse-Led Rehabilitation Services for Patients With Stroke After Discharge to Improve Self-Care Management in Bangladesh: Pilot Randomized Controlled Trial (e88808)
Salma Akhter, KATM Huq, Shota Date, Mahmudul Islam, Farida Chhobi, SK Sayeed, Md Islam, Mahabuba Afrin, Kana Kazawa, Michiko Moriyama. . . . 607

Experiences of Accessing Physical Rehabilitation Services and Assistive Products Among Orthosis and Wheelchair Users: Qualitative Study (e84735)
Charlotte Spurway, Andrew Power, Maggie Donovan-Hall, Andrew Channon. . . . 623

Basic Technology to Support Stimulation Access, Activity, and Correct Posture in Individuals With Intellectual and Multiple Disabilities: Development and Usability Study (e92873)
Giulio Lancioni, Gloria Alberti, Valeria Chiariello, Elia Zacchini, Nirbhay Singh, Mark O'Reilly, Jeff Sigafoos, Serafino Buono. . . . 634

Usability and Functional Outcomes of the Obi3 Robotic Feeding Device Among People With Disabilities, Caregivers, and Providers: Mixed Methods Study (e86338)
Andrea Fairman, Ryan Osal, Heather Keeton, Jonathan Dekar. . . . 652

AI Scribe in Specialized Rehabilitation: Usability Study of Perceived Usability and Usefulness Among Physicians and Psychologists (e90442)
Harald Engan, Ingeborg Landstad, Kari Sand, Ståle Evenshaug, Einar Henjum Halsnes. . . . 668

Robotic Rollator Assistance for Sit-to-Stand Transfers in Older Adults with Mobility Challenges: Within-Participant Feasibility Study (e101303)
Anas Mahdi, Theresa Buchner, Christian Werner, Yue Hu, Katja Mombaur. . . . 687

Usability and Acceptance of a Social Robot in People With Multiple Sclerosis: Comparative Cross-Sectional Study (e94068)
Jessica Podda, Alice Bollini, Alessia Susini, Domenico Presenza, Keven Kearney, Ester Cecchella, Erica Grange, Mario Battaglia, Andrea Tacchino, Giampaolo Brichetto. . . . 705

Effects of Using a Lumbar-Type Hybrid Assistive Limb by Older Adults With Frailty on Exercise Self-Efficacy: Retrospective Analysis Based on the Kanagawa Prospective ME-BYO Cohort Study (e79358)
Makoto Nagasawa, Sho Nakamura, Yoshinobu Saito, Yuko Oguma, Takashi Kasukawa, Kaname Watanabe, Hiroto Narimatsu. . . . 724

Concurrent Validity of Isokinetic Hip and Knee Strength Measurements Using the UGO Exoskeleton in Healthy Young Adults: Cross-Sectional Validation Study (e99564)
Hao Xie, Ming-Hui Ding, Pei-Ming Chen, Shamay Ng, Chu-Huai Wang, Jiang-Li Zhao. . . . 736

Economic, Organizational, and Research Impact Evaluation of Introducing Robotic Rehabilitation in a Medium-Sized Hospital: Scenario-Based Comparative Analysis (e93572)
Alessandro Scano. . . . 748

From Metrics to Meaning in Neurological Rehabilitation: Clinicians' Perspectives on Digital Metrics of Upper Limb Functioning—A Focus Group Study (e87339)
Johannes Pohl, Laura Mayrhuber, Olivier Lambercy, Chris Easthope Awai. . . . 764

Smartphone App–Delivered Mindfulness-Based Intervention for Concussion in Adolescents (MBI-4-mTBI): Feasibility Randomized Controlled Trial (e84623)
Veronik Sicard, Molly Cairncross, Roger Zemek, Noah Silverberg, Gary Goldfield, Nick Reed, Bechara Saab, Andra Smith, Andrée-Anne Ledoux. . . . 781

User Experience of Extended Reality Treatment for Visuospatial Neglect Among Patients and Informal Caregivers: Qualitative Interview Study (e80136)
Eileen Bousché, Hanne Huygelier, Tanja Nijboer, Knowledge Broker Neglect Study Group. . . . 800

XSL•FO
RenderX

Digital Platform to Provide Health Data Feedback for Neurorehabilitation Patients: User-Centered Development and Proof-of-Concept Usability Study (e85072)
Nadine Domnik, Katarzyna Krasnopolska, Ramona Sylvester, Jens Bansi, Jaeyong Song, Roman Gonzenbach, Olivier Lambercy. 815

Therapeutic Use of Virtual Reality for Patients With Fibromyalgia and Chronic Neck Pain: Randomized Controlled Trial (e81158)
Edurne Úbeda-D'Ocasar, Yaiza Moreno-Crespo, Eduardo Cimadevilla-Fernández-Pola, Juan Hernández-Lougedo, Álvaro Navas-Mosqueda, Mario Caballero-Corella, Noemí Mayoral-Gonzalo, Blanca Pedauyé-Rueda, María Fernández-Aceñero, Juan Hervás-Pérez, Cristina Ojedo-Martín. 836

Gaze Allocation and Performance Across Task-Demand Conditions During Squat-Based Exergaming: Pilot Study Using Eye Tracking (e81667)
Keiichi Takei. 850

Designing an mHealth App for Stroke Rehabilitation in Indonesia: Mixed Methods Design Science Research Study (e91464)
Sabrina Shania, Putu Handayani, Kamila Imanuddin. 863

Deep Learning–Based Automated Clinical Gait Assessment From Kinematic Data in People With Stroke: Development and Validation Study (e94031)
Zhexuan Gu, Jiaqi Li, Wenbin Zhou, Kenneth Fong, Lin Wang, Eng Lua, Patrick Kwong, Yancheng Yuan. 881

Efficacy of High-Intensity Laser Therapy Combined With Plantar Fascia Stretching Exercises in the Treatment of Plantar Fasciitis: Randomized, Double-Blind, Sham-Controlled Trial (e77419)
Nantaporn Jitpimolmard, Phonlawat Ouemphancharoen, Preeda Arayawichanon. 895

Validity and Reliability of an Immersive Virtual Reality System for Multidimensional Assessment of Cervical Sensorimotor Control: Cross-Sectional Study (e88498)
Ya-Lan Chiu, Pei-Yun Lee, Pooi-Ling Lim, Kai-Chia Cheng, Zong-Xian Yin, Yi-Ju Tsai. 904

Measurement Properties of the European Portuguese Michigan Hand Outcomes Questionnaire (MHQ) in Paper and Electronic Formats: Instrument Validation Study (e83563)
Eugénio Moita Gonçalves, Maria Silva, Henrique Ribeiro, Mariana Almeida, Bernardo Filipe, Sandra Branco, Filipa Henriques, Lilibeth Campos, Ana Campolargo, Pedro Cantista, Miguel Bernardes, Cristina Jácome. 918

Using Natural Language Prompts With AI Models for Low-Cost Assistive Software Design: Exploratory Comparative Evaluation (e86786)
Francesc Bañuls-Lapuerta, Vicent Marti-Miralles, Rómulo Gónzalez-García, Gabriel Martínez-Rico. 1021

User Satisfaction With Assistive, Rehabilitation, and Training Technologies: Questionnaire Development and Content Validation Study (e88182)
Oskar Cvjeti anin, Ivana Mari , Nejc Šarabon. 1270

Analysis of Health Care Providers' Attitudes and Expectations on Intradialytic Exergaming: Qualitative SWOT Study (e106283)
Aurel Zelko, Marianna Antoniadou, Jaroslav Rosenberger, Zuzana Dankulincova Veselska, Andrea Madarasova Geckova, Leonie Klompstra, Tiny Jaarsma. 1284

Appropriateness and Impact of a Vocal Cord Vibration Switch for Children with Complex Communication Needs: Case Series (e75626)
Leslie Mumford, Denise Guerriere, Tom Chau. 1302

Technology-Driven Group Exercise Program Implementation in an Underserved Community: Multimethod Retrospective Evaluation Study (e79598)
Whitney Neal, Laurie Malone. 1317

XSL•FO
RenderX

Adherence to a Digital Knee Rehabilitation Platform Among Patients With Knee Osteoarthritis and Anterior Cruciate Ligament Reconstruction in Hong Kong: Qualitative Study (e87471)
Nipuna Cooray, Stacie Powell, Mingqian Yu, Michael Ong, Julie Brown. 1331

Digital Aids for Social Media in Adults With Traumatic Brain Injury: Longitudinal Evaluation Study (e83639)
Dakota Sullivan, Nirav Patel, Lisa Kakonge, Callie Kim, Yaxin Hu, Hailey Johnson, Catalina Toma, Melissa Duff, Lyn Turkstra, Bilge Mutlu. . .
1 3 4 4

The SmartGlove BISINDO Wearable Sensor and Large Language Model–Based Sign Language–to-Speech System for Users With Communication Disorders: Design Science Research Study (e96258)
Suhartono Suhartono, Bambang Yulianto, Burhanuddin Arafah, Kusubakti Andajani, Abdul Kholiq, Muhammad Latief, Bambang Prastio, Diding Rohaedi, Deshinta Dewi, Taqdir Taqdir. 1366

Toward a Digitally Informed Knitted Prosthetic Interface With Graded Stiffness to Enhance Comfort in Transtibial Amputees: Proof-of-Concept Case Study (e91396)
Trevor Binedell, Pei Chia, Ying Tan, Hong Low. 1383

Multigesture Electromyographic Control Complexity in Upper Limb Prostheses Actuated via Single Sensor Input Contraction Magnitude: Qualitative Study for Evaluating Performance and Cognitive Load (e88528)
Abrianna Lalle, Samantha Migliore, Jeffrey Stevenson, Ethan Bell, Maanya Pradeep, Delaney Gunnell, Sophie Bennett, Viviana Rivera, Peter Smith, Matt Dombrowski, John Sparkman, Albert Manero II. 1400

Specialist Rehabilitation Providers' Experiences With an Online Self-Compassion Course: Reflexive Thematic Analysis (e81706)
Eva Cohen, Kristina Kokorelias, Dorothy Luong, McKyla McIntyre, Marina Wasilewski, Sander Hitzig, Sarah Munce, Rosalie Steinberg, Anthony Feinstein, Mark Bayley, Lawrence Robinson, Carolyn Steele Gray, Robert Simpson. 1415

Understanding the Origins and Factors of Burnout in Physical Medicine and Rehabilitation: Grounded Theory Analysis (e80499)
Robert Simpson, Eva Cohen, Stephanie Posa, Marina Wasilewski, Anthony Feinstein, Mark Bayley, Larry Robinson, Sarah Munce, Carolyn Steele Gray, Kristina Kokorelias. 1430

Barriers and Facilitators in the Implementation of Digital, Blended, and Analog Evidence-Based Osteoarthritis Care: International Qualitative Expert Interview Study (e85941)
Theresa Busse, Julia Nitsche, Ina Otte, Nino Chikhradze, Barbara Jömann, Horst Vollmar, Jennifer Bosompem. 1445

Design Recommendations for Virtual Reality–Based Upper Limb Exercises From People With Tetraplegia and Spinal Cord Injury Rehabilitation Specialists: Focus Group Study (e66832)
Andrew Goodsell, Mariel Purcell, Matthieu Poyade, Louise Cownie, Co-designers, Lorna Paul. 1457

Usability and Satisfaction of a Multidomain Information and Communication Technology–Based Rehabilitation Management Program in Patients With Postacute Stroke: Multicenter Formative Evaluation (e98098)
Won Chang, Yun-Sun Jung, Ji-Soo Choi, Won-Seok Kim, Min Sohn, Sungju Jee, Yong-Il Shin, Sung-Hwa Ko, Nam-Jong Paik. 1472

Validation of Single-Camera MediaPipe BlazePose for Knee Joint Angle Measurement Across Multiple Functional Tasks: Concurrent Validity, Interrater Reliability, and Exploratory Test–Retest Reliability Study (e102399)
Xiaotian Yue, YiXuan Gu, Bo Yu. 1487

Integrating Functional Cognitive Rehabilitation With Activity-Based Stroke Telerehabilitation: Intervention Design and Development Study (e84627)
Patricia Finetto, Michelle Woodbury, Julianne Laura, Kelly Rishe, Stephanie Garner, Lisa McTeague, Deena Schwen Blackett, Elizabeth Twamley, Stephanie Aghamoosa. 1502

XSL•FO
RenderX

Telerehabilitation Following Stroke: Development of Training Content and Evaluation of an App-Based Training Program (e77090)
Carina Ziller, Szabina Gäumann, Silya Lüscher, Nele Paulissen, Frank Behrendt, Zorica Suica, Björn Crüts, Luana Gamerschlag, Katrin Parmar, Hans Gerth, Leo Bonati, Corina Schuster-Amft. . . . 1513

Emotional Training via Telerehabilitation After Surgical Treatment for Facial Palsy: Prospective, Assessor-Blinded, 2-Arm Pilot Cohort Study (e79520)
Matteo Guidetti, Silvia Cupello, Jacopo Reali, Natale Maiorana, Sara Marceglia, Rossella Pagani, Federico Biglioli, Alberto Priori, Antonino Previtera. . . . 1530

Goal Setting During Telerehabilitation in Japan: Qualitative Study of Occupational Therapists' and Physical Therapists' Experiences Using Reflexive Thematic Analysis (e101355)
Yuho Okita, Kounosuke Tomori, Minoru Okita. . . . 1547

User Acceptance of Remote Care Assist, a Telecare System for Home Care Among Care and Nursing Staff: Cross-Sectional Pilot Study (e80514)
Friedrich Ebner, Ulrike Schneider, Cornelia Schneider, Birgit Trukeschitz. . . . 1562

Iterative Development and Usability Evaluation of an Open-Source Rehabilitation Follow-Up Platform: User-Centered Development Study (e100614)
Noora Angelva, Franziska Gurschler, Michael Single, Valerie Frischknecht, Dominik Näf, Noel Kampus, Eva Cioffi, Thimo Marcin, Branislav Savic, Matthias Wilhelm, Anke Scheel-Sailer, Anna Martin-Niedecken, Urs Mosimann, Tobias Nef. . . . 1576

A Visual Telerehabilitation Program in Virtual Reality for Age-Related Macular Degeneration: Randomized Feasibility and Proof-of-Concept Trial (e87596)
Yulia Pyatova, Bill Zhang, Mariana Misawa, Samuel Markowitz, Atri Sen, Lora Appel, Kyle Cheung, Pi Nasir, Danielle Tchao, Eduardo Garcia-Giler, Monica Daibert-Nido, Michael Reber. . . . 1594

Telerehabilitation Trends in Australian Physiotherapy and an Exploration of Factors That Influence Use After COVID-19 Restrictions: Qualitative Content Analysis (e81008)
Megan Ross, Joshua Simmich, Belinda Lawford, Kim Bennell, Rana Hinman, Trevor Russell. . . . 1613

Linguistic Validation and Cross-Cultural Adaptation of the Shoulder Telehealth Assessment Tool for Filipino Patients with Musculoskeletal Shoulder Condition: Cross-Sectional Study (e67974)
Jeffrey Arboleda, Sharon Ignacio, Jose Mojica, Carl Leochico. . . . 1628

Assessing the Role of Medical Caption Technology to Support Physician-Patient Communication for Patients With Hearing Loss: Mixed Methods Pilot Study (e79073)
Sarah Hughes, Liang-Yuan Wu, Lindsay Ma, Dhruv Jain, Michael McKee. . . . 1668

A Therapeutic Conversational Agent (Solace) for Management of Chronic Pain: Acceptability and Usability Study (e87689)
P Slepian, Stephanie Buryk-Iggers, Anna Lomanowska, Binh Nguyen, Tahir Janmohamed, Hance Clarke, Joel Katz, Nils Niederstrasser. . . 1 6 8 0

Barriers to Adoption of Electronic Low Vision Aids Among Eye Care Professionals in Jordan: Descriptive Cross-Sectional Study (e87685)
Areej Okasheh-Otoom. . . . 1705

A Smart Textile Biofeedback Training System for Upper Limb Rehabilitation After Stroke: Co-Design Development and Evaluation Study (e77999)
Maria Munoz-Novoa, Li Guo, Anna Björkquist, Morten Kristoffersen, Peiman Khorramshahi, Leif Sandsjö, Margit Alt Murphy. . . . 1724

Voice-Assisted Technology for People With Parkinson's Disease Experiencing Speech and Voice Difficulties: Co-Designing Solutions Using Design Thinking (e84364)
Jodie Mills, George Kernohan, Katy Pedlow, Orla Duffy. . . . 1739

XSL•FO
RenderX

Delivering Virtual Group Interventions for Youth With Persisting Symptoms After Concussion (Move&Connect-Youth) and Their Caregivers (Move&Connect-Caregivers): Fidelity and Engagement Evaluation (e88771)
Andrea Hickling, Hiba Al-Hakeem, Andrew Lovell, Sophie Madge-Nyren, Kylie Mallory, Jennifer Ryan, Christine Provvidenza, Ruth Wilcock, Patricia Knapp, Brenda Knapp, Shari Wade, Keith Yeates, Shannon Scratch. . . . 1758

Asynchronous Broadcasting of Audiovisual Content as a Telerehabilitation Strategy for Patients in Rural Areas: Development and Usability Study (e66343)
Lilia Aparicio Pico, Roberto Ferro Escobar, Paulo Coronado Sánchez. . . . 1782

Exploring the Influence of Digitalization on Multidisciplinary Poststroke Rehabilitation Practice: Qualitative Study (e77753)
Ann Hestetun-Mandrup, Charlotta Hamre, Anne Lund, Anne Martinsen, Hong-Gu He, Minna Pikkarainen. . . . 1798

Experiences With Technology Among Adults Aging With HIV Engaged in an Online Community–Based Exercise Intervention Study: Longitudinal Qualitative Descriptive Study and Secondary Data Analysis (e86785)
Julia Mucha, Rana Hamdy, Melina Marini, Reda Aasem, Chung Duong, Tai-Te Su, Soo Chan Carusone, Kelly O'Brien. . . . 1817

Fidelity of an Online Telecoaching Community-Based Exercise Intervention Among Adults Living With HIV in Toronto, Ontario, Canada: Observational Longitudinal Study (e84902)
Marfy Abousifein, Julia Nathanson, Tizneem Jiancaro, Ada Tang, Puja Ahluwalia, Ahmed Bayoumi, Lisa Avery, George Da Silva, Francisco Ibáñez-Carrasco, Ivan Ilic, Kiera McDuff, Mona Loutfy, Annamaria Furlan, Zoran Pandovski, Helen Trent, Soo Chan Carusone, Kelly O'Brien.
1 8 3 4

Real-World Use and Effectiveness of Internet-Based Vestibular Rehabilitation Following Nationwide Implementation: Prospective Observational Cohort Study (e106458)
Hà Ngo, Otto Maarsingh, Marco Blanker, Jettie Bont, Vincent van Vugt. . . . 1852

Reviews

Leveraging AI and Machine Vision in Telephysiotherapy for Promoting Physical Activity and Reducing Sedentary Behavior: Scoping Review (e80459)
Liva Redliha, Guna Semjonova, Liva Araka. . . . 393

Functions and Sensors of Smart Walkers From 2015 to 2024: Scoping Review (e78480)
Nicole Strutz, Hanna Brodowski, Stephan Schulze. . . . 439

Digital Technologies in Postsurgical Orthopedic Rehabilitation: Scoping Review Mapping Interventions to the CAMS Framework (2020 2025) (e100548)
Ngan Lai, Hai Huang, Tet Chen, Muhammad Nor'a, Jeffery Sijore, Wei Chen, YongSheng Luo. . . . 936

Advances in Machine Learning Paradigms for Reliable Swimming Pool Drowning Detection: Systematic Review of Methods and Implementations (e100939)
Mercy Chepkorir, Idah Orowe, Elisha Abade, Timothy Kamanu. . . . 959

Methods and Effectiveness of Auditory-Cognitive Training for Older Adults: Scoping Review and Meta-Analysis (e101632)
Mengyuan Li, Chifen Ma, Piers Dawes, Yanni Yang, Weiwei Liu. . . . 977

Impact of Emerging Technologies on Outpatient Care Quality: Mixed Methods Study (e81847)
Inas Al Khatib, Mahmoud Awad, Abdulrahim Shamayleh. . . . 996

XSL•FO
RenderX

Eye-Tracking Technologies for Cognitive Assessment After Acquired Brain Injury: Systematic Review (e81276)
Andrea Calderone, Rosaria De Luca, Francesco Corallo, Rosalia Calapai, Alessio Mirabile, Angelo Quartarone, Alessandro De Nunzio, Carmela Casella, Rocco Calabrò. . . . 1035

Patient-Perceived Factors Influencing Physical Activity Sensor Use in Stroke Prevention and Rehabilitation: Systematic Review of Qualitative Studies Using Thematic Synthesis (e86915)
Paul Harris, Ingrid Maine. . . . 1060

Use of Wearable Devices to Augment Traditional Measurements of Postoperative Outcomes Following Total Joint Arthroplasty: Systematic Review (e84671)
Robert Hall III, Alexander Hamel, Pamela Chan, Hua Zheng, David Ayers. . . . 1076

The Use of 3D Printing Technology in Rehabilitation for Adults Living With Neurological Conditions: Scoping Review (e81782)
Salena Aggerwal, Zahra Lalani, Jennifer O'Neil. . . . 1114

State-of-the-Art Knee Exoskeleton for Gait Assistance and Rehabilitation: Systematic Review (e94916)
Dilnoza Karibzhanova, Aibek Niyetkaliyev, Amna Khawaja, Aray Zhaisanbek, Zeeshan, Ardana Utepbergen, Prashant Jamwal, Kassymbek Ozhikenov, Aiman Ozhikenova, Chingiz Alimbayev, Zhadyra Alimbayeva, Zhanar Bigaliyeva, Lashin Bazarbay, Serik Didar. . . . 1128

Affective Computing in Serious Games for Physical Rehabilitation: Scoping Review (e81344)
María Beristain-Colorado, Patricia Batres-Mendoza, Erick Guerra-Hernández, José Cano-Pérez, Christian Perezcampos-Mayoral, Marciano Vargas-Treviño, Jaime Gutiérrez-Gutiérrez, Jorge Ambros-Antemate. . . . 1152

Digital Technologies for Children With Hearing Impairments to Support Language Learning: Scoping Review (e85066)
Amarria Sari, Abdullah Al Mahmud, Johanna Octavia, Shivani Tyagi. . . . 1168

Ecological Momentary Assessment of Daily Activities in Individuals With Neurological Disorders: Scoping Review (e92797)
Eva Delooz, Barbara Piškur, Aki Rintala, Bruno Bonnechère, Annemie Spooren. . . . 1187

Leisure-Time Physical Activity Interventions for Health Promotion in Individuals With Rett Syndrome: Scoping Review (e91550)
Christen Mendonca, James Rimmer, Ashley Wright, Raven Young, Amitha Ananth, Byron Lai. . . . 1208

Ankle-Foot Orthosis and Footwear Interactions in Adult Healthy and Patient Populations: Systematic Review of Biomechanical Outcomes (e93423)
Alexandra Millar, Christian Greve, Juha Hijmans. . . . 1224

Effects of Telerehabilitation on Quality of Life, Physical Capacity, and Symptom Burden Within the Post–COVID-19 Condition: Systematic Review and Meta-Analysis of Randomized Controlled Trials (e95368)
Xiaojing Huang, Yidan Tang, Dongsen Liu, Ryan Burns, Wonwoo Byun. . . . 1642

Viewpoint

Evaluating the Feasibility of Technology-Based Interventions in Disability and Rehabilitation: Definitions, Considerations, and Dimensions (e79026)
Josh Shore, Amy Hwang, Rosalie Wang. . . . 1246

XSL•FO
RenderX

Corrigenda and Addendas

Correction: Smartphone App–Delivered Mindfulness-Based Intervention for Concussion in Adolescents (MBI-4-mTBI): Feasibility Randomized Controlled Trial (e110902)
Veronik Sicard, Molly Cairncross, Roger Zemek, Noah Silverberg, Gary Goldfield, Nick Reed, Bechara Saab, Andra Smith, Andrée-Anne Ledoux. . . . 1697

Addendum: Managing Musculoskeletal Pain in Older Adults Through a Digital Care Solution: Secondary Analysis of a Prospective Clinical Study (e97495)
Anabela Areias, Dora Janela, Maria Molinos, Robert Moulder, Virgílio Bento, Vijay Yanamadala, Steven Cohen, Fernando Correia, Fabíola Costa. . . . 1699

Letters to the Editor

Crossover Analysis Is Needed in Smartphone-Supported Vestibular Rehabilitation (e107851)
Jin Yan, Kaiyuan Yang, Xixin Wang, Boyan Su. . . . 1701

Clarifying Methodological Principles and Preliminary Trends in Digital Vestibular Rehabilitation: Authors' Reply (e109396)
Azriel Kaplan, Liran Kalderon, Amit Wolfovitz, Yoav Gimmon, Shelly Levy-Tzedek. . . . 1703

Research Letters

Can GPT-5 Support Licensing Examination Preparation? Analysis of Accuracy, Reasoning, and Semantic Similarity Across Rehabilitation Disciplines (e91019)
Christy Muasher-Kerwin, M Hughes, Aida Sanatizadeh. . . . 1714

Remote Therapeutic Monitoring Use Among Commercially Insured and Medicare Advantage Patients (e103600)
Joseph Joo, Jessica Billig, Nadia Lieu, Lingmei Zhou, Joshua Liao. . . . 1718

XSL•FO
RenderX

Validity and Reliability of the Track-UL Algorithm Compared With Kinovea Software for Measuring Upper-Limb Functional Range of Motion in People After Stroke: Cross-Sectional Observational Study

Hatem Lazem[1,2,3], MSc; David Harris[4], PhD; Abi Hall[1], PhD; Thomas C Richards[1,3], PhD; Phaedra Leveridge[3,4], BSc; Maedeh Mansoubi[1,3], PhD; Xiaohan Xu[4], PhD; Paul Newell[1,3], BSc, MSc; Sarah E Lamb[1,3], MA, MSc, DPhil; Helen Dawes[1,3], PhD

[1]Medical School, Faculty of Health and Life Sciences, University of Exeter, St Luke's Campus, Exeter, United Kingdom

[2]Basic Science Department, Faculty of Physical Therapy, Cairo University, Giza, Egypt

[3]Faculty of Health and Life Sciences, NIHR Exeter Biomedical Research Centre, University of Exeter, Exeter, United Kingdom

[4]School of Public Health and Sport Science, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom

Corresponding Author:
Hatem Lazem, MSc
Medical School, Faculty of Health and Life Sciences, University of Exeter, St Luke's Campus, Exeter, United Kingdom

Abstract

Background: Approximately 70% of survivors of stroke have problems with arm function. Physiotherapists assess arm functional range of motion (ROM) using either a goniometer or functional questionnaires, which lack objective accuracy and require a skilled physiotherapist. We developed the Track-UL algorithm based on a markerless motion capture system to measure arm ROM.

Objective: This study aimed to measure the agreement between our novel Track-UL algorithm and Kinovea software in assessing arm ROM during functional tasks in the laboratory and home settings.

Methods: Videos were recorded while 27 survivors of chronic stroke performed 4 functional tasks (forward reaching, arm abduction, moving the hand toward the mouth, and moving the hand toward the head) in the laboratory and at home. The videos were analyzed by 2 independent raters using the Track-UL algorithm and Kinovea software. The limits of agreement and intraclass correlation coefficients were calculated.

Results: We found no clinically significant systematic bias in shoulder and elbow angle, with good agreement between the Track-UL algorithm and Kinovea software (assessed via Bland-Altman plots). The 95% limits of agreement were –3.18 to 6.41 degrees for the shoulder joint and −5.35 to 8.78 degrees for the elbow joint in the laboratory setting, and –6.21 to 3.62 degrees for the shoulder joint and −4.06 to 2.53 degrees for the elbow joint in the home setting. There was excellent absolute agreement between the measurement tools across all tasks and joints; intraclass correlation coefficient values ranged from 0.97 (95% CI 0.97-0.99) to 0.99 (95% CI 0.99-0.99; P<.001 for both laboratory and home measurements).

Conclusions: The novel Track-UL algorithm is an accurate, valid, and easy tool that can be used to assess upper-limb ROM in survivors of stroke at clinics and potentially at home. This will support physiotherapists in remotely monitoring and adapting rehabilitation programs.

(JMIR Rehabil Assist Technol 2026;13:e87128) doi:10.2196/87128

KEYWORDS

stroke; telemonitoring; range of motion; MediaPipe; markerless tracking; validity; reliability

Introduction

Stroke is the second leading cause of death and the third leading cause of long-term disability for approximately 70% of people who experience a stroke, affecting their quality of life [1]. Approximately 40% of people with stroke experience upper-limb problems at the chronic stage of recovery [2]. Following a stroke, people can experience muscle weakness [3], spasticity [4], and changes in the pattern of arm mobility on the affected side [5].

According to the National Institute for Health and Care Excellence guidelines, physiotherapists routinely design rehabilitation programs to improve upper-limb function and, therefore, need to efficiently and effectively assess and monitor

XSL•FO
RenderX

improvements in the functional range of motion (ROM) in people with stroke [6]. In clinical settings, physiotherapists typically use a clinical goniometer [7] or validated clinical questionnaires such as the Fugl-Meyer Assessment for the upper extremity (FMA-UE) [8] to evaluate and monitor the upper-limb ROM in people with stroke.

A goniometer is the current standard tool for assessing ROM in physiotherapy practice [9]. However, this requires an adequate level of expertise from the practitioner to execute, so its precision can be influenced by clinical skills and experience [10,11]. In addition, the presence of compensation during upper-limb movements, such as shoulder and trunk rotation, and the disruption of elbow-shoulder coordination can affect the ease with which goniometers can be used to measure ROM [12]. A limitation of upper-limb clinical assessment questionnaires such as the FMA-UE scale is the time required to complete them, which can be up to 25 minutes. They also rely on a Likert-based rating system in which scores of 0, 1, and 2 indicate no movement, partial movement, and full movement, respectively, which is discrete and bounded, depends on the raters' experience, and lacks the precision in quantifying progression and monitoring small changes in arm mobility over time [8].

An alternative approach is the use of affordable and free 2D markerless motion analysis tools such as the Kinovea software that allow therapists to measure the upper-limb kinematics from video footage [13]. However, such tools are better suited to research as the time and expertise of an experienced health care professional required for operation make them less accessible to physiotherapists in clinics. Computer vision algorithms based on markerless motion capture and pose estimation frameworks offer a potential solution for faster and easier measurement of the ROM without the need for an experienced health care professional [14]. So far, computer vision–based algorithms have been validated for people with ankylosing spondylitis [15] and following knee replacement [16] but are still in need of validation among people with stroke, who have different movement complexities such as compensatory movement patterns and reduced ROM, which differ from the normative datasets on which most algorithms are trained, which can introduce the possibility of tracking errors [12,17,18].

Therefore, based on the gap in the current literature on the validity and reliability of the pose estimation algorithms in stroke, the aim of this study was to assess the concurrent validity and test-retest reliability of a novel Track-UL algorithm based on MediaPipe, a pose estimation framework, as an objective and easy-to-use tool for measuring and monitoring upper-limb mobility in people with stroke in the laboratory and home settings.

The main objectives of this study were as follows:

- To investigate the concurrent validity of the Track-UL algorithm based on a 2D pose estimation framework compared to the Kinovea software for measuring shoulder and elbow joint functional active ROM
- To investigate the test-retest reliability of the Track-UL algorithm based on a 2D pose estimation framework for measuring shoulder and elbow joint functional active ROM
- To explore the potential for using the Track-UL algorithm as a telemonitoring tool for measuring upper-limb functional active ROM from videos recorded by people with stroke in their own homes

Methods

Ethical Considerations

This study was conducted in accordance with the ethical standards of the responsible institutional and national committees on human experimentation and with the World Medical Association Declaration of Helsinki (1975, as revised in 2013). This project was reviewed and approved by an independent National Health Service research ethics committee and the Health Research Authority (reference 24/LO/0434). Written informed consent was obtained from all participants prior to their inclusion in the study. The privacy and confidentiality of participants' data and identities were strictly maintained. Personally identifiable information was stored securely and separately from research data, with access limited to authorized members of the research team. Participants received compensation for travel expenses associated with study participation.

Participants

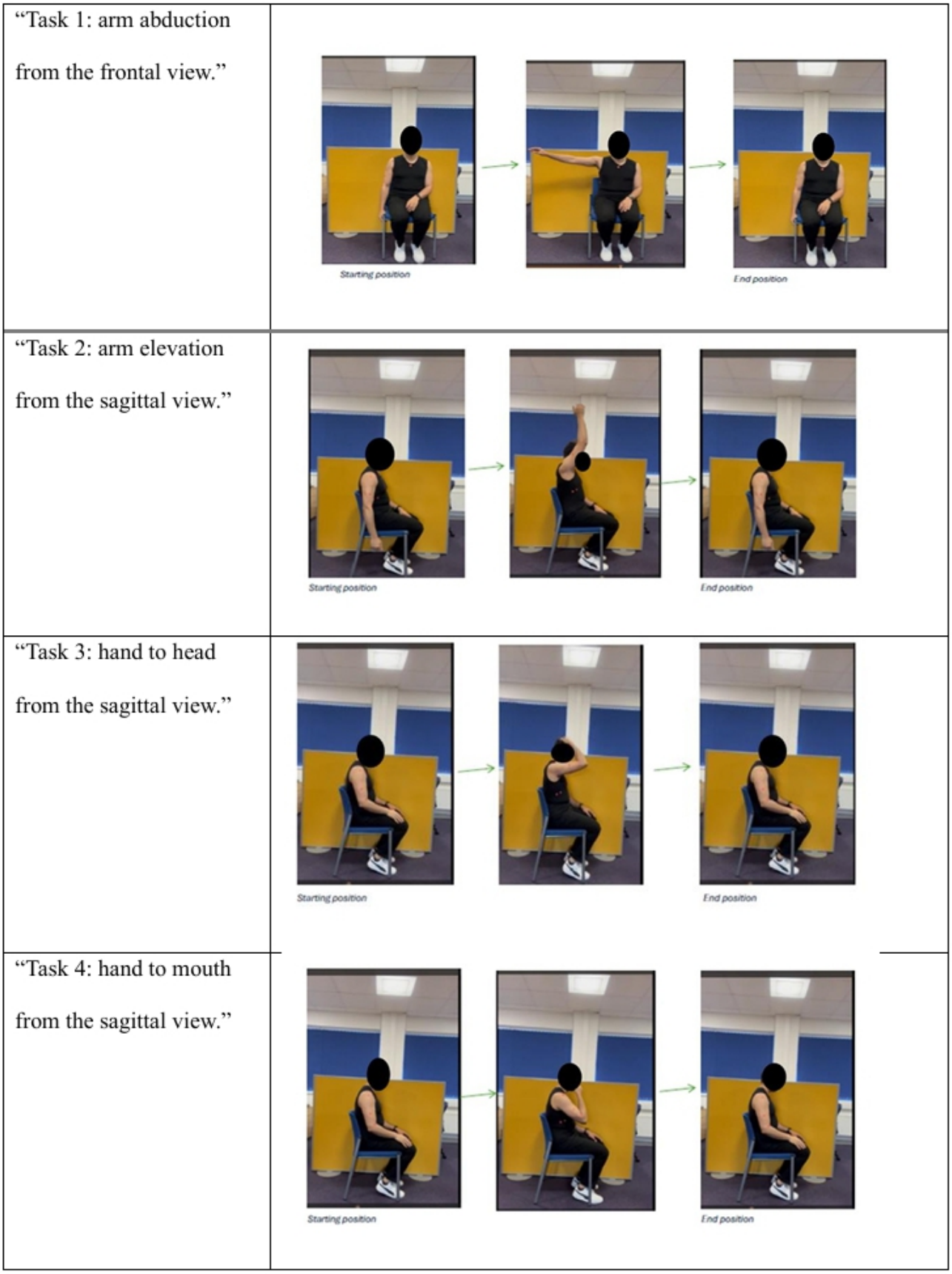
This was a cross-sectional observational study in which a total of 27 people with stroke participated. Data collection was conducted in the laboratory by the researcher (a physiotherapist) at the University of Exeter and participants' homes. We recruited people with stroke aged 18 years or older who were more than 6 months after stroke with a score of 2 or lower on the Modified Ashworth Scale and any degree of upper-limb impairment resulting from stroke (FMA-UE score<57). Survivors of stroke were excluded if they had severe cognitive deficits and could not follow the instructions given by the researcher or had severe spasticity (score of >2 on the Modified Ashworth Scale). The reason for these exclusions was that these patient groups are not able to achieve active movements, which are a requirement of the tasks this study aimed to validate.

Procedure

All participants completed a single face-to-face laboratory testing session at the University of Exeter with a physiotherapist, which lasted up to 60 minutes, including collection of demographic data and assessment of upper-limb ROM, with videos recorded while participants performed 4 functional upper-limb tasks (Figure 1). Participants were then asked to repeat the same 4 tasks alone at home with or without support from their carers following written instructions provided by the research team while recording videos of the tasks within 3 days following the initial assessment.

XSL•FO
RenderX

Figure 1. The 4 functional complex upper-limb tasks.



To familiarize participants with the tasks, each movement was demonstrated to them, who then practiced it twice before starting. Participants were asked to wear light clothes with sleeveless shirts to aid identification of bony landmarks on the upper limb. Before recording the videos, colored circular stickers (markers) were applied to predefined bony landmarks of the upper limb (acromion process, anterior and midthoracic line, lateral epicondyle, midshaft of the humerus, midforearm, and

XSL•FO
RenderX

midwrist). The use of these markers enabled measurement of shoulder and elbow range of movements during the 4 tasks (Figure 1). For the videos recorded at home by the participants or their carers, we used virtual markers implemented in the Kinovea software.

All videos were imported into a laptop and analyzed using both the Kinovea software and PyCharm Community Edition (version 2022.3.1; JetBrains) using the Track-UL algorithm, a function programmed in PyCharm. This function calculates the angle of joints of the human body based on the joint coordinates using the 2D vector dot product formula. All videos were recorded by 1 physiotherapist. Two physiotherapists (one of whom recorded the videos) who were blinded to each other's measurements then independently analyzed the video recordings: one using the Kinovea software and the other using the Track-UL algorithm.

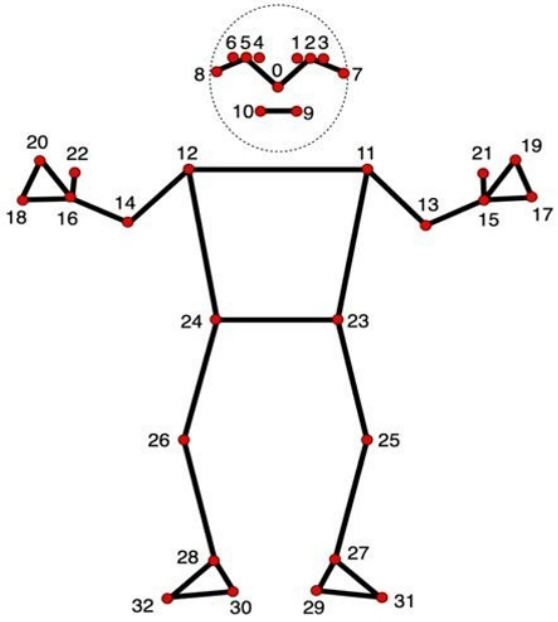
For home testing, we asked the participants themselves or their carers to record the videos only for the affected side within 3 days of the laboratory test and then send the videos back to the researcher.

Instruments

Kinovea is free and open-source analysis software used by health care professionals in research and clinical settings. It is considered to be a valid and reliable tool for measuring ROM using a digital goniometer, which is widely used in clinical and sports biomechanics for 2D kinematic analysis [19,20]. While the gold standard for measuring ROM is the 3D motion capture system, using Kinovea allows for a more applicable comparison aligned with the intended use of this algorithm in real-world environments such as a clinic or the home setting, in which stroke assessment and telerehabilitation usually occur. Analysis of the captured videos was carried out using Kinovea (version 0.9.5; Joan Charmant). Circular colored markers were placed on the bony landmarks of the upper limb prior to capture to facilitate calculation of upper-limb joint angles using a digital goniometer.

MediaPipe is an open-source pose estimation model developed by Google for machine learning applications [21] (Figure 2 [22]). The model infers 33 pose landmarks from the whole body; offers real-time processing and compatibility with standard video input from commonly available devices such as mobile phones; and offers a practical balance between accuracy and usability [23] compared to other pose estimation software such as OpenPose [24] and OpenCap [25] that require high computational resources and more complex setup, which can limit their applicability in a clinical setting or at home. The Track-UL algorithm was developed using PyCharm for automatic calculation of both shoulder and elbow joint angles from videos, which are then presented to the user. To calculate joint angles, coordinates for bony landmarks (hip, shoulder, elbow, and wrist) were extracted from video frames using the MediaPipe pose estimation framework. Each landmark provides 2D coordinates (x, y). Shoulder and elbow joint angles were computed by treating the connected rigid body segments (trunk, upper arm, and forearm) as 2D vectors (Figures 2 and 3 [22]).

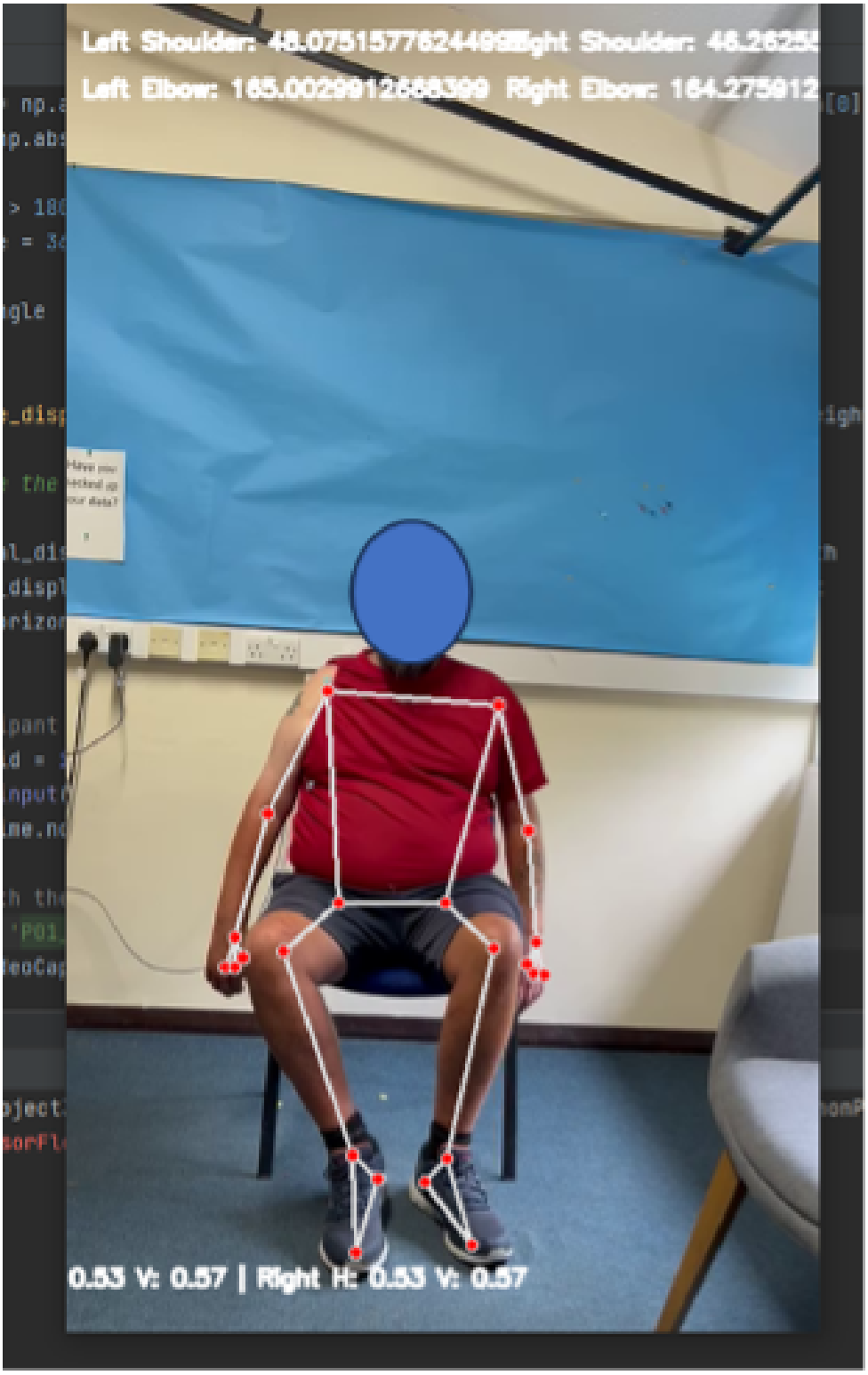
Figure 2. BlazePose model of 33 human poses [22].



0. Nose	17. Left pinky
1. Inner left eye	18. Right pinky
2. Left eye	19. Left index
3. Outer left eye	20. Right index
4. Inner right eye	21. Left thumb
5. Right eye	22. Right thumb
6. Outer right eye	23. Left hip
7. Left ear	24. Right hip
8. Right ear	25. Left knee
9. Left mouth	26. Right knee
10. Right mouth	27. Left ankle
11. Left shoulder	28. Right ankle
12. Right shoulder	29. Left heel
13. Left elbow	30. Right heel
14. Right elbow	31. Left foot index
15. Left wrist	32. Right foot index
16. Right wrist	

XSL•FO
RenderX

Figure 3. How the video appears to the physiotherapist when we apply the Track-UL algorithm.



A 2D vector dot product formula was used to calculate shoulder joint angle (term B in the following formulas) using coordinates of the shoulder joint and the adjacent hip (term A) and elbow (term C) joints:

$$\theta = \arccos A - B \cdot C - B A - B C - B$$

$$\theta = \arccos\left(\frac{(A_x - B_x)(C_x - B_x) + (A_y - B_y)(C_y - B_y)}{\sqrt{(A_x - B_x)^2 + (A_y - B_y)^2}\sqrt{(C_x - B_x)^2 + (C_y - B_y)^2}}\right)$$

Video Processing

The researcher recorded 8 videos of 4 different upper-limb functional tasks performed on each side (affected and unaffected side) using a mobile phone camera (iPhone 13 Pro Max [Apple Inc] with a 12-MP main camera with sensor-shift stabilization, 1.9-µm pixels, and a 26-mm equivalent f/1.5-aperture lens) and a tripod. The camera was placed 1.5 m from the participants on a tripod at a height of 90 cm oriented with an angle of 90 degrees perpendicular to the chair. Participants were then asked to record

XSL•FO
RenderX

the same videos at home following standardized instructions (Multimedia Appendix 1) within the following 3 days and only for the affected side to lower the risk of fatigue from repeated movements. Each task was repeated 3 times independently in each video with 10 seconds of rest in between each repetition and one 1-minute rest in between each task. During analysis, there were no apparent challenges regarding landmark detection by the Track-UL algorithm.

This study measured the functional active ROM for both the shoulder joint (flexion and abduction) and the elbow joint (flexion and extension) during 4 different functional complex tasks, which are routinely used in the clinic as part of the FMA-UE assessment. We measured both the affected and unaffected sides to understand the extent to which the Track-UL algorithm can estimate the functional active ROM accurately on the unaffected side as well as the affected side, which can include some substitute movements due to the motor impairments compared to the unaffected side.

Analysis

We used SPSS Statistics (version 29; IBM Corp) for data analysis. To assess the Track-UL algorithm's concurrent validity against Kinovea, absolute agreement was evaluated using Bland-Altman analysis [26] to obtain the 95% limit of agreement (LoA) and mean bias metrics using the following equation: LoA = mean difference ± 1.96 × SD difference.

To measure the absolute agreement between our algorithm and Kinovea, we calculated the intraclass correlation coefficient (ICC) based on a 2-way random-effect model for a single measurement with an absolute agreement model (ICC(2,1)). This model accounts for both participant variability and systematic differences between the 2 assessment tools.

To assess the test-retest reliability for the Track-UL algorithm, we also calculated the ICC(2,1) for the repeated measures, the SE of measurement, and the minimum detectable change using the following equations [27,28]:

$$SEM = StDev \cdot \sqrt{1 - ICC}$$

$$MDC = SEM \cdot \sqrt{2} \times 1.96$$

ICC values between 0.81 and 1.0 are interpreted as very good or excellent, ICC values between 0.61 and 0.80 are interpreted as good, ICC values between 0.41 and 0.60 are interpreted as moderate, ICC values between 0.21 and 0.40 are interpreted as fair, and ICC values below 0.20 are interpreted as poor [29].

Results

Overview

The demographic data of the participants are listed in Table 1.

Table . Participants' demographic data (N=27).

Characteristics	Participants
Age (years), mean (SD; range)	60.56 (13.82; 26-88)
Time since stroke (months), mean (SD; range)	42.93 (33.07; 6-126)
FMA-UE[a] score, mean (SD; range)	32.15 (16.21; 10-55)
MAS[b] score—shoulder, mean (SD; range)	1.22 (0.42; 1-2)
MAS score—elbow, mean (SD; range)	1.44 (0.51; 1-2)
MAS score—hand, mean (SD; range)	1.41 (0.50; 1-2)
Sex, n (%)	
Male	21 (77.8)
Female	6 (22.2)
Type of stroke, n (%)	
Ischemic	19 (70.4)
Hemorrhagic	8 (29.6)
Affected side, n (%)	
Right	16 (59.3)
Left	11 (40.7)

[a]FMA-UE: Fugl-Meyer Assessment for the upper extremity.

[b]MAS: Modified Ashworth Scale.

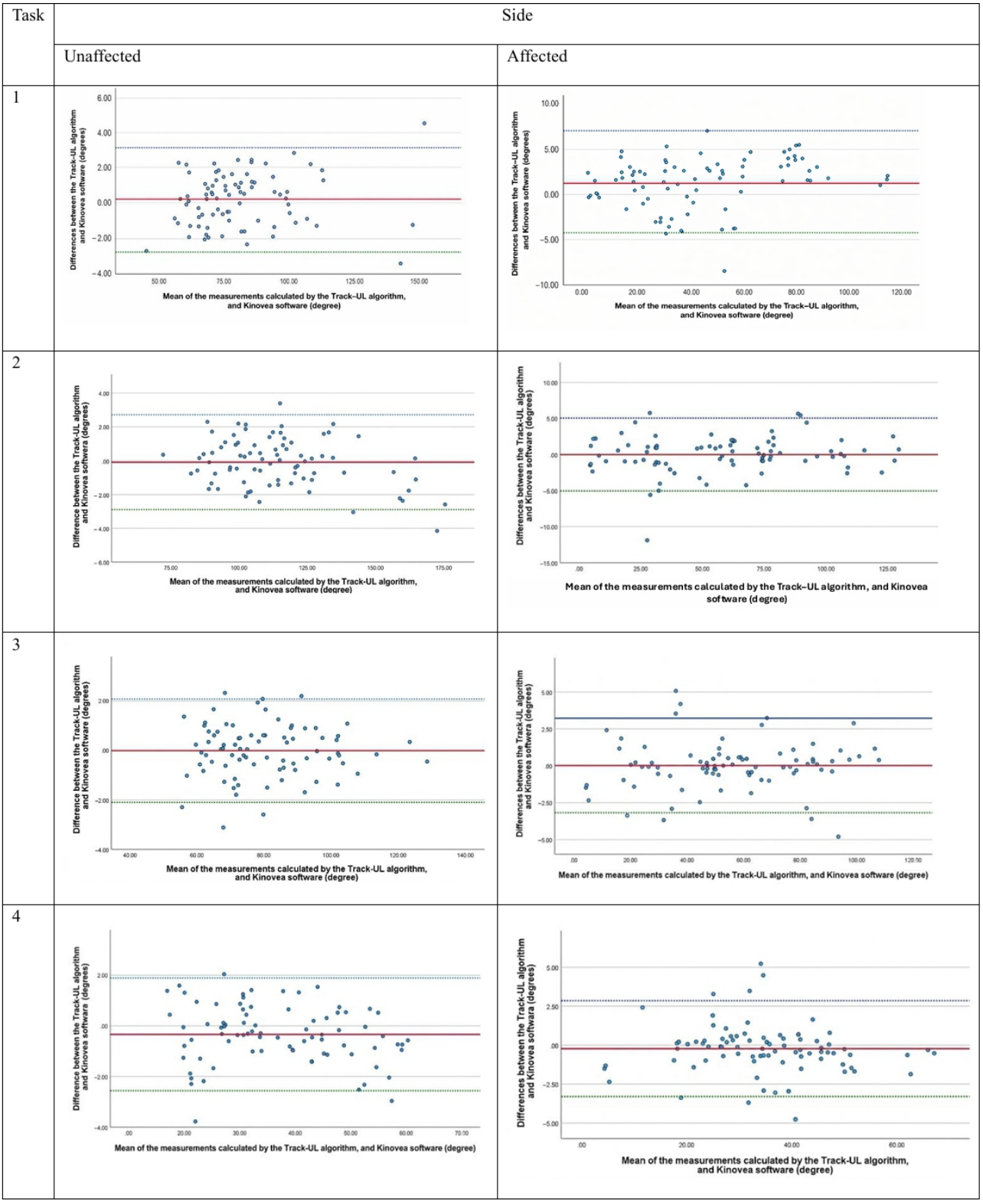
Concurrent Validity for Laboratory Measurement

The Bland-Altman plots (Figures 4 and 5) reveal that the mean difference in shoulder and elbow joint angles between the Track-UL algorithm and Kinovea was close to 0, with a mean value ranging from −0.34 (SD 1.13) to 0.76 (SD 1.25) degrees for the unaffected side and −1.24 (SD 3.84) to 1.17 (SD 2.67) degrees for the affected side, indicating no systematic bias. The 95% LoA for joint measurements on the unaffected side ranged from −2.88 to 3.03 degrees for the shoulder and from −3.41 to 3.21 degrees for the elbow (Table 2). On the affected side, the 95% LoA ranged from −3.18 to 6.41 degrees for the shoulder

XSL•FO
RenderX

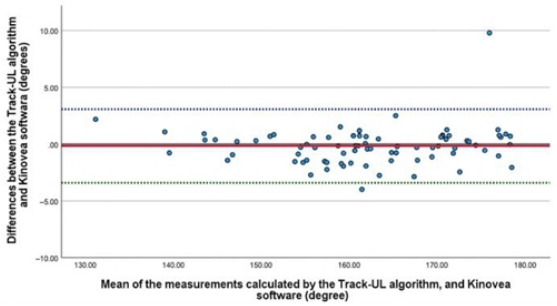
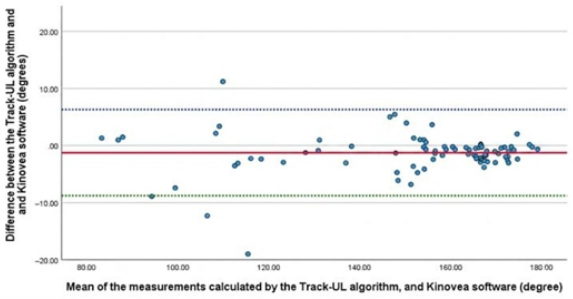
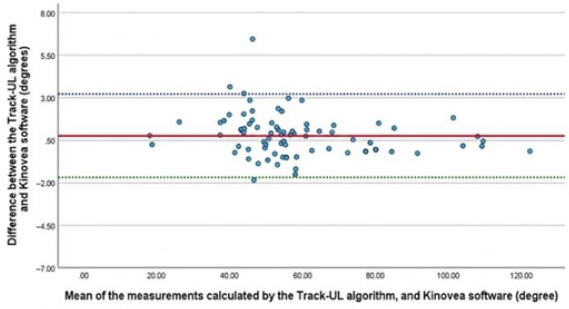
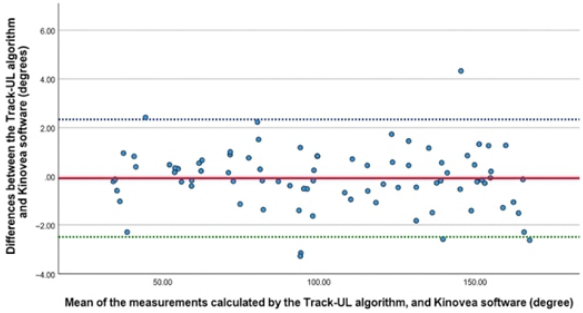
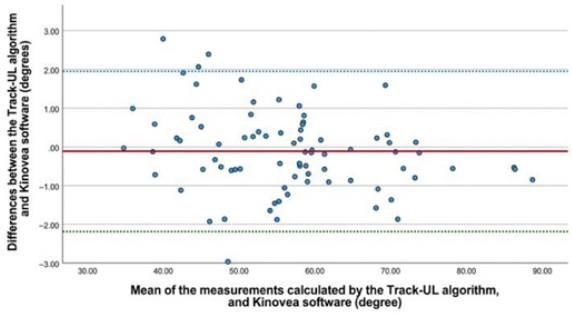
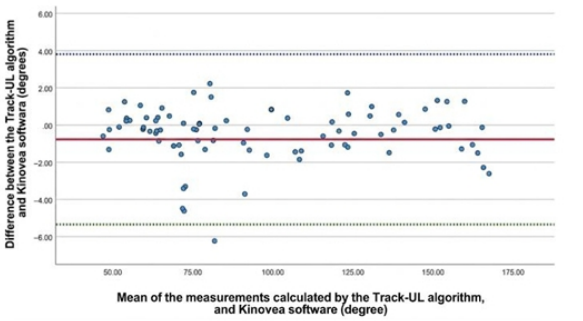
and −5.35 to 6.29 degrees for the elbow (Table 2). This difference is not considered clinically significant [30]. There are a few outliers in some of the graphs; however, there is no clear evidence of a trend or heteroscedasticity in the plots except for the shoulder joint during shoulder abduction (task 1) on the affected side. In this case, a proportional bias was observed: as the abduction angle increased, the positive difference also increased. However, this positive difference of up to 6.41 degrees does not have a clinically significant impact on the measurement of the shoulder functional active ROM [30].

Figure 4. Bland-Altman plots for comparing the Track-UL algorithm and Kinovea software for the shoulder joint in both the unaffected and affected side in the laboratory setting.



XSL•FO
RenderX

Figure 5. Bland-Altman plots for comparing the Track-UL algorithm and Kinovea software for the elbow joint in both the unaffected and affected side in the laboratory setting.

Task	Side	
	Unaffected	Affected
2		
3		
4		

XSL•FO
RenderX

Table . Mean values and limit of agreement (LoA) between the Track-UL algorithm and the Kinovea software based on the laboratory measurements.

Tested side and task number	Elbow joint				Shoulder joint			
	Observations, n	Values, mean (SD; 95% CI)	Lower LoA	Upper LoA	Observations, n	Values, mean (SD; 95% CI)	Lower LoA	Upper LoA
Unaffected side								
1	—[a]	—	—	—	81	0.21 (1.45; –0.12 to 0.52)	−2.64	3.03
2	81	−0.16 (1.65; –0.52 to 0.21)	−3.41	3.08	80	−0.09 (1.43; –0.40 to 0.23)	−2.88	2.71
3	81	0.76 (1.25; 0.48 to 1.03)	−1.69	3.21	81	−0.02 (1.06; –0.26 to 0.21)	−2.10	2.06
4	81	−0.11 (1.05; –0.34 to 0.11)	−2.19	1.95	81	−0.34 (1.13; –0.59 to –0.09)	−2.56	1.87
Affected side								
1	84	—	—	—	84	1.17 (2.67; 0.59 to 1.75)	−4.06	6.41
2	84	−1.24 (3.84; –2.07 to –0.40)	−8.78	6.29	84	−0.01 (2.58; –0.57 to 0.55)	−5.07	5.05
3	84	−0.07 (1.23; –0.34 to 0.18)	−2.49	2.33	84	0.02 (1.63; –0.34 to 0.37)	−3.18	3.22
4	84	−0.77 (2.33; –1.27 to –0.26)	−5.35	3.81	84	−0.23 (1.56; –0.57 to 0.11)	−3.31	2.84

[a]Not applicable.

Absolute Agreement Analysis for Laboratory Measurement

We used the ICC to assess the absolute agreement between the Track-UL algorithm and the Kinovea software when measuring ROM from videos recorded in the laboratory. The ICC(2,1) for the absolute agreement was calculated, showing excellent agreement across all tasks, joints, and sides (unaffected and affected). The ICC values ranged from 0.98 to 0.99 for the unaffected side and from 0.98 to 0.99 for the affected side (Table 3). This indicates that measurements were very consistent between the 2 assessment tools.

XSL•FO
RenderX

Table . Intraclass correlation between the Track-UL algorithm and the Kinovea software (laboratory setting).

Side and task number	Joint	Intraclass correlation coefficient (95% CI)	*F* test (*df*)	*P* value
Unaffected side				
1	Shoulder	0.99 (0.99-0.99)	714.557 (80, 80)	<.001
2	Shoulder	0.98 (0.98-0.99)	156.212 (80, 80)	<.001
2	Elbow	0.98 (0.98-0.99)	156.212 (80, 80)	<.001
3	Shoulder	0.99 (0.99-0.99)	879.210 (80, 80)	<.001
3	Elbow	0.99 (0.99-0.99)	1027.538 (80, 80)	<.001
4	Shoulder	0.99 (0.99-0.99)	473.902 (80, 80)	<.001
4	Elbow	0.99 (0.99-0.99)	481.433 (80, 80)	<.001
Affected side				
1	Shoulder	0.99 (0.99-0.99)	455.606 (83, 83)	<.001
2	Shoulder	0.99 (0.99-0.99)	677.202 (83, 83)	<.001
2	Elbow	0.98 (0.98-0.99)	164.610 (83, 83)	<.001
3	Shoulder	0.99 (0.99-0.99)	1001.732 (83, 83)	<.001
3	Elbow	0.99 (0.98-0.99)	216.907 (83, 83)	<.001
4	Shoulder	0.99 (0.98-0.99)	260.500 (83, 83)	<.001
4	Elbow	0.99 (0.99-0.99)	959.581 (83, 83)	<.001

Track-UL Algorithm Test-Retest Reliability for Laboratory Measurement

The Track-UL algorithm reliability results for both the shoulder and elbow joints on both sides are shown in Table 4. The ICC showed that the Track-UL algorithm had excellent reliability (all ICC values>0.81).

Table . Track-UL algorithm test-retest reliability for the laboratory measurements (N=27).

Tested side and task number	Joint	ICC[a]	SEM[b]	MDC[c]
Affected side				
2	Elbow	0.95	5.32	14.77
3	Elbow	0.91	11.78	32.65
4	Elbow	0.95	7.41	20.55
1	Shoulder	0.97	4.11	11.41
2	Shoulder	0.98	4.75	13.19
3	Shoulder	0.93	6.56	18.19
4	Shoulder	0.88	4.20	11.65
Unaffected side				
2	Elbow	0.84	3.83	10.62
3	Elbow	0.91	5.82	16.15
4	Elbow	0.93	2.90	8.04
1	Shoulder	0.92	5.19	14.39
2	Shoulder	0.89	7.20	19.96
3	Shoulder	0.87	5.43	15.07
4	Shoulder	0.97	1.87	5.20

[a]ICC: intraclass correlation coefficient.

[b]SEM: SE of measurement.

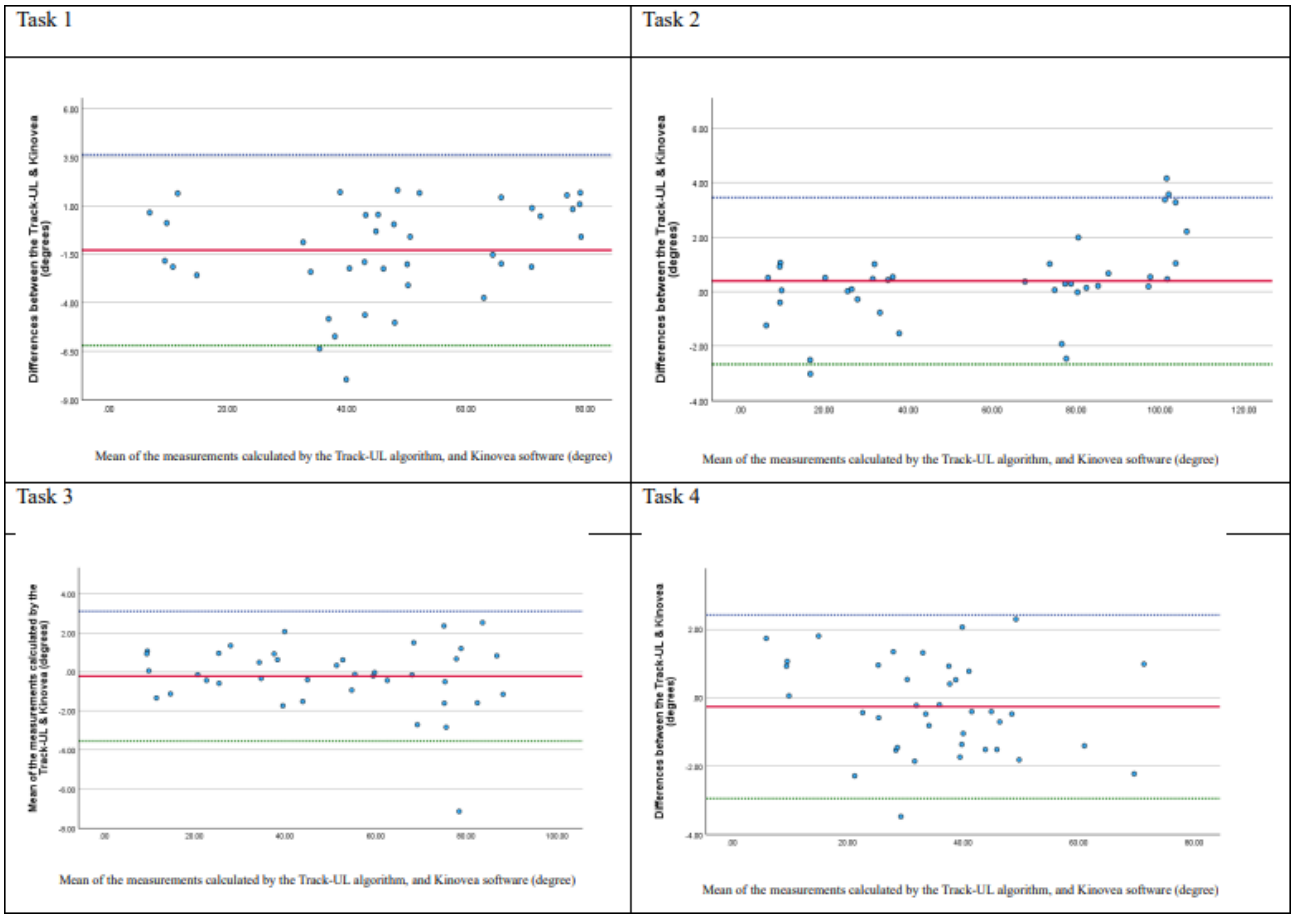
[c]MDC: minimum detectable change.

XSL•FO
RenderX

Concurrent Validity for Home Measurement

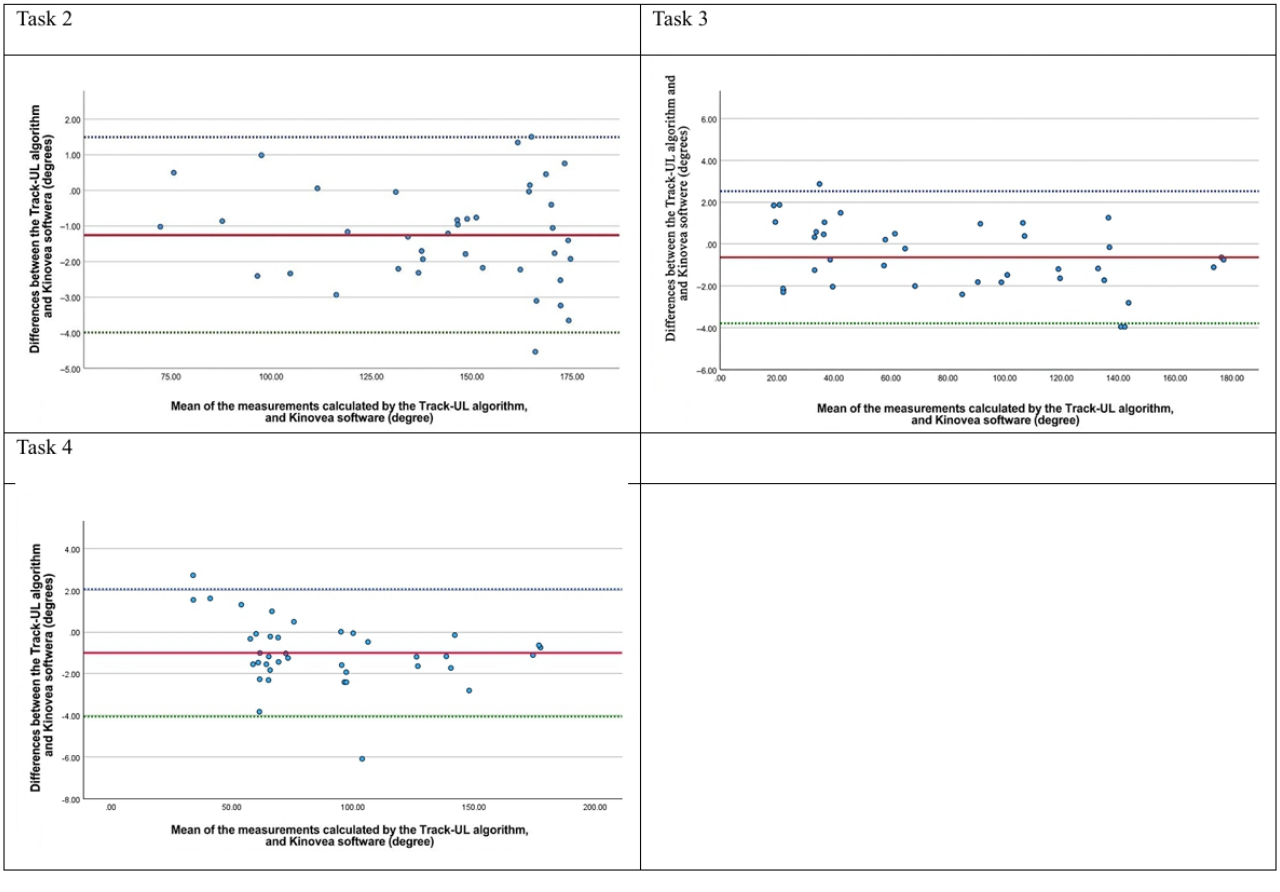
Only 48.1% (13/27) of the participants were able to record videos at home that were fit for analysis. For the acceptable videos, the Bland-Altman plots (Figures 6 and 7) indicated that the mean difference in the shoulder and elbow joint angles between the Track-UL algorithm and Kinovea software was close to 0; the mean values ranged from −1.29 to 0.39 degrees, indicating no systematic bias. The 95% LoA ranged from −6.21 to 3.62 degrees for shoulder joint measurements and from −4.06 to 2.53 degrees for elbow joint measurements. This difference is not considered clinically significant [30] (Table 5).

Figure 6. Bland-Altman plots for comparing the Track-UL algorithm and Kinovea software for the shoulder joint in the affected side in the home setting.



XSL•FO
RenderX

Figure 7. Bland-Altman plots for comparing the Track-UL algorithm and Kinovea software for the elbow joint in the affected side in the home setting.



There was no clear evidence of a trend or heteroscedasticity on the plots except for the shoulder joint during shoulder elevation (task 2), which displayed a proportional bias whereby the positive difference increased at higher angles of elevation and the negative difference increased at lower angles. However, the positive difference of up to 3.45 degrees and the negative difference of up to −2.66 degrees had no substantial clinical influence on measuring shoulder functional active ROM [30] (Table 5).

Table . The limit of agreement (LoA) between the Track-UL algorithm and the Kinovea software from videos recorded at home.

Tested side and task number	Elbow joint				Shoulder joint			
	Observations, n	Mean (SD; 95% CI)	Lower LoA	Upper LoA	Observations, n	Mean (SD; 95% CI)	Lower LoA	Upper LoA
Affected side								
1	—[a]	—	—	—	39	−1.29 (2.51; −2.11 to −0.47)	−6.21	3.62
2	39	−1.25 (1.39; −1.71 to −0.81)	−3.99	1.47	39	0.39 (1.56; −0.11 to 0.89)	−2.66	3.45
3	39	−.63 (1.61; −1.51 to −0.11)	−3.79	2.53	39	−0.22 (1.69; −0.77 to 0.32)	−3.54	3.09
4	39	−1.01 (1.56; −1.51 to −0.49)	−4.06	2.05	39	−0.27 (1.36; −0.71 to 0.17)	−2.95	2.41

[a]Not applicable.

XSL•FO
RenderX

Absolute Agreement Analysis for Home Measurement

The absolute agreement between the 2 measurement tools (Track-UL algorithm and Kinovea software) for videos recorded at home was assessed using the ICC. The results demonstrated excellent reliability across all evaluated tasks and joints. ICC values ranged from 0.97 to 0.99, indicating a high degree of consistency between the 2 assessment tools (Table 6).

Table . Intraclass correlation between the Track-UL algorithm and the Kinovea software (home setting).

Side and task number	Joint	Intraclass correlation coefficient (95% CI)	*F* test (*df*)	*P* value
Affected side				
1	Shoulder	0.99 (0.97-0.99)	284.85 (38, 38)	<.001
2	Shoulder	0.99 (0.99-0.99)	2091.99 (38, 38)	<.001
2	Elbow	0.99 (0.98-0.99)	1772.51 (38, 38)	<.001
3	Shoulder	0.97 (0.95-0.98)	89.19 (38, 38)	<.001
3	Elbow	0.99 (0.98-0.99)	260.63 (38, 38)	<.001
4	Shoulder	0.99 (0.99-0.99)	467.71 (38, 38)	<.001
4	Elbow	0.99 (0.99-1.00)	2508.41 (38, 38)	<.001

Track-UL Algorithm Test-Retest Reliability for Home Measurement

To assess the test-retest reliability of the Track-UL algorithm from the videos recorded at home, we collected 3 independent measures for each task. The ICC showed that the Track-UL algorithm had good to excellent reliability results for both the shoulder and elbow joints on the affected side (all ICC values≥0.778; Table 7).

Table . Track-UL algorithm test-retest reliability for the home measurement (n=13).

Tested side and task number	Joint	ICC[a]	SEM[b]	MDC[c]
Affected side				
2	Elbow	0.77	13.07	36.24
3	Elbow	0.98	5.94	16.48
4	Elbow	0.94	9.12	25.28
1	Shoulder	0.97	3.23	8.95
2	Shoulder	0.99	2.29	6.35
3	Shoulder	0.96	4.52	12.54
4	Shoulder	0.91	4.30	11.92

[a]ICC: intraclass correlation coefficient.

[b]SEM: SE of measurement.

[c]MDC: minimum detectable change.

Discussion

Principal Findings

The main aims of this study were to measure the concurrent validity, absolute agreement, and test-retest reliability of the newly developed Track-UL algorithm. This algorithm uses a 2D markerless motion capture system based on MediaPipe, a pose estimation framework, to measure shoulder and elbow functional active ROM from recorded videos in people with stroke. While most algorithms that use MediaPipe pose estimation models have been tested and implemented in healthy populations for a variety of applications [31-33], such as assessment for spinal diseases and frozen shoulder [34], among people with tremors [35], and for telerehabilitation purposes [36,37], they have not yet been validated in individuals with stroke using complex functional tasks. The use of computer vision–based assessment of functional movement in people with stroke has the potential to improve the accessibility and scalability of movement assessment by providing accurate, low-cost, and markerless measurement outside of specialist laboratory settings. Therefore, we assessed shoulder and elbow functional active ROM during 4 complex functional tasks in both laboratory and home settings to explore the efficacy and effectiveness of this approach for individuals with stroke. We also present the possibilities of using this technique for telemonitoring in stroke rehabilitation.

For the videos recorded by the researcher in the laboratory, we observed a positive agreement between the Track-UL algorithm and the Kinovea software in the measurement of upper-limb function in survivors of stroke in 4 different tasks. The LoA was below the minimum clinically significant difference, which supports the use of the Track-UL algorithm in clinical practice

XSL•FO
RenderX

to evaluate and monitor the progression of upper-limb kinematics. The LoA was slightly higher in certain movements on the affected side relative to the unaffected side, potentially due to the compensatory shoulder joint movements (internal rotation during arm elevation) and elbow joint movements (flexion and extension), which influenced the process of analysis.

We observed a significant degree of absolute agreement between the Track-UL algorithm and the Kinovea software in videos captured by both the physiotherapist and the participants. This finding supports the interchangeability of tools as the Track-UL algorithm performs as an intuitive and convenient tool for assessing arm mobility in both clinical and remote settings. We support the integration of this tool into the clinical assessment of upper-limb kinematics in individuals with stroke. This tool offers significant advantages, particularly for those who experience difficulties commuting to clinics or who reside in rural areas, by enabling remote follow-up and monitoring of upper-limb recovery. Furthermore, it provides more convenient solutions for carers, reducing their time and financial burdens as a result of traveling to clinics with their relatives with stroke. We can integrate this tool with any telerehabilitation model to provide a telerehabilitation and telemonitoring tool for people to use at home.

The test-retest reliability of the Track-UL algorithm demonstrated excellent consistency and robustness in repeated clinical assessments both in laboratory and home environments. The findings indicate that the Track-UL algorithm can deliver highly reliable kinematic measurements for recorded videos of people with stroke, rendering it appropriate for both in-clinic and remote assessments in stroke rehabilitation. This can be clinically beneficial, particularly for individuals reporting fatigue, allowing us to request that they perform the task only once rather than multiple times.

We calculated the minimum detectable change of the Track-UL assessment tool and determined that it was slightly elevated for the elbow joint compared to the shoulder joint, more pronounced on the affected side. This is potentially attributable to personal variability and the complexity of the tasks, which emphasize functional movements rather than simple joint movements. Additionally, the use of a 2D markerless motion capture system limits measurements to a single plane, making the results more sensitive to compensatory movements such as whole-arm rotation. Therefore, if this algorithm is to be used for monitoring upper-limb kinematics over time as a telemonitoring tool, the data should be interpreted with caution in clinical practice.

Our findings align with those of Latreche et al [31], who reported good agreement and reliability of an algorithm using the MediaPipe pose estimation framework in comparison to a goniometer; however, their study focused on simple shoulder movements. In our study, we explored complex functional arm movements involving the shoulder and elbow joints in people with stroke, which were close to what we usually evaluate in clinical practice. Another study measured upper-limb kinematics using sensorized gloves equipped with electromagnetic sensors capable of assessing 6 df during functional reach and grasp tasks [38]; however, the Track-UL algorithm–based measurement tool has the potential to be more user-friendly, requiring no additional equipment, which increases its suitability for clinical use.

Many participants were either unable to accurately record videos at home (4/27, 14.8%), potentially because restricted home environments, such as small houses, made it difficult to achieve the required camera angle, or they were unable to record the videos at all (10/27, 37%) due to technical difficulties in following the video recording instructions and then sending the videos back to the research team, particularly among older survivors. Consequently, we could not analyze the videos to determine whether the setup was incorrect, whether the entire movement was captured within the camera's frame, or whether the camera was positioned arbitrarily, all of which could affect pose estimation analysis. All of these factors affected the ability to record the videos at home and the ability of the algorithm to analyze the joint angles from the videos recorded independently by the participants or their carers at home. We recommend integrating this algorithm into an application that can offer a step-by-step guide for video recording alongside real-time feedback on the video quality, such as the camera angle, the distance from the camera, and frame coverage, to enhance the feasibility of using this algorithm as a telemonitoring tool for people with stroke.

For the current application, which aimed to validate 4 functional active ROM tasks, 2D analysis is appropriate. However, there are known limitations, such as lack of depth information and issues with out-of-plane movement detection. These could potentially be addressed using 3D with 2D recorded videos to yield further depth of information that could be useful in other contexts. We also recommend incorporating more functional tasks into the assessment and cross-validating them with functional questionnaires (FMA-UE). In addition, time-series analysis could yield extra layers of useful clinical information, including qualitative considerations such as movement smoothness, which could add further insight in future studies.

Conclusions

The Track-UL algorithm can be adopted as a cost-effective, fast, and user-friendly alternative to traditional motion analysis software such as Kinovea. This allows for its use in both clinical and potentially in remote settings. This technology has the potential to benefit people with stroke through monitoring and motor performance assessments. We recommend developing an application with an easy-to-use interface that enables survivors of stroke to securely record and analyze their arm movement both in clinics and at home. Automated reports can provide physiotherapists with objective, measurable kinematic data to help design and adjust personalized rehabilitation programs.

XSL•FO
RenderX

Acknowledgments

This study was supported by the Egyptian Ministry of Higher Education and Scientific Research as part of a funded PhD project. SEL is supported by the Dennis and Mireille Gillings Foundation. The research was carried out at the National Institute for Health and Care Research Exeter Biomedical Research Centre.

Funding

This study was funded by the Egyptian government and supported by the National Institute for Health and Care Research Biomedical Research Centre at the University of Exeter. HL has received funding for his PhD from the Egyptian Cultural and Educational Bureau (ID: MM48/21). HL, TCR, PL, HD, MM, and SEL are supported by the National Institute for Health and Care Research Exeter Biomedical Research Centre. HD is supported by National Institute for Health and Care Research HealthTech Research Centre.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HL, DH, AH, SEL, MM, and HD conceptualized the methods and designed the study. HL collected the data from the participants. HL, PL, XX, PN, and TCR analyzed and interpreted the data. HL and HD wrote the initial manuscript, and all authors reviewed, edited, and approved the final manuscript.

Conflicts of Interest

MM is a Co-Founder and Director of DigiTherapix. HD is codirector of DigiTherapix. DigiTherapix is not related to the subject matter of the manuscript, and no company resources or funds were used in this work. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Instructions for recording the videos.
[PDF File, 358 KB - rehab_v13i1e87128_app1.pdf]

References

1. Donkor ES. Stroke in the 21st century: a snapshot of the burden, epidemiology, and quality of life. Stroke Res Treat 2018;2018:3238165. [doi: 10.1155/2018/3238165] [Medline: 30598741]
2. Raghavan P. Upper limb motor impairment after stroke. Phys Med Rehabil Clin N Am 2015 Nov;26(4):599-610. [doi: 10.1016/j.pmr.2015.06.008] [Medline: 26522900]
3. Colebatch JG, Gandevia SC. The distribution of muscular weakness in upper motor neuron lesions affecting the arm. Brain 1989 Jun;112(Pt 3):749-763. [doi: 10.1093/brain/112.3.749] [Medline: 2731028]
4. O'Dwyer NJ, Ada L, Neilson PD. Spasticity and muscle contracture following stroke. Brain 1996;119(Pt 5):1737-1749. [doi: 10.1093/brain/119.5.1737] [Medline: 8931594]
5. Levin MF, Selles RW, Verheul MH, Meijer OG. Deficits in the coordination of agonist and antagonist muscles in stroke patients: implications for normal motor control. Brain Res 2000 Jan 24;853(2):352-369. [doi: 10.1016/s0006-8993(99)02298-2] [Medline: 10640634]
6. Tang E, Moran N, Cadman M, et al. Stroke rehabilitation in adults: summary of updated NICE guidance. BMJ 2024 Mar 22;384:q498. [doi: 10.1136/bmj.q498] [Medline: 38519084]
7. de Jong LD, Dijkstra PU, Stewart RE, Postema K. Repeated measurements of arm joint passive range of motion after stroke: interobserver reliability and sources of variation. Phys Ther 2012 Aug;92(8):1027-1035. [doi: 10.2522/ptj.20110280] [Medline: 22577062]
8. Gladstone DJ, Danells CJ, Black SE. The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair 2002 Sep;16(3):232-240. [doi: 10.1177/154596802401105171] [Medline: 12234086]
9. Hanks J, Myers B. Validity, reliability, and efficiency of a standard goniometer, medical inclinometer, and builder's inclinometer. Int J Sports Phys Ther 2023;18(4):989-996. [doi: 10.26603/001c.83944] [Medline: 37547826]
10. Lea RD, Gerhardt JJ. Range-of-motion measurements. J Bone Joint Surg Am 1995 May;77(5):784-798. [doi: 10.2106/00004623-199505000-00017] [Medline: 7744906]
11. Akizuki K, Yamaguchi K, Morita Y, Ohashi Y. The effect of proficiency level on measurement error of range of motion. J Phys Ther Sci 2016 Sep;28(9):2644-2651. [doi: 10.1589/jpts.28.2644] [Medline: 27799712]
12. Cirstea MC, Levin MF. Compensatory strategies for reaching in stroke. Brain 2000 May;123(Pt 5):940-953. [doi: 10.1093/brain/123.5.940] [Medline: 10775539]

XSL•FO
RenderX

13. Elrahim RMA, Embaby EA, Ali MF, Kamel RM. Inter-rater and intra-rater reliability of Kinovea software for measurement of shoulder range of motion. Bull Fac Phys Ther 2016 Dec;21(2):80-87. [doi: 10.4103/1110-6611.196778]
14. Hellsten T, Karlsson J, Shamsuzzaman M, Pulkkis G. The potential of computer vision-based marker-less human motion analysis for rehabilitation. Rehabil Process Outcome 2021;10:11795727211022330. [doi: 10.1177/11795727211022330] [Medline: 34987303]
15. Hannink E, Mansoubi M, Cronin N, et al. Validity and feasibility of remote measurement systems for functional movement and posture assessments in people with axial spondylarthritis. Healthc Technol Lett 2022 Dec;9(6):110-118. [doi: 10.1049/htl2.12038] [Medline: 36514477]
16. Mansoubi M, Leveridge P, Smith M, et al. Knee4Life: empowering knee recovery after total knee replacement through digital health protocol. Sensors (Basel) 2024 Nov 17;24(22):7334. [doi: 10.3390/s24227334] [Medline: 39599111]
17. Wade L, Needham L, McGuigan P, Bilzon J. Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. PeerJ 2022;10:e12995. [doi: 10.7717/peerj.12995] [Medline: 35237469]
18. Aprile I, Rabuffetti M, Padua L, Di Sipio E, Simbolotti C, Ferrarin M. Kinematic analysis of the upper limb motor strategies in stroke patients as a tool towards advanced neurorehabilitation strategies: a preliminary study. Biomed Res Int 2014;2014:636123. [doi: 10.1155/2014/636123] [Medline: 24868536]
19. Puig-Diví A, Escalona-Marfil C, Padullés-Riu JM, Busquets A, Padullés-Chando X, Marcos-Ruiz D. Validity and reliability of the Kinovea program in obtaining angles and distances using coordinates in 4 perspectives. PLoS One 2019;14(6):e0216448. [doi: 10.1371/journal.pone.0216448] [Medline: 31166989]
20. Fernández-González P, Koutsou A, Cuesta-Gómez A, Carratalá-Tejada M, Miangolarra-Page JC, Molina-Rueda F. Reliability of Kinovea® software and agreement with a three-dimensional motion system for gait analysis in healthy subjects. Sensors (Basel) 2020 Jun 2;20(11):3154. [doi: 10.3390/s20113154] [Medline: 32498380]
21. Lugaresi C, Tang J, Nash H, et al. MediaPipe: a framework for perceiving and processing reality. Presented at: Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR) 2019; Jun 17, 2019 URL: https://research.google/pubs/mediapipe-a-framework-for-perceiving-and-processing-reality/ [accessed 2026-04-21]
22. Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, Grundmann M. BlazePose: on-device real-time body pose tracking. arXiv. Preprint posted online on Jun 17, 2020. [doi: 10.48550/arXiv.2006.10204]
23. Akturk S, Derdiyok FB, Serbest K. Markerless joint angle estimation using MediaPipe with a rapid setup for joint moment calculation. Multimed Tools Appl 2026;85:38. [doi: 10.1007/s11042-026-21256-z]
24. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell 2021 Jan;43(1):172-186. [doi: 10.1109/TPAMI.2019.2929257] [Medline: 31331883]
25. Uhlrich SD, Falisse A, Kidziński Ł, et al. OpenCap: human movement dynamics from smartphone videos. PLoS Comput Biol 2023 Oct;19(10):e1011462. [doi: 10.1371/journal.pcbi.1011462] [Medline: 37856442]
26. Martin Bland J, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb;327(8476):307-310. [doi: 10.1016/S0140-6736(86)90837-8]
27. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. Phys Ther 1996 Oct;76(10):1109-1123. [doi: 10.1093/ptj/76.10.1109] [Medline: 8863764]
28. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris Questionnaire. Phys Ther 1996 Apr;76(4):359-365. [doi: 10.1093/ptj/76.4.359] [Medline: 8606899]
29. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
30. Mullaney MJ, McHugh MP, Johnson CP, Tyler TF. Reliability of shoulder range of motion comparing a goniometer to a digital level. Physiother Theory Pract 2010 Jul;26(5):327-333. [doi: 10.3109/09593980903094230] [Medline: 20557263]
31. Latreche A, Kelaiaia R, Chemori A, Kerboua A. Reliability and validity analysis of MediaPipe-based measurement system for some human rehabilitation motions. Measurement 2023 Jun;214:112826. [doi: 10.1016/j.measurement.2023.112826]
32. Garg S, Saxena A, Gupta R. Yoga pose classification: a CNN and MediaPipe inspired deep learning approach for real-world application. J Ambient Intell Human Comput 2023 Dec;14:16551-16562. [doi: 10.1007/s12652-022-03910-0]
33. Hisham MN, Hassan MF, Ibrahim N, Zin ZM. Mono camera-based human skeletal tracking for squat exercise abnormality detection using double exponential smoothing. Int J Adv Comput Sci Appl 2022;13(7). [doi: 10.14569/IJACSA.2022.0130709]
34. Zhang W, Li Y, Cai S, et al. Combined MediaPipe and YOLOv5 range of motion assessment system for spinal diseases and frozen shoulder. Sci Rep 2024;14:15879. [doi: 10.1038/s41598-024-66221-8]
35. Friedrich MU, Roenn AJ, Palmisano C, et al. Validation and application of computer vision algorithms for video-based tremor analysis. NPJ Digit Med 2024 Jun 21;7(1):165. [doi: 10.1038/s41746-024-01153-1] [Medline: 38906946]
36. Pornpipatsakul K, Chuengwutigool W, Chaichaowarat R, Foongchomcheay A. Bridging exercise monitoring system using RGB camera for stroke rehabilitation. Presented at: 2023 IEEE R10 Conference (TENCON); Oct 31 to Nov 3, 2023. [doi: 10.1109/TENCON58879.2023.10322445]
37. Kushnir A, Kachmar O, Bonnechère B. STASISM: a versatile serious gaming multi-sensor platform for personalized telerehabilitation and telemonitoring. Sensors (Basel) 2024 Jan 6;24(2):351. [doi: 10.3390/s24020351] [Medline: 38257442]

XSL•FO
RenderX

38. Khanna P, Oppenheim T, Tu-Chan A, Abrams G, Ganguly K. Measuring arm and hand joint kinematics to estimate impairment during a functional reach and grasp task after stroke. Neurorehabil Neural Repair 2023 Jun;37(6):409-417. [doi: 10.1177/15459683231179173] [Medline: 37300318]

Abbreviations

FMA-UE: Fugl-Meyer Assessment for the upper extremity
ICC: intraclass correlation coefficient
LoA: limit of agreement
ROM: range of motion

Edited by S Munce; submitted 27.Nov.2025; peer-reviewed by A ELMelhat, J Son; revised version received 19.Mar.2026; accepted 09.Apr.2026; published 11.May.2026.

Please cite as:
Lazem H, Harris D, Hall A, Richards TC, Leveridge P, Mansoubi M, Xu X, Newell P, Lamb SE, Dawes H
Validity and Reliability of the Track-UL Algorithm Compared With Kinovea Software for Measuring Upper-Limb Functional Range of Motion in People After Stroke: Cross-Sectional Observational Study
JMIR Rehabil Assist Technol 2026;13:e87128
URL: https://rehab.jmir.org/2026/1/e87128
doi:10.2196/87128

© Hatem Lazem, David Harris, Abi Hall, Thomas C Richards, Phaedra Leveridge, Maedeh Mansoubi, Xiaohan Xu, Paul Newell, Sarah E Lamb, Helen Dawes. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 11.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Intelligent Robot-Aided, Task-Oriented Physiotherapy Method for Upper Limb Rehabilitation: Development and Usability Study

Piotr Falkowski[1], PhD; Kajetan Jeznach[1]; Jan Oleksiuk[1]; Tomasz Osiak[1]; Maciej Pikuliński[1,2]; Andrzej Zakręcki[1,3], PhD; Krzysztof Zawalski[1]; Piotr Kołodziejski[1]; Daniel Śliż[1,4], MBA, MD, PhD

[1]Section of Biorobotics and Medical Devices, Łukasiewicz Research Network—Industrial Research Institute for Automation and Measurements PIAP, Al Jerozolimskie 202, Warsaw, Mazovia, Poland

[2]Institute of Aeronautics and Applied Mechanics, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Warsaw, Poland

[3]Department of Manufacturing Systems, Faculty of Mechanical Engineering and Robotics, AGH University of Science and Technology in Cracow, Krakow, Poland

[4]3rd Department of Internal Diseases and Cardiology, Medical University of Warsaw, Warsaw, Poland

Corresponding Author:
Piotr Falkowski, PhD
Section of Biorobotics and Medical Devices, Łukasiewicz Research Network—Industrial Research Institute for Automation and Measurements PIAP, Al Jerozolimskie 202, Warsaw, Mazovia, Poland

Abstract

Background: Task-oriented rehabilitation supported by exoskeletons has the potential to increase therapy intensity, personalization, and accessibility. However, to achieve fully automatic treatment, robotized systems need to analyze therapy in a more complex way than only based on reference trajectories following.

Objective: This study aimed to investigate the effects of an intelligent, context-aware control algorithm for an upper limb rehabilitation exoskeleton on patients' musculoskeletal engagement, compared with constant-admittance robot-assisted therapy, and conventional physiotherapist-guided treatment.

Methods: A single-session experimental study was conducted with 34 adult participants performing 6 activities of daily living under 3 therapy modes: robot-assisted therapy with constant admittance, robot-assisted therapy with an intelligent assist-as-needed algorithm, and physiotherapist-guided therapy. Muscle activity was assessed using surface electromyography of 8 upper limb muscle groups, while joint kinematics were recorded using inertial measurement units. Metrics included electromyography power, muscle activation time, joint range of motion, and Burst Duration Similarity Indices. Statistical comparisons were performed using the *t* test and the Mann-Whitney *U* test depending on data normality.

Results: Results indicate that the intelligent control strategy engages the musculoskeletal system at least as effectively as constant-admittance control across all exercises. At the same time, more motion control is given to the patient, which is consistent with the principles of neuroplastic motor learning. Compared with physiotherapist-guided therapy, robot-assisted treatment with intelligent control elicited significantly higher and more consistent muscular engagement. Intelligent assistance also modified joint-level motion patterns by reducing compensatory movements, particularly in shoulder-elbow coupling, while maintaining functional task execution. Muscle activation timing patterns during intelligent robot-assisted therapy were more consistent with robotic control than with manual therapy, reflecting altered movement strategies.

Conclusions: These findings demonstrate that context-aware, intelligent control in rehabilitation exoskeletons can promote active patient participation, reduce compensatory behaviors, and maintain physiologically meaningful muscle engagement. The proposed approach exceeds the results of recent similar studies, being a promising step toward effective, minimally supervised, and task-oriented rehabilitation.

(JMIR Rehabil Assist Technol 2026;13:e92898) doi:10.2196/92898

KEYWORDS

electromyography; exoskeleton; home therapy; intelligent algorithms; minimally supervised treatment; personalized medicine; rehabilitation robotics

XSL•FO
RenderX

Introduction

Background

Upper extremity impairment is among the most common consequences of neurological conditions, including stroke. Moreover, it is severely affecting functional independence in activities of daily living (ADLs) and requires intensive, long-term physiotherapy, which significantly burdens the health care systems [1]. Despite the recognized clinical importance of rehabilitation, the global supply of trained physiotherapy professionals remains critically insufficient relative to demand, even 10-fold in less developed European countries [2]. This structural deficit is even more serious when combined with geographic inequalities in the reachability of specialist centers and place-related qualified personnel ready to perform physically demanding work [3]. Therefore, this paper is focused on the proposition of a technological approach to robot-aided task-oriented treatment, mitigating the problem of medical staff shortages.

Robot-Aided Task-Oriented Treatment

Task-oriented rehabilitation constitutes the foundation of contemporary physiotherapy, and its primary objective is to restore patients' ability to perform ADLs with the highest possible level of independence. From the physiotherapist's perspective, the key factors are not merely the repetitiveness of movement but, above all, its functional nature, appropriate engagement of the musculoskeletal system, and the continuous adjustment of the level of assistance to the patient's current capabilities [4].

The literature indicates that robot-assisted rehabilitation may improve motor function and quality of life in patients with neurological impairments to a degree comparable with that achieved with conventional therapies, particularly with respect to movement repetitiveness, training intensity, and the facilitation of neuroplasticity [5]. Beyond the aspects of repetitiveness and functionality, current technology enables a key element of therapy—personalization. Tailoring the therapy to the individual patient may significantly enhance its effectiveness [6].

Moreover, automation enhances physiotherapists' capabilities as telemedicine becomes a clinical standard in many countries [7]. This is particularly important due to the shortages in medical staff—robotization of physical therapy can bring a significant benefit of early treatment at high doses to patients. Nevertheless, this requires safe operation, which is guaranteed by continuously adjusting the device-human interaction based on the analyzed operation. Apart from the risk of physical damage, potential discomfort and unexpected situations during treatment can decrease trust in the robot. This can become a significant obstacle in encouraging people to use robotized treatment instead of human-conducted treatment—even despite the presented benefits [8].

In these aspects, solutions that enable dynamic, context-aware adjustment of exoskeleton assistance become particularly important, analogous to the decisions a physiotherapist makes during manually guided therapy. The integration of the assessment of novel task-oriented exercises with intelligent control algorithms may represent a key factor in enhancing the effectiveness of robot-assisted rehabilitation by promoting more effective engagement of the musculoskeletal system and supporting the patient's active participation in the rehabilitation process [6].

Control Strategies

Contemporary rehabilitation robots, for example, end-effectors or wearable exoskeletons, seek to deliver repeatable therapy while controlling interaction forces and preserving active patient participation [9]. A prevalent control architecture consists of a fast inner loop enforcing joint- or task-space motion and an outer impedance or admittance loop shaping the dynamics through a virtual mass–spring–damper driven by measured interaction forces, balancing tracking performance and contact sensation [10,11].

A key challenge is that admittance stability and transparency are determined by the complete closed-loop chain rather than by the virtual dynamics alone. Stability margins depend strongly on force sensing and filtering, computation or communication delays, and the bandwidth of the inner motion loop [12]. Excessive force filtering introduces phase lag and can destabilize human-robot interaction, especially when rendering low virtual inertia to encourage the patient's motion. Practical guidelines, therefore, emphasize minimizing force filtering, injecting sufficient virtual damping, and explicitly preserving passivity of the coupled system [12,13]. Clinically, this matters because excessive mechanical resistance reduces the patient's voluntary contribution to movement, significant for neurorehabilitation.

Within this framework, recent developments focus on adapting assistance to the specific requirements of the rehabilitation task. Rather than using fixed interaction dynamics, controllers are increasingly designed to adjust the assistance based on inferred user intent and on task-dependent objectives. Yu et al [14] proposed a proportional-integral-derivative controller–type admittance and highlighted implementation aspects, including gravity compensation and sensor fusion for reference generation. At the same time, task-space formulations have been shown to reduce unnecessary energy exchange during exercises [15]. Beyond time-domain tuning, admittance shaping may specify target natural frequency, damping, and low-frequency gain for the human-robot pair [16]. Also, variable admittance strategies adapt parameters online based on interaction force, tracking error, or collaboration state; for example, a collaboration observer can detect unsafe interaction and temporarily raise virtual inertia and damping to recover stability [17]; a task-oriented variable admittance can modulate assistance across different exercise phases [18]. Assist-as-needed (AAN) formulations embedded in variable admittance controllers are emerging [19], with analogous concepts explored in lower limb devices, for example, strength-index–based AAN for a hip exoskeleton [20]. From a rehabilitation perspective, AAN strategies are motivated by the principle that the optimal therapy should be focused on providing just enough assistance to enable successful task completion while remaining the other part to the patient's motor system.

XSL•FO
RenderX

Various metrics, in addition to standard robotics indices such as tracking errors, are used to assess the efficacy of these control approaches; for example, the Fugl-Meyer assessment, electromyography (EMG), or electroencephalography [21]. Results from such trials yield important conclusions, as shown in Aurich-Schuler et al [22], where rehabilitation with different control settings of the Lokomat showed that more permissive modes can increase muscle activation and promote more physiological patterns. Crucially, this finding highlights that control-level choices, not just exercise selection, directly determine the biomechanical profile of the therapy session.

The rule-based admittance adaptation can be further personalized. Learning-augmented admittance uses recurrent neural networks to fine-tune virtual inertia and damping for improved tracking [23]. Others incorporate physiological metrics, such as muscular manipulability, to direct assistance to where the patient's capability is low [24]. An important venue is dynamic compensation. Because rehabilitation motions are often slow, many systems emphasize gravity compensation rather than full inertial dynamics, using various learning-based methods [25,26].

A critical question is then whether an algorithm that autonomously adapts support based on contextual assessment of task performance, rather than manual parameter tuning, can produce measurably different musculoskeletal engagement compared with both fixed-admittance robotic control and conventional physiotherapy.

Aim of the Study

Taken together, these results motivate context-aware tuning of assistance settings and evaluation of whether such settings measurably improve the rehabilitation process. Despite the advances, selecting and updating controller parameters are still most often manual, task-specific processes, and controllers that perform well in one exercise may not generalize when the task context or patient capability changes. Moreover, improvements in stability or tracking do not necessarily translate into better rehabilitation outcomes, which even strengthens the case for algorithms that assist based on contextual assessment and for outcome-oriented indices.

Prior research conducted by the authors showed that using an intelligent algorithm to set support for exoskeleton-aided therapy, based on a contextual assessment of new exercises rather than constant support settings, results in more accurate exercise performance, although more control is given to the patient. This shows that providing more support upon detecting significant errors and less support during the rest of the operation increases motion quality. Moreover, it remains in line with the principles of neuroplastic exercising, as the patient is more independent while deciding on motion patterns. Despite confirmation of superiority in terms of patient participation in therapy and reduced need for constant support, the impact on the patient's musculoskeletal system must be analyzed.

This study is a natural continuation of the investigation process, focusing on the biomechanics of patients. It aims to examine whether the developed algorithms engage patients at a higher level than conventional control algorithms and physiotherapist-conducted therapy. This is determined based on statistical tests performed using metrics of muscular activity and joint range of motion (ROM). The test hypotheses aim to assess whether EMG- and joint motion (JM)–related parameters remain higher for the intelligent mode than for the control of the same robot with constant-admittance active therapy than the manual physiotherapy. The precise formulation of these is presented in the Metrics subsection of the Methods section.

The results of the experiment will enable an evidence-based comparison of the developed approach to providing therapy with exoskeletons with alternative, well-established approaches. This is particularly significant, as the proposed therapy methodology also enables easy treatment scaling and personalization and, as a consequence, brings robot-aided therapy closer to a homelike environment with minimal specialist supervision (also remote).

Methods

Ethical Considerations

The experiments were carried out under the KB/132/2024 approval of the Bioethical Committee of the Medical University of Warsaw. Written informed consent was obtained from all of the participants involved in this study. It involved a group of 34 volunteers (with no compensation) who had given written informed consent to participate. All of the personal data were anonymized as described within the participant information card. The group was heterogeneous in terms of age (23 - 65 years, μ=36.2), sex, and medical conditions (cardiological, neurological, and orthopedical). An a priori power analysis conducted in G*Power software (Heinrich-Heine-Universität Düsseldorf) indicated that, assuming a within-subject effect size of Cohen d=0.4, α=.05, power=0.95, 2-group tests used within this paper, and 2-tailed analysis, 34 participants with 5 repetitions were required [27].

Experimental Procedure

The experimental phase consisted of performing a standardized set of 6 exercises with 5 repetitions each, all in 3 modes dependent on the support of the therapy, also with the rehabilitation exoskeleton. Every exercise reference motion was recorded by the device during physiotherapist-guided exercises, uniquely for each patient during a physiotherapist-guided familiarization session prior to the experimental protocol. Thanks to this, individual limitations and natural motion patterns were taken into account. The modes used in the study included the following:

1. Mode 1: active therapy with the support of the SmartEx-Home rehabilitation exoskeleton controlled with the admittance algorithm. The device adapted to the intended motion of a patient based on measurements from multiaxial force sensors integrated with the exoskeleton and corrected the motion proportionally to the trajectory deviation. The admittance is constantly set to 0.5.
2. Mode 2: active therapy with the support of the SmartEx-Home rehabilitation exoskeleton controlled with the intelligent algorithm. The algorithm assesses reference trajectories (not preprogrammed) for potential functional

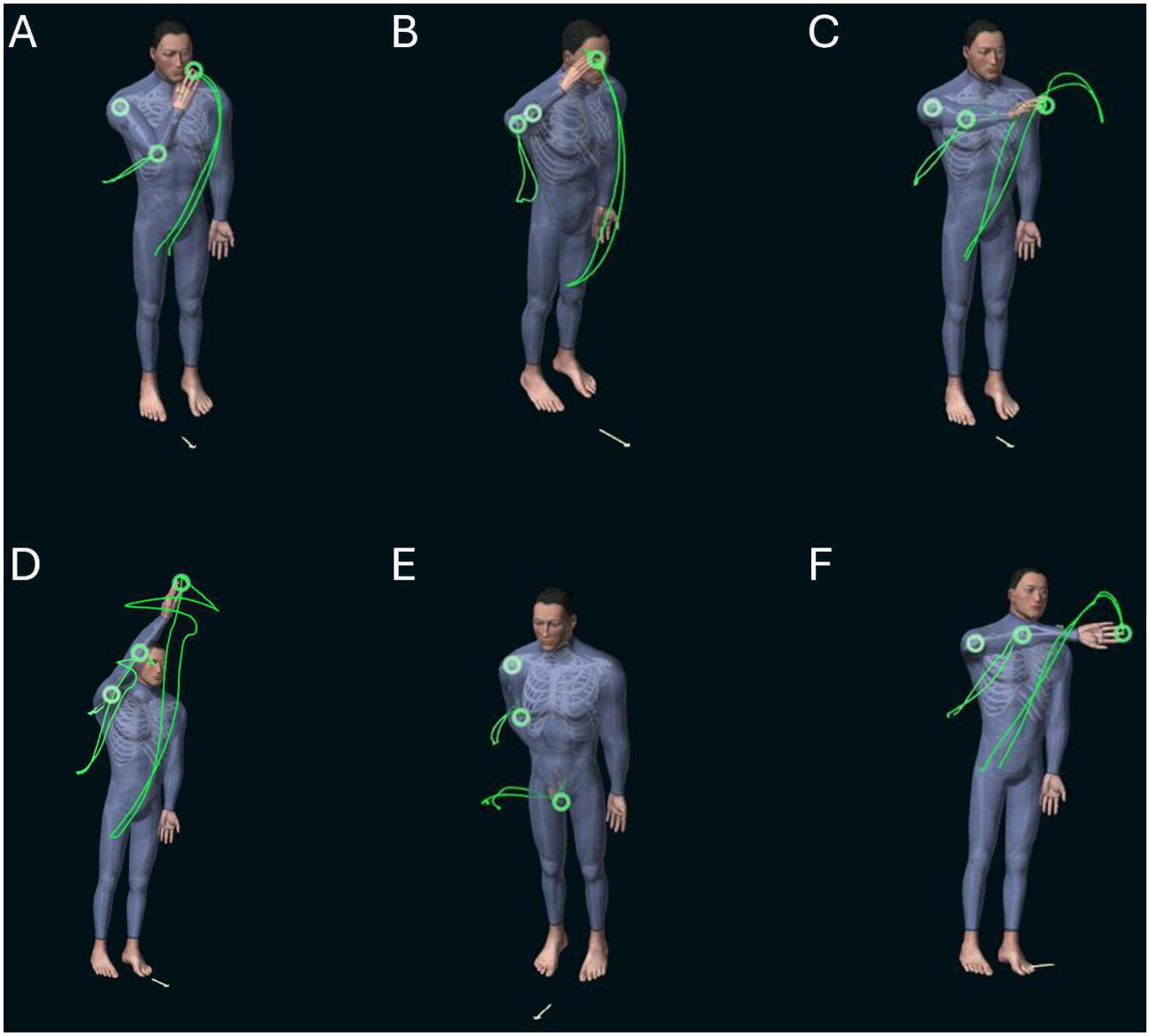
XSL•FO
RenderX

errors and anatomical compensations and assists the patient as needed to minimize the risk of their occurrence. This enabled the patient to modify the motions as within mode 1, but the corrections were connected not only to the trajectory deviations but also to the contextually analyzed errors. The precise operation of this mode is described in the complementary paper (Falkowski P and colleagues, unpublished data, 2026).

3. Mode 3: active therapy with the support of the physiotherapist and correction of motions as within real-life treatment.

The exercises used in the procedure are visualized in Figure 1 and include (A) exercise 1: snacking, (B) exercise 2: answering the phone call, (C) exercise 3: carrying the shopping bags, (D) exercise 4: shutting the curtains, (E) exercise 5: wiping buttocks, and (F) exercise 6: pouring a drink.

Figure 1. Visual presentation of exercise trajectories (top: A, B, and C; bottom: D, E, and F).

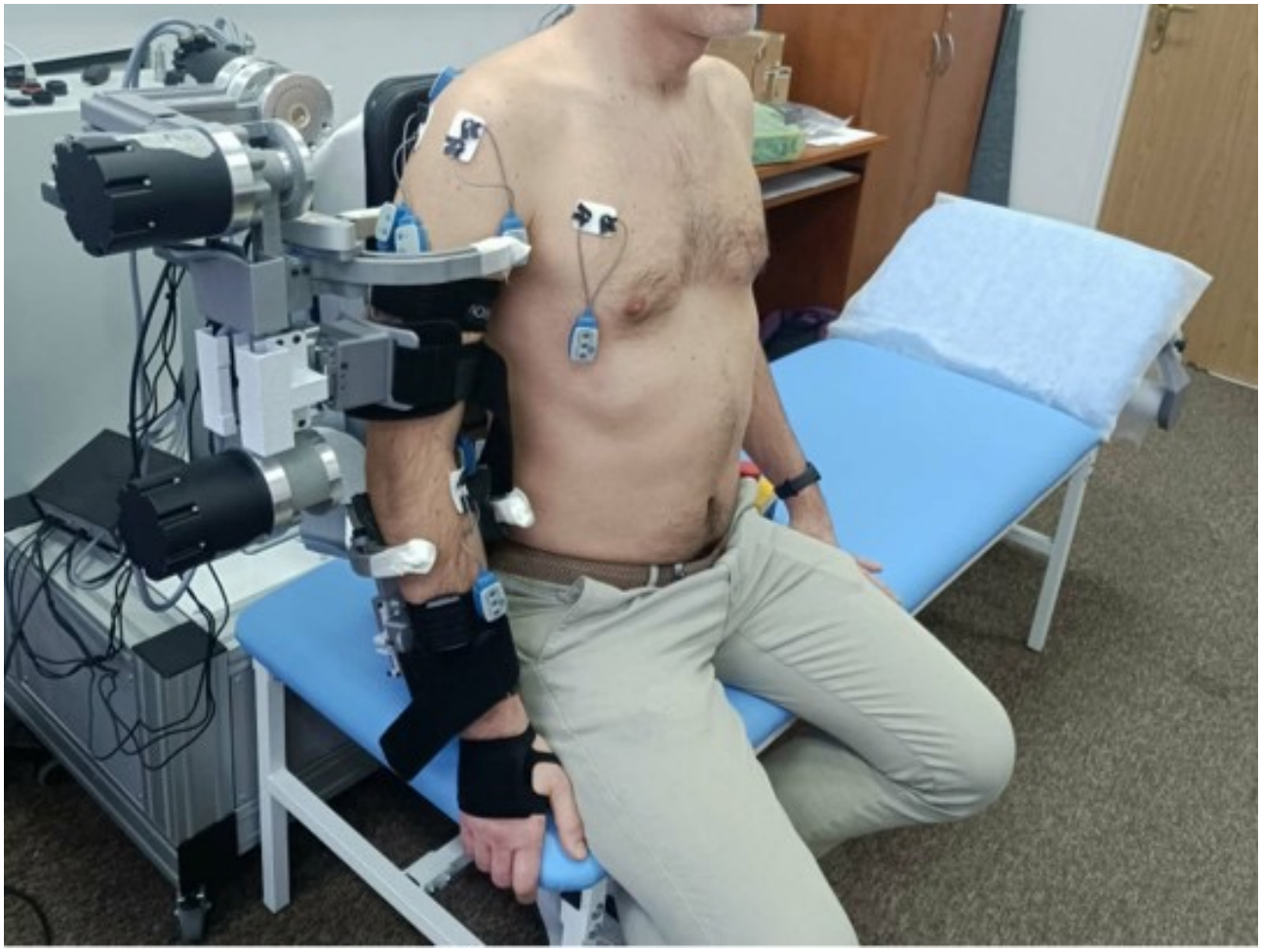


SmartEx-Home Exoskeleton

The SmartEx-Home exoskeleton for the upper extremity on a mobile platform was used in the study [28,29]. It has 5 degrees of freedom—3 driven (shoulder abduction or adduction, shoulder flexion or extension, and elbow flexion or extension) and 2 passive—open sliding bearings [30] (shoulder and radioulnar rotations). The wrist motion remains free. A patient is connected to the robot's construction at the torso, arm, and forearm with elastic straps, with the 2 extremity attachments equipped with 3-axis force sensors. The exoskeleton can provide passive or active treatment, also in AAN mode. The last one involves an intelligent algorithm, which can assess any exercise reflecting the ADL contextually. This includes finding similarities between new exercises and preprogrammed ADLs [31]. Then it assesses trajectory following, anatomical compensations [32], and functional errors. Assistance is set based on the patients' performance. A patient performing an experimental routine with the exoskeleton is presented in Figure 2. While exercising, patients can see a visual of a human model on screen—1 representing their current position, and 1 representing the desired position.

XSL•FO
RenderX

Figure 2. Experimental setup.



Measuring

The study aimed to assess the impact on the patients' musculoskeletal system. Therefore, metrics related to muscle activity and the involved ranges of motion were selected. To monitor them, the Noraxon Ultium system with inertial moment units (IMU sensors) and EMG sensors was used to measure the process. Additionally, 2 integrated cameras were recording exercises to facilitate the possibility of validating analyses and dividing datasets into repetitions.

Four IMU sensors were attached to the following regions of every patient's body according to the system's documentation: (1) upper thoracic: below C7 in line with the spinal column but high enough to not be affected by upper trapezius muscle movement; (2) upper arm: midway between the shoulder and elbow joints, lateral to the bone axis; (3) forearm: posterior and distal, where there is a low amount of muscle tissue; and (4) hand: dorsal.

The calibration process was performed according to the system's documentation. The calibration position was static—standing or seated for patients with severe disabilities of the lower extremities. This involves setting all joints to the zero-degree position (or 90 degrees for those with disabilities) and activating the calibration mode. The standing calibration position requires all joints to be in the neutral, zero position, with the palms aligned with the sagittal plane. The seated calibration position requires a patient to be seated with the elbow, hip, and knee at 90-degree flexion.

As a result, the following 5 JMs were computed as time series in degrees regarding base position: shoulder abduction or adduction (JM 1), shoulder flexion or extension (JM 2), shoulder rotations (JM 3), elbow flexion or extension (JM 4), and wrist radial rotations (JM 5). The computation was performed automatically in MR 4.2 software (Noraxon) based on the registered signals for IMU sensors. However, each time series also underwent an unwrapping signal process to eliminate potential discontinuities resulting from the assumed ranges of motion while capturing the original sensors' Euler angles [33].

Eight EMG sensors were attached with surface gel electrodes according to *Noraxon* [34] and *SENIAM* [35] guidelines on the following muscles: middle deltoid (EMG 1), posterior deltoid (EMG 2), pectoralis major (EMG 3), teres major (EMG 4), latissimus dorsi (EMG 5), biceps brachii (EMG 6), triceps brachii (medial head; EMG 7), and brachioradialis (EMG 8). EMG signals were collected at 2000 Hz using Noraxon's Ultium EMG system and processed offline following the manufacturer's guidelines and according to analog experimental practice [36,37]. The processing chain included the following steps: (1) high-pass filter: applied a 10-Hz IIR Butterworth filter to remove motion artifacts and baseline drift (standard setting for chosen tasks that are not very dynamic when performed with a robot), (2) low-pass filter: applied a 500-Hz IIR Butterworth filter to

XSL•FO
RenderX

keep relevant EMG content and remove high-frequency noise, (3) rectification: converted the bipolar signal to positive values only using full-wave rectification, (4) smoothing: used a 200-ms moving average to create a smooth envelope (recommended for our type of motion recordings), and (5) normalization: divided by the functional muscular voluntary contraction level [38] from an automatically chosen maximum value over a 1000-ms window. The gathered data were then statistically analyzed with distinction between modes, exercises, and channels (different for EMG and IMU). This enabled drawing global conclusions and considering the significant muscle activity and JM during various exercises.

Metrics

The theses from the introduction were analyzed using four metrics [39]:

- *maxEMG*: measure reflecting maximum activity generated with the muscle within the exercise, computed as the maximum of the EMG signal.
- *iEMG*: measure reflecting power (midpoint integral) of the muscular activity, computed according to the formula (1), where *EMG* is the value of EMG signal, and *t* is time.

$$(1) iEMG = 12 \sum_{n=2}^{N} (EMG(n) + EMG(n-1)) \cdot (t(n) - t(n-1))$$

- T_{act}: measure reflecting the time within the exercise when the muscle is active, computed according to the formula (2), where N is the number of samples in the signal, and *act*(*n*) is the activation function calculated according to the formula (3) with the levels set based on literature sources [40].

$$(2) Tact = \sum_{n=2}^{N} act(n) \cdot (t(n) - t(n-1))$$

$$(3) act(n) = \{0 \text{ if } EMG(n) \< 15\% \cdot fMVC \; 1 \text{ if } EMG(n) \geq 15\% \cdot fMVC$$

- ROM: measure reflecting the range of a JM used within the exercise, computed as the minimal value of the joint position subtracted from its maximum value.

For statistical analysis, the metrics were grouped by combinations of modes, exercises, and channels (8 muscular groups for EMG-based metrics and 5 JMs for ROM). These were treated as separate population groups. First of all, each of these was validated for normality using the Shapiro-Wilk test. As the groups were relatively large (all >150 samples), the 1-tailed *t* test was used for those that were normally distributed. For the nonnormal distributions, the Mann-Whitney *U* test was used [41]. Both test types were conducted using MATLAB (MethWorks) functions that enable validation of 1-sided alternative hypotheses.

In fact, 2 subsequent statistical tests were conducted for each dataset. To prove the relation with the constant-admittance therapy, the first hypothesis (null and alternative) was formed:

- Hypothesis 0: Effectiveness metrics for the intelligent algorithm are statistically *comparable to or lower* than the ones for the nonintelligent algorithm.
- Hypothesis 1: Effectiveness metrics for the intelligent algorithm are statistically *higher* than those for the nonintelligent algorithm.

Then the second hypothesis (null and alternative) was formed:

- Hypothesis 0: Effectiveness metrics for the intelligent algorithm are statistically *comparable to or greater* than the ones for the nonintelligent algorithm.
- Hypothesis 1: Effectiveness metrics for the intelligent algorithm are statistically *lower* than those for the nonintelligent algorithm.

If hypothesis 0 is rejected during the first test, it is assumed that the compared metrics are greater than the reference metrics (further marked G). For this statistical test result, the *P* value from the first test is presented. If it fails to reject hypothesis 0 in the first test and hypothesis 0 is rejected in the second test, it is assumed that the compared metrics are lower than the reference metrics (further marked L). If it fails to reject hypothesis 0 in both tests, it is assumed that the compared metrics are statistically equal to the reference metrics (meaning that no significant difference is observed; further marked E). For these 2 statistical test results, the *P*-value from the second test is presented.

Analogous hypotheses were formed for comparison with the conventional physiotherapy, manually guided by a physiotherapist. To validate the hypothesis, all the presented metrics beforehand reflecting muscular activity and joint mobility were used. No family-wise error correction was applied; results should be interpreted in an exploratory context, particularly for comparisons. However, additional ANOVA comparison was performed for the whole groups of results for different modes, not dividing them based on EMG channels, JMs, or exercises.

Additionally, BDSI (Burst Duration Similarity Index) values were computed for comparison of activity of muscles during mode 2 with mode 1 and mode 2 with mode 3 according to the formula (4), including activity functions computed for 2 compared muscular groups according to formula (3) [39]. Before, the time series for the compared EMG signals were approximated and normalized to have the same number of elements, as the comparison is connected to the phase of the realized motion, and their durations varied.

$$(4) BDSI = \frac{1}{N} \sum_{n=1}^{N} XNOR(act1(n), act2(n))$$

BDSI indices were then grouped by muscle group and exercise (without any other distinguishing factors). The mean and standard deviation were then computed for them. This enabled analysis of differences in the moments of muscle activation across motions and between therapy modes.

Results

Overview

The results in this section are presented with the distinction between muscular activity, ROM, and motion patterns comparisons. They are analyzed for exercises and either EMG channels or single JMs combined with exercises separately to draw the precise initial conclusions on the impact of the intelligent mode on performing treatment. However, to assess it at the global scale, the comparison for all the classes combined was additionally performed.

XSL•FO
RenderX

Muscular Activity

Each exercise represents a different ADL. Therefore, the motions differ significantly, as shown by the activity of different muscular groups and the range of JM. Based on the results from mode 3, the most critical items for each exercise were identified. Analysis for muscular groups was performed based on mean values and standard deviations for EMG values and the time of activation during exercises. Analysis for JMs was performed considering the mean ranges of motions used in the exercise, compared with the ranges of motion used for all of the exercises.

The first conclusion from this analysis was that the latissimus dorsi was contracted at a stable level close to the functional muscular voluntary contraction during most exercises to stabilize the upper body of patients. Also, pectoralis major and teres major were relatively active during the exercises, with the most significant activity during answering the phone and carrying the shopping bags for ADLs, respectively. Posterior deltoid, the muscle acting antagonistically to it, was continuously active at a significantly lower level. Similarly, biceps brachii and brachioradialis were active for long periods but with stronger contractions. The latter was also outstandingly active during answering the phone call exercise, due to the rotation of the hand close to the ear. The biceps brachii were relatively inactive during exercises, except for the wiping-the-buttocks movement.

In terms of joint kinematics, patients used mainly abduction or adduction and flexion or extension of the shoulder, along with flexion or extension of the elbow. A relatively low amount of shoulder flexion or extension was observed only for wiping buttocks, while low elbow flexion or extension was observed for pouring a drink during ADL. Shoulder rotations were activated mainly during exercises 2, 3, and 4, while wrist radial rotation was used mainly for exercises 5 and 6.

In terms of muscular activity, 2 different metrics were used to compare modes of exercise. *iEMG* was used to represent the power exerted by every monitored muscular group during exercises. Comparison of intelligent control in the constant-admittance mode shows that, for most combinations, the results are statistically comparable, although the metric means during intelligent treatment are typically higher (Table 1). In a few cases for EMG 2, and EMG 6-8 muscular groups, the results are statistically better, while for EMG 1, the results are lower in exercise 4. T_{act} was used as a measure reflecting the time of the muscles' activity. For compared modes 2 and 1, the times are statistically equal for most of the exercises and muscular groups and are higher during intelligently controlled treatment for 5 combinations of those (Table 2). This metric was not greater for the constant-admittance mode for any of the considered cases.

The outcomes prove that the robot controlled with intelligent algorithms examined in this study tends to engage the muscular system of a patient at least at the same level as the one controlled with the constant-admittance mode. At the same time, previous research evidenced that the former gives more control over the motion to the patient, which is applicable for neuroplastic rehabilitation [42]. Moreover, using the former results in more accurate following of the recorded trajectories. The results from the presented analysis confirm that the newly developed method is at least as effective as the well-established approach to active therapy with rehabilitation robots [2].

Moreover, the *iEMG* and T_{act} parameters were compared between modes 2 and 1 for the whole gathered dataset, without distinction on EMG channels and exercises. The results of statistical analysis rejected the hypothesis that the effectiveness metrics for the intelligent algorithm are statistically comparable with or lower than the ones for the nonintelligent algorithm for both parameters (μ=105.42, μ =99.57, P<.001, and μ=3.39, μ =3.14, P<.001, respectively, for *iEMG* and T_{act}).

Analogical analyses were performed to compare mode 2 with physiotherapist-provided treatment. The results are significantly better for the therapy provided by the robot (Tables 3 and 4). While validating the outcomes, it was observed that exercises provided by a physiotherapist without the device were significantly faster during short sessions than those designed with the exoskeleton. Nevertheless, this naturally affected not only $_t$ parameters but also *iEMG*, which, as reflecting muscle work, should not be time-dependent. As observed, patients feel more responsible for the motions performed while exercising with the exoskeleton. Hence, they engage more physically and cognitively in the treatment.

Also, the *iEMG* and $_t$ parameters were compared between modes 2 and 3 for the whole gathered dataset, without distinction on EMG channels and exercises. The results of statistical analysis rejected the hypothesis that the effectiveness metrics for the intelligent algorithm are statistically comparable with or lower than the ones for the nonintelligent algorithm for both parameters (μ=105.42, μ =35.27, P<.001, and μ=3.39, μ =1.08, P<.001, respectively, for *iEMG* and T_{act}).

Moreover, physiotherapists tend to become physically tired during their work [3]. Even if they focus on faster motions of patients, they tend to take longer breaks between them. Long-term results regarding the time to muscle activation are expected to be significantly lower than those observed during robot-aided treatment with the proposed rehabilitation exoskeleton.

XSL•FO
RenderX

Table . Statistical results for mode 2 versus mode 1; metric: *iEMG*[a,b].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
EMG[c] 1	E (*P*=.91, μ=108.35, μ =101.14)	E (*P*=.13, μ=94.66, μ =102.86)	E (*P*=.90, μ=81.19, μ =82.03)	L (*P*=.04, μ=162.51, μ =178.58)	E (*P*=.49, μ=43.54, μ =37.62)	E (*P*=.21, μ=98.71, μ =95.00)
EMG 2	G (*P*=.01, μ=96.60, μ =78.51)	E (*P*=.94, μ=81.20, μ =69.86)	G (*P*=.005, μ=85.67, μ =68.36)	E (*P*=.58, μ=117.61, μ =105.54)	E (*P*=.36, μ=83.28, μ =80.21)	E (*P*=.93, μ=91.02, μ =77.47)
EMG 3	E (*P*=.93, μ=154.21, μ =140.95)	E (*P*=.83, μ=141.66, μ =130.08)	E (*P*=.75, μ=133.83, μ =128.89)	E (*P*=.86, μ=181.94, μ =167.09)	E (*P*=.94, μ=114.52, μ =100.90)	E (*P*=.81, μ=153.98, μ =145.06)
EMG 4	E (*P*=.61, μ=141.42, μ =139.53)	E (*P*=.49, μ=128.04, μ =128.99)	E (*P*=.41, μ=114.66, μ =117.88)	E (*P*=.45, μ=172.11, μ =173.52)	E (*P*=.93, μ=112.52, μ =96.47)	E (*P*=.67, μ=129.44, μ =123.30)
EMG 5	E (*P*=.76, μ=134.83, μ =128.56)	E (*P*=.31, μ=120.47, μ =127.44)	E (*P*=.63, μ=111.12, μ =113.91)	E (*P*=.59, μ=157.19, μ =156.54)	E (*P*=.65, μ=103.69, μ =103.41)	E (*P*=.72, μ=125.55, μ =120.19)
EMG 6	E (*P*=.94, μ=91.72, μ =81.08)	G (*P*=.03, μ=95.96, μ =78.16)	E (*P*=.62, μ=85.03, μ =82.51)	E (*P*=.55, μ=107.30, μ =103.72)	E (*P*=.82, μ=69.98, μ =62.28)	E (*P*=.59, μ=86.06, μ =86.54)
EMG 7	G (*P*=.05, μ=78.80, μ =71.89)	E (*P*=.71, μ=67.00, μ =66.45)	E (*P*=.80, μ=62.31, μ =62.84)	E (*P*=.78, μ=92.30, μ =84.77)	E (*P*=.59, μ=60.38, μ =57.02)	E (*P*=.83, μ=63.38, μ =58.93)
EMG 8	E (*P*=.78, μ=93.34, μ =82.14)	E (*P*=.25, μ=81.55, μ =88.39)	G (*P*=.03, μ=83.20, μ =72.14)	E (*P*=.88, μ=95.61, μ =88.24)	E (*P*=.93, μ=73.04, μ =56.43)	E (*P*=.90, μ=86.01, μ =73.02)

[a]*iEMG:* integrated EMG.

[b]E: statistically equal for both modes; L: statistically lower for mode 2 compared with mode 1; G: statistically greater for mode 2 compared with mode 1.

[c]EMG: electromyography.

Table . Statistical results for mode 2 versus mode 1; metric: T_{act}[a].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
EMG[b] 1	G (*P*=.03, μ=3.50, μ =3.06)	E (*P*=.34, μ=3.19, μ =3.27)	G (*P*=.009, μ=2.96, μ =2.57)	E (*P*=.34, μ=4.52, μ =4.68)	E (*P*=.84, μ=1.12, μ =0.98)	E (*P*=.38, μ=2.81, μ =2.88)
EMG 2	G (*P*=.02, μ=3.06, μ =2.52)	E (*P*=.95, μ=2.56, μ =2.14)	G (*P*=.003, μ=2.79, μ =2.14)	E (*P*=.62, μ=3.77, μ =3.66)	E (*P*=.37, μ=2.58, μ =2.42)	E (*P*=.93, μ=2.94, μ =2.32)
EMG 3	E (*P*=.95, μ=4.74, μ =4.19)	E (*P*=.69, μ=4.27, μ =4.16)	E (*P*=.87, μ=4.15, μ =3.80)	E (*P*=.91, μ=5.43, μ =4.94)	E (*P*=.92, μ=3.52, μ =3.07)	E (*P*=.74, μ=4.64, μ =4.32)
EMG 4	E (*P*=.91, μ=4.57, μ =4.31)	E (*P*=.81, μ=4.36, μ =4.21)	E (*P*=.77, μ=4.08, μ =3.97)	E (*P*=.788, μ=5.30, μ =5.00)	E (*P*=.83, μ=3.88, μ =3.62)	E (*P*=.63, μ=4.54, μ =4.42)
EMG 5	E (*P*=.58, μ=5.07, μ =4.81)	E (*P*=.76, μ=4.81, μ =4.55)	E (*P*=.77, μ=4.58, μ =4.36)	E (*P*=.59, μ=5.60, μ =5.46)	E (*P*=.62, μ=4.09, μ =3.95)	E (*P*=.64, μ=4.76, μ =4.69)
EMG 6	E (*P*=.77, μ=2.51, μ =2.40)	G (*P*=.02, μ=2.40, μ =1.98)	E (*P*=.69, μ=2.16, μ =2.07)	E (*P*=.90, μ=2.91, μ =2.63)	E (*P*=.66, μ=1.62, μ =1.62)	E (*P*=.59, μ=2.32, μ =2.32)
EMG 7	G (*P*=.03, μ=2.34, μ =2.05)	E (*P*=.95, μ=2.32, μ =2.01)	E (*P*=.81, μ=2.11, μ =2.01)	E (*P*=.81, μ=3.20, μ =2.84)	E (*P*=.74, μ=1.88, μ =1.62)	E (*P*=.60, μ=1.94, μ =1.75)
EMG 8	E (*P*=.86, μ=3.05, μ =2.60)	E (*P*=.42, μ=2.59, μ =2.56)	E (*P*=.92, μ=2.60, μ =2.33)	E (*P*=.71, μ=3.05, μ =2.93)	E (*P*=.79, μ=2.14, μ =1.86)	E (*P*=.79, μ=2.78, μ =2.54)

[a]E: statistically equal for both modes; L: statistically lower for mode 2 compared with mode 1; G: statistically greater for mode 2 compared with mode 1.

[b]EMG: electromyography.

XSL•FO
RenderX

Table . Statistical results for mode 2 versus mode 3; metric: *iEMG*[a,b].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
EMG[c] 1	G (*P*<.001, μ=108.35, μ =26.00)	G (*P*<.001, μ=94.66, μ =32.60)	G (*P*<.001, μ=81.19, μ =38.97)	G (*P*<.001, μ=162.51, μ =79.37)	E (*P*=.37, μ=43.54, μ =33.97)	G (*P*<.001, μ=98.71, μ =39.68)
EMG 2	G (*P*<.001, μ=96.60, μ =14.22)	G (*P*<.001, μ=81.20, μ =14.28)	G (*P*<.001, μ=85.67, μ =20.39)	G (*P*<.001, μ=117.61, μ =39.19)	E (*P*=.64, μ=83.28, μ =71.33)	G (*P*<.001, μ=91.02, μ =21.88)
EMG 3	G (*P*<.001, μ=154.21, μ =36.18)	G (*P*<.001, μ=141.66, μ =39.84)	G (*P*<.001, μ=133.83, μ =45.47)	G (*P*<.001, μ=181.94, μ =57.27)	G (*P*<.001, μ=114.52, μ =35.86)	G (*P*<.001, μ=153.98, μ =44.77)
EMG 4	G (*P*<.001, μ=141.42, μ =25.08)	G (*P*<.001, μ=128.04, μ =34.48)	G (*P*<.001, μ=114.66, μ =33.94)	G (*P*<.001, μ=172.11, μ =48.10)	G (*P*<.001, μ=112.52, μ =37.13)	G (*P*<.001, μ=129.44, μ =33.78)
EMG 5	G (*P*<.001, μ=134.83, μ =29.27)	G (*P*<.001, μ=120.47, μ =32.66)	G (*P*<.001, μ=111.12, μ =32.12)	G (*P*<.001, μ=157.19, μ =40.41)	G (*P*<.001, μ=103.69, μ =33.55)	G (*P*<.001, μ=125.55, μ =32.43)
EMG 6	G (*P*<.001, μ=91.72, μ =35.05)	G (*P*<.001, μ=95.96, μ =39.43)	G (*P*<.001, μ=85.03, μ =40.09)	G (*P*<.001, μ=107.30, μ =49.40)	G (*P*<.001, μ=69.98, μ =30.59)	G (*P*<.001, μ=86.06, μ =31.62)
EMG 7	G (*P*<.001, μ=78.80, μ =14.16)	G (*P*<.001, μ=67.00, μ =18.27)	G (*P*<.001, μ=62.31, μ =22.26)	G (*P*<.001, μ=92.30, μ =33.76)	G (*P*<.001, μ=60.38, μ =38.49)	G (*P*<.001, μ=63.38, μ =19.59)
EMG 8	G (*P*<.001, μ=93.34, μ =35.17)	G (*P*<.001, μ=81.55, μ =46.51)	G (*P*<.001, μ=83.20, μ =30.33)	G (*P*<.001, μ=95.61, μ =39.59)	G (*P*<.001, μ=73.04, μ =36.56)	G (*P*<.001, μ=86.01, μ =27.18)

[a]*iEMG:* integrated EMG.

[b]E: statistically equal for both modes; L: statistically lower for mode 2 compared with mode 3; G: statistically greater for mode 2 compared with mode 3.

[c]EMG: electromyography.

Table . Statistical results for mode 2 versus mode 3; metric: t[a].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
EMG[b] 1	G (*P*<.001, μ=3.50, μ =0.88)	G (*P*<.001, μ=3.19, μ =1.16)	G (*P*<.001, μ=2.96, μ =1.27)	G (*P*<.001, μ=4.52, μ =1.68)	L (*P*<.001, μ=1.12, μ =1.23)	G (*P*<.001, μ=2.81, μ =1.26)
EMG 2	G (*P*<.001, μ=3.06, μ =0.39)	G (*P*<.001, μ=2.56, μ =0.38)	G (*P*<.001, μ=2.79, μ =0.63)	G (*P*<.001, μ=3.77, μ =1.19)	E (*P*=.94, μ=2.58, μ =1.60)	G (*P*<.001, μ=2.94, μ =0.66)
EMG 3	G (*P*<.001, μ=4.74, μ =1.14)	G (*P*<.001, μ=4.27, μ =1.20)	G (*P*<.001, μ=4.15, μ =1.38)	G (*P*<.001, μ=5.43, μ =1.51)	G (*P*<.001, μ=3.52, μ =1.10)	G (*P*<.001, μ=4.64, μ =1.41)
EMG 4	G (*P*<.001, μ=4.57, μ =1.02)	G (*P*<.001, μ=4.36, μ =1.30)	G (*P*<.001, μ=4.08, μ =1.22)	G (*P*<.001, μ=5.30, μ =1.41)	G (*P*<.001, μ=3.88, μ =1.23)	G (*P*<.001, μ=4.54, μ =1.18)
EMG 5	G (*P*<.001, μ=5.07, μ =1.10)	G (*P*<.001, μ=4.81, μ =1.19)	G (*P*<.001, μ=4.58, μ =1.15)	G (*P*<.001, μ=5.60, μ =1.34)	G (*P*<.001, μ=4.09, μ =1.19)	G (*P*<.001, μ=4.76, μ =1.18)
EMG 6	G (*P*<.001, μ=2.51, μ =1.01)	G (*P*<.001, μ=2.40, μ =1.15)	G (*P*<.001, μ=2.16, μ =1.12)	G (*P*<.001, μ=2.91, μ =1.31)	G (*P*<.001, μ=1.62, μ =0.83)	G (*P*<.001, μ=2.32, μ =1.07)
EMG 7	G (*P*<.001, μ=2.34, μ =0.38)	G (*P*<.001, μ=2.32, μ =0.48)	G (*P*<.001, μ=2.11, μ =0.65)	G (*P*<.001, μ=3.20, μ =0.94)	E (*P*=.800, μ=1.88, μ =1.05)	G (*P*<.001, μ=1.94, μ =0.56)
EMG 8	G (*P*<.001, μ=3.05, μ =0.96)	G (*P*<.001, μ=2.59, μ =1.24)	G (*P*<.001, μ=2.60, μ =0.91)	G (*P*<.001, μ=3.05, μ =1.15)	G (*P*=.01, μ=2.14, μ =1.24)	G (*P*<.001, μ=2.78, μ =0.98)

[a]E: statistically equal for both modes; L: statistically lower for mode 2 compared with mode 3; G: statistically greater for mode 2 compared with mode 3.

[b]EMG: electromyography.

Ranges of Motion

Apart from positive outcomes, the potential drawbacks in the approach were also spotted. The main areas of those are the ranges of motion actively used by patients during the treatment. One of the 2 most important targets of the treatment, along with increasing muscular activity, is to increase individuals' range of mobility. While comparing modes of treatment provided by

XSL•FO
RenderX

the robot, it can be assumed that the motion patterns differ in terms of pattern for some exercises (particularly exercises 1, 2, and 6; Table 5). Intelligent algorithms used for control generally result in increased shoulder flexion or extension and radioulnar rotation. At the same time, the range of elbow flexion or extension decreases. Theoretically, for patients undergoing therapy of whole extremities, especially neurological, there is a tendency for major problems with the shoulder joint. Moreover, the problem refers specifically to coupled shoulder-elbow flexion motion and ends in unanatomical compensations [1]. As the intelligent mode acts to prevent undesired compensation, the results presented in Table 5 can be treated as a confirmation of the expected system's operation.

The case is more complex in terms of comparing the developed rehabilitation system with the therapy provided by a human. Definitely, the motion patterns differ in terms of patterns across all the exercises (Table 6). However, the pattern of changes is not as clear as the one described for mode 2 compared with mode 1. Therefore, no definite conclusion about the preferred mechanism can be drawn. It was observed that some of the motions were limited by the construction of the device. This is not different from the problems of other similar constructions [43].

Moreover, it can be partially explained by the fact that the robot in the intelligent mode always reacts to the user's intention, while the physiotherapist often operates near the pain limits [44]. The second is aimed at maximizing the effects of treatment by increasing ROM quickly. On the contrary, it cannot be applied to the minimally supervised therapy provided by the device, as the reaction time to pain reflexes may be too long, increasing the risk of soft-tissue damage. Given this, it is necessary to specifically focus on changes in the joint ROMs of patients using the system during further long-term effects investigations. This should enable validating whether the robotized therapy can result in slower progress than human-provided and whether the difference can be mitigated by higher-dose treatment [45].

Table . Statistical results for mode 2 versus mode 1; metric: ROM[a,b].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
JM[c] 1	L (*P*=.005, μ=43.73, μ =49.99)	E (*P*=.13, μ=45.40, μ =49.22)	E (*P*=.29, μ=36.57, μ =41.47)	E (*P*=.55, μ=44.51, μ =45.70)	E (*P*=.78, μ=29.57, μ =29.80)	G (*P*=.006, μ=37.35, μ =38.10)
JM 2	G (*P*<.001, μ=73.30, μ =67.23)	G (*P*<.001, μ=76.04, μ =62.91)	G (*P*<.001, μ=58.20, μ =48.24)	E (*P*=.67, μ=110.51, μ =125.59)	G (*P*<.001, μ=32.71, μ =25.80)	G (*P*=.001, μ=65.78, μ =55.24)
JM 3	E (*P*=.94, μ=73.01, μ =55.31)	E (*P*=.07, μ=52.80, μ =57.95)	E (*P*=.80, μ=40.24, μ =30.61)	E (*P*=.32, μ=183.74, μ =188.28)	E (*P*=.31, μ=13.75, μ =13.21)	E (*P*=.13, μ=42.91, μ =34.33)
JM 4	L (*P*=.002, μ=67.04, μ =75.07)	L (*P*=.02, μ=72.82, μ =78.52)	E (*P*=.23, μ=54.02, μ =55.25)	E (*P*=.50, μ=52.80, μ =52.68)	E (*P*=.23, μ=42.22, μ =43.99)	L (*P*<.001, μ=50.52, μ =60.50)
JM 5	G (*P*=.02, μ=27.54, μ =25.23)	E (*P*=.54, μ=26.73, μ =28.28)	G (*P*=.002, μ=27.12, μ =23.21)	G (*P*<.001, μ=29.78, μ =24.74)	G (*P*<.001, μ=21.22, μ =17.26)	E (*P*=.90, μ=24.93, μ =23.60)

[a]ROM: range of motion.

[b]E: statistically equal for both modes; L: statistically lower for mode 2 compared with mode 1; G: statistically greater for mode 2 compared with mode 1.

[c]JM: joint motion.

XSL•FO
RenderX

Table . Statistical results for mode 2 versus mode 3; metric: ROM[a,b].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
JM[c] 1	G (*P*<.001, μ=43.73, μ =31.43)	G (*P*<.001, μ=45.40, μ =34.29)	L (*P*<.001, μ=36.57, μ =42.67)	G (*P*=.001, μ=44.51, μ =35.82)	G (*P*<.001, μ=29.57, μ =24.45)	L (*P*<.001, μ=37.35, μ =47.25)
JM 2	G (*P*<.001, μ=73.30, μ =47.81)	G (*P*<.001, μ=76.04, μ =60.02)	E (*P*=.53, μ=58.20, μ =58.07)	G (*P*<.001, μ=110.51, μ =99.38)	L (*P*<.001, μ=32.71, μ =52.80)	G (*P*<.001, μ=65.78, μ =51.96)
JM 3	G (*P*<.001, μ=73.01, μ =35.53)	L (*P*=.002, μ=52.80, μ =62.94)	L (*P*<.001, μ=40.24, μ =61.96)	L (*P*<.001, μ=183.74, μ =211.47)	L (*P*<.001, μ=13.75, μ =25.18)	E (*P*=.50, μ=42.91, μ =42.55)
JM 4	L (*P*<.001, μ=67.04, μ =82.23)	L (*P*<.001, μ=72.82, μ =89.31)	E (*P*=.70, μ=54.02, μ =51.39)	L (*P*<.001, μ=52.80, μ =61.28)	G (*P*<.001, μ=42.22, μ =32.10)	G (*P*<.001, μ=50.52, μ =25.33)
JM 5	G (*P*<.001, μ=27.54, μ =15.16)	G (*P*<.001, μ=26.73, μ =18.52)	G (*P*=.001, μ=27.12, μ =22.09)	G (*P*<.001, μ=29.78, μ =20.87)	E (*P*=.10, μ=21.22, μ =22.77)	G (*P*<.001, μ=24.93, μ =21.40)

[a]ROM: range of motion.

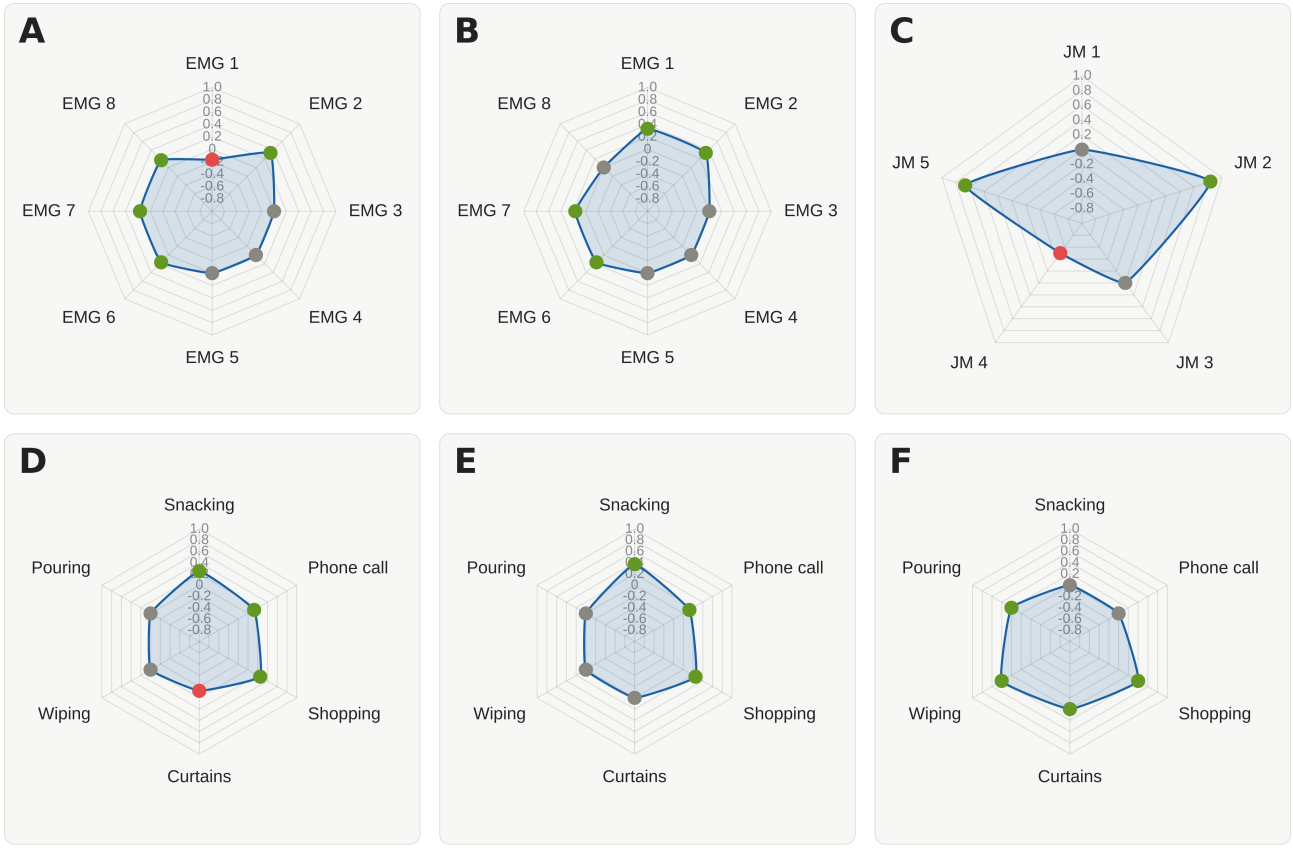
[b]E: statistically equal for both modes; L: statistically lower for mode 2 compared with mode 3; G: statistically greater for mode 2 compared with mode 3.

[c]JM: joint motion

Additionally, the ROM was compared between modes 2 and 1 or 3 for the whole gathered dataset, without distinction on JMs and exercises. The result of statistical analysis comparing mode 2 and 3 rejected the hypothesis that this effectiveness metric for the intelligent algorithm is statistically comparable with or lower than the ones for the human-performed trials (μ=47.48, μ =46.65; *P*<.001). However, analogical hypothesis was accepted for comparison of modes 2 and 1 (μ=47.48, μ =48.13; *P*=.93). Therefore, additional statistical test was performed and simultaneously rejected the hypothesis that this effectiveness metric for the intelligent algorithm is statistically comparable with or higher than the ones for the nonintelligent algorithm (μ=47.48, μ =48.13; *P*=.07). Finally, it is assumed that the ROMs engaged within the treatment in both modes 2 and 1 are comparable at the statistical level. The results of EMG- and JMs-related parameters are visualized in Figures 3A-3F and 4A-4F.

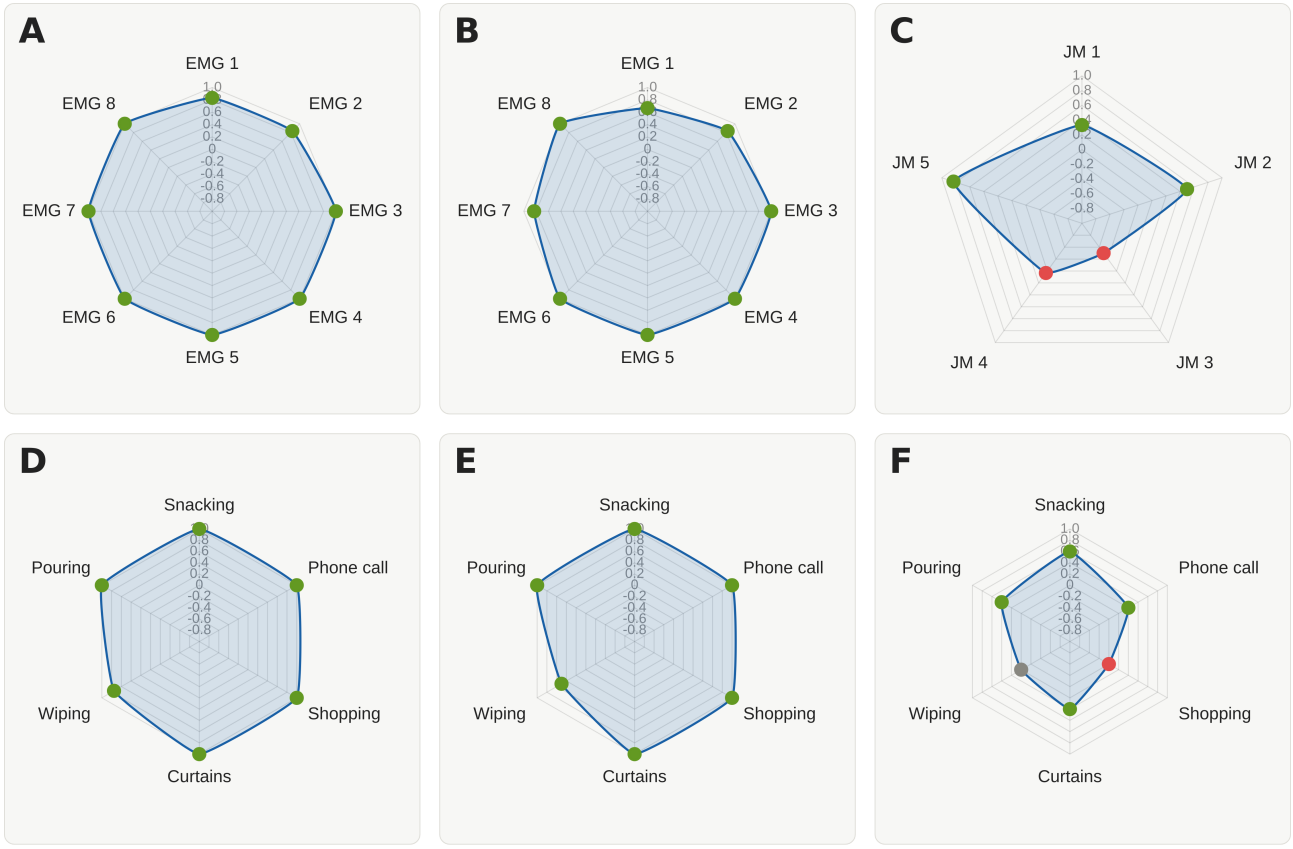
XSL•FO
RenderX

Figure 3. Statistical analysis results for mode 2 versus mode 1 (presented by channel or joint and averaged for all exercises: *iEMG* (A), T_{act} (B), and ROM (C), presented by exercise and averaged for all channels or joints: *iEMG* (D), T_{act} (E), and ROM (F); 1 represents results for mode 2 statistically higher for all channels or exercises; −1 represents results for mode 2 statistically lower for all channels or exercises). EMG: electromyography; iEMG: integrated EMG; JM: joint motion; ROM: range of motion; T_{act}: activation time.



XSL•FO
RenderX

Figure 4. Statistical analysis results for mode 2 versus mode 3 (presented by channel or joint and averaged for all exercises: *iEMG* (A), T_{act} (B), and ROM (C), presented by exercise and averaged for all channels or joints: *iEMG* (D), T_{act} (E), and ROM (F); top; 1 represents results for mode 2 statistically higher for all channels or exercises; −1 represents results for mode 2 statistically lower for all channels or exercises). EMG: electromyography; iEMG: integrated EMG; JM: joint motion; ROM: range of motion; T_{act}: activation time.



Movement Patterns

Additionally, the BDSI metrics were computed for both comparisons between modes. They reflect how similar muscular activation patterns are when performing the same activity. The similarity of engaging muscles for the robot-assisted treatment modes (Table 7) is higher than for the intelligent mode and exercises supported by a physiotherapist (Table 8). The former are typically in the range of 0.65 - 0.8, which is a relatively high result for the different approach [39]. No muscular group nor exercise significantly stands out from the observed pattern.

On the contrary, *BDSI* metrics have significantly more variable values while comparing modes 2 and 3. They range from below 0.5 to 0.8. The first 2 analyzed muscles, the middle and posterior deltoids, differ most in terms of activation when comparing human-provided and intelligent robot-assisted treatment. This is observed across all exercises. Therefore, it is proof that the movements performed during robot-aided treatment differ in terms of mechanisms of muscular control. As this aspect is critical for recalling healthy, functional patterns, further research on long-term effects on patients' independent functionality should be conducted. The BDSI computed for all the trials resulted in values 0.753 and 0.656 while comparing mode 2 with mode 1 and mode 3, respectively.

XSL•FO
RenderX

Table . Statistical results for mode 2 versus mode 1; metric: BDSI[a].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
EMG[b] 1	0.668	0.721	0.711	0.770	0.739	0.813
EMG 2	0.673	0.679	0.664	0.721	0.696	0.705
EMG 3	0.739	0.743	0.753	0.792	0.761	0.741
EMG 4	0.787	0.788	0.791	0.818	0.790	0.760
EMG 5	0.774	0.798	0.793	0.816	0.793	0.727
EMG 6	0.736	0.737	0.736	0.745	0.729	0.722
EMG 7	0.745	0.772	0.750	0.768	0.779	0.750
EMG 8	0.746	0.797	0.780	0.797	0.788	0.752

[a]BDSI: Burst Duration Similarity Index.

[b]EMG: electromyography.

Table . Statistical results for mode 2 versus mode 3; metric: BDSI[a].

	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	Exercise 6
EMG[b] 1	0.569	0.525	0.492	0.567	0.515	0.535
EMG 2	0.605	0.638	0.618	0.565	0.611	0.506
EMG 3	0.686	0.697	0.680	0.718	0.696	0.636
EMG 4	0.740	0.755	0.744	0.727	0.767	0.743
EMG 5	0.760	0.783	0.765	0.783	0.768	0.759
EMG 6	0.623	0.565	0.587	0.584	0.587	0.607
EMG 7	0.717	0.703	0.705	0.642	0.719	0.609
EMG 8	0.692	0.661	0.679	0.657	0.685	0.599

[a]BDSI: Burst Duration Similarity Index.

[b]EMG: electromyography.

Discussion

Principal Findings

The results of this study demonstrate that the proposed rehabilitation robotic system constitutes an effective tool for task-oriented neurorehabilitation, with advantages over conventional therapy in several physiotherapeutic relevant aspects. First, the previously observed increase in active patient engagement indicates that the system promotes voluntary motor involvement rather than passive movement execution. By dynamically adapting assistance based on task execution quality, compensatory behaviors, and functional errors, the robot facilitates motor learning, a fundamental process for functional recovery. The biomechanical profile observed during intelligent robot-assisted therapy is consistent with the theoretical requirements of neuroplastic motor rehabilitation—the system provides an adjustable assistance that preserves voluntary motor demand rather than replacing it. This is reflected in the pattern of *iEMG* and T_{act} results, where mode 2 produces at least equivalent muscular engagement to mode 1, while giving the patient greater autonomy over the motion trajectory. From a motor learning standpoint, this configuration aligns with the paradigm, in which task-relevant errors are tolerated up to the threshold of compensation, and assistance is reserved for preventing significant functional or hazardous anatomical compensations [46].

The comparison between control modes 1 and 2 revealed limited statistically significant differences in overall muscular participation, indicating that both modes effectively support active engagement of the musculoskeletal system. However, for some muscular groups and exercises, an intelligent approach outperforms the constant-admittance system. Nevertheless, results suggest that the intelligent control strategy primarily benefits by increasing patients' volunteer participation and not by increasing total muscle work. The absence of observable statistical *iEMG* differences between modes 1 and 2 at the per-channel level should not be interpreted as a negative effect. The intelligent mode achieves comparable muscular engagement with a fundamentally different control philosophy by allowing voluntary motion and intervening only contextually.

Compared with conventional physiotherapy, robot-assisted training resulted in significantly greater muscular participation across the analyzed muscle groups. This finding indicates that the robotic system can elicit more consistent, task-relevant muscle activation. Moreover, the use of robots reduces therapist-dependent variability and patient passivity that may occur during manual therapy. The higher mean of *iEMG* values observed in mode 2 versus mode 3 requires clinical

XSL•FO
RenderX

contextualization. This most likely reflects a combination of 2 mechanisms—the longer robot-aided exercise duration, along with the postural demand imposed by the exoskeleton's mechanical constraints. This is evidenced by persistently high activation of the latissimus dorsi and pectoralis major across exercises in modes 1 and 2, consistent with their roles in trunk and proximal stabilization. Whether this additional muscular demand is therapeutically beneficial should be verified in longitudinal clinical trials planned as the follow-up to this investigation.

Notably, distinct motion patterns were observed between the different robot control modes. These changes suggest that the robot not only modifies the level of assistance but also influences patients' movement strategy, which is a critical aspect of recalling motor capabilities. However, joint ROMs during robot-assisted therapy were generally lower than those mobilized in conventional treatment. While reduced ROM could be interpreted as a limitation, it may also reflect a more controlled, functionally relevant movement execution, thereby reducing compensatory strategies. This is particularly the case while comparing mode 1 with mode 2. From a clinical perspective, the reduction in elbow flexion or extension within mode 2 in the context of increased shoulder flexion or extension may be clinically favorable for patients poststroke, in whom pathological shoulder-elbow flexion synergy is a primary recovery obstacle [1]. If the intelligent algorithm's compensation detection is effectively preventing this pattern, the ROM changes observed here may represent a functionally meaningful impact, which should be verified with the standardized functional assessment within longitudinal trials. In conventional human-provided treatment, this is due to physiotherapists operating closer to the pain threshold.

The *BDSI* values for mode 2 with mode 1 indicate high temporal similarity between the 2 robotic modes. The intelligent algorithm alters assistance intensity contextually but preserves the fundamental activation sequencing established by the constant-admittance reference. On the contrary, analyses of EMG synchronization during exercises confirmed significantly altered muscular activation patterns in EMG 1 and EMG 2 during mode 2 compared with conventional muscular therapy. These changes indicate a redistribution of muscle activity toward more correct motions in terms of their functionality. On the other hand, more significant changes between mode 2 and mode 3 might result from possible motion limitations in robotic devices, which are typical for such structures. Overall, the robotic system proved to be an effective therapeutic tool, reducing compensations and functional errors commonly performed by patients, while maintaining active participation and movement quality.

Nevertheless, several confounding factors should be acknowledged when interpreting the mode 2 versus mode 3 comparison: exercise duration differed systematically between conditions, with physiotherapist-guided sessions being consistently shorter; the level of physical guidance provided by the physiotherapist varied across participants and sessions; the mechanical constraints of the exoskeleton may have restricted some degrees of freedom not constrained during manual therapy; participant heterogeneity may have introduced variability not fully accounted for in the group-level analyses. However, mode 3 and recording of the reference trajectories for mode 1 and mode 2 were performed by the same specialist for all the cases. This aimed to minimize variability from different assistance during treatment from an individual work approach. Despite this, the mentioned factors are acknowledged as limitations and will be addressed in future controlled trials.

Comparison With Other Works

The obtained results are consistent with and extend previous research on variable admittance control in rehabilitation robotics. Studies comparing variable admittance, which correlated mainly with synergic joint kinematics, with constant admittance have demonstrated increased patient participation, although often with similar EMG effort distributions across conditions [19]. In contrast, our approach to variable admittance, explicitly linked to trajectory deviations, compensatory movements, and functional errors, yielded improved outcomes in selected cases, particularly in movement quality and muscle activation patterns.

Other studies investigating AAN strategies regulated through muscle entropy have reported positive effects primarily in patients with mild impairments (Fugl-Meyer scores between 58 and 65) [21]. In this study, the proposed method was evaluated in a heterogeneous group of participants, and no significant differences in impairment severity were observed, suggesting robustness across varying functional levels.

Furthermore, recent meta-analyses have demonstrated that EMG-based robot-assisted rehabilitation is significantly superior to conventional rehabilitation techniques [47]. Notably, the results obtained in this study are superior, despite the control system not relying on EMG signals for real-time control. This advantage likely arises from prioritizing user intention during treatment, with robotic intervention applied only when hazardous errors or unacceptable compensations occur. This strategy reduces unnecessary assistance and supports autonomous motor control, aligning with contemporary principles of neuroplastic motor rehabilitation.

Limitations and Planned Continuation

Several limitations of this study should be acknowledged. First, the experimental protocol was limited to a single-session trial, precluding conclusions about long-term therapeutic effects or functional outcome improvements. Therefore, it has also not included a standardized functional pre- or postassessment. Future studies will therefore focus on longitudinal assessments to evaluate sustained motor recovery and gained independence in performing ADLs. This is planned as a follow-up with outcome measures based on the Fugl-Meyer Assessment and Action Research Arm Test.

Second, the relatively small and heterogeneous participant group limits the further applicability and strength of the findings. Future research will focus specifically on neurological patient populations, enabling more targeted analysis of impairment-specific responses to the therapy.

From a technical perspective, limitations related to the mobility of the robotic structure were identified. Planned construction modifications, including removing the developed open bearings,

XSL•FO
RenderX

aim to reduce friction and limit the bulkiness of the device. As a result, patients should be able to perform ADLs in almost all natural ROM.

Finally, differences in exercise duration observed in mode 3 highlight the influence of therapist involvement and potential fatigue on the number of repetitions for exercises performed. Future research will investigate how physiotherapist fatigue affects therapy delivery in real-life clinical settings and whether increased treatment dosage or high-intensity protocols can compensate for longer execution times.

Final Conclusions

The findings of this study confirm that task-oriented rehabilitation can be effectively supported by robotic systems, leading to improved patient engagement, reduced compensatory behaviors, and enhanced movement quality. The proposed intelligent control approach benefits from context-based assessment of task execution, enabling physiotherapists to design individualized exercises that reflect any ADLs. As such, rehabilitation robots have the potential not only to augment conventional therapy but also to provide structured, repeatable, and fully automatic therapy. Thanks to this, the capabilities of physiotherapists can be multiplied, and patients can receive the required care quickly and in higher doses.

Acknowledgments

The experiments were carried out under the KB/132/2024 approval of the Bioethical Committee of the Medical University of Warsaw. Written informed consent was obtained from all of the participants involved in this study. The authors declare the use of Grammarly EDU in the writing process for language corrections and Claude Sonnet 4.6 for creating interactive visuals with charts. Responsibility for the final manuscript lies entirely with the authors. Generative AI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This paper is based on the results of the “Development of a universal and lightweight construction of rehabilitation exoskeleton with a control algorithm dedicated to remote, home and task-oriented rehabilitation” SmartEx-Home project, financed in 2024 - 2026 (1,793,900 PLN), within the scope of scientific research and development works by the National Center for Research and Development (LIDER XIV, contract number LIDER14/0196/2023).

Data Availability

The medical data gathered for analyses have not been made available. The remaining data can be shared upon request.

Authors' Contributions

Conceptualization: PF
Methodology: PF, KJ, TO, DŚ
Software: MP, JO, KZ, KJ
Mechanical construction of the setup: PF, JO, AZ, PK, KZ
Validation: PF, TO, DŚ
Formal analysis: PF
Investigation: PF, MP, TO, KJ, DŚ
Resources: PF
Data curation: TO, PF, KJ
Writing – original draft preparation: PF, MP, TO, KJ
Writing – review and editing: PF, MP, TO, KJ
Visualization: PF, KJ
Supervision: PF, DŚ
Project administration: PF
Funding acquisition: PF

Conflicts of Interest

None declared.

Checklist 1
GUIDED checklist.
[PDF File, 282 KB - rehab_v13i1e92898_app1.pdf]

References

XSL•FO
RenderX

1. Raghavan P. Upper limb motor impairment after stroke. Phys Med Rehabil Clin N Am 2015 Nov;26(4):599-610. [doi: 10.1016/j.pmr.2015.06.008] [Medline: 26522900]
2. Krug E, Cieza A. Strengthening health systems to provide rehabilitation services. Am J Phys Med Rehabil 2017 Jun;96(6):438-439. [doi: 10.1097/PHM.0000000000000753] [Medline: 28520593]
3. Cornwell L, Doyle H, Stohner M, Hazle C. Work-related musculoskeletal disorders in physical therapists attributable to manual therapy. J Man Manip Ther 2021 Apr;29(2):92-98. [doi: 10.1080/10669817.2020.1793470] [Medline: 32697156]
4. Halford E, Jakubiszak S, Krug K, Umphress A. What is task-oriented training? A scoping review. Scand J Occup Ther 2024;4(1):1-23. [doi: 10.46409/001.DPYW4980]
5. Banyai AD, Bri an C. Robotics in physical rehabilitation: systematic review. Health Care (Don Mills) 2024 Aug;12(17):1720. [doi: 10.3390/healthcare12171720]
6. Mashud G, Hasan SK, Alam N. Advances in control techniques for rehabilitation exoskeleton robots: a systematic review. Actuators 2025;14(3):108. [doi: 10.3390/act14030108]
7. van Westerhuis C, Sanders AF, Aarden JJ, et al. Capabilities for using telemonitoring in physiotherapy treatment: exploratory qualitative study. JMIR Rehabil Assist Technol 2024 Oct 24;11(1):e56432. [doi: 10.2196/56432] [Medline: 39447150]
8. Reicherzer L, Scheermesser M, Kläy A, Duarte JE, Graf ES. Barriers and facilitators to the use of wearable robots as assistive devices: qualitative study with older adults and physiotherapists. JMIR Rehabil Assist Technol 2024 Aug 9;11(1):e52676. [doi: 10.2196/52676] [Medline: 39132691]
9. Jarrassé N, Proietti T, Crocher V, et al. Robotic exoskeletons: a perspective for the rehabilitation of arm coordination in stroke patients. Front Hum Neurosci 2014;8:947. [doi: 10.3389/fnhum.2014.00947] [Medline: 25520638]
10. Cao J, Zhang J, Wang C, et al. Variable admittance control of high compatibility exoskeleton based on human–robotic interaction force. Chin J Mech Eng 2024;37(1):119. [doi: 10.1186/s10033-024-01113-6]
11. He B, Zhou X, Huai Y, et al. Admittance control based hemiplegic exoskeleton motion control. Presented at: 2022 IEEE International Conference on Real-time Computing and Robotics (RCAR); Jul 17-22, 2022; Guiyang, China p. 331-336. [doi: 10.1109/RCAR54675.2022.9872219]
12. Ranatunga I, Lewis FL, Popa DO, Tousif SM. Adaptive admittance control for human–robot interaction using model reference design and adaptive inverse filtering. IEEE Trans Contr Syst Technol 2017;25(1):278-285. [doi: 10.1109/TCST.2016.2523901]
13. Mujica M, Crespo M, Benoussaad M, Junco S, Fourquet JY. Robust variable admittance control for human–robot co-manipulation of objects with unknown load. Robot Comput Integr Manuf 2023 Feb;79:102408. [doi: 10.1016/j.rcim.2022.102408]
14. Yu W, Rosen J, Li X. PID admittance control for an upper limb exoskeleton. 2011 Presented at: Proceedings of the 2011 American Control Conference; Jun 29 to Jul 1, 2011. [doi: 10.1109/ACC.2011.5991147]
15. Miller LM, Roldan JR, Rosen J, Kim H, Li Z. Admittance control of an upper limb exoskeleton—reduction of energy exchange. 2012 Presented at: 2012 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC; Aug 28 to Sep 1, 2012. [doi: 10.1109/EMBC.2012.6347475]
16. Aguirre-Ollinger G, Nagarajan U, Goswami A. An admittance shaping controller for exoskeleton assistance of the lower extremities. Auton Robot 2016 Apr;40(4):701-728. [doi: 10.1007/s10514-015-9490-8]
17. Kim H, Yang W. Variable admittance control based on human–robot collaboration observer using frequency analysis for sensitive and safe interaction. Sensors 2021;21(5):1899. [doi: 10.3390/s21051899]
18. Topini A, Sansom W, Secciani N, Bartalucci L, Ridolfi A, Allotta B. Variable admittance control of a hand exoskeleton for virtual reality-based rehabilitation tasks. Front Neurorobot 2021;15:789743. [doi: 10.3389/fnbot.2021.789743] [Medline: 35095457]
19. Zarrin RS, Zeiaee A, Langari R. A variable-admittance assist-as-needed controller for upper-limb rehabilitation exoskeletons. IEEE Robot Autom Lett 2024;9(6):5903-5910. [doi: 10.1109/LRA.2024.3398565]
20. Naghavi N, Akbarzadeh A, Tahamipour-Z. SM, Kardan I. Assist-as-needed control of a hip exoskeleton based on a novel strength index. Rob Auton Syst 2020 Dec;134:103667. [doi: 10.1016/j.robot.2020.103667]
21. Arantes AP, Bressan N, Borges LR, McGibbon CA. Evaluation of a novel real-time adaptive assist-as-needed controller for robot-assisted upper extremity rehabilitation following stroke. PLoS One 2023;18(10):e0292627. [doi: 10.1371/journal.pone.0292627] [Medline: 37819932]
22. Aurich-Schuler T, Grob F, van Hedel HJA, Labruyère R. Can Lokomat therapy with children and adolescents be improved? An adaptive clinical pilot trial comparing guidance force, path control, and FreeD. J Neuroeng Rehabil 2017 Jul 14;14(1):76. [doi: 10.1186/s12984-017-0287-1] [Medline: 28705170]
23. Lian P, He Y, Ma Y, Liu J, Wu X. Adaptive admittance control of human-exoskeleton system using RNN optimization. 2021 Presented at: 2021 IEEE International Conference on Real-time Computing and Robotics (RCAR); Jul 15-19, 2021; Xining, China p. 584-589. [doi: 10.1109/RCAR52367.2021.9517398]
24. Petrič T, Peternel L, Morimoto J, Babič J. Assistive arm-exoskeleton control based on human muscular manipulability. Front Neurorobot 2019;13:30. [doi: 10.3389/fnbot.2019.00030] [Medline: 31191289]

XSL•FO
RenderX

25. Moubarak S, Pham MT, Moreau R, Redarce T. Gravity compensation of an upper extremity exoskeleton. 2010 Presented at: Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Aug 31 to Sep 4, 2010; Buenos Aires, Argentina p. 4489-4493. [doi: 10.1109/IEMBS.2010.5626036]
26. Zhang L, Yu H, Li D. Reducing upper-limb muscle effort with model-based gravity compensation during robot-assisted movement. Sensors 2025;25(10):3032. [doi: 10.3390/s25103032]
27. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
28. Falkowski P, Kołodziejski P, Zawalski K, et al. Measurement-based optimization of a lightweight upper-extremity rehabilitation exoskeleton for task-oriented treatment. Sensors (Basel) 2026 Mar 15;26(6):1849. [doi: 10.3390/s26061849] [Medline: 41902019]
29. Falkowski P, Zawalski K. Multibody dynamics simulation of upper extremity rehabilitation exoskeleton during task-oriented exercises. Actuators 2025;14(9):426. [doi: 10.3390/act14090426]
30. Falkowski P, Leczkowski B. The design of 3D-printed open bearings for human assisting robots. J Autom Mobile Rob Intell Syst 2025;19(3):13-24. [doi: 10.14313/jamris-2025-021]
31. Falkowski P, Pikuliński M, Osiak T, et al. Similarity analysis of upper extremity's trajectories in activities of daily living for use in an intelligent control system of a rehabilitation exoskeleton. Actuators 2025;14(7):324. [doi: 10.3390/act14070324]
32. Falkowski P, Pikuliński M, Osiak T, Jeznach K, Zawalski K, Kołodziejski P, et al. Analysis of anatomical compensations of activities of daily living for automatic exoskeleton-aided task-oriented physiotherapy. In: Lecture Notes in Networks and Systems: Springer; 2026, Vol. 1687:165-174. [doi: 10.1007/978-3-032-08359-3_15]
33. Sánchez-Moreno J, Moreno-Salinas D, Álvarez-Ortiz JC. Use of inertial measurement units for detection of the support phases in discus throwing. Sensors (Basel) 2025 Oct 3;25(19):6095. [doi: 10.3390/s25196095] [Medline: 41094918]
34. Konrad P. The ABC of EMG: a practical introduction to kinesiological electromyography. Noraxon Inc. 2006. URL: https://www.noraxon.com/article/abcs-of-emg-pdf/ [accessed 2026-02-04]
35. Stegeman D, Hermens H. Standards for surface electromyography: the European project surface EMG for non-invasive assessment of muscles (SENIAM). Roessingh Research and Development. 2007. URL: https://www.researchgate.net/publication/228486725 [accessed 2026-02-04]
36. EMG signal processing: key techniques and practical recommendations. Noraxon. URL: https://www.noraxon.com/article/emg-signal-processing-key-techniques-and-practical-recommendations/ [accessed 2026-02-04]
37. Rodriguez-Tapia B, Soto I, Martinez DM, Arballo NC. Myoelectric interfaces and related applications: current state of EMG signal processing—a systematic review. IEEE Access 2020;8:7792-7805. [doi: 10.1109/ACCESS.2019.2963881]
38. Burden A. How should we normalize electromyograms obtained from healthy participants? What we have learned from over 25 years of research. J Electromyogr Kinesiol 2010 Dec;20(6):1023-1035. [doi: 10.1016/j.jelekin.2010.07.004] [Medline: 20702112]
39. Androwis GJ, Pilkar R, Ramanujam A, Nolan KJ. Electromyography assessment during gait in a robotic exoskeleton for acute stroke. Front Neurol 2018;9:630. [doi: 10.3389/fneur.2018.00630] [Medline: 30131756]
40. Johanson ME, Radtka SA. Amplitude threshold criteria improve surface electrode specificity during walking and functional movements. Gait Posture 2006 Dec;24(4):429-434. [doi: 10.1016/j.gaitpost.2005.09.011] [Medline: 16423526]
41. Stojanović M, Andjelković - Apostolović M, Milošević Z, Ignjatović A, Public Health Institute Niš, Center for Informatics and Biostatistics in Health Care, Niš, Serbia, Public Health Institute Niš, Center for Informatics and Biostatistics in Health Care, Niš, Serbia. Parametric versus nonparametric tests in biomedical research. Acta Med Median 2018 Mar 15;57(2):75-80. [doi: 10.5633/amm.2018.0212]
42. Turner DL, Ramos-Murguialday A, Birbaumer N, Hoffmann U, Luft A. Neurophysiology of robot-mediated training and therapy: a perspective for future use in clinical populations. Front Neurol 2013 Nov 13;4:184. [doi: 10.3389/fneur.2013.00184] [Medline: 24312073]
43. Lee SH, Park G, Cho DY, et al. Comparisons between end-effector and exoskeleton rehabilitation robots regarding upper extremity function among chronic stroke patients with moderate-to-severe upper limb impairment. Sci Rep 2020;10(1). [doi: 10.1038/s41598-020-58630-2]
44. Dias D, Neto MG, Sales SDS, et al. Effect of mobilization with movement on pain, disability, and range of motion in patients with shoulder pain and movement impairment: a systematic review and meta-analysis. J Clin Med 2023 Nov 29;12(23):7416. [doi: 10.3390/jcm12237416] [Medline: 38068468]
45. Hsieh YW, Wu CY, Lin KC, Yao G, Wu KY, Chang YJ. Dose-response relationship of robot-assisted stroke motor rehabilitation: the impact of initial motor status. Stroke 2012 Oct;43(10):2729-2734. [doi: 10.1161/STROKEAHA.112.658807] [Medline: 22895994]
46. Levac D, Driscoll K, Galvez J, Mercado K, O'Neil L. OPTIMAL practice conditions enhance the benefits of gradually increasing error opportunities on retention of a stepping sequence task. Hum Mov Sci 2017 Dec;56(Pt B):129-138. [doi: 10.1016/j.humov.2017.10.017] [Medline: 29128736]
47. Huo Y, Wang X, Zhao W, Hu H, Li L. Effects of EMG-based robot for upper extremity rehabilitation on post-stroke patients: a systematic review and meta-analysis. Front Physiol 2023;14:1172958. [doi: 10.3389/fphys.2023.1172958] [Medline: 37256069]

XSL•FO
RenderX

Abbreviations

AAN: assist-as-needed
ADL: activity of daily living
BDSI: Burst Duration Similarity Index
EMG: electromyography
IMU: inertial moment unit
JM: joint motion
ROM: range of motion

Edited by S Munce; submitted 04.Feb.2026; peer-reviewed by D Mosconi, M Chakit; revised version received 18.Jun.2026; accepted 22.Jul.2026; published 01.Sep.2026.

Please cite as:
Falkowski P, Jeznach K, Oleksiuk J, Osiak T, Pikuliński M, Zakręcki A, Zawalski K, Kołodziejski P, Śliż D
Intelligent Robot-Aided, Task-Oriented Physiotherapy Method for Upper Limb Rehabilitation: Development and Usability Study
JMIR Rehabil Assist Technol 2026;13:e92898
URL: https://rehab.jmir.org/2026/1/e92898
doi:10.2196/92898

© Piotr Falkowski, Kajetan Jeznach, Jan Oleksiuk, Tomasz Osiak, Maciej Pikuliński, Andrzej Zakręcki, Krzysztof Zawalski, Piotr Kołodziejski, Daniel Śliż. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 1.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Scientific and Technological Developments in Brain-Computer Interfaces: Dual Bibliometric Analysis of Articles and Patents

Junhui Wang*, PhD; Yakun Yuan*, MSc; Hang Xu*, MSc; Guannan Luan*, PhD

Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, 3 Yabao Road, Chaoyang District, Beijing, China

*all authors contributed equally

Corresponding Author:
Guannan Luan, PhD
Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, 3 Yabao Road, Chaoyang District, Beijing, China

Abstract

Background: Brain-computer interface (BCI) technology is undergoing rapid translation from laboratory research to clinical applications, heralding a fundamental restructuring of human-machine relationships, with profound societal implications. Existing bibliometric analyses of this domain have exclusively relied on scholarly article databases, critically overlooking patent data that is essential for capturing the full spectrum of technological innovation in this highly translational field.

Objective: This study aims to investigate the overall scientific and technological trajectories, key research drivers (journals, institutions, countries, and funding agencies), as well as research topics and trends in the BCI field through a comprehensive analysis of scientific and technical literature (articles and patents).

Methods: BCI-related articles (11,346) and granted patents (1551) published from 2015 to 2025 were retrieved from the Web of Science (WOS) and the incoPat database, respectively. VOSviewer, CiteSpace (Dr Chaomei Chen), Microsoft Excel, and InCites (Clarivate) were used to summarize bibliometric features. Additionally, the Disruptive Index was calculated to characterize the developmental trajectory of scientific and technological advancements in BCIs.

Results: The number of BCI articles and patents has been continuously increasing over the past decade. The *Journal of Neural Engineering* published the largest number of BCI articles. The Chinese Academy of Sciences ranked first in article output, while the University of California System achieved the greatest citation impact, and Tianjin University from China led in patent filings. At the country level, China dominated article output and patent filings, whereas the United States attained the highest article citation impact and the most extensive international patent portfolios. The National Natural Science Foundation of China (NSFC) was the most prolific funding agency for BCI articles. Disruptive Index analysis revealed that BCI scientific research (article-based) maintained sustained growth in disruptiveness, whereas technological development (patent-based) demonstrated a pattern of fluctuation rather than consistent growth during the observation period. A total of 5 major research topics were identified, with neural interfaces and motor control attracting the greatest attention. In parallel, the leading technology categories were computer input or output interface devices (IPC: G06F3) and diagnostic measurement and human identification (IPC: A61B5). Citation analysis revealed an average knowledge transfer lag of 8.8 years, alongside limited bidirectional article-patent linkages.

Conclusions: This study reveals a rapidly expanding yet strategically differentiated global BCI landscape dominated by China and the United States, with the former leading in output volume and the latter achieving the highest citation impact and international patent portfolios. BCI scientific research maintains active disruptive potential, whereas technological development lacks a commensurate upward trajectory; this asynchrony, compounded by prolonged knowledge transfer lag and weak article-patent linkages, points to translational challenges that require strengthened mechanisms for converting scientific discovery into technical innovations.

(JMIR Rehabil Assist Technol 2026;13:e95902) doi:10.2196/95902

KEYWORDS

Bibliometric; brain–computer interface; topic analysis; category normalized citation impact; disruptive index

Introduction

Brain-computer interface (BCI) has undergone substantial development over recent decades, establishing itself as a multidisciplinary domain spanning neuroscience, engineering, computer science, and rehabilitation medicine. Pioneering experiments, exemplified by Delgado's remote brain stimulation via a "stimoceiver" [1], contributed to the conceptual origins of modern BCI research. BCI technology initially developed as a tool for elucidating neural mechanisms, subsequently enabling

XSL•FO
RenderX

direct communication between the brain and external devices by translating neural activity into commands for controlling computers, prosthetic limbs, or other assistive systems, and later advancing to bidirectional interfaces that both record neural signals and deliver sensory feedback, enabling closed-loop systems capable of restoring motor function in paralyzed individuals through bypassing damaged neural pathways [2].

The field is currently experiencing extensive integration of AI, including deep learning and neural networks, to enhance BCI performance and customize systems to individual users [3,4]. These advances are driving the technology's evolution from basic scientific experiments to versatile applications spanning medical rehabilitation, assistive devices, and entertainment. Continued advances in AI, coupled with sustained interdisciplinary collaboration, are anticipated to enhance both the functionality and accessibility of BCI systems.

Bibliometric analysis, as a mature quantitative methodology, enables systematic investigation of research participants, thematic evolution, and developmental trajectories within specific fields [5,6], and has been increasingly applied to the BCI field [7-15]. A state-of-the-art literature review reveals that existing bibliometric studies of this domain have predominantly focused on specific applications or subfields. Examples include disease-specific investigations such as stroke [7,8], spinal cord injury [9], and Alzheimer disease [10]; signal processing approaches including electroencephalography [11], motor imagery, and steady-state visually evoked potentials [12]; and specialized subfields such as BCI software platforms [13], emotion research [14], and brain-computer music interfaces [15].

A limited number of studies have examined the overall development of the BCI field using bibliometric methods. Hu et al [16] analyzed 100 highly cited BCI papers published from 2006 - 2015, while Yin et al [2] provided a comprehensive overview of BCI research from 1990 to 2020. However, all aforementioned studies relied exclusively on article data from the Web of Science (WOS) Core Collection, PubMed, or Scopus, with no incorporation of patent data. Given the pronounced technological orientation of the BCI field, patent analysis constitutes an essential component for understanding its comprehensive development trajectory. To address this gap, this study adopts a dual-source approach integrating both articles and patents to provide a holistic characterization of scientific and technological development in the BCI domain.

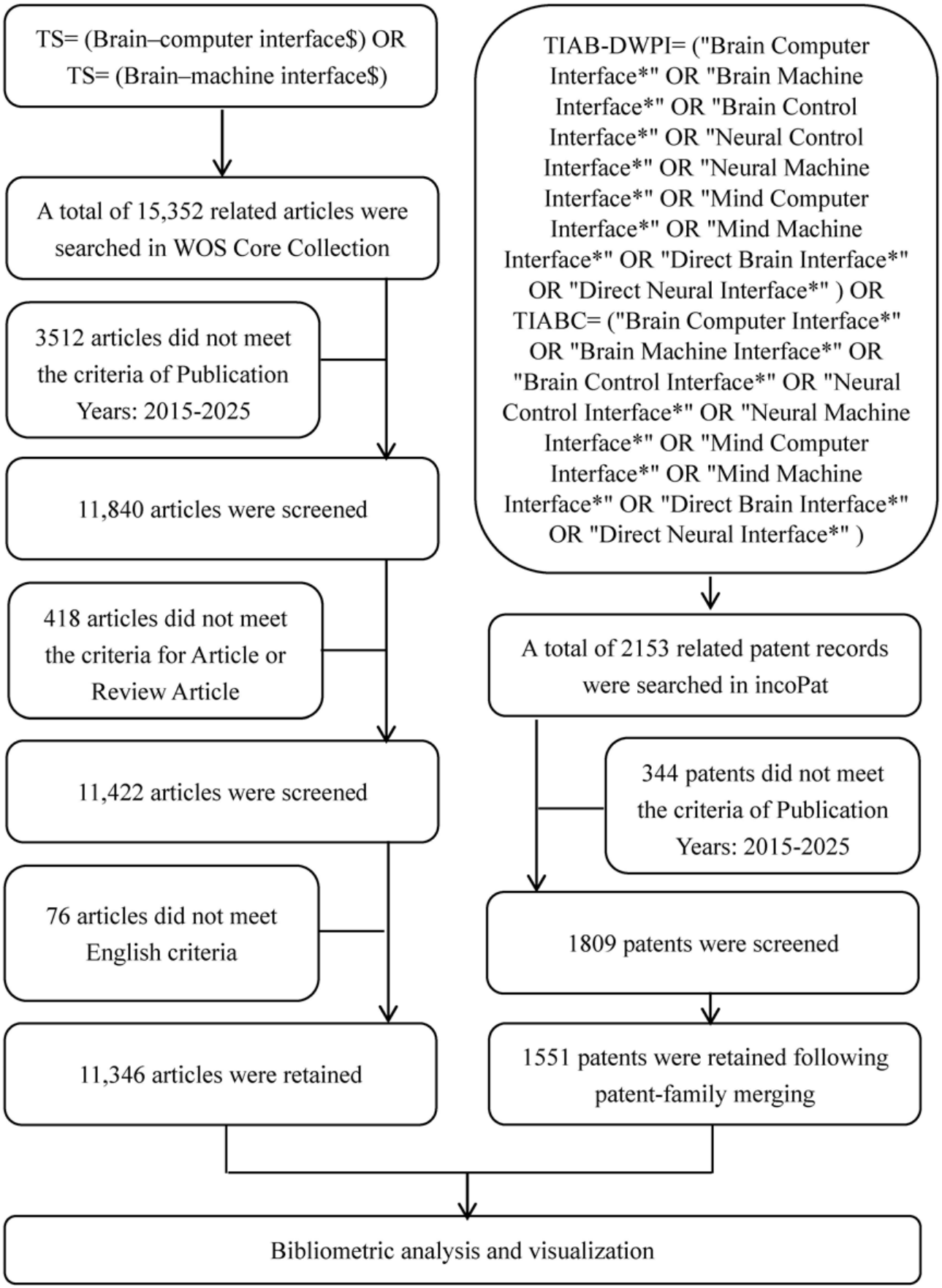
Methods

Data Source and Search Strategy

This study uses both articles and patents in the field of BCI as complementary analytical data sources, with the publication timeframe restricted to 2015-2025. Specifically, article data were retrieved from the WOS Core Collection database, with the search executed on February 1, 2026, yielding a total of 11,346 BCI-related articles. Patent data were exclusively sourced from the incoPat database (Clarivate) and limited to granted patents to ensure the focus on legally protected innovations [17]; the search was conducted on June 1, 2026, retrieving 1809 patents that were later reduced to 1551 records via patent-family merging (grouping documents from identical simple patent families). The search strategies for both articles and patents are detailed in Figure 1.

XSL•FO
RenderX

Figure 1. Flowchart of articles and patents screening.



XSL•FO
RenderX

Bibliometric Analysis

Bibliometric mapping and visualization were conducted using VOSviewer (version 1.6.20; Leiden University) and CiteSpace (version 6.3.1; Dr Chaomei Chen). VOSviewer generated keyword cooccurrence networks with color-coded clusters, where internode distances indicate thematic relatedness, sizes reflect occurrence frequency, and density represents connection strength—enabling identification of research trends and hotspots. CiteSpace detected burst keywords as indicators of cutting-edge topics, visualized along a horizontal blue timeline with red segments denoting the start year, end year, and duration of each keyword [18]. Additionally, relevant descriptive analysis charts were created using Microsoft Excel (version 2021) and InCites (Clarivate).

Bidirectional citation analysis was employed to explore the translational dynamics from scientific research to technological innovation. On one hand, "backward citation analysis" was conducted at the patent level to identify granted patents that cite scientific articles, with simultaneous analysis of the time lag between patent grant years and the publication years of the cited papers. On the other hand, "forward citation analysis" was conducted at the article level to identify highly cited papers (HCPs; top 1% by citations for their category) that receive patent citations. To effectively capture the citation relationships between patents and articles, the Lens database was used as a supplementary source [19], which provides comprehensive cross-reference information for both articles and patents, in addition to the incoPat database.

Disruptive Index

The Disruptive Index, first proposed by Funk and Smith [20], has emerged as a crucial bibliometric metric for evaluating the development trend of science and technology. This innovative metric helps researchers, policymakers, and institutions to assess how groundbreaking new studies change the direction of science and technology. The core logic of the Disruption Index is to characterize the consolidating or disruptive nature of science and technology by analyzing the changes in the citation network of articles and patents. The Disruptive Index of a focal work (article or patent) is calculated as:

$$D=n_i-n_j+n_i+n_j+n_k$$

Where n_i is the number of subsequent works that cite the focal work, n_j is the number of subsequent works that cite both the focal work and its references, and n_k is the number of subsequent works that only cite the focal work's references. To ensure comparability with prior research, this study used the Five-Year Disruptive Index (CD5) proposed by Park et al [21], restricting citation counts to the five-year window following each focal work. Consequently, CD5 values were calculated only for publications up to 2020, ensuring complete observation periods. To assess the current state and trajectory of scientific and technological development in the field of BCI, the CD5 was computed separately for articles and patents.

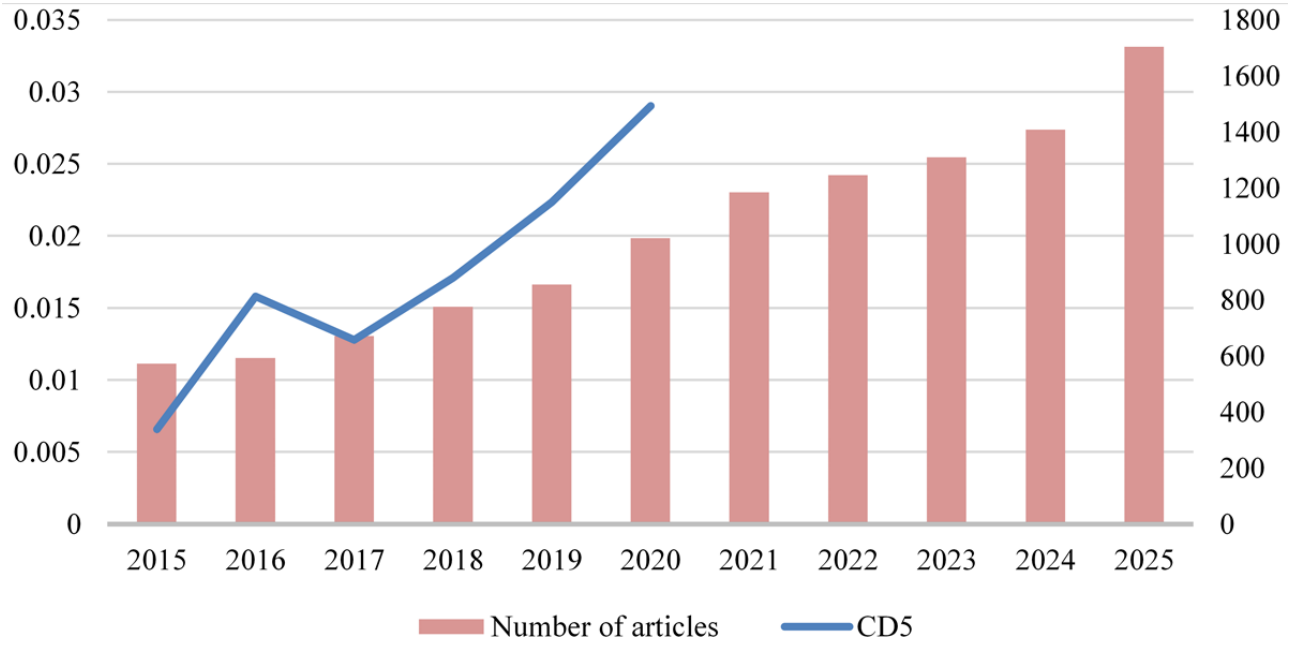
Results

Annual Publication Trends

The annual trends in article and patent outputs in the BCI field are illustrated in Figures 2 and 3. Both datasets demonstrate a sustained growth trajectory over the past decade, reaching their respective peaks in 2025 (1705 articles and 371 patents). The volume of both publication types is projected to continue increasing in 2026.

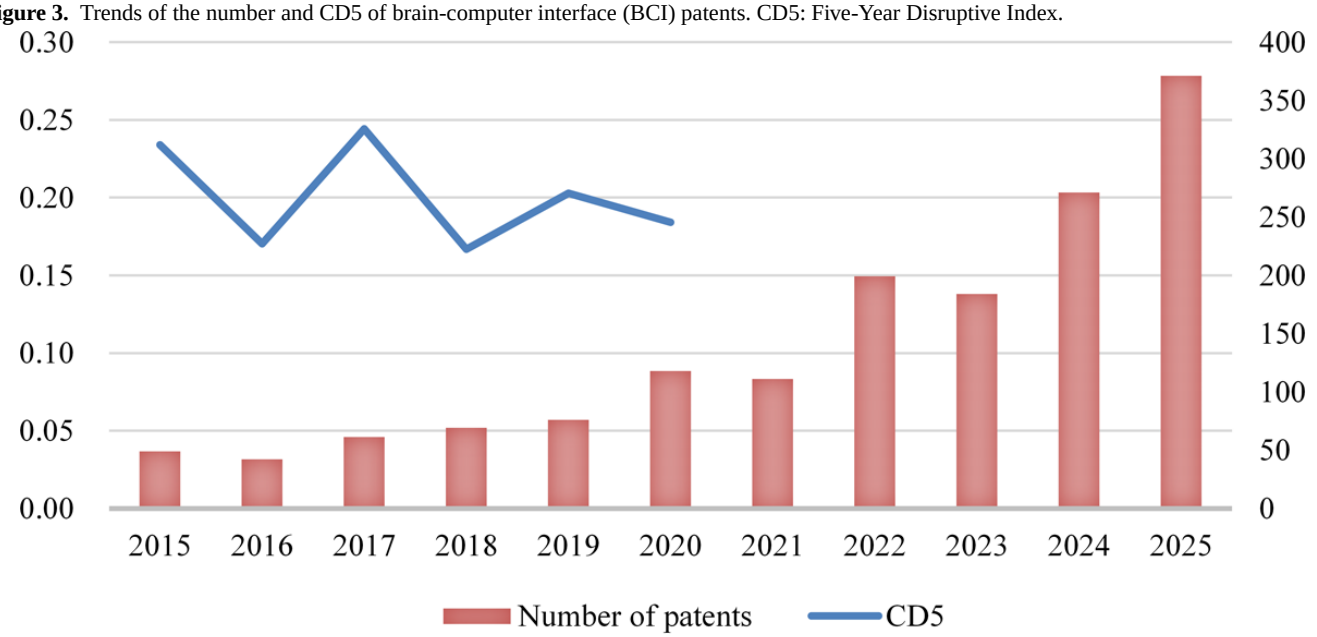
Granted BCI patents increased from 49 in 2015 to over 100 annually from 2020 onward, with notable accelerations during 2021 - 2022 and 2024 - 2025. These surges potentially reflect the influence of supportive policies and strategic plans worldwide, exemplified by China's 13th Five-Year Plan for brain science (2016) and the US BRAIN Initiative 2.0 (2019), which may have propelled research and development efforts in BCI technology.

Figure 2. Trends of the number and CD5 of BCI articles. CD5: five-year Disruptive Index.



XSL•FO
RenderX

Figure 3. Trends of the number and CD5 of brain-computer interface (BCI) patents. CD5: Five-Year Disruptive Index.

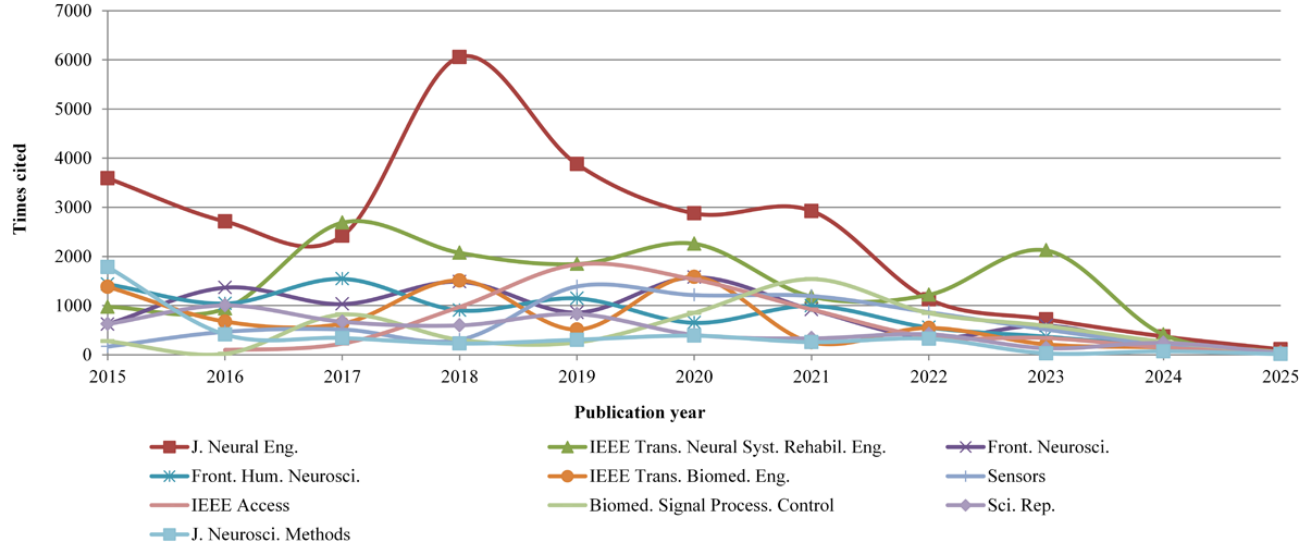


Journal and Citation Analysis

Multimedia Appendix 1 presents the top 10 journals with the highest publication output in the BCI field. The *Journal of Neural Engineering* ranks first with 908 publications, accounting for 22.1% (908/4101) of the total among these top 10 journals, followed by *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (577/4101 articles, 14.1%) and *Frontiers in Neuroscience* (489/4101, 11.9% articles). The remaining journals each published fewer than 400 articles. Among these 10 journals, 7 are classified as quartile 2 journals, while the other 3 are quartile 1 journals.

Figure 4 illustrates the annual citation trends of the top 10 most cited journals in the BCI field. The *Journal of Neural Engineering* accumulated the highest total citations, peaking in 2018 before entering a sustained decline. This downward trajectory intensified after 2021 across almost all journals due to insufficient citation accumulation, with the notable exception of *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, which demonstrated exceptional growth and reached a peak in 2023.

Figure 4. Annual citation trends of the top 10 most cited journals in the brain-computer interface (BCI) field.



According to the WOS database, 157 HCPs and one Hot Paper (published within the past 2 years and ranked top 0.1% by citations for its category) were identified within our retrieved article dataset. The 3 journals with the highest number of HCPs are the *Journal of Neural Engineering* (12 papers), *IEEE Transactions on Biomedical Engineering* (8 papers), and *Nature Communications* (6 papers). The single Hot Paper, which was also identified as an HCP, was published in *Advanced Materials* in 2024 by researchers affiliated with a Chinese institution, presenting a flexible hydrogel electronics designed for smart BCI applications [22].

Institution Analysis

Multimedia Appendix 2 demonstrates a comprehensive comparison of the top 10 influential institutions in the BCI field by article output, total citation frequency, and category

XSL•FO
RenderX

normalized citation impact (CNCI). The CNCI of an article is calculated by dividing the actual count of citing items by the expected citation rate for articles with the same article type, year of publication, and subject area. The CNCI of a set of articles, for example, the collected articles of an individual, institution, country, or region, is the average of the CNCI values for all the articles in the set. The University of California System ranks first in total citation frequency, with 416 articles and 16,607 total citations. The institution with the largest number of BCI articles is the Chinese Academy of Sciences (504 articles and 14,975 citations), representing the only institution exceeding 500 articles. Stanford University achieves the highest CNCI among all institutions, yet produces fewer articles, indicating a quality-over-quantity strategy in BCI research. Among the top 10 institutions, 4 are based in the United States: the University of California System, Stanford University, Harvard University, and the Pennsylvania Commonwealth System of Higher Education (PCSHE). China and Switzerland are each represented by 2 institutions—the Chinese Academy of Sciences and Tsinghua University from China, and the Swiss Federal Institutes of Technology Domain and the École Polytechnique Fédérale de Lausanne from Switzerland.

Multimedia Appendix 3 presents a list of the primary applicants for patents related to BCI. Among the top 10 applicants, 9 are from China and one from South Korea. In terms of institutional types, 9 of the top 10 applicants are higher education institutions, and only one is an enterprise. Notably, Tianjin University ranks first on the list with over 50 granted patents. The focus areas of Tianjin University's patent applications include "brain-computer interface systems based on motor imagery brain signal features," "analysis of EEG signals based on deep or integrated learning," and "brain-computer interfaces for controlling devices such as mice or wheelsets."

Country Analysis

Multimedia Appendix 4 presents a comprehensive comparison of the top 10 countries by article output, total citation frequency, and citation rate (% articles cited). Point radius in the visualization corresponds to citation rate, with larger radii indicating higher citation rates. It can be seen that both Germany and Switzerland exhibit exceptional citation rates, each exceeding 94%. The United States and China demonstrate distinct performance profiles that differentiate them from other leading countries. The United States ranks second in article output but attains the highest total citation frequency, indicating substantial aggregate research impact. Conversely, China leads in article output and ranks second in total citation frequency; however, it demonstrates the lowest citation rate among the top 10 countries, suggesting a productivity-impact disconnect.

Multimedia Appendix 5 shows that the United States has the largest number of international collaboration articles in the BCI field, whereas China exhibits the greatest number of domestic collaboration articles. Both nations occupy the top tier, substantially surpassing all other countries. Additionally, the number of international collaboration articles in Germany and the United Kingdom is much higher than that of domestic collaboration articles.

Table 1 ranks the top 10 countries by the number of BCI-related patents. China and the United States emerge as the primary technical contributors, with China leading by a substantial margin. The United States follows with 200+ patents, while South Korea, ranking third, has a patent volume approximately half that of the United States. The remaining seven countries each account for fewer than 100 patents.

The patent landscape across jurisdictions reflects distinct regional strategies of patent applicants. As shown in Table 2, patents originating from China are highly concentrated in its domestic market, with merely a small number deployed in the United States and South Korea. The United States conducts extensive patent layouts in the European Patent Office (EPO), China, South Korea, and Japan, showing remarkable characteristics of globalized patent deployment. For South Korea and Japan, their overseas patent filings are predominantly concentrated in the United States. By contrast, France has completed patent arrangements in the EPO and all the other 4 target countries.

Table . Top 10 countries by number of patents in the BCI[a] field from 2015 to 2025.

Rank	Applicant country	Number of patents, n
1	China	960
2	United States	251
3	South Korea	127
4	Japan	29
5	France	28
6	Israel	14
7	India	14
8	Canada	13
9	Germany	13
10	Russia	13

[a]BCI: brain-computer interface.

XSL•FO
RenderX

Table . Patent disclosure layouts of the top 5 countries in the BCI[a] field from 2015 - 2025[b].

Country	Publishing country or organization					
	EPO[c]	China	United States	South Korea	Japan	France
China	0	946	20	1	0	0
United States	23	13	217	12	18	0
South Korea	0	0	28	119	3	0
Japan	4	2	12	1	24	0
France	14	6	18	6	2	11

[a]BCI: brain-computer interface.

[b]Statistics based on nonmerged patent family counts.

[c]EPO: European Patent Office.

Funding Agency Analysis

Bibliometric analysis revealed that the National Natural Science Foundation of China (NSFC), the United States Department of Health & Human Services, and the National Institutes of Health (NIH) ranked as the top 3 funding agencies in the BCI field by both article output and total citation frequency (Table 3). NSFC demonstrated marked dominance in article output, with a substantial lead over the second-ranked United States Department of Health & Human Services. Among the top 10 funding agencies, 4 are based in the United States, while 2 each are from China, Germany, and the European Union. The United States Department of Defense (DoD) and the National Key Research & Development Program of China (NKRDPC) have the highest percentages of articles in the top 1% by citations (% articles in top 1%), both exceeding 6%. Notably, among Chinese agencies, both NSFC and NKRDPC exhibit citation rates (% articles cited) below 90%.

Table . Statistics of the top 10 funding agencies of the brain-computer interface (BCI) field from 2015 to 2025.

Name	WOS[a] articles	Times cited	% Articles Cited	% Articles in top 1%	HCPs[b] count	HCPs cited by patents
National Natural Science Foundation of China	2081	46672	87.7 (1825)	3.9 (81)	44	9
United States Department of Health & Human Services	718	24735	93.9 (674)	3.2 (23)	21	7
National Institutes of Health- United States	710	24394	93.9 (667)	3.2 (23)	21	6
National Science Foundation	489	20569	95.1 (465)	5.3 (26)	19	6
United States Department of Defense	279	18791	97.1 (271)	6.5 (18)	13	6
German Research Foundation	233	10712	96.6 (225)	3.9 (9)	7	4
European Union	337	9803	94.4 (318)	2.4 (8)	8	3
National Key Research & Development Program of China	403	9648	88.1 (355)	7.4 (30)	18	1
Federal Ministry of Education & Research	170	9309	98.2 (167)	5.3 (9)	6	4
European Research Council	184	8030	94.0 (173)	4.4 (8)	9	4

[a]WOS: Web of Science.

[b]HCPs: highly cited papers.

XSL•FO
RenderX

Multimedia Appendix 6 illustrates the performance of the top 10 funding agencies across two normalized impact indicators: CNCI and impact relative to world (IRW). The DoD and the Federal Ministry of Education & Research (BMBF) demonstrated exceptional performance, with both agencies exceeding the threshold of 2.0 on both metrics. While the remaining agencies exhibited variable performance, the NSFC was the sole entity with an IRW value below 1.0, indicating below-world-average citation impact.

Analysis of the HCPs revealed that the top 10 funding agencies collectively supported 89 of the 157 HCPs. Among these, the NSFC contributed the most, funding 44 HCPs. Forward citation analysis showed that only 28 of these 89 HCPs were cited by patents, and merely 5 of them had citing patents from institutions ranked among the top 10 applicants (Multimedia Appendix 3). The funding agencies for these 5 articles were NSFC (3 articles) and the DoD (2 articles).

Disruptive Trend Analysis

Following the methodology outlined in the Methods Section, we constructed citation networks and calculated CD5 values for articles and patents published up to 2020, ensuring complete 5-year observation windows. Annual averages were then derived (Figures 2 and 3, respectively). According to the calculation formula of the disruption index, the final value ranges from −1 (consolidating) to 1 (disruptive). Given that all CD5 values in Figures 2 and 3 are positive, both scientific research and technological innovation in the BCI field exhibited disruptive characteristics during the observation period, with new research and development practices continuously challenging and expanding existing knowledge boundaries.

Comparative analysis of Figures 2 and 3 reveals distinct patterns. Article CD5 values generally increased alongside publication volume, reaching a peak in 2020, the final year of complete observation. This trajectory suggests sustained growth in scientific disruptiveness. In contrast, patent CD5 values exhibited fluctuation within a bounded range rather than consistent growth. The divergence between rising scientific disruptiveness and the absence of a corresponding upward trend in patent disruptiveness raises the possibility that scientific paradigm shifts in BCI have not yet fully permeated the patent landscape, implying a potential translation lag. However, given the 5-year observation window and the likelihood of longer diffusion cycles, this interpretation remains speculative and requires validation through extended temporal analysis and direct science-to-patent citation tracking.

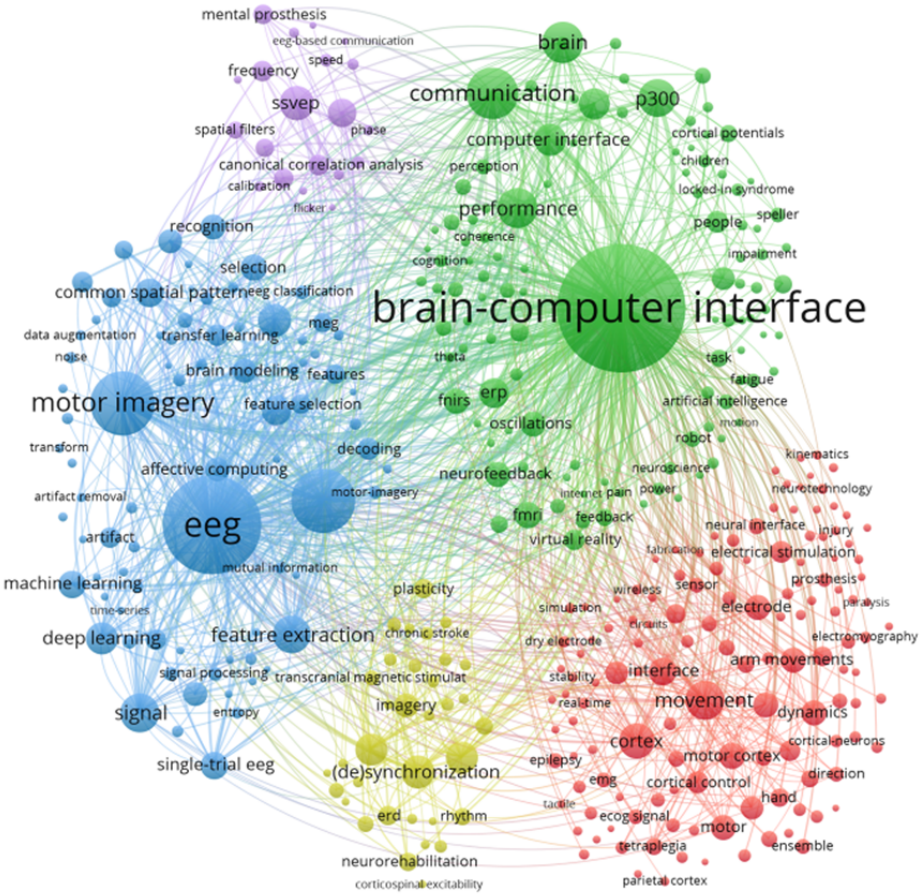
Research Topics Analysis

Overview

Keywords with a cooccurrence frequency exceeding 30 were identified as high-frequency terms. This threshold was set to prioritize well-established and field-representative keywords, thereby revealing the core research structure of the field. Keywords were derived from author-supplied terms, with synonym normalization performed prior to analysis. For example, “BCI” and “brain-computer interface” were unified as “brain-computer interface”; “eeg” and “electroencephalogram” were unified as “eeg.” After synonym normalization, the high-frequency keywords were imported into VOSviewer for bibliometric mapping, where the link strength normalization method was set as association strength and clustering was configured with a minimum of 20 keywords per cluster, yielding five distinct clusters (Figure 5). These clusters centered on the following research topics: neural interfaces and motor control (red), brain-computer interface and cognitive neuroscience (green), electroencephalogram (EEG) signal processing and machine learning (blue), neurorehabilitation and stroke recovery (yellow), and steady-state visually evoked potentials (SSVEP) & system optimization (purple).

XSL•FO
RenderX

Figure 5. Keyword clusters of the brain-computer interface (BCI) field from 2015 to 2025.



Cluster 1: Neural Interfaces and Motor Control

This cluster (represented in red) comprised 133 keywords, with prominent nodes including movement, cortex, motor cortex, neurons, neural interface, spinal-cord-injury, and assistive technology. The thematic focus centered on motor control mechanisms and the application of neural interface technologies, particularly the decoding and modulation of neural activity in the motor cortex. Key methodologies involved signal acquisition via electrodes or microelectrode arrays to record neuronal activity, enabling precise control of arm and hand movements or functional recovery in spinal-cord-injury patients. Keywords such as brain-machine interface and neuroprosthesis highlighted applications in assistive technology and tetraplegia. Further investigations extended to neuromodulation therapies (eg, deep brain stimulation) for neurological disorders like Parkinson disease.

Cluster 2: BCI and Cognitive Neuroscience

This cluster (represented in green) encompasses 111 keywords, including prominent nodes such as BCI, communication, performance, P300, functional magnetic resonance imaging (fMRI), and amyotrophic lateral sclerosis (ALS), as well as smaller yet significant nodes like AI, virtual reality, human-computer interaction, augmented reality, and neurophysiology. Thematic analysis reveals that this cluster centers on BCI, emphasizing their applications in communication and disease intervention.

High-frequency keywords such as P300 and Event-Related Potential (ERP) reflect a concentrated focus on EEG signal decoding, while functional near-infrared spectroscopy (fNIRS) and fMRI indicate the integration of multimodal neuroimaging techniques. Key research objectives within this direction encompass restoring interaction capabilities for patients with ALS and locked-in syndrome, as well as exploring the synergy between virtual reality and neurofeedback technologies. Additionally, cognitive mechanisms such as working memory and attention are incorporated into studies aimed at optimizing BCI performance.

Cluster 3: EEG Signal Processing and Machine Learning

This cluster (represented in blue) encompasses 80 keywords, with central nodes including EEG, classification algorithms, motor imagery, deep learning, convolutional neural network (CNN), neural networks, and machine learning, complemented by specialized terms such as data augmentation, genetic algorithm, principal component analysis (PCA), and ensemble learning. The thematic focus centers on algorithmic processing and pattern recognition of EEG signals.

Key research areas include motor imagery and single-trial EEG analysis, with widely used methodologies such as CNN and common spatial pattern (CSP). Keywords such as artifact removal and feature extraction highlight the critical role of signal preprocessing, while emerging analytical approaches (eg, Riemannian geometry and wavelet transform) reflect ongoing

XSL•FO
RenderX

methodological advancements. The research scope extends to translational domains such as affective computing and seizure detection, emphasizing the interdisciplinary integration of machine learning in neural engineering.

Cluster 4: Neurorehabilitation and Stroke Recovery

This cluster (represented in yellow) consists of 34 keywords, including significant nodes such as stroke, rehabilitation, plasticity, neurorehabilitation, and transcranial magnetic stimulation, as well as smaller nodes like sensorimotor rhythms, chronic stroke, direct-current stimulation, and induced movement therapy. The thematic focus of this cluster is centered on motor recovery and neuroplasticity mechanisms following a stroke. Key therapeutic approaches encompass functional electrical stimulation and exoskeleton-assisted therapy, with rehabilitation efficacy evaluated through EEG-derived features such as event-related desynchronization (ERD) and mu-rhythm modulation. Notably, terms such as chronic stroke and upper-limb highlight specific patient subpopulations, while neuromodulation techniques (eg, transcranial magnetic stimulation [TMS] and direct-current stimulation) are used to regulate cortical excitability. The prevailing research paradigm integrates motor execution and motor imagery to promote functional reorganization of neural circuits.

Cluster 5: SSVEP and System Optimization

This cluster (represented in purple) consists of 21 keywords, with central nodes including SSVEP, potentials, frequency, canonical correlation analysis (CCA), and evoked potentials, complemented by specialized terms such as hybrid BCI and enhancing detection. Thematic analysis indicates that this cluster primarily revolves around SSVEPs, with a focus on high-frequency visual stimuli (eg, flicker frequency) to elicit neural responses. CCA serves as the primary decoding algorithm, used to enhance the signal-to-noise ratio (SNR). Keywords such as hybrid BCI and cursor control highlight applications in interactive systems, while topography and phase reflect explorations into spatial and temporal signal characteristics. The research objectives within this direction include developing rapid calibration methods and detection enhancement techniques to optimize real-time BCI systems.

Research Topics Evolutionary Analysis

Based on the burst detection function of CiteSpace, a longitudinal analysis of burst keywords in the BCI field reveals that research topics over the past decade can be broadly categorized into 3 distinct stages (Figure 6).

- Fundamental neural mechanism exploration (2015-2017): "Brain-machine interface" ranked first with burst strength of 51.73, cooccurring with “motor cortex” and “amyotrophic lateral sclerosis,” indicating a research focus on motor control mechanisms and neurodegenerative disease applications dominated by classical physiological paradigms.
- Algorithmic transition (2018-2021): The successive emergence of “restoration” (17.29) and “deep learning” (22.02) marked a shift from conventional signal processing to AI algorithms, with machine learning techniques such as “support vector machine” gaining traction.
- Deep learning dominance (2022-2025): Terms including “domain adaptation,” “convolutional neural networks,” and “data augmentation” demonstrated concentrated bursts during 2022-2023, demonstrating that CNN architectures and data-driven approaches have become the current research frontier. The sustained burst of “task analysis” (2020-2025) further confirms the deepening expansion of application scenarios.

International Patent Classification Analysis

The International Patent Classification (IPC) codes, established by the World Intellectual Property Organization (WIPO), served as the analytical framework for mapping patent landscapes. BCI-related patents were categorized according to IPC classifications, and the top 10 technical topics ranked by the number of patents were selected. As shown in Table 4, key technical areas include input devices for transforming data into a computer-processable format or output devices for transferring data from a processor (G06F3), medical diagnostic measurements (A61B5), pattern recognition methods or devices (G06F18 and G06K9), and biologically modeled computer systems (G06N3).

XSL•FO
RenderX

Figure 6. Top 25 strongest citation burst keywords in the brain-computer interface (BCI) field from 2015 to 2025. BCI: brain-computer interface.

Keywords	Year	Strength	Begin	End	2015 - 2025
brain-machine interface	2015	51.73	**2015**	2017	
grasp	2015	25.08	**2015**	2019	
chronic stroke	2015	18.63	**2015**	2017	
reach	2015	16.97	**2015**	2019	
amyotrophic lateral sclerosis	2015	16.76	**2015**	2016	
potentials	2015	16.44	**2015**	2019	
cortical control	2015	16.39	**2015**	2019	
motor cortex	2015	14.51	**2015**	2016	
humans	2015	14.2	**2015**	2018	
algorithms	2015	13.44	**2015**	2019	
independent component analysis	2016	21.57	**2016**	2019	
primary motor cortex	2017	19.94	**2017**	2019	
support vector machine	2017	15.15	**2017**	2018	
restoration	2018	17.29	**2018**	2021	
task analysis	2020	45.78	**2020**	2025	
frequency recognition	2020	15.2	**2020**	2023	
domain adaptation	2022	22.98	**2022**	2025	
convolutional neural network (cnn)	2021	15.6	**2022**	2025	
memory	2022	12.94	**2022**	2025	
convolutional neural networks	2020	26.66	**2023**	2025	
deep learning	2018	22.02	**2023**	2025	
data models	2023	17.39	**2023**	2025	
data augmentation	2023	16.46	**2023**	2025	
convolutional neural network	2019	16.26	**2023**	2025	
perception	2023	16.13	**2023**	2025	

Table . Top 10 IPC[a] numbers of the BCI[b] field from 2015 to 2025.

IPC (main group)	Number of patents, n
G06F3 (Input devices for transforming data to a form capable of being processed by a computer; output devices for transferring data from a processor to output equipment, eg, interface devices)	737
A61B5 (Measuring for diagnostic purposes; identification of humans)	734
G06F18 (Pattern recognition)	285
G06N3 (Computer systems based on specific computational models)	238
G06K9 (Methods or devices for recognizing patterns)	105
A61N1 (Electrotherapy; apparatus therefor)	102
G06N20 (Machine learning)	62
A61F2 (Filters to be implanted into blood vessels; prostheses, ie, artificial substitutes or replacements for parts of the human body; appliances for connecting prostheses to the human body; devices for providing openings in body tissues or for preventing collapse of tubular structures in the human body, eg, stents)	52
G06F17 (The specialized digital computing or data processing equipment or methods for specific functions)	44
B25J9 (Programmed manipulators)	42

[a]IPC: international patent classification.

[b]BCI: brain-computer interface.

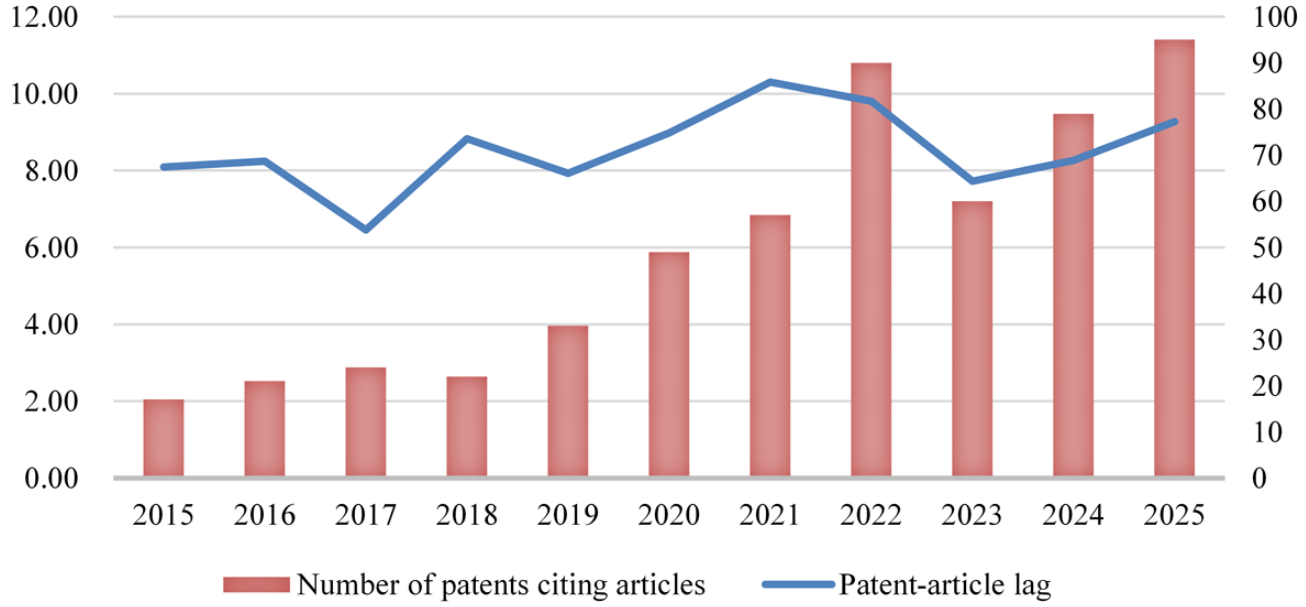
XSL•FO
RenderX

Dual Citation Analysis

Backward citation analysis revealed that 547 of the 1551 granted patents (35.3%) collectively cited 2726 articles; however, 397 of these 547 patents (72.6%) cited only one article each, indicating a pronounced long-tail distribution. The mean time lag between the patent grant year and the publication year of the cited articles was 8.8 years. Further analysis of the annual citation time lags (Figure 7) revealed that the number of patents citing articles increased in tandem with the total number of granted patents; however, the year lag between the patent grant year and the publication year of the cited articles fluctuated within the range of 6.5 - 10.3, showing no discernible unidirectional trend.

Forward citation analysis revealed that 50 of the 157 HCPs were cited by patents after publication. Simultaneous examination of the article publication institutions and patent application institutions revealed that for only 10 of the 157 HCPs, the publishing institutions were also the applicants of the patents that cited them. These 10 institutions were: Ecole Polytechnique Federale de Lausanne, Guangdong University of Technology, Harvard University, Huazhong University of Science & Technology, Imperial College London, King Abdullah University of Science & Technology, Massachusetts Institute of Technology, Neuralink Corp, Shenzhen University, and University of Pittsburgh. Similar to the top 10 applicants of the granted patents (Multimedia Appendix 3), 9 of these 10 institutions were higher education institutions, with Neuralink Corp being the only nonacademic institution.

Figure 7. Trends of the number of patents citing articles and patent-article lags.



Discussion

Principal Findings

The sustained growth in both BCI articles and patents over the past decade reflects the development of this interdisciplinary field, bridging neuroscience, engineering, and computer science. The parallel upward trajectories of scientific output and technological protection suggest a robust innovation ecosystem where fundamental research advances are translated into proprietary technologies. Promoters of scientific research and technological innovation at different levels, such as journals, institutions, countries, and funding agencies, have collectively facilitated this expansion; however, their respective contributions and impacts are different.

The concentration of BCI articles in specialized engineering and neuroscience journals reflects the field's strong methodological orientation. The *Journal of Neural Engineering*, despite its leading productivity, exhibited sharply declining citations after 2018. In contrast, *IEEE Transactions of Neural Systems and Rehabilitation Engineering* demonstrated exceptional growth, peaking in 2023 while others declined due to insufficient citation accumulation. This divergence suggests that the latter's publications have attracted disproportionate disciplinary attention, potentially reflecting timely responses to emerging methodological demands. The sole Hot Paper, published in *Advanced Materials* by Chinese researchers, underscores the emerging convergence between materials science and BCI development, signaling a potential shift toward translational applications requiring advanced hardware solutions.

The institutional landscape reveals distinct strategies in BCI research performance. Stanford University achieves the highest CNCI despite modest article output, indicating a quality-over-quantity approach that prioritizes foundational impact. Conversely, both the Chinese Academy of Sciences and the University of California System have gained substantial influence through scaled production: the former leads in article output, while the latter generates the greatest total citation frequency. This divergence suggests that elite institutions can achieve research leadership through alternative pathways, selective high-impact contributions or comprehensive engagement, reflecting differing mission orientations and resource allocation strategies within the field.

The divergent performance profiles of the United States and China illustrate contrasting national strategies in BCI research.

XSL•FO
RenderX

The United States achieves the highest aggregate citation impact despite ranking second in article output, coupled with the largest number of international collaborations articles—suggesting a network-centric approach that leverages global partnerships to amplify research influence. China, while dominating in article output and domestic collaboration intensity, exhibits the lowest citation rate among leading nations, indicating a potential quantity-quality trade-off. This productivity-impact disconnect may reflect differing evaluation incentives or developmental stages of research ecosystems. The substantial gap between these 2 nations and all others underscores their central positioning in the global BCI research architecture, yet their divergent pathways imply distinct trajectories for shaping the field's future directions.

Similar to the analysis results of the articles, China and the United States form a distinct top tier in global BCI patenting, substantially exceeding all other nations. However, their internationalization strategies diverge markedly. While both prioritize domestic patent filings, the United States demonstrates systematic expansion across EPO and foreign national markets—indicating a deliberate strategy to secure global market access. In contrast, China's limited extraterritorial presence, despite its leadership of patent output, suggests an orientation toward domestic exploitation rather than international competition. Meanwhile, patent analysis at the institutional level reveals remarkable concentration in higher education institutions, with Tianjin University ranked first, reflecting systemic incentives for academic engagement in technology transfer. However, the near-absence of nonacademic applicants in the top tier indicates a university-centric commercialization ecosystem, potentially constraining rapid market translation compared to industry-led innovation models.

The funding landscape reveals distinct agency approaches to BCI research investment. China's NSFC operates as a high-volume engine, substantially outpacing all other agencies in article output, yet exhibits below-world-average citation impact (IRW<1.0) and a citation rate below 90%. This volume-impact disconnect suggests that scale-oriented funding strategies may not automatically translate into commensurate research influence. Conversely, the DoD and Germany's BMBF achieve exceptional normalized impact (CNCI and IRW both exceeding 2.0), indicating highly selective, high-return investment approaches. Intranational comparison reveals that China's NKRDPC, tasked with strategic technology development, achieves the highest percentage of articles in the top 1% by citations (7.4%) among all funding agencies, outperforming NSFC (3.9%) in impact concentration. This intranational divergence (NKRDPC vs NSFC) suggests that mission-oriented funding with clear technological targets may yield higher per-unit impact compared to investigator-driven schemes, offering a pattern with implications for science policy design globally. Additionally, forward citation analysis of HCPs revealed that, although mainstream funding agencies made substantial contributions to the output of HCPs, the overall efficiency of translating these high-impact research achievements into patents remained low, with only a limited number of such outputs attracting the attention of top-ranking patent applicants. This indicates that the linkage mechanism between research funding orientation and technology transfer still warrants further strengthening.

Regarding the disruptive trends in the BCI field, although the observation period of this study is limited, preliminary comparisons can still be drawn with existing research. Park et al [21] analyzed 25 million articles (1945 - 2010) and 3.9 million patents (1976 - 2010) using the Disruptive Index and demonstrated a universal decline in disruptiveness across science and technology. The sustained growth in BCI scientific disruptiveness through 2020 suggests that this interdisciplinary field continues to generate novel conceptual frameworks that challenge existing knowledge boundaries. However, the fluctuating and directionless trajectory of patent disruptiveness indicates that recent scientific paradigm shifts have not yet fully permeated the patent landscape. This asynchrony, rising scientific disruptiveness alongside the absence of an upward trend in patent disruptiveness, is consistent with potential translational friction between academic advances and technological outputs, warranting continued monitoring as the field evolves.

The thematic evolution of BCI research demonstrates a clear trajectory from neural mechanism exploration through algorithmic transition to application deepening. Initially centered on classical physiological paradigms investigating motor control and neurodegenerative diseases, the field underwent a fundamental shift toward AI-driven approaches, with deep learning architectures now dominating both research and technological development. This scientific progression aligns closely with the patent landscape, where pattern recognition methods (G06F18 and G06K9) and biologically modeled computing systems (G06N3) feature prominently alongside medical diagnostic measurements (A61B5) and human-computer interaction technologies (G06F3). The emergence of pattern recognition methods as core technical categories particularly underscores the critical role of algorithmic innovation in enabling reliable neural decoding. The parallel emphasis on algorithmic sophistication in both research bursts and patent classifications suggests substantial potential for advancing deep learning applications within the BCI field.

Due to variations in developmental stages and technological application environments, the time lag for knowledge transfer from articles to patents differs across fields [23]. Emerging fields generally exhibit shorter translation timelines, typically estimated at 3 - 4 years [24,25]. This study found that the time lag between the patent grant year and the publication year of the cited articles in the BCI field was 8.8 years. Combined with the percentages obtained from the bidirectional citation analysis—such as 1004 of 1551 patents (64.7%) citing no articles, 107 of 157 HCPs (68.2%) receiving no patent citations, and the limited number and type of institutions achieving internal translation (concentrated within higher education institutions)—these findings collectively suggest that knowledge translation from scientific research to technological innovation in this field faces certain barriers. Our findings are partially consistent with prior research on human-computer interaction (HCI), a closely related and broader field, where the time lag between the patent grant date and the publication date of cited papers was reported as 10.5 years, leading to the conclusion

XSL•FO
RenderX

that research and practice in the HCI domain may not be efficiently connected [26].

Conclusions

This study provides a comprehensive scientometric characterization of the BCI field through dual analysis of articles and patents. The findings reveal a rapidly expanding yet strategically differentiated global landscape: the United States and China constitute a distinct top tier in both scientific output and technological innovation, China leads in article output and patent quantity with intensive domestic collaboration, while the United States achieves the highest citation impact and most extensive international patent portfolios through globalized research networks. Institutional and funding-level analyses further demonstrate that research leadership can be achieved through alternative strategies, whether selective quality-focused investment or comprehensive volume-driven engagement, with mission-oriented funding demonstrating superior per-unit impact compared to investigator-driven schemes. However, the linkage mechanism between research funding orientation and technology transfer still warrants further strengthening.

A critical finding emerges from the divergent disruptiveness trajectories: while BCI scientific research maintains active disruptive potential, technological development exhibits no synchronous upward trajectory. This asynchrony, combined with prolonged knowledge transfer lag, a high proportion of HCPs receiving no patent citations and patents citing no prior articles, and an academic-dominated patent landscape, collectively suggests certain barriers to knowledge translation in this field. The field's thematic evolution from neural mechanisms through AI integration to application-specific deployment traces a discernible disciplinary trajectory; however, bridging the gap between scientific capability and technological implementation requires strengthened mechanisms to facilitate translation from scientific advances into technical innovations.

Strengths and Limitations

This study simultaneously selected articles and patents in the BCI field for synchronous bibliometric and visualized analysis, which can obtain relatively objective and comprehensive analysis results, and thereby draw reliable conclusions. Moreover, the application of the Disruptive Index can achieve nuanced characterization of the development trends of scientific research and technological innovation in the BCI field. However, limitations also exist in our study. First, WOS Core Collection was selected as the only data source for articles, so publications indexed exclusively in other databases such as PubMed or Scopus might be left out. Second, keyword inclusion thresholds for topic analysis resulted in limited coverage of emerging domain concerns, such as neural data ownership, ethics, and privacy security. Third, the CD5 index requires a 5-year forward citation window, restricting this analysis to publications through 2020. As a result, post-2020 disruptive trends, including the impact of the AI boom on BCI, fall outside the scope of this study. In the future, we will conduct longitudinal tracking of the BCI field based on the analytical framework established in this study.

Acknowledgments

During the preparation of this work, the authors used KIMI to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This study was supported by the Humanity and Social Science Youth Foundation of Ministry of Education of China under Grant number 21YJC870016.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: JW and GL
Data curation: YY and HX
Formal analysis: JW, YY, and HX
Funding acquisition: JW
Investigation: JW and GL
Methodology: GL and HX
Project administration: JW
Resources: JW
Software: GL, HX, and JW
Supervision: GL
Validation: JW and GL
Visualization: GL and HX
Writing – original draft: JW, GL, HX, and YY

XSL•FO
RenderX

Writing – review & editing: JW and GL

Conflicts of Interest

None declared.

Multimedia Appendix 1
Top 10 journals by number of brain-computer interface–related articles from 2015 to 2025.
[PNG File, 109 KB - rehab_v13i1e95902_app1.png]

Multimedia Appendix 2
Top 10 most influential institutions in the brain-computer interface field from 2015 to 2025.
[PNG File, 611 KB - rehab_v13i1e95902_app2.png]

Multimedia Appendix 3
Top 10 institutions by number of granted patents in the brain-computer interface field from 2015 to 2025.
[PNG File, 116 KB - rehab_v13i1e95902_app3.png]

Multimedia Appendix 4
Top 10 most influential countries in the brain-computer interface field from 2015 to 2025.
[PNG File, 282 KB - rehab_v13i1e95902_app4.png]

Multimedia Appendix 5
Top 10 countries by number of international and domestic collaboration articles in the brain-computer interface field from 2015 to 2025.
[PNG File, 128 KB - rehab_v13i1e95902_app5.png]

Multimedia Appendix 6
Impact of the top 10 funding agencies of the brain-computer interface field from 2015 to 2025.
[PNG File, 240 KB - rehab_v13i1e95902_app6.png]

References

1. Delgado JM. Instrumentation, working hypotheses, and clinical aspects of neurostimulation. Appl Neurophysiol 1977;40(2-4):88-110. [doi: 10.1159/000102436] [Medline: 101139]
2. Yin Z, Wan Y, Fang H, et al. Bibliometric analysis on brain-computer interfaces in a 30-year period. Appl Intell 2023 Jun;53(12):16205-16225. [doi: 10.1007/s10489-022-04226-4]
3. Tekin U, Dener M. A bibliometric analysis of studies on artificial intelligence in neuroscience. Front Neurol 2025;16:1474484. [doi: 10.3389/fneur.2025.1474484] [Medline: 40040909]
4. Sun J, Lin PJ, Zhai X, et al. Multimodal behavioral data predict stroke patient's response to BCI treatment through explainable AI. IEEE Trans Neural Syst Rehabil Eng 2026;34:1-9. [doi: 10.1109/TNSRE.2025.3625222] [Medline: 41134943]
5. Arntz A, Weber F, Handgraaf M, et al. Technologies in home-based digital rehabilitation: scoping review. JMIR Rehabil Assist Technol 2023 Jul 27;10:e43615. [doi: 10.2196/43615] [Medline: 37253381]
6. Lei R, Yue F, Yue C, et al. Characterization of global research trends and prospects on prone positioning in respiratory failure: bibliometric analysis. Interact J Med Res 2025 Jun 20;14:e67276. [doi: 10.2196/67276] [Medline: 40540635]
7. Chen H, Yun G. Efficacy of brain-computer interface therapy for upper limb rehabilitation in chronic stroke: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2026 Jan 28;28:e79132. [doi: 10.2196/79132] [Medline: 41605490]
8. Li F, Zhang D, Chen J, Tang K, Li X, Hou Z. Research hotspots and trends of brain-computer interface technology in stroke: a bibliometric study and visualization analysis. Front Neurosci 2023;17:1243151. [doi: 10.3389/fnins.2023.1243151]
9. Feng J, Gao S, Hu Y, Sun G, Sheng W. Brain-computer interface for patients with spinal cord injury: a bibliometric study. World Neurosurg 2024 Dec;192:170-187. [doi: 10.1016/j.wneu.2024.08.163] [Medline: 39245135]
10. Wen H, Xu M, Cui S. Global research trends in brain-computer interfaces for Alzheimer's disease: a bibliometric perspective. Int J Surg 2026 Feb 1;112(2):5518-5520. [doi: 10.1097/JS9.0000000000004000] [Medline: 41222907]
11. Angulo Medina AS, Aguilar Bonilla MI, Rodríguez Giraldo ID, Montenegro Palacios JF, Cáceres Gutiérrez DA, Liscano Y. Electroencephalography-based brain-computer interfaces in rehabilitation: a bibliometric analysis (2013-2023). Sensors (Basel) 2024 Nov 6;24(22):7125. [doi: 10.3390/s24227125] [Medline: 39598903]

XSL•FO
RenderX

12. Chio N, Quiles-Cucarella E. A bibliometric review of brain-computer interfaces in motor imagery and steady-state visually evoked potentials for applications in rehabilitation and robotics. Sensors (Basel) 2024 Dec 30;25(1):154. [doi: 10.3390/s25010154] [Medline: 39796947]
13. Stegman P, Crawford CS, Andujar M, Nijholt A, Gilbert JE. Brain–computer interface software: a review and discussion. IEEE Trans Human-Mach Syst 2020;50(2):101-115. [doi: 10.1109/THMS.2020.2968411]
14. Yan W, Liu X, Shan B, Zhang X, Pu Y. Research on the emotions based on brain-computer technology: a bibliometric analysis and research agenda. Front Psychol 2021;12:771591. [doi: 10.3389/fpsyg.2021.771591] [Medline: 34790157]
15. Torres-Cardona HF, Aguirre-Grisales C. Brain-computer music interface, a bibliometric analysis. Brain-Comput Interfaces 2022 Oct 2;9(4):226-245. [doi: 10.1080/2326263X.2022.2109313]
16. Hu K, Chen C, Meng Q, Williams Z, Xu W. Scientific profile of brain-computer interfaces: bibliometric analysis in a 10-year period. Neurosci Lett 2016 Dec 2;635:61-66. [doi: 10.1016/j.neulet.2016.10.022] [Medline: 27751787]
17. Incopat flagship. Incopat. URL: https://www.incopat.com/login?locale=en [accessed 2026-06-01]
18. Liu X, Zhao S, Tan L, et al. Frontier and hot topics in electrochemiluminescence sensing technology based on CiteSpace bibliometric analysis. Biosensors and Bioelectronics 2022 Apr;201:113932. [doi: 10.1016/j.bios.2021.113932]
19. Explore global science and technology knowledge. Lens.org. URL: https://www.lens.org/ [accessed 2026-06-01]
20. Funk RJ, Owen-Smith J. A dynamic network measure of technological change. Manage Sci 2017 Mar;63(3):791-817. [doi: 10.1287/mnsc.2015.2366]
21. Park M, Leahey E, Funk RJ. Papers and patents are becoming less disruptive over time. Nature 2023 Jan;613(7942):138-144. [doi: 10.1038/s41586-022-05543-x] [Medline: 36600070]
22. Wang W, Zhou HL, Xu ZS, Li Z, Zhang L, Wan P. Flexible conformally bioadhesive MXene hydrogel electronics for machine learning-facilitated human-interactive sensing. Adv Mater 2024 Aug;36(31):e2401035. [doi: 10.1002/adma.202401035] [Medline: 38552161]
23. Zhang G, Feng Y, Yu G, Liu L, Hao Y. Analyzing the time delay between scientific research and technology patents based on the citation distribution model. Scientometrics 2017 Jun;111(3):1287-1306. [doi: 10.1007/s11192-017-2357-3]
24. Dorta-González P, Rodríguez-Caro A, Dorta-González MI. Linking science and industry: influence of scientific research on technological innovation through patent citations. Res Eval 2024 Dec 27;34:rvaf019. [doi: 10.1093/reseval/rvaf019]
25. Finardi U. Time relations between scientific production and patenting of knowledge: the case of nanotechnologies. Scientometrics 2011 Oct;89(1):37-50. [doi: 10.1007/s11192-011-0443-5]
26. Cao H, Lu Y, Deng Y, Mcfarland D, Bernstein MS. Breaking out of the ivory tower: a large-scale analysis of patent citations to HCI research. Presented at: CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; Apr 23-28, 2023. [doi: 10.1145/3544548.3581108]

Abbreviations

ALS: amyotrophic lateral sclerosis
BCI: brain-computer interface
BMBF: Federal Ministry of Education & Research
CCA: canonical correlation analysis
CD5: five-year disruptive index
CNCI: category normalized citation impact
CNN: convolutional neural network
CSP: common spatial pattern
DoD: United States Department of Defense
EEG: electroencephalogram
EPO: European Patent Office
ERD: event-related desynchronization
ERP: event-related potential
fMRI: functional magnetic resonance imaging
fNIRS: functional near-infrared spectroscopy
HCI: human-computer interaction
HCP: highly cited papers
IPC: International Patent Classification
IRW: impact relative to world
NIH: National Institutes of Health
NKRDPC: National Key Research & Development Program of China
NSFC: National Natural Science Foundation of China
PCA: principal component analysis
PCSHE: Pennsylvania Commonwealth System of Higher Education
SNR: signal-to-noise ratio

XSL•FO
RenderX

SSVEP: steady-state visually evoked potential
TMS: transcranial magnetic stimulation
WIPO: World Intellectual Property Organization
WOS: Web of Science

Edited by S Munce; submitted 23.Mar.2026; peer-reviewed by N Hong, Z Shao; revised version received 18.Jun.2026; accepted 14.Jul.2026; published 10.Sep.2026.

Please cite as:
Wang J, Yuan Y, Xu H, Luan G
Scientific and Technological Developments in Brain-Computer Interfaces: Dual Bibliometric Analysis of Articles and Patents
JMIR Rehabil Assist Technol 2026;13:e95902
URL: https://rehab.jmir.org/2026/1/e95902
doi:10.2196/95902

©Junhui Wang, Yakun Yuan, Hang Xu, Guannan Luan. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 10.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digitally Assisted Training System for Home-Based Scoliosis Therapy: Proof-of-Concept and Technical Verification Study

Florian Günther[1,2], M.Sc.; Sandra Hunger[1,2], Dr. rer. med.; Michael Werner[3], Dipl.-Ing.; Nikolaus Landerer[4], Dipl.-Volksw., M.Sc.; Christoph-Eckhard Heyde[2], Prof. Dr. med.; Welf-Guntram Drossel[5,6], Prof. Dr.-Ing.

[1]Functionalization Technologies, Fraunhofer Institute for Machine Tools and Forming Technology (IWU), Nöthnitzer Straße 44, Dresden, Germany

[2]Department of Orthopaedic, Hospital of Orthopaedic Surgery, Traumatology and Plastic Surgery, University Hospital Leipzig, Leipzig, Germany

[3]Acoustical Engineering, Fraunhofer Institute for Machine Tools and Forming Technology (IWU), Dresden, Germany

[4]Fuzz Tech IT Solutions GmbH, Leipzig, Germany

[5]Professorship Adaptronics and Lightweight Design, Chemnitz University of Technology, Chemnitz, Germany

[6]Scientific Field Functional and System Integration, Fraunhofer Institute for Machine Tools and Forming Technology (IWU), Dresden, Germany

Corresponding Author:
Florian Günther, M.Sc.
Functionalization Technologies, Fraunhofer Institute for Machine Tools and Forming Technology (IWU), Nöthnitzer Straße 44, Dresden, Germany

Abstract

Background: Schroth therapy is one of the conservative methods used to treat scoliosis. It includes not only therapeutic exercises in physiotherapy facilities but also concepts for independent training at home. Consistent and correct performance of this training has a major impact on the success of the therapy. However, without guidance and supervision by a therapist, patient motivation may decline. In addition, incorrect training techniques may be used, which can have a negative impact on the success of the therapy.

Objective: This study aimed to develop a digitally assisted training system for home-based scoliosis therapy and to verify it technically on a laboratory scale.

Methods: On the basis of training devices used in Schroth therapy, such as mirrors, balance boards, wall bars, and exercise sticks, we developed a concept for a modular, digitally assisted scoliosis training system (iScoolio). The development process included a requirements specification, iterative design drafts, and the transfer of the concepts into a functional demonstrator. Finally, a theoretical evaluation of the use cases and a technical verification of the measurement functions were carried out using specifically developed test scenarios.

Results: A demonstrator with tracking, training, and modification functions was developed, consisting of 4 modules. Technical verification of the integrated measurement functions showed that the angle measurement in motion tracking (n=50) had a mean deviation of 2.29° (SD 1.66°) compared to the ARAMIS 3D optical reference system. Respiration tracking enabled the detection of all simulated respiratory cycles (n=40) and the significant distinction between shallow and deep breathing (t_{39}=−27.4; P<.001). The investigation of the 2 force sensor concepts in the handles (χ^2_2=80.0; P<.001) and supporting feet (χ^2_2=50.0; P<.001) showed a significant differentiation between 3 force levels, but approximately linear drift behavior in each case. When tracking the stance width (n=50), the mean value was 15.4 (SD 0.54) cm. Verification of the center of pressure measurement function (n=50) showed that 48% of the measurements had a deviation of ≤10 mm, and 82% of the measurements had a deviation of ≤20 mm.

Conclusions: In this paper, we have presented the iScoolio concept, which has the potential to support home-based scoliosis therapy through adaptable devices and tracking elements. Integrated sensors enable monitoring of movements, breathing, forces, and pressure distribution during training. An associated web application visualizes the recorded data and serves as a guide for performing the exercises. Technical verification indicates that the integrated tracking methods are functional; however, there is still room for optimization in terms of hardware and software. Further validation is required before conducting clinical trials under real-world conditions.

(JMIR Rehabil Assist Technol 2026;13:e93721) doi:10.2196/93721

KEYWORDS

scoliosis; Schroth therapy; physiotherapeutic scoliosis-specific exercises; PSSE; training device; exercise system; assistive devices; test bench; spine model; breathing simulation; motion tracking

XSL•FO
RenderX

Introduction

Background

In the field of conservative scoliosis therapy, therapeutically guided physiotherapeutic scoliosis-specific exercises (PSSE) are supplemented by training units that patients perform independently at home. One of the 7 major schools in the field of PSSE is Schroth therapy, which has established itself worldwide and especially in Germany. Developed in 1920 by Katharina Schroth, this therapeutic method follows five principles: (1) auto-elongation (detorsion), (2) deflection, (3) derotation, (4) rotational angular breathing (RAB), and (5) stabilization. On the basis of diagnostic findings, Schroth also divides the human body conceptually into a 3- or 4-block model (Appendix Chapter 1, Multimedia Appendix 1), which then serves as the basis for the therapy plan. The change in shape of these blocks from a rectangle to a trapezoid and the horizontal positioning visualize the scoliotic trunk deformation and lateral displacement [1].

Another special feature of Schroth therapy compared to other PSSE is RAB. This uses the connection between the ribs and spine and the increase in volume of the chest during breathing to correct the concave and convex areas of the torso through targeted contraction of the torso muscles. This means that every breath can also be a corrective exercise [1]. Some of the frequently used exercises in Schroth therapy, which involve RAB, are “Between Two Poles” (Zwischen zwei Stäben), “50X” (50X Pezziball), and “Sail” (Segel) [1,2]. Further information on these exercises can be found in Multimedia Appendix 2.

These exercises require patients to be able to understand the complex movements cognitively and motorically. This often requires intensive therapy assistance from a therapist or physician, which is not available in the home setting [3]. The lack of safety when training independently and the lack of suitable therapy equipment therefore often lead to reduced motivation, which in turn has a negative effect on the success of the therapy. Various possible solutions are described in the literature: Portable textile systems such as the “Wearable Sensor Array” by Caviedes et al [4], the “Scoliosis Rehabilitation Assistant” by Li et al [5], the “SpineCurer” by Mai et al [6], or the “Wearable Scoliosis Monitoring System” [7] use stretch or inertial sensors placed on the back to track movement. Pattern recognition and linked apps can then be used to monitor scoliosis exercises and provide training instructions [4-6]. Stationary training systems such as the “DISC4SPINE-System” [8,9] or the “rehabilitation exercise system” enable exercises in a seated position [10] or while standing and monitoring of therapy progress using various sensors [8,9]. In addition, several apps and web-based systems are presented, such as “ScolioTrack,” “ScolioMeter,” “ScolioDetector,” “Scoliosis Tracker,” or the “Scoliosis Manager,” which determine the progression of spinal curvature via smartphone, remind patients of appointments, or assist in the management of scoliosis rehabilitation [11,12].

All these wearables, training devices, and apps offer approaches and partial solutions for improving and supporting home-based scoliosis therapy. However, the current literature has not yet identified a holistic system that enables a wide range of exercises and guides and monitors them and, at the same time, increases motivation for training.

Study Aims

In a previous study [13], we investigated the potential for optimizing home-based scoliosis therapy by conducting online surveys of patients with scoliosis and therapists. The results showed that, according to their own statements, patients exercise significantly less at home than what their therapists recommend. The reasons for this are a lack of motivation and uncertainty when performing complex exercises, a lack of awareness of the importance of training, and the absence of supervision by therapists [14,15].

The patients surveyed stated that their motivation could be increased by training together with friends and acquaintances, by a supportive therapy system, and by a digital profile. We also asked about the popularity of various training devices. It turned out that balance boards, wall bars, mirrors, and exercise sticks are among the most popular devices, but that patients use them relatively rarely at home (≤55.6%). Regarding digital tools, patients attached great importance to an app, video support, and vibration feedback. Therapists saw particular advantages in increasing motivation, improving home-based therapy, monitoring therapy progress, and communicating therapy instructions [13].

On the basis of the results of our survey and the lack of devices on the market, the aim of this study was to develop a novel concept for a holistic and digitally assisted scoliosis training system. Consequently, this study focuses on the technical verification of the demonstrator, serving as a prerequisite for future clinical evaluation and formal medical device development. To this end, we have formulated the following hypotheses:

- H_0: The integrated sensor framework of the new home-based scoliosis therapy concept provides valid and accurate measurements against technical references.
- H_1 (motion tracking): The demonstrator yields motion measurements that closely align with high-precision optical reference data across various test bench positions.
- H_2 (respiration tracking): The demonstrator accurately detects simulated breathing cycles and variations in breathing depth generated by a test bench.
- H_3 (force tracking): The demonstrator provides clear, load-proportional raw sensor deflections in response to applied forces, confirming its technical capability to capture force variations.
- H_4 (stance width and center of pressure tracking): The demonstrator reliably determines stance width and center of pressure (COP) using 2 test weights to simulate a 2-foot stance.

Methods

Study Design

This study was designed as a proof-of-concept and technical verification study. It describes the development process and the resulting demonstrator, its functionalities using exemplary

XSL•FO
RenderX

Schroth exercises, and the technical verification of its measurement functions. The study's findings correspond to Technological Readiness Level 4.

Ethical Considerations

As part of the study, no human participants or animals were involved, and no clinical trials were conducted. Images showing people together with the demonstrator are for illustrative purposes only. All devices containing electrical components were deactivated at the time.

After consultation with the Ethics Commission in Leipzig (Germany), and in accordance with section 15 of the professional code of conduct of the Saxony State Medical Association, no ethics approval was required for the work described.

Proof-of-Concept Development

In the first phase of the development process, we identified requirements for a training system with digital assistance based on the wishes and challenges associated with home-based scoliosis therapy [13]. To address the lack of therapeutic support, low motivation, and limited space, the system should have the following features: a digital profile, exercise instructions, measurement functions to monitor exercise execution, low space requirements, modifiability for holistic training, and the integration of established but rarely used at-home training devices such as balance boards, wall bars, mirrors, and exercise sticks. We have transferred these requirements into a modular approach, which combines conventional devices with sensor functions, and further specified them (Table 1).

Table . Requirements specification for the 4 modules of the new training system.

Specification	Module 1	Module 2	Module 3	Module 4
Device	Mirror	Wall bar and exercise stick	Balance board	Mirror
Form	Smart textile and wearable	Modifiable training device	Balance training device	Sensor hub and web interface
Function	Monitoring exercise performance	Versatile training with minimal space requirements	Training proprioception, self-correction, and compensating for imbalances	Preparation and visualization of measurement data and output of exercise instructions
Tracking	Breathing (chest and abdomen) and movement (torso and head)	Pressure and gripping forces	Inclination and pressure distribution	None

For each of these modules, the names, designs, and sensor concepts were created in a second development phase. The first module was named "Wearable-Module" due to its design as a portable, functional textile that is primarily used for tracking. Modules 2 and 3 represent the training components, with the second module named "Bars-Module" because of its rod-shaped design and the third module named "Balance-Module" due to its resemblance to conventional balance boards. The fourth module is used to visualize all sensor data and acts as an interface between the user and the new system, which is why it was named "Interface-Module." For the entire system, the acronym "iScoolio" was chosen, which combines the terms "scoliosis" and "school" and describes the process of "I school myself," as the guiding principle of home-based scoliosis therapy.

In the final development phase, we implemented a demonstrator for each module, which is presented as the "iScoolio concept" in the *Results* section. A key component of this work was the development of a custom circuit board (iScoolio board). The board consists of a microcontroller (BlueNRG-355MT; STMicroelectronics N.V.), inertial sensors (BHI260AB and BMM150; Bosch), and a Wheatstone quarter bridge (R=10 kΩ), and it serves as a basis for all sensor functions. Powered by a lithium battery (LIR2032; AccuCell) and housed in an enclosure (70×34×17 mm), each circuit board can be used as a stand-alone unit. For wireless communication, each board has its own ID and radio channel and transmits its data every 6 milliseconds via its own 2.4 GHz frequency channel. This interval was chosen as a compromise between transmission speed and power consumption. The corresponding receiver scans each channel for 13 milliseconds to ensure data acquisition in case the first transmission fails. This resulted in a sampling frequency of 4 Hz.

Theoretical Evaluation of Use Case Scenarios

A theoretical use case evaluation was conducted using the developed concept for the exemplary Schroth exercises "Between Two Poles," "50X," and "Sail" (Multimedia Appendix 2). Of particular interest were the variability, usability, and tracking capabilities of the system.

Technical Verification

Overview

To test H_0, a technical verification of the demonstrator's measurement functions was conducted. This included motion tracking (H_1) and respiration tracking (H_2) of the Wearable-Module, force tracking in the handles and supporting feet of the Bars-Module (H_3), and measurement of COP and detection of the stance width in the Balance-Module (H_4). Test scenarios were developed for each of these functions, where the number of measurements was derived from the information for advanced users in the Schroth exercise manual (Multimedia Appendix 2). As identical hardware is used multiple times within a single measurement function, only a subset of the sensors used in the Wearable-Module and Bars-Module was tested on a sample basis. The data were measured using the iScoolio boards, with the Interface-Module serving as a wireless receiver. All

XSL•FO
RenderX

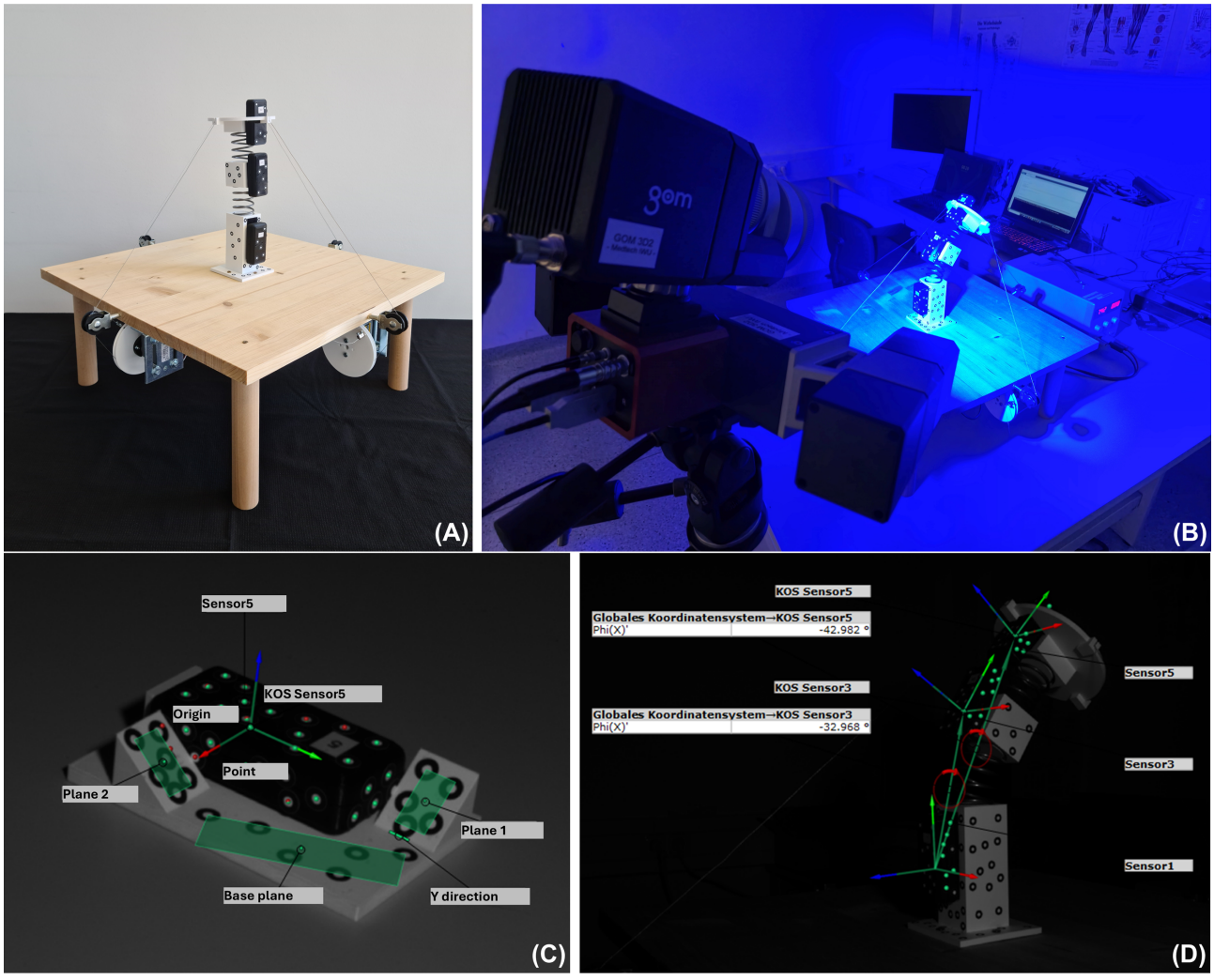
measured values were recorded and stored in a CSV file using a data logging function of the Interface-Module. The data were analyzed and prepared graphically using the OriginPro software (version 2023; OriginLab Corporation).

Wearable-Module

Motion Tracking

To verify H_1, a test bench was developed that can move to a specified position with repeatable accuracy. This consisted of a wooden table, 4 servo motors (2000 Series Dual Mode Servo [25 - 2, Torque]; goBILDA) with 3D-printed pulleys and deflection rollers, a microcontroller (Uno R3; Arduino), and a distribution board (Servo Power Distribution Board [8 Channel]; goBILDA). Additional components were additively manufactured for the top of the table to serve as holders for 3 iScoolio boards. The boards could be arranged in a tower shape using a Velcro connection. The bottom board was fixed in place, while the top 2 were mounted on pressure springs for mobility. The concept thus resembled a greatly simplified spine with fixed elements (vertebrae/3D-printed components) and flexible elements (intervertebral discs/compression springs; Figure 1A). With the help of servo motors arranged at a 90° angle and wire ropes mounted at the top of the tower, the simplified spine structure could be moved 3D in space within the volume of a hemisphere.

Figure 1. Verification of the Wearable-Module's motion tracking. Actuator test bench designed as a simplified spinal model for reproducible motion of 3 iScoolio boards (A) in the neutral position and (B) during optical measurement with the ARAMIS 3D system. (C) Screenshot from the GOM Correlate Pro 2022 software illustrating virtual placement of a local coordinate system (KOS Sensor5) at the inertial measurement unit position on an iScoolio circuit board. (D) Screenshot showing the coordinate systems of sensors 3 and 5 during motion in the sagittal plane, including the calculated rotation angles relative to sensor 1.



Three iScoolio boards (1, 3, and 5) were attached to the test bench to verify motion tracking. These boards were selected to mirror the planned vertical arrangement and because sensor 1 serves as the reference for all the others. On the basis of the "50X" exercise, 2 positions in the frontal plane (yaw) and 2 positions in the sagittal plane (pitch) were then approached 50 times, with an intermediate return to the zero position, and held for 10 seconds each. In the frontal plane, the "spine" was tilted once to the left and once to the right, and in the sagittal plane, once anteriorly and once posteriorly. At the same time, the sensor data (Euler angles) were recorded.

A marker-based optical system (ARAMIS 3D camera, 600 mm measuring volume; Carl Zeiss GOM Metrology GmbH) served as a reference for this verification (Figure 1B). After covering all relevant surfaces with 3 mm measuring points, a manual

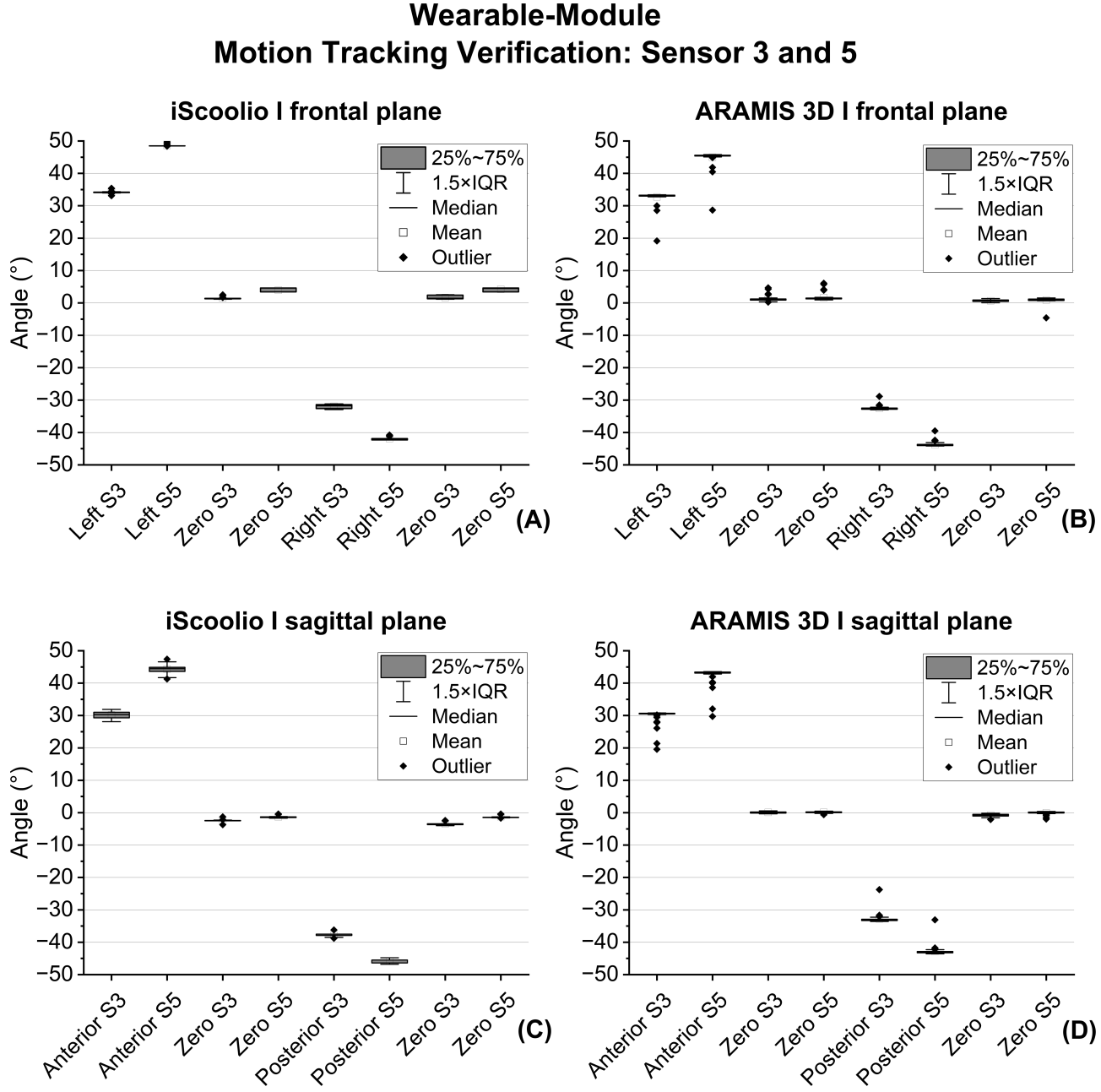
XSL•FO
RenderX

image was taken at each position approached during the test run using a manual trigger.

As a basis for evaluating the recorded images, the actual sensor position of the inertial measurement unit (IMU) within the iScoolio board housing had to be determined by software. For this reason, an additional adapter was additively manufactured to enable the calculation (Figure 1C). For each of the 3 sensors, the graphically determined points were then defined as the origin of a local coordinate system in the 3D measurement software GOM Correlate Pro 2022 (Carl Zeiss GOM Metrology GmbH) associated with the ARAMIS 3D system.

Using these 3 coordinate systems, the inclination for the approached positions could be determined for sensors 3 and 5 in relation to sensor 1 (Figure 1D). In Figure 2, the angle measurements from both systems are depicted as box plots.

Figure 2. Verification of the Wearable-Module's motion tracking. (A, B) Box plots of the measured angles of sensors 3 and 5 during repeated movement to 2 positions in the frontal plane, including return to the 0 position. (C, D) Box plots of the measured angles during repeated movement to 2 positions in the sagittal plane, including return to the 0 position. (A, C) Measured values of the Wearable-Module. (B, D) Measurements taken with the ARAMIS 3D system.

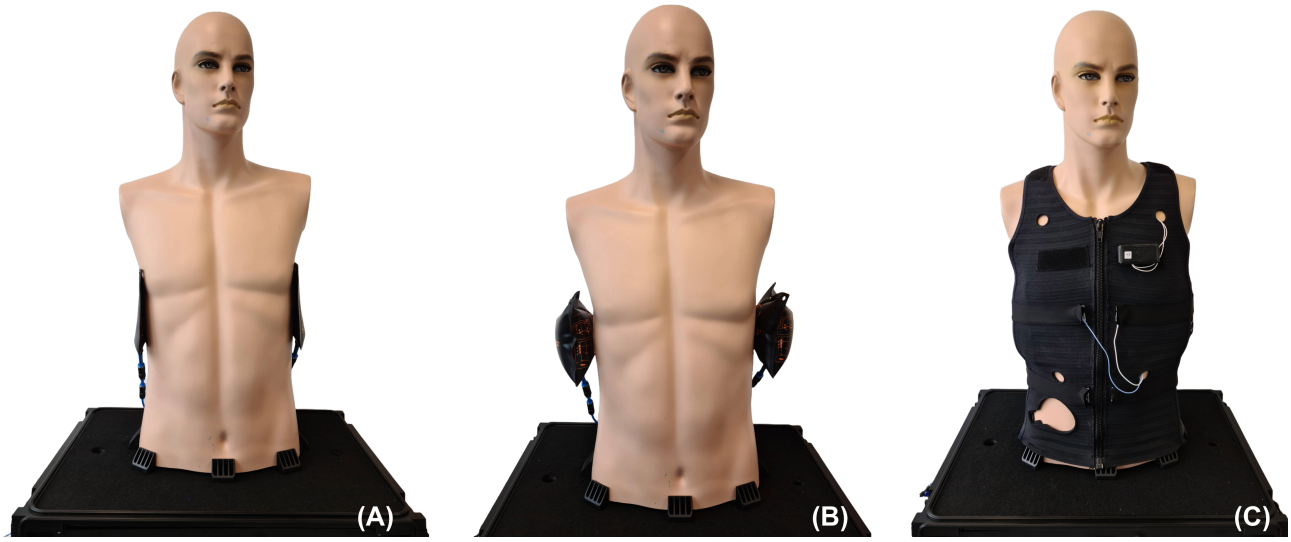


Respiration Tracking

A torso test bench with breathing simulation was set up for evaluating H_2. To simulate breathing, 2 pneumatic mounting cushions were attached laterally to a mannequin, which generated a stroke of 75 mm at a pressure of 0.5 bar. This was intended to simulate the lifting of the chest during inspiration (Figure 3). The mounting cushions were filled and emptied using a proportional pressure control valve (VEAB-L-26-D07-04-A4-1R1; Festo), which was supplied with 3 bar of compressed air and controlled via an Arduino circuit.

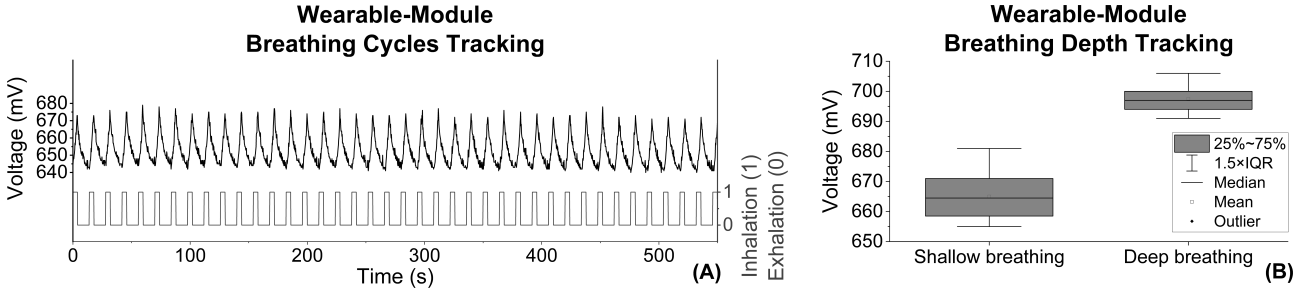
XSL•FO
RenderX

Figure 3. Verification of the Wearable-Module's respiration tracking. A torso model with laterally applied air cushions served as the test device. (A) Expiration. (B) Inspiration without the Wearable-Module. (C) Inspiration with the Wearable-Module donned.



Using this test bench, 2 measurements with the Wearable-Module attached were carried out. The first test, "frequency measurement," was designed to examine the detection of breathing cycles. On the basis of the "Sail" exercise, 40 breathing cycles (5×8 breaths) were simulated while recording the sensor data. In the second test, "amplitude measurement," it was checked whether the depth of breathing could be determined. For this purpose, 40 cycles of shallow breathing (4 seconds of filling and 10 seconds of emptying) and deep breathing (6 seconds of filling and 15 seconds of emptying) were simulated while logging the measured values. The results of both experiments are shown in Figure 4. For the frequency measurement, the trigger curve was plotted in time synchronization with the measured values. For the amplitude measurement, 2 box plots were calculated. This calculation was based on the measured maxima for shallow and deep breathing in each of the 40 cycles. After checking the normal distribution of the differences between these 2 datasets, a 2-tailed paired *t* test (α=.05) was used to evaluate whether the difference was significant.

Figure 4. Verification of the Wearable-Module's respiration tracking. (A) Measured values for 40 simulated breaths over time, including the trigger signal. (B) Measured values for simulated shallow and deep breathing in the form of 2 box plots.



Bars-Module

Force Tracking: Handles

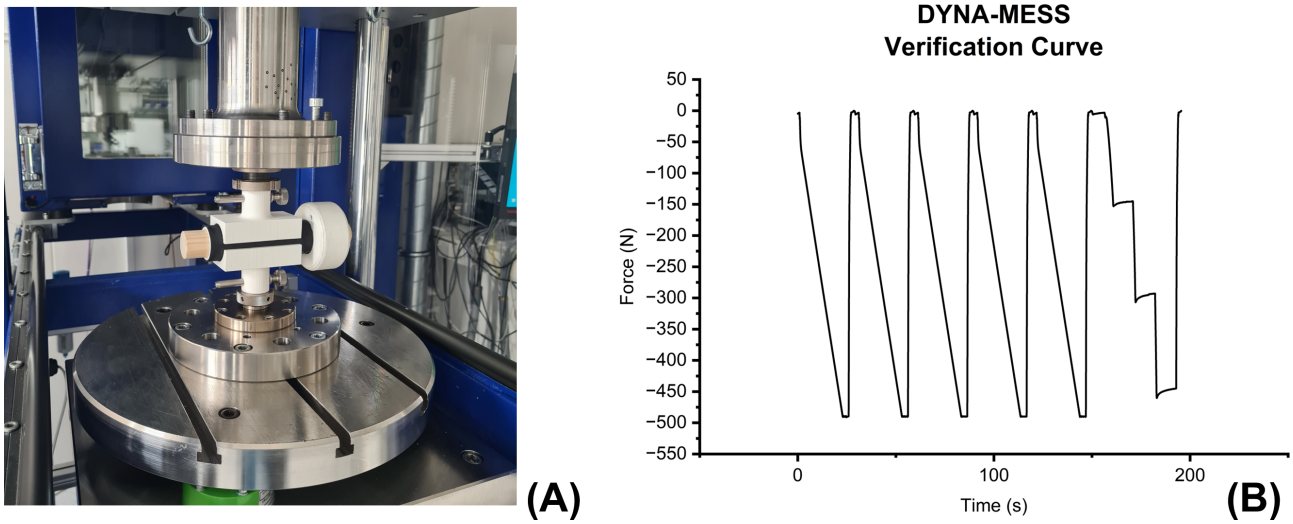
To verify H_3, a test setup was designed for the handles of the Bars-Module that used a servo-electric testing machine (tensile/compression and torsion testing machine, 10 kN/200 Nm; DYNA-MESS Prüfsysteme GmbH) as a reference. The verification was conducted at 3 evenly spaced load levels to fully cover the sensor's measuring range. The following test protocol was implemented for the testing machine for verification purposes:

1. Preconditioning: Apply 110% of the maximum force (484 N) to the sample, hold for 3 seconds, and then release. This procedure was repeated 5 times.
2. First force level: Apply one-third of the maximum force (147 N) to the test specimen and hold for 10 seconds, then release.
3. Second force level: Apply two-thirds of the maximum force (296 N) to the sample and hold for 10 seconds, then release.
4. Third force level: Apply the maximum force (444 N) to the test specimen and hold for 10 seconds, then release.
5. Repeat the cycle 40 times in total.

The cycle count of 40 is attributable to the performance of the "50X" exercise. To be able to mount the iScoolio handpiece on the testing machine, 2 half-shell adapters, designed to closely resemble a grasping hand, were additively manufactured. As a replacement for the rod-shaped base body of the Bars-Module, a round wooden profile (diameter 28 mm and length 19 cm) was also inserted into the handpiece to provide a counterbearing. After placing the handpiece in the testing machine (Figure 5), the test program was started, and the sensor data were recorded.

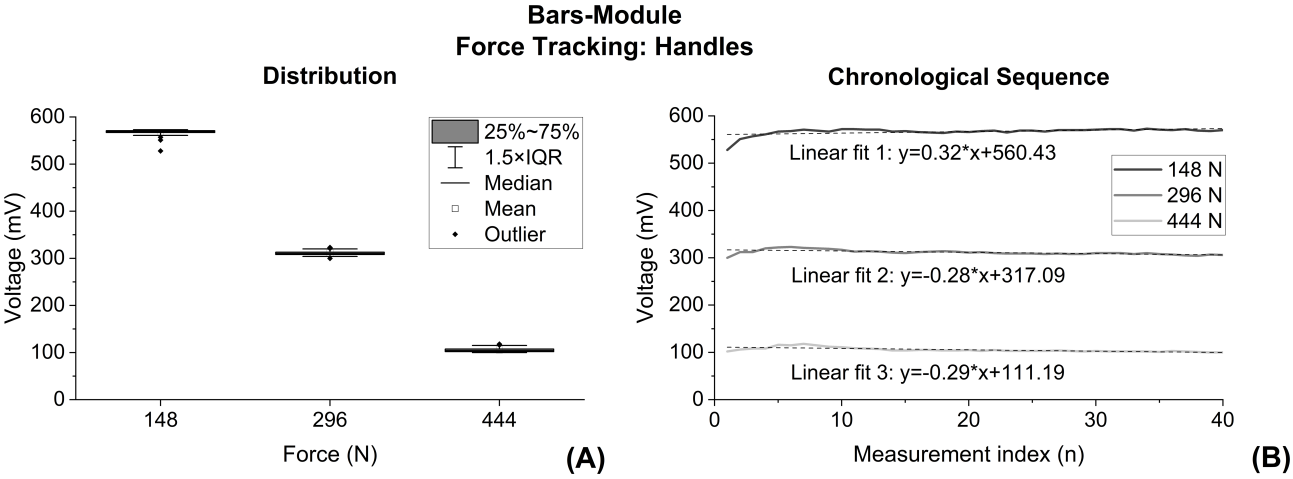
XSL•FO
RenderX

Figure 5. Verification of the Bars-Module's force tracking (handles). (A) Measurement setup. (B) Verification curve of the testing machine.



To examine the differentiability of the force levels, repeatability, and drift, the measurement data were evaluated in the form of 2 diagrams. First, 3 box plots were created for the measured values of the 3 force levels. These were based on the calculated mean values of the 10-second holding phases, excluding the first and last seconds, to eliminate overswinging of the testing machine. Second, these calculated mean values were visualized in chronological order in a line diagram (Figure 6). To assess significant differences among the 3 force levels in a within-subject design (N=40, k=3), we used the Friedman test. The Dunn test was subsequently applied for post hoc pairwise comparisons following a significant Friedman result. Significance was set at an α of .05. The choice of this procedure reflects the within-subject structure (3 related samples) and the nonnormality of the differences among the 3 conditions, as indicated by the Shapiro-Wilk test (α=.05).

Figure 6. Verification of the Bars-Module's force tracking (handles). Measured values for the 3 force levels are visualized in the form of (A) 3 box plots and (B) a line diagram.



Force Tracking: Supporting Feet

Similar to the measurements of the iScoolio handles, the supporting feet were also verified using the servo-electric testing machine, contributing to the investigation of H_3 (Figure 7). Due to the slightly larger force range of the sensors installed in the bases, the test protocol for the handpieces was adjusted to a maximum force of 445 N. In addition, the number of cycles was reduced to 25, corresponding to the repetitions (5×5) for the "Between Two Poles" exercise. The analyzed measurement data are shown in Figure 8.

A Friedman test for related samples (N=25, k=3) assessed differences across the 3 force levels, and the Dunn test was subsequently applied for post hoc pairwise comparisons, with an α of .05. The test procedure was selected in the same way as for the handles.

XSL•FO
RenderX

Figure 7. Verification of the Bars-Module's force tracking (supporting feet). The test setup is shown in 2 different views (A and B).

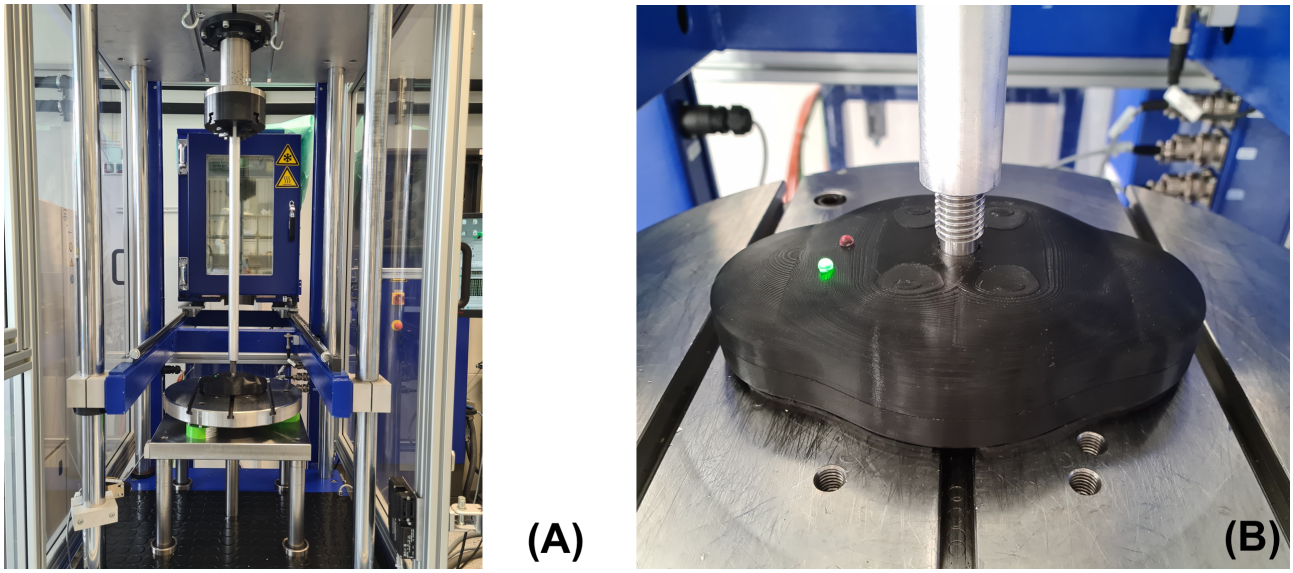
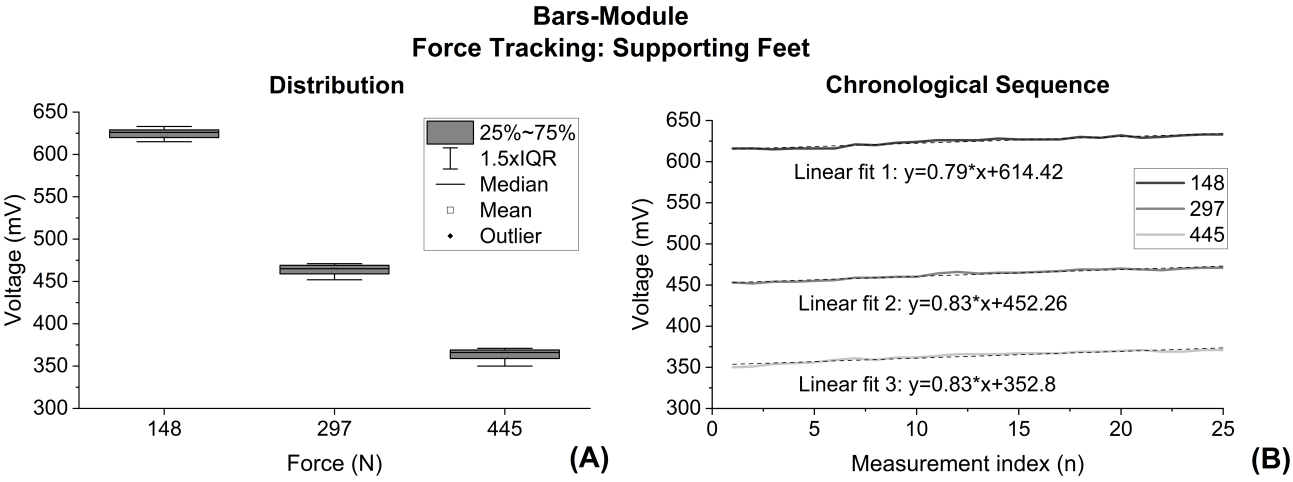


Figure 8. Verification of the Bars-Module's force tracking (supporting feet). The measured values for the 3 force levels are visualized in the form of (A) 3 box plots and (B) a line diagram.



Balance-Module

Stance Width

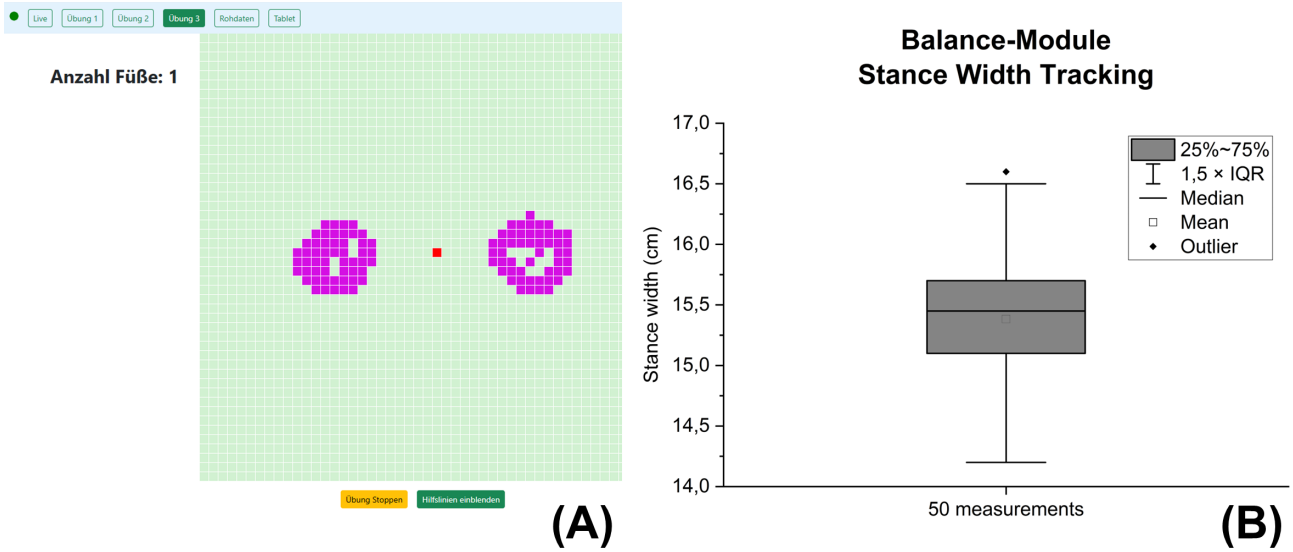
The evaluation of H_4 involved examining 2 measurement functions of the Balance-Module. To assess the first function, stance width measurement, the board was locked in the neutral position using the base feet. In addition, the measurement surface was covered with a transparent film so that reversible markings could be made. For the verification, 2 test weights (5 kg) were placed on the board so that they touched each other and were then manually pushed apart until the Interface-Module indicated a color change for the minimum distance being reached (Figure 9). The positions of the weights were then marked on the transparent film in color, and their shortest distance was measured. This test procedure was performed 50 times. The corresponding box plot is shown in Figure 10.

XSL•FO
RenderX

Figure 9. Verification of the Balance-Module's stance width tracking. The test setup consists of the locked Balance-Module with applied film and 2 test weights, as well as the Interface-Module.



Figure 10. Verification of the Balance-Module's stance width tracking. (A) Screenshot of the Interface-Module with test weights that are too close together, as indicated by the same color appearing twice. (B) Box plot for measuring stance width over 50 passes.

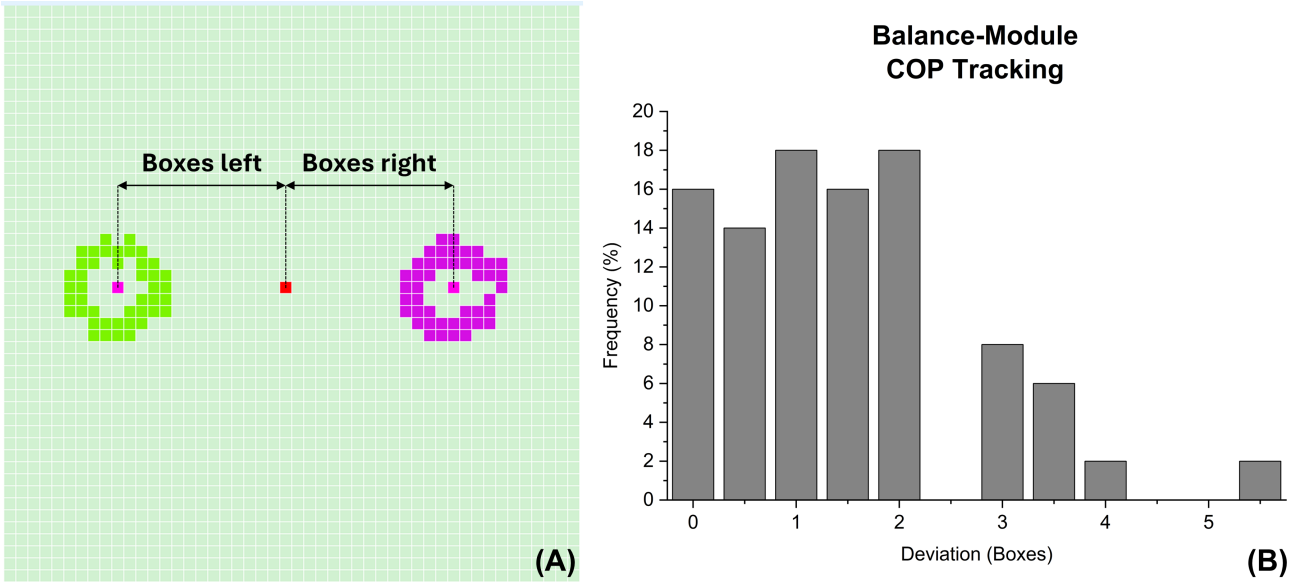


Center of Pressure

To verify the second measurement function, the COP measurement, the board was also locked in a neutral position, and the transparent film was marked with horizontal and vertical lines spaced 4 cm apart. This distance was determined by the diameter of the 5-kg test weights. While maintaining the minimum distance, 2 of these test weights were placed at 50 different positions along the grid on the board, always in a horizontal alignment to simulate a normal 2-foot stance. A screenshot was then taken, and the boxes between the local COPs and the global COP were counted (Figure 11A). These were subtracted from each other and halved to determine the deviation from the target value in terms of boxes. Due to the equal weight of the 2 test weights, the global COP should be exactly in the middle between the local COPs. The histogram for the 50 measurements is shown in Figure 11B.

XSL•FO
RenderX

Figure 11. Verification of Balance-Module's center of pressure (COP) tracking. The figure shows a screenshot of the Interface-Module (A) from 1 of the 50 measurements with 2 local COPs of the clusters and a global COP of the entire measurement area (red dot). The double arrows show the distances between the COPs. The histogram (B) visualizes the frequency of the deviations determined in boxes of the global COP (direction unspecific) from the target value. The value 0 thus describes the perfectly centered alignment of the global COP between the 2 local COPs.



Results

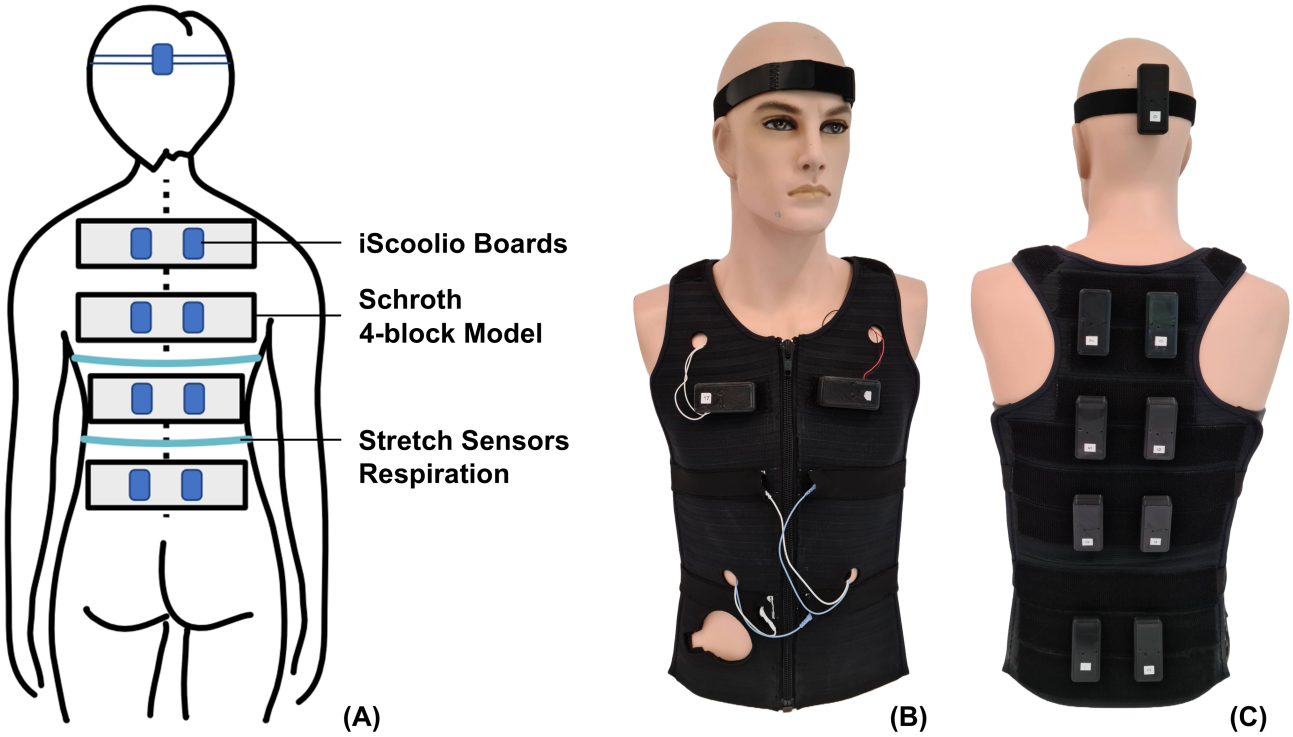
iScoolio Concept

Wearable-Module

The Wearable-Module consists of a headband and a vest with elastic elements on the sides to adapt to the wearer's body shape (Figures 12B and 12C). Two resistive textile strain sensors, each connected to an iScoolio board, record the circumferential changes of the chest and abdomen during breathing. This allows both inspiration and expiration, as well as the relative depth of breathing, to be detected. For motion tracking, the module contains a total of 9 iScoolio boards, 1 on the headband and 8 distributed across the back. The positioning on the back was based on the 4-block model used in Schroth therapy. For each block, a total of 2 iScoolio boards were placed parallel at the same height on the back (Figure 12A). The movement data are derived based on quaternions. iScoolio board 1, which was placed on the left side of the hip-pelvic block, serves as a reference for all other motion tracking sensors, as there is less movement in this position compared to the other positions further up. The Wearable-Module thus serves as a tracking component for movement and breathing.

XSL•FO
RenderX

Figure 12. (A) Schematic placement of the iScoolio boards and strain sensors in the Wearable-Module. (B, C) Transfer to a functional prototype based on the 4-block model according to Lehnert-Schroth [1].



Bars-Module

The Bars-Module consists of 2 bars with attachments and combines the functions of exercise sticks and wall bars. Both bars are made of an aluminum profile and have an internal thread on both ends. Using a left-hand and a right-hand thread and the corresponding threaded rods, the Bars-Module can be braced in the door frame to simulate a wall bar (Figures 13A and 13B). It can also be used as an exercise stick by changing its length with an extension piece (Figure 13C). To simulate the wall bars, supporting feet with an antislip coating have been designed, which provide additional support when clamped in the door frame (Figure 13D). The Bars-Module is completed by sliding handles, which are also equipped with a special antislip coating to ensure a secure grip. Resistive force sensors were integrated into both the handles (FlexiForce A502; Tekscan) and the supporting feet (FlexiForce A201; Tekscan). These enable touch-specific (eg, “reach for the rungs”) as well as force- and intensity-specific training instructions (eg, “push both bars into the floor as you exhale”) to be followed. In addition, the clamping force of the Bars-Module can be measured when it is clamped into the door frame. The sensors are read, and the data are transmitted wirelessly via an iScoolio board, which has been integrated into each attachment. As an additional safety feature, the Bars-Module also offers the option of providing visual feedback on the clamping force via an LED integrated into the supporting feet (Figure 13D). It thus functions as a configurable and versatile training device for the new therapy system.

Figure 13. Use of the Bars-Module as a (A) single rung with a gymnastic band, (B) wall bar replacement with 2 rungs, and (C) long exercise sticks. (D) The green and red LEDs inside the supporting feet can be used to provide visual feedback to the user.

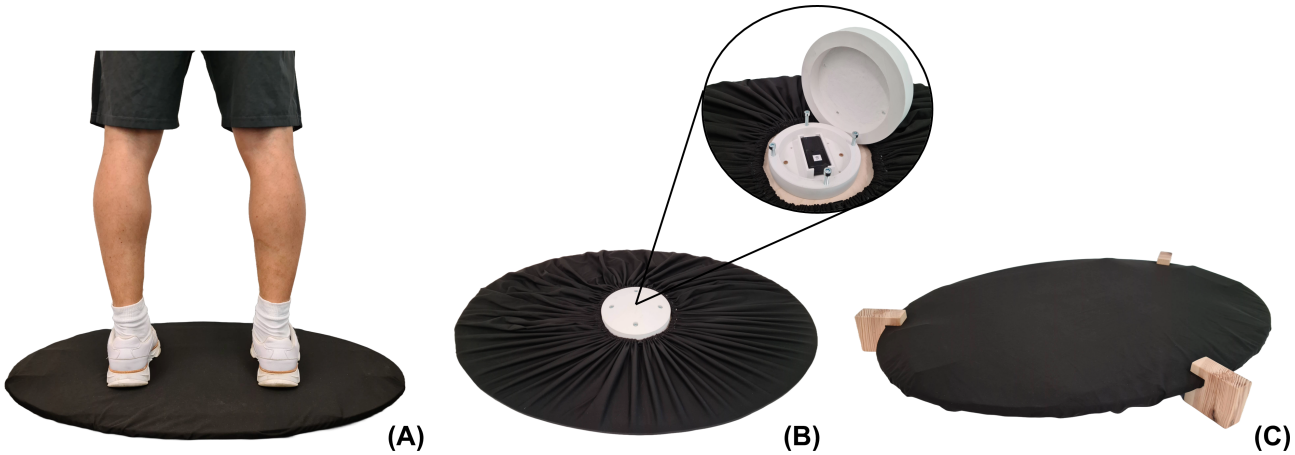


XSL•FO
RenderX

Balance-Module

The Balance-Module was based on the design of existing balance boards. To achieve the desired unstable surface, an oval-shaped base plate was made of wood and fitted with a 2-part, 3D-printed plastic element at the center of the underside. This consisted of an adapter plate with an iScoolio board mount and a screw-on ball segment (Figures 14A and 14B).

Figure 14. (A) Top and (B) bottom of the Balance-Module with (C) feet for targeted decoupling of mobility. The iScoolio board is integrated via the additively manufactured ball segment, which was designed to be modular and with an associated socket (B, detailed view).



The typical balance training device was expanded to include sensor components, which are of central importance for guiding and monitoring exercises in our development. On the basis of the structural and functional requirements of a balance board, 2 monitoring parameters were selected: the inclination of the board and the pressure distribution on the standing surface. The former was achieved by integrating the iScoolio board into the adapter plate. For the latter, a pressure measurement mat with 2304 sensor elements (Health Mat Dev Kit 1.9; Sensing Tex) was applied to the base plate. Specific algorithms were implemented to measure the number of feet on the board, the foot pressure distribution, the stance width, and the COP. When the sensor surface is loaded, adjacent sensor elements are combined into a cluster by the software while maintaining a specified maximum distance. A local COP is then calculated for each of these clusters, and the pressure distribution is visualized. On the basis of the intended application scenario, a detected cluster is recognized as a single foot. When a second cluster is detected, an additional algorithm checks the distance between the two. If this exceeds a specified value, the output indicates that a second foot is on the board. With the help of this parameter, the minimum stance width can be determined and tracked. Finally, an algorithm was implemented that calculates and visualizes a global COP for the entire sensor surface. To prevent the board from moving during these measurements, wooden feet were developed that can be attached to lock the Balance-Module in a horizontal position (Figure 14C). Conversely, it is also possible to shift the COP in a targeted manner as part of a PSSE exercise. To protect it from external influences and enhance its appearance, the Balance-Module was finally encased in a cover made of black functional jersey (PA 6.6). Alongside the Bars-Module, it is the second training device of the new system.

Interface-Module

In terms of hardware, the Interface-Module consists of 2 physically separate systems: a data processing unit (DPU; MeLE Mini PC Quieter 3Q Celeron N5105; Shenzhen MeLE Digital Technology Ltd) with a separate Bluetooth receiver (BlueNRG-LP; STMicroelectronics N.V.) and an output device, such as a standard mobile phone or tablet as a portable display device. Communication between the 2 components takes place via a wireless local area network data connection, while communication with the other modules takes place via Bluetooth. The software developed for the Interface-Module consists architecturally of several logical individual components, with data processing carried out in Python (Python Software Foundation) and the implementation of the user interface based on TypeScript (Microsoft Corp) and the Angular (Google) framework. A dashboard was set up to visualize all sensor data (Figure 15).

XSL•FO
RenderX

Figure 15. Dashboard of the iScoolio Interface-Module.



To implement therapy, the user selects an exercise from the catalog and starts it. Exercise performance is monitored by the DPU, and sensor data received via Bluetooth are evaluated against target parameters. The user receives real-time visual feedback on body position and exercise-specific instructions. A basic safety monitor protects against harm from system or application errors. If a sensor malfunctions, data transmission is interrupted, or tolerances are exceeded, the system enters an error mode with audible and visual warnings.

Theoretical Evaluation of Use Case Scenarios

In a thought experiment, the potential applications and limitations of the iScoolio concept were examined using the 3 exemplary Schroth exercises, “Between Two Poles,” “50X,” and “Sail.” As a result, Table 2 lists which modules can be used for the respective exercises and what functions they fulfill. Figure 16 shows what the use of the system might look like during the performance of the exercises.

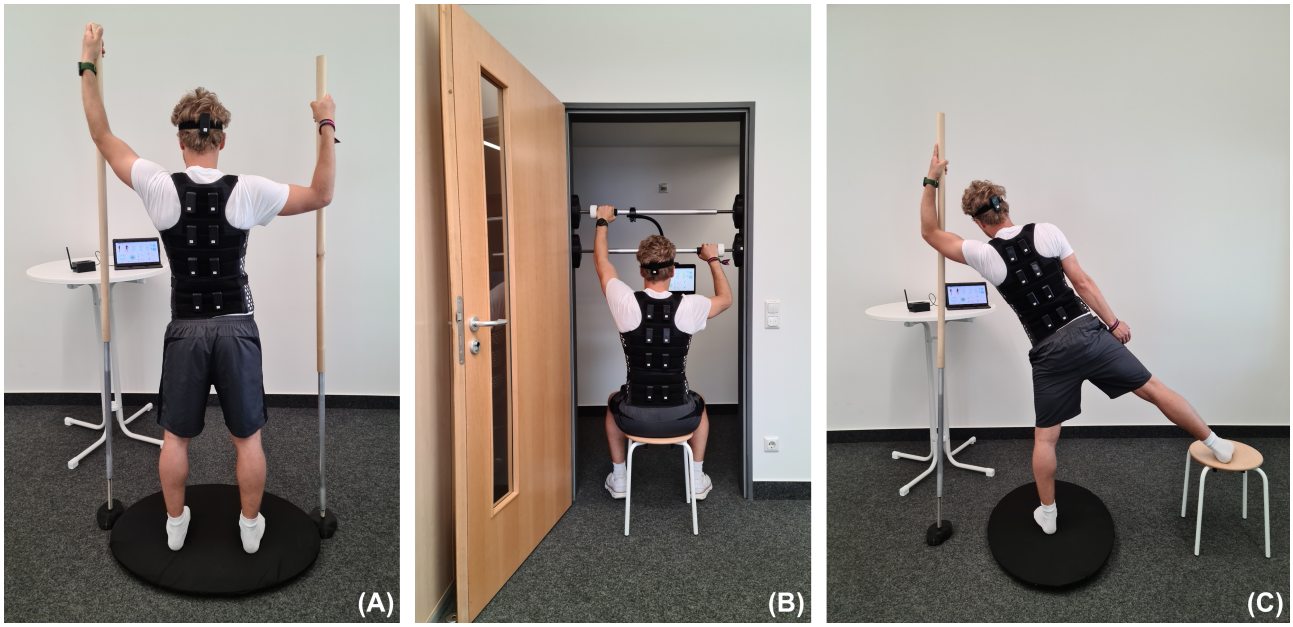
With the current state of development, the following limitations would apply to the intended application scenario: a lack of motion tracking for the arms and legs; a low sampling rate of 4 Hz; a lengthy setup process due to the modular design and the fact that all iScoolio boards must be activated and calibrated separately; and unsuitability for exercises performed in the supine position due to the sensor housings on the back.

Table . Possible use cases of the iScoolio modules for 3 exemplary Schroth exercises.

iScoolio modules	Between Two Poles	50X	Sail
Wearable-Module	Tracking: upright posture and breathing	Tracking: alignment of the back and breathing	Tracking: upper body and head position and breathing
Bars-Module	Function: support elements; tracking: support force	Function: wall bars, tablet holder; tracking: grip strength	Function: support element; and tracking: support force
Balance-Module	Function: unstable surface; tracking: balancing and number or position of feet	Optional function: wobbly seat pad; tracking: balancing	Function: unstable surface; and tracking: balancing and number or position of feet
Interface-Module	Providing exercise instructions, visualizing measured values, and providing feedback on exercise performance	Providing exercise instructions, visualizing measured values, and providing feedback on exercise performance	Providing exercise instructions, visualizing measured values, and providing feedback on exercise performance

XSL•FO
RenderX

Figure 16. Potential applications of the iScoolio modules for the following sample exercises: (A) “Between Two Poles,” (B) “50X,” and (C) “Sail.”



Verification of the Measurement Functions in the Laboratory

Wearable-Module

Motion Tracking

For all 4 positions approached, the mean values showed an absolute deviation between iScoolio and ARAMIS 3D of 2.29° (SD 1.66°). The largest deviations were observed in the posterior movement on the sagittal plane. The mean values of the zero positions were positive for both measuring systems in the frontal plane (iScoolio: 2.28°, SD 1.44° and ARAMIS 3D: 1.09°, SD 0.42°) and negative in the sagittal plane (iScoolio: −2.11°, SD 1.03° and ARAMIS 3D: −0.16°, SD 0.43°), with larger deviations occurring in the iScoolio demonstrator. The maximum SD in the iScoolio system was 1.48° (sagittal, S5, anterior; Multimedia Appendix 3). The measured values in the frontal plane were more accurate than those in the sagittal plane. The ARAMIS 3D reference system produced more pronounced outliers, which always affected sensor 3 and sensor 5 equally in the corresponding measurements (Figure 2).

Respiration Tracking

From the examination of the breathing cycle plot, the 40 simulated cycles can be clearly extracted from the sensor data provided by the Wearable-Module. For each cycle, there is a jagged curve with a rising flank during inhalation, a falling flank during exhalation, and a maximum at the change between the 2 states in this order. Fluctuations and outliers occur in particular when the filling volume of the artificial lung is low and during the transition between exhalation and inhalation.

For breathing depth, 2 datasets were obtained from the measured maxima: the mean for shallow breathing was 665.0 (SD 7.4) mV, and the mean for deep breathing was 697.4 (SD 3.7) mV. A paired *t* test of 40 paired measurements yielded a *t* value of −27.4 (df=39; *P*<.001), indicating a 2-tailed significant difference between the maxima. Despite all this, there is a high degree of variation within the 2 groups, which decreases with increasing breathing depth and thus increasing stretching of the sensor (Figure 4).

Bars-Module

Force Tracking: Handles

The diagrams show 3 clearly separated measurement sets, corresponding to the 3 force levels: mean 567.1 (SD 7.5) mV at 148 N; mean 331.4 (SD 5.0) mV at 296 N; and mean 105.2 (SD 4.6) mV at 444 N. A Friedman test indicated significant differences among these levels: χ^2_2=80.0; *P*<.001. Dunn post hoc tests revealed significant pairwise differences: level 1 vs level 2 (*Z*=4.47; *P*<.001), level 1 vs level 3 (Z=8.94; *P*<.001), and level 2 vs level 3 (*Z*=4.47; *P*<.001). Within the box plots, only a few outliers are present. Furthermore, for the 3 equally spaced force levels (ΔF constant), the measured sensor data are not proportional, suggesting a nonlinear sensor response. Moreover, concerning the line diagrams, after a settling phase, a slight drift is observed during the approximately 2-hour and 13-minute test procedure, with a load-dependent drift direction, as indicated by the following slopes *m* and correlation coefficients *r*: *m*=0.32 and *r*=0.50 at 148 N; *m*=−0.28 and *r*=−0.64 at 296 N; and *m*=−0.29 and *r*=−0.74 at 444 N (Figure 6).

Force Tracking: Supporting Feet

The measurements exhibited a pattern similar to the handles. Across the 3 reference forces, no linearity was present in the 148 to 445 N range. The corresponding values for the 3 force levels are as follows: mean 624.7 (SD 6.0) mV at 148 N; mean 463.0 (SD 6.3) mV at 297 N; and mean 363.6 (SD 6.4) mV at 445 N. A Friedman test indicated significant differences among these levels: χ^2_2=50.0; *P*<.001. Dunn post hoc tests revealed significant pairwise differences: level 1 vs level 2 (*Z*=3.54; *P*<.01), level 1 vs level 3 (*Z=7.07*; *P*<.001), and level 2 vs level 3 (*Z*=3.54; *P*<.01). The box plots show no outliers with an IQR of 1.5. Compared to the handles, higher measuring ranges can also be observed for the 3 forces. The time series shows a

XSL•FO
RenderX

roughly linear drift during the approximately 1-hour and 23-minute test procedure, with a positive drift direction, as indicated by the following slopes *m* and correlation coefficients *r*: *m*=0.79 and *r*=0.96 at 148 N; *m*=0.83 and *r*=0.97 at 297 N; and *m*=0.83 and *r*=0.95 at 445 N (Figure 8).

Balance-Module

Stance Width

When examining the stance width, the 50 measurements resulted in an average value of 15.4 (SD 0.54) cm. The box plot also shows a narrow IQR (<1 cm) and only 1 outlier at 1.5× IQR. The range of the measurement is 2.4 cm (Figure 10).

Center of Pressure

During technical verification of the Balance-Module's COP measurement function, the 50 measurements revealed an average deviation of 1.5 (SD 1.2) boxes. The histogram shows that the deviation was ≤1 box in 48% of the measurements and ≤2 boxes in 82% of the measurements. With a hardware center distance of 10 mm between 2 sensor elements or boxes, 82% of the measurements had a COP deviation of ≤20 mm. The largest deviation was 5.5 boxes. Furthermore, it has been shown that the edge areas of the sensor surfaces in particular lead to measurement errors when determining the COP. If the statistical measures were calculated for all measurements in which at least one test weight touched the edge of the sensor surface (n=18), the result was a mean value of 2.6 (SD 1.3) boxes. For all other measured values (n=32), the mean value was 0.9 (SD 0.7) boxes (Figure 11).

Discussion

Principal Findings

This paper introduces a concept for a home-based scoliosis training system, presented as a demonstrator and verified technically. The concept builds on findings from a previous online survey on scoliosis therapy, with a particular focus on independent training at home. Consisting of 4 modules, it enables varied training in a space-saving design, including sensor-based monitoring of exercise execution and the output of instructions without the need for additional conventional training devices. Through dedicated test scenarios and test benches, we verified the implemented measurement functions. The results confirm that the sensor framework yields valid and accurate measurements compared to technical references, in line with H_0. The development can therefore be classified at Technological Readiness Level 4. A comparison with other approaches to scoliosis therapy found in the literature is presented in Table 3. The following sections discuss the selected subcomponents of the iScoolio concept.

XSL•FO
RenderX

Table . Comparison of the iScoolio concept with other systems described in the literature.

Study ID	Motion tracking	Respiration tracking	Force tracking upper limb	BPD[a]	App or web interface	Training concept
"iScoolio"	9×IMU[b]	2×stretch sensor	4×force sensor	Pressure mat	ME[c], EG[d], DS[e]	FT/IE[f]
"Wearable Scoliosis Monitoring System" [7]	3×IMU	—[g]	—	—	—	FT/EE[h]
"Scoliosis Rehabilitation Assistant" [5]	3×IMU	—	—	—	ME, EG, GF[i], DS	FT/EE
"SpineCurer" [6]	3×IMU	—	—	—	ME, EG, GF, DS	FT/EE
"Wearable Sensor Array" [4]	3×stretch sensor	—	—	—	ME, EG, DS	FT/EE
"Disc4Spine" [8,9]	1×time-of-flight camera, 1×Empatica E4 (Empatica Inc), pelvis belt (N/A)	—	—	Force plate	ME, EG, DS	ST/FA[j]
"Rehabilitation Exercise System" [10]	—	—	—	Pressure mat	—	ST/FA
Scoliosis Apps [12]	—	—	—	—	ScolioTrack, ScolioMeter, APECS[k], CobbMeter, ScolioDetector, Scoliosis Tracker, Scoliosis Manager (web) or functions: measurement of spinal curvature (ATR[l] or Cobb angle), posture tracking, exercise instructions, remote monitoring, documentation, compliance tracking/smartphone hardware used: IMU and camera	—

[a]BPD: body pressure distribution.

[b]IMU: inertial measurement unit.

[c]ME: monitoring of exercise execution.

[d]EG: exercise guidance.

[e]DS: data storage.

[f]FT/IE: free training, including training equipment.

[g]Not applicable.

[h]FT/EE: free training, excluding training equipment.

[i]GF: gamification.

[j]ST/FA: stationary training on fixed apparatus.

[k]APECS: Artificial Intelligence Posture Evaluation and Correction System.

[l]ATR: angle of trunk rotation.

Motion Tracking and Respiration Tracking

Motion tracking is one of our core components for monitoring exercise execution. To achieve this, we have integrated 9 inertial sensors into a textile wearable, which also allow for monitoring of posture. During the technical verification of our motion tracking, we examined 3 representative samples of the 9 iScoolio boards. Compared to the optical reference system (ARAMIS 3D), this resulted in an average absolute deviation of 2.29° (SD 1.66°), supporting H_1. It should be noted that the significant outliers in the reference system contributed to an increase in the deviation. These are most likely due to delays or errors in manual image capture. Similar deviations between IMU in wearable applications and an optical system (OptiTrack) were

XSL•FO
RenderX

found, for example, by Li et al [16] with 2.00° (SD 2.65°), Teufl et al [17] with ≤2.40°, and Bell et al [18] with ≤3.00°. In terms of repeatability, the iScoolio boards achieved good results with a maximum SD of 1.48°. Studies by Fennema et al [19], Hamidon et al [20], and Beange et al [21] also show that IMUs can exhibit high repeatability and be suitable for motion tracking concepts. We have thus created a basis for monitoring scoliosis exercises.

Unlike the wearables for home-based scoliosis training described in the literature, we have integrated breathing tracking into our system. In many Schroth exercises, the number of breaths taken is used as a counting variable for repetitions. Common measurement systems for respiration that can also be worn on the body include inertial sensors [22,23], stretch sensors [23-26], and bioimpedance sensors [23,27]. Due to their good integrability, we opted for textile strain sensors. Although this measurement principle is susceptible to motion artifacts [26], it is nevertheless very well suited to our application due to the classic sequence of many Schroth exercises: (1) adopting a corrective posture and (2) targeted breathing.

For the technical verification of the sensor concept, we examined 2 respiratory parameters relevant to scoliosis therapy: respiratory rate and respiratory depth. Both can be extracted from the raw data measured using the strain sensors we integrated. This enabled us to verify H_2. However, automated evaluation requires the integration of additional data processing. Low-pass filters and finite impulse response filters, as well as window size [28], maximum-minimum, zero-crossing, and threshold detection algorithms [29], can be used to detect respiratory rate. For the detection of breathing depth, it is sufficient to differentiate between the measured values, as the increasing filling volume of the lungs is accompanied by an increasing expansion of the sensor.

By supplementing motion tracking and posture monitoring with components for measuring respiratory parameters, we have created a technical basis for recording RAB. Further software development and the integration of machine learning approaches are required for accurate tracking.

Force Tracking Upper Limbs

To provide the most versatile workout possible, we have combined the features of wall bars and exercise sticks into a single training device, the Bars-Module. The supporting feet and handles of this module can be used to measure forces in the upper limbs. The technical verification showed that the demonstrator's raw sensor signals responded clearly and proportionally to applied forces, which substantiates H_3. On the basis of the 3 selected force levels, nonlinear behavior was observed when considering the entire measuring range, as well as drift behavior. These characteristics, supplemented by the uninvestigated hysteresis, reflect the underlying resistive measuring principle [30,31]. However, these negative effects can be significantly reduced through the use of appropriate algorithms [31]. In addition to the software options, optimization through hardware adaptation would also be conceivable. Due to the available installation space, the force-sensing resistor sensors in the supporting feet could be replaced by high-quality load cells, which in turn would mean a significantly higher financial outlay.

Thanks to the integrated measurement functions of the Bars-Module, there are various options for monitoring the exercise execution. In addition to traditional grip strength, which correlates with scoliosis [32], it is possible to measure isometric force during an exercise. In combination with the posture tracking feature of the Wearable-Module, it is possible to determine whether the user maintains their corrective posture throughout a set. Furthermore, the ability to distinguish between the left and right arms allows for the determination of symmetries or asymmetries in force distribution. In combination with the other modules, this enables an analysis of how the user assumes their corrective posture.

Body Pressure Distribution

To further increase the variety of exercises, we have added a second training device to our concept. We chose a balance board, which plays an important role in many PSSE exercises. The basis for this is the impairment of the cortical centers of the brain in scoliosis, which are responsible for controlling balance [2]. In their study involving 195 participants (adolescent idiopathic scoliosis [AIS]: n=114, 58.5%; control: n=81, 41.5%), Le Berre et al were able to show that dynamic proprioception is selectively altered in AIS [33]. Targeted training of balance reactions is therefore intended to improve axial, static, and dynamic balance, which in turn helps to intensify self-correction during everyday activities [2]. As an extension of the training functions of conventional balance boards, we have also equipped our module with measuring functions. Integrated sensors can be used to determine both the inclination of the board and the pressure distribution on the standing surface. This makes it possible to monitor the execution of training instructions. We have also created the hardware basis for investigating parameters such as foot pressure distribution and COP, both of which are influenced by scoliosis [34-37]. In their study, Horng et al found a statistically significant linear relationship between the Cobb angle and maximum foot pressure on the hallux [35]. Regarding COP, the study by Larni et al showed a decline in COP parameters in adolescents with AIS compared to the healthy control group in tests using a Kistler force plate [37]. In addition to these standing measurement applications, balance boards equipped with sensors also offer various application possibilities when sitting. Jung et al were able to use integrated acceleration sensors to measure the asymmetrical sitting postures caused by scoliosis and pelvic deformity in patients with AIS [38]. To record all these parameters in detail, further development steps are necessary for our system.

Regarding the detection of stance width, our measurements showed slight absolute deviations from the mean, with a maximum of 1.2 cm. Comparing this value with the intervals described in the literature for targeted changes in the training effect, such as a narrow stance of 31.7 (SD 2.7) cm and a wide stance of 59.8 (SD 5.1) cm [39], it can be seen that these variations are sufficiently small and therefore have no significant negative effect on the measurements.

The verification of our COP measurement function showed that 48% of the measurements had a deviation of ≤1 box and 82%

XSL•FO
RenderX

had a deviation of ≤2 boxes. With a sensor element resolution of 10 mm, 82% of the measured values had a COP shift of ≤20 mm. It should also be noted that a single sensor element has a diameter of 6 mm. With a center-to-center distance of 10 mm between these sensor elements, this results in horizontal and vertical dead zones of 4 mm and diagonal dead zones of 8.14 mm. Depending on the positioning of the test weights, this may have resulted in real pressure differences between the 2 clusters, which would also cause a real shift in the COP. The increased edge errors may be attributed to material transitions, higher conductor resistances, or calibration limitations.

These measurement functions allow the user a specific stance width to be set and tracked, which affects the difficulty level of the balance exercises. In addition, the COP can be measured during static standing or targeted balance training. Finally, during seated exercises in which the Balance-Module is used as a support surface, it is possible to measure how hip displacement and seat pressure distribution are used during the exercise.

App or Web Interface

The Interface-Module completes our system with a digital human-technology interface. On the one hand, it serves to communicate, bundle, and process all sensor data. On the other hand, it contains an exercise catalog and outputs visualizations of the measurement data and therapy instructions via an associated web application. In the targeted therapeutic setting, it functions as a scoliosis-specific software tool, which is represented in the literature with various application scenarios.

Compared to the scoliosis software solutions described by Bottino et al [12], the Interface-Module focuses more on the processing and visualization of sensor data as a basis for exercise monitoring and guidance. Nevertheless, there are minor similarities in terms of the selection of suitable exercises or the provision of information for compliance.

Limitations

In its current state of development, the iScoolio concept still has some room for improvement. One of the biggest limitations is that motion tracking is currently limited to the torso and head and that all sensor data are recorded at a frequency of only 4 Hz. Furthermore, all iScoolio boards are independent components, which therefore require a separate charging and calibration process. This results in a time-consuming setup process. In addition, exercises in the supine position are not possible due to the size of the current boards. For the training modules, the aim should be to reduce the size of the Balance-Module. In addition, different force sensors, a faster extension mechanism with less backlash, and a locking function for the supporting feet would be useful for the Bars-Module. All these questions will be addressed in further work.

There were also some limitations regarding technical verification. These included the fact that all measurements were limited to repetitions of a specific Schroth exercise.

For the motion tracking investigation, we limited ourselves to 3 representative iScoolio boards. In addition, only points and angles on the frontal and sagittal planes were examined; mixed movements were not. Nevertheless, these data enable us to draw conclusions about the accuracy of the measurement system, which was the aim of the investigation. Another limitation regarding motion tracking resulted from the manual triggering of the ARAMIS 3D system.

Regarding respiration tracking, the test was likewise limited to a single representative sensor. In addition, the filling and deflation speed of the air cushions was limited by the valve installed in the test setup. Furthermore, the rising and falling of the chest was simulated by 2 air cushions, which only produced a selective increase in volume rather than a full-surface increase. Finally, it should be noted that no classic reference measurement system was used for this measurement, but rather an actuator-based approach with high repeatability.

When verifying the handpieces of the Bars-Module, the greatest limitations were caused by the 3D-printed half-shells. These only represent a hand in a highly simplified form and had inaccuracies due to additive manufacturing. Furthermore, the handles were verified with a round wooden profile dummy instead of the metal rod. Both limitations led to a deviation from the real load but represented a suitable approach for integrating a testing machine with regard to the measurement function to be tested.

In the case of the Balance-Module, the limiting factors during verification were, on the one hand, the test weights, which did not accurately represent a human foot in terms of their shape and weight. However, the smaller outline of the weights enabled a more detailed examination of the sensor surface.

Another limitation is that verification was only carried out in a technical setting and without the involvement of human participants. Therefore, no clinical conclusions can be drawn regarding the impact on the effects of scoliosis, such as pain, reduced quality of life, or depression [40,41].

Conclusions

In this paper, we have presented the iScoolio concept, which has the potential to support home-based scoliosis therapy through adaptable devices and tracking elements. Integrated sensors enable monitoring of movements, breathing, forces, and pressure distribution during training. An associated web application visualizes the recorded data and serves as a guide for performing the exercises. Using the new approach could increase motivation, reduce concerns about incorrect exercise execution, and enable adherence monitoring and teletherapy. Technical verification indicates that the integrated tracking methods are functional; however, there is still room for optimization in terms of hardware and software. The system's greatest areas for improvement include the low sampling frequency (4 Hz), the lack of motion tracking for arms and legs, the long setup process, and measurement inaccuracies in some sensor designs. Future work will address these limitations. Before conducting clinical trials under real-world conditions, further validation is required.

XSL•FO
RenderX

Acknowledgments

We would like to express our gratitude to Joshua Stoppel, Thomas Wegner, Jens Unglaube, Dr Christian Ludwig, and Dr Steffen Derlien for their contributions to the development of the iScoolio concept. Special thanks go to Sophia Elsner, who played a crucial role in initiating the research project and in finding a suitable acronym. Additionally, we appreciate Christian Rotsch for his valuable suggestions regarding the technical verification. We also thank the Asklepios Katharina-Schroth-Klinik in Bad Sobernheim for supplying an exercise catalog with the Schroth exercises mentioned in the publication. Our thanks extend to Nicole Klemm for reproducing the Schroth exercises and to Max Pohl for his support with the motion tracking test bench. Finally, we would like to thank Sporlastic GmbH for providing us with textile samples for the development process.

Generative AI Use

This manuscript was translated with the assistance of AI-based language tools (DeepL) and was subsequently reviewed and edited by the authors.

Funding

This work was partially funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) based on the decision by the German Bundestag (ZIM, 16KN051667) and by the German Federal Ministry of Research, Technology and Space (BMFTR, 16SV9351).

Authors' Contributions

FG, SH, and NL developed the iScoolio prototype. FG was responsible for the creation of the verification concepts, test benches, and test scenarios. MW set up the DYNA-MESS and ARAMIS 3D testing systems for the measurements and operated the GOM Correlate Pro software. NL provided technical support for the iScoolio software. FG performed the measurements, evaluated the collected data, created the diagrams, and wrote the original draft. FG and SH designed additional graphics. MW, CEH, SH, NL, and WGD reviewed and edited the manuscript. CEH and WGD supervised the project. All authors have read and approved the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Schroth body blocks.
[PDF File, 169 KB - rehab_v13i1e93721_app1.pdf]

Multimedia Appendix 2
Schroth exercises.
[PDF File, 231 KB - rehab_v13i1e93721_app2.pdf]

Multimedia Appendix 3
Characteristics of motion tracking analysis.
[PDF File, 186 KB - rehab_v13i1e93721_app3.pdf]

References

1. Lehnert-Schroth C, Auner-Gröbl P. Dreidimensionale Skoliosebehandlung: Atmungs-Orthopädie System Schroth: Ein Physiotherapeutisches Konzept Zur Verbesserung von Wirbelsäulenverkrümmungen, 9th edition: Elsevier; 2021.
2. Berdishevsky H, Lebel VA, Bettany-Saltikov J, et al. Physiotherapy scoliosis-specific exercises - a comprehensive review of seven major schools. Scoliosis Spinal Disord 2016;11(1):20. [doi: 10.1186/s13013-016-0076-9] [Medline: 27525315]
3. Nisser J, Smolenski U, Śliwiński G, et al. Skoliosespezifische Physiotherapie bei Patienten mit idiopathischer Adoleszentenskoliose (AIS) – ein narratives Review. Phys Med Rehab Kuror 2018 Apr;28(2):88-102. [doi: 10.1055/s-0043-124472]
4. Caviedes JE, Li B, Jammula VC. Wearable sensor array design for spine posture monitoring during exercise incorporating biofeedback. IEEE Trans Biomed Eng 2020 Oct;67(10):2828-2838. [doi: 10.1109/TBME.2020.2971907] [Medline: 32031929]
5. Li J, Mai E, Hu D, Yang Z, Liang J, Zhao B. Scoliosis rehabilitation assistant: a real-time scoliosis training system. Presented at: 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC); Oct 9-12, 2022; Prague, Czech Republic p. 978-971. [doi: 10.1109/SMC53654.2022.9945516]

XSL•FO
RenderX

6. Mai E, Hu D, Li J, Yang Z. SpineCurer: an inertial measurement unit based scoliosis training system. In: Froehlich J, Shinohara K, Ludi S, editors. Presented at: The 24th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 23-26, 2022; Athens, Greece p. 1-4. [doi: 10.1145/3517428.3550393]
7. Voinea GD, Mogan G. Development of a wearable scoliosis monitoring system using inertial sensors. AMM 2015;811:353-358. [doi: 10.4028/www.scientific.net/AMM.811.353]
8. Romaniszyn-Kania P, Pollak A, Danch-Wierzchowska M, et al. Hybrid system of emotion evaluation in physiotherapeutic procedures. Sensors (Basel) 2020 Nov 6;20(21):6343. [doi: 10.3390/s20216343] [Medline: 33172146]
9. Szurmik T, Bibrowicz K, Lipowicz A, Mitas AW. Methods of therapy of scoliosis and technical functionalities of DISC4SPINE (D4S) diagnostic and therapeutic system. In: Pietka E, Badura P, Kawa J, Wieclawek W, editors. Information Technology in Biomedicine: Springer International Publishing; 2021, Vol. 1186:201-212.
10. Jung JY, Yang CM, Kim JJ. Effectiveness of combined stretching and strengthening exercise using rehabilitation exercise system with a linear actuator and MR damper on static and dynamic sitting postural balance: a feasibility study. Applied Sciences 2021;11(16):7329. [doi: 10.3390/app11167329] [Medline: 36003951]
11. Ahmad F, Wang MY. Spinal deformity? There's an app for that!. World Neurosurg 2013 Nov;80(5):446-447. [doi: 10.1016/j.wneu.2013.09.011] [Medline: 24035992]
12. Bottino L, Settino M, Promenzio L, Cannataro M. Scoliosis management through apps and software tools. Int J Environ Res Public Health 2023 Apr 14;20(8):5520. [doi: 10.3390/ijerph20085520] [Medline: 37107802]
13. Günther F, Schober F, Hunger S, et al. Improving home-based scoliosis therapy: findings from a web-based survey. JMIR Rehabil Assist Technol 2023 Aug 4;10:e46217. [doi: 10.2196/46217] [Medline: 37540557]
14. Palazzo C, Klinger E, Dorner V, et al. Barriers to home-based exercise program adherence with chronic low back pain: patient expectations regarding new technologies. Ann Phys Rehabil Med 2016 Apr;59(2):107-113. [doi: 10.1016/j.rehab.2016.01.009] [Medline: 27050664]
15. Essery R, Geraghty AWA, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil 2017 Mar;39(6):519-534. [doi: 10.3109/09638288.2016.1153160] [Medline: 27097761]
16. Li R, Jumet B, Ren H, Song W, Tse ZTH. An inertial measurement unit tracking system for body movement in comparison with optical tracking. Proc Inst Mech Eng H 2020 Jul;234(7):728-737. [doi: 10.1177/0954411920921695] [Medline: 32419605]
17. Teufl W, Miezal M, Taetz B, Fröhlich M, Bleser G. Validity, test-retest reliability and long-term stability of magnetometer free inertial sensor based 3D joint kinematics. Sensors (Basel) 2018 Jun 21;18(7):1980. [doi: 10.3390/s18071980] [Medline: 29933568]
18. Bell KM, Onyeukwu C, McClincy MP, et al. Verification of a portable motion tracking system for remote management of physical rehabilitation of the knee. Sensors (Basel) 2019 Feb 28;19(5):1021. [doi: 10.3390/s19051021] [Medline: 30823373]
19. Fennema MC, Bloomfield RA, Lanting BA, Birmingham TB, Teeter MG. Repeatability of measuring knee flexion angles with wearable inertial sensors. Knee 2019 Jan;26(1):97-105. [doi: 10.1016/j.knee.2018.11.002] [Medline: 30554906]
20. Hamidon N, Azaman A, Mahmod SR, Abdullah NHO. Validity and repeatability of inertial measurement unit for measuring walking gait parameter of patients with non-specific low back pain. MJMHS 2022:45-51. [doi: 10.47836/mjmhs.18.s6.9]
21. Beange KHE, Chan ADC, Beaudette SM, Graham RB. Concurrent validity of a wearable IMU for objective assessments of functional movement quality and control of the lumbar spine. J Biomech 2019 Dec 3;97:109356. [doi: 10.1016/j.jbiomech.2019.109356] [Medline: 31668717]
22. Cesareo A, Nido SA, Biffi E, Gandossini S, D'Angelo MG, Aliverti A. A wearable device for breathing frequency monitoring: a pilot study on patients with muscular dystrophy. Sensors (Basel) 2020 Sep 18;20(18):5346. [doi: 10.3390/s20185346] [Medline: 32961986]
23. Hussain T, Ullah S, Fernández-García R, Gil I. Wearable sensors for respiration monitoring: a review. Sensors (Basel) 2023 Aug 30;23(17):7518. [doi: 10.3390/s23177518] [Medline: 37687977]
24. Min SD, Yun Y, Shin H, et al. Simplified structural textile respiration sensor based on capacitive pressure sensing method. IEEE Sensors J 2014;14(9):3245-3251. [doi: 10.1109/JSEN.2014.2327991]
25. Chu M, Nguyen T, Pandey V, et al. Respiration rate and volume measurements using wearable strain sensors. NPJ Digit Med 2019;2:8. [doi: 10.1038/s41746-019-0083-3] [Medline: 31304358]
26. Yamamoto A, Nakamoto H, Bessho Y, et al. Monitoring respiratory rates with a wearable system using a stretchable strain sensor during moderate exercise. Med Biol Eng Comput 2019 Dec;57(12):2741-2756. [doi: 10.1007/s11517-019-02062-2] [Medline: 31734768]
27. Shen CL, Huang TH, Hsu PC, et al. Respiratory rate estimation by using ECG, impedance, and motion sensing in smart clothing. J Med Biol Eng 2017;37(6):826-842. [doi: 10.1007/s40846-017-0247-z] [Medline: 30220900]
28. Kim JS, Kim J. An accurate respiratory rate estimation algorithm for a rubber fiber respiratory sensor. Journal of Industrial Textiles 2024 Jan;54. [doi: 10.1177/15280837241248181]
29. Vanegas E, Igual R, Plaza I. Sensing systems for respiration monitoring: a technical systematic review. Sensors (Basel) 2020 Sep 22;20(18):5446. [doi: 10.3390/s20185446] [Medline: 32972028]
30. Giovanelli D, Farella E. Force sensing resistor and evaluation of technology for wearable body pressure sensing. J Sens 2016;2016:1-13. [doi: 10.1155/2016/9391850]

XSL•FO
RenderX

31. Velásquez EIG, Gómez V, Paredes-Madrid L, Colorado HA. Error compensation in force sensing resistors. Sens Biosensing Res 2019 Nov;26:100300. [doi: 10.1016/j.sbsr.2019.100300]
32. Yagci G, Aydin Ozcan D, Ayhan C, Demirkiran G, Yakut Y, Yazici M. Evaluation of upper extremity function and its relation to curve pattern in female adolescents with idiopathic scoliosis: a pilot study. Spine Deform 2020 Dec;8(6):1175-1183. [doi: 10.1007/s43390-020-00148-8] [Medline: 32474902]
33. Le Berre M, Guyot MA, Agnani O, et al. Clinical balance tests, proprioceptive system and adolescent idiopathic scoliosis. Eur Spine J 2017 Jun;26(6):1638-1644. [doi: 10.1007/s00586-016-4802-z] [Medline: 27844226]
34. Ca an L, Cerbu S, Amaricai E, et al. Assessment of static plantar pressure, stabilometry, vitamin D and bone mineral density in female adolescents with moderate idiopathic scoliosis. Int J Environ Res Public Health 2020 Mar 24;17(6):2167. [doi: 10.3390/ijerph17062167] [Medline: 32214036]
35. Horng J, Liu XC, Thometz J, Tassone C, Duey-Holtz A. Evaluation of plantar pressures and center of pressure trajectories in adolescent idiopathic scoliosis. Stud Health Technol Inform 2021 Jun 28;280:131-135. [doi: 10.3233/SHTI210451] [Medline: 34190074]
36. Yi L, Houwei L, Lin W, et al. Evaluation of correlation between sagittal balance and plantar pressure distributions in adolescent idiopathic scoliosis: a pilot study. Clin Biomech (Bristol) 2021 Mar;83:105308. [doi: 10.1016/j.clinbiomech.2021.105308] [Medline: 33662651]
37. Larni Y, Mohsenifar H, Ghandhari H, Salehi R. Investigation of static balance differences between adolescents with idiopathic scoliosis and healthy age-matched adolescents: a cross-sectional study. JIMC 2023. [doi: 10.18502/jimc.v6i3.12859]
38. Jung JY, Yoo CI, Kim KA, et al. Evaluation of sitting balance in adolescent idiopathic scoliosis patients with pelvic rotation using balance board system with accelerometer. Presented at: 2014 International Conference on IT Convergence and Security (ICITCS); Oct 28-30, 2014; Beijing, China p. 978-971. [doi: 10.1109/ICITCS.2014.7021769]
39. Larsen S, Kristiansen E, Helms E, van den Tillaar R. Effects of stance width and barbell placement on kinematics, kinetics, and myoelectric activity in back squats. Front Sports Act Living 2021;3:719013. [doi: 10.3389/fspor.2021.719013] [Medline: 34541522]
40. Kaya MH, Erbahçeci F, Alkan H, et al. Factors influencing of quality of life in adolescent idiopathic scoliosis. Musculoskelet Sci Pract 2022 Dec;62:102628. [doi: 10.1016/j.msksp.2022.102628] [Medline: 35872563]
41. Yetiş M, Yildiz NT, Canli M, et al. Determination of predictors associated with pain in non-surgically treated adults with idiopathic scoliosis. J Orthop Surg Res 2024 Jul 16;19(1):406. [doi: 10.1186/s13018-024-04912-8] [Medline: 39014368]

Abbreviations

AIS: adolescent idiopathic scoliosis
COP: center of pressure
DPU: data processing unit
IMU: inertial measurement unit
PSSE: physiotherapeutic scoliosis-specific exercises
RAB: rotational angular breathing

Edited by S Munce; submitted 18.Feb.2026; peer-reviewed by H Kocaman, X Qin; revised version received 15.Aug.2026; accepted 28.Aug.2026; published 29.Oct.2026.

Please cite as:
Günther F, Hunger S, Werner M, Landerer N, Heyde CE, Drossel WG
Digitally Assisted Training System for Home-Based Scoliosis Therapy: Proof-of-Concept and Technical Verification Study
JMIR Rehabil Assist Technol 2026;13:e93721
URL: https://rehab.jmir.org/2026/1/e93721
doi:10.2196/93721

© Florian Günther, Sandra Hunger, Michael Werner, Nikolaus Landerer, Christoph-Eckhard Heyde, Welf-Guntram Drossel. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 29.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Benchmarking Framework for Real-Time 3D Pose Estimation in Motor Rehabilitation: Development and Proof-of-Concept Study

Emanuel Alexander Lorenz[1], PhD; Beatrix Vereijken[2], PhD; Kerstin Bach[1], PhD; Xiaomeng Su[1], PhD

[1]Department of Computer Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, Sem Sælands vei 9, Trondheim, Trøndelag, Norway

[2]Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Trøndelag, Norway

Corresponding Author:
Emanuel Alexander Lorenz, PhD
Department of Computer Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, Sem Sælands vei 9, Trondheim, Trøndelag, Norway

Abstract

Background: Accurate motor assessment is crucial for diagnosing and monitoring neurological and musculoskeletal disorders. However, traditional manual and semiautomated methods can be subjective, time-consuming, or expensive. Real-time 3D human pose estimation (HPE) is emerging as a low-cost, noninvasive alternative. However, its clinical adoption is limited due to a lack of thorough, standardized evaluation of its accuracy and robustness in real-world clinical settings.

Objective: This study aims to introduce a technical pipeline for evaluating the applicability of real-time 3D HPE in simulated clinical rehabilitation settings.

Methods: The framework facilitates the evaluation of an HPE model's spatial and spatiotemporal accuracy, real-time capability, and robustness against simulated common setup errors and camera position variations, using a marker-based motion capture system as a gold-standard baseline. To address the technical challenge of interdataset comparison, the framework includes a morphing model that harmonizes differing joint definitions.

Results: Applying the framework to a popular, accessible real-time algorithm (MediaPipe) on 4 healthy participants (N=127,828 frames) showed that simulated underexposure and defocus blur had negligible effects (rank-biserial correlation [RBC] ≤0.02), whereas body segment occlusion (mean-per-joint angular error [MPJAE] increase of 1.5, SD 10 degrees; +24.3%) and insufficient multiocular calibration (MPJAE increase of 7.8, SD 14 degrees; +123.4%, RBC=0.95) significantly degraded tracking accuracy. The framework also showed significant sensitivity to nonfrontal camera placements, both side- and back-facing views (RBC 0.41 - 0.49 and bias up to 2.2 degrees/37.1%), relative to stable frontal placements (RBC ≤0.03). The skeleton morphing module reduced Procrustes-aligned mean per joint position error (PA-MPJPE) by 55.29 mm (N=79,423 frames, 57.2% reduction) and its SD by 24.69 mm (N=79,423 frames, 50.7%).

Conclusions: This study establishes a technical foundation for the standardized evaluation of markerless tracking for clinical applications, like motor rehabilitation. The morphing model enables interdataset comparison, and the proposed framework facilitates domain-specific interpretation of HPE performance. While further validation on pathological populations is required, this work provides the necessary engineering foundation for future validation studies.

(JMIR Rehabil Assist Technol 2026;13:e91360) doi:10.2196/91360

KEYWORDS

motion capture; human pose estimation; motor rehabilitation; benchmark; skeleton morphing

Introduction

Background

In neurological and musculoskeletal disorders, accurate assessment of a patient's motor abilities is crucial for both initial diagnosis and subsequent monitoring of progress [1]. Nevertheless, traditional manual and semiautomated assessments have limitations that hinder frequent and extensive assessments [2-4]. To overcome these limitations, new deep learning methods, including human pose estimation (HPE), are being developed for real-time, 3D evaluation of motor function in both clinical and nonclinical settings. However, as a new technology, its applicability for motor assessment in clinical settings has not yet been thoroughly evaluated. Thus, its adoption to date has been limited [5]. To address these concerns, we introduce a comprehensive framework for benchmarking HPE algorithms in various simulated clinical scenarios to evaluate not only their spatial and spatiotemporal accuracy but also their robustness to common setup errors and variabilities.

XSL•FO
RenderX

Unlike existing 3D HPE benchmarks, which evaluate accuracy on curated datasets under controlled laboratory conditions, our framework targets the specific requirements of motor rehabilitation, shifting the evaluation from theoretical accuracy on curated datasets toward clinical deployability. Its specific contributions are listed at the end of this section. We explicitly frame this study as a technical proof-of-concept in healthy participants to establish the evaluation pipeline and identify necessary methodological considerations for future clinical validation.

Clinical Motor Assessment

Traditionally, motor assessments largely depend on the manual evaluation of specific movements and tasks via standardized tests, which are well established for different pathologies and populations [3,6]. While manual assessments allow for personalization and increased accessibility, even in resource-limited settings, they are often time-consuming, subjective, and dependent on the presence of experienced personnel [2,3,7].

Thus, semiautomated assessments have been developed to address these limitations, offering objective, precise, and real-time measurements using optical systems (eg, marker-based motion capture systems for assessing movement quality in weight-shifting exercises [8]), mechanical systems (eg, pressure sensors for gait analysis [9] and inertial measurement units for upper limb assessment [10]), and robotic systems (eg, exoskeletons for assessing reaching movements [11]). However, despite their advantages, they are often expensive, intrusive in their application, and also require trained personnel [4]. Additionally, they may suffer from artifacts caused by biological and environmental variables [12,13] or by sensor placement errors [14], which can affect their accuracy and reliability.

HPE in Motor Assessment

Recent advances in machine learning have enabled markerless joint position tracking from generic image data using HPE. Thus, low-cost camera systems can be used in diverse environments to track movements in real time. Consequently, HPE has emerged as a technology for clinical and nonclinical applications, enabling continuous, noninvasive monitoring of motor function, including gait, balance, and range of motion [5]. The field has advanced rapidly in recent years. Transformer-based lifters such as MotionBERT [14] elevate 2D pose sequences to 3D by modeling spatiotemporal joint dependencies, while one-stage real-time detectors such as RTMPose [15] and the YOLO-Pose family [16] jointly detect people and regress keypoints at high frame rates. Against this backdrop, MediaPipe is best characterized not as the most accurate model, but as a widely used, well-documented, and readily deployable one: comparative gait studies report that earlier models such as OpenPose can match or exceed it on clinically relevant joints [17], and its former advantages in speed and ease of use are increasingly shared by newer lightweight architectures. Although current offline systems achieve angular errors (AEs) below the maximum permissible AE of 5 degrees for clinical applications [18,19], they cannot provide immediate feedback during rehabilitation exercises. This limits their applicability for real-time correction and adaptation of movement patterns during activities such as exergaming [20]. Online systems, however, do not yet achieve the same spatial accuracy [21]. Furthermore, HPE still struggles with real-world environments, diverse clothing, and setup errors, with occlusion now the principal remaining challenge; recent architectures address it explicitly, for example, through diffusion- and transformer-based designs robust to occluded and noisy keypoints [22], though it is not yet fully solved [5,18,23].

While some HPE approaches use red, green, blue-depth (RGB-D) sensors, standard RGB-based HPE relies solely on 2D image data, with no depth information provided directly to the model. However, while state-of-the-art models can infer 3D poses from a single 2D image frame, this remains an ambiguous problem, as a single 2D pose representation can correspond to multiple 3D representations [24]. Thus, multiocular systems, which capture data from various camera angles, are still advised, as triangulating 2D poses across different viewpoints can improve the accuracy, reliability, and robustness of 3D reconstructions by providing additional information. However, the benefits of multiocular setups are associated with increased system complexity and higher computational demands [5].

Benchmarking of HPE Models

The adoption of HPE in clinical rehabilitation depends on its accuracy, reliability, and robustness [5]. Health care providers are cautious about adopting new technologies due to concerns about patient safety, data privacy, and resource reallocation [25]. While general evaluation pipelines for HPE models exist, such as MMPose [26], PoseTrack [27], and other comparative analyses [24], they primarily evaluate HPE models' accuracy on generic movements and metrics like mean-per-joint position error (MPJPE) and percentage of correct keypoints [28]. While these metrics are valuable for benchmarking HPE models, they lack clinical relevance, particularly for rehabilitation applications that demand higher precision, specifically for angular changes and contextual robustness.

Recent studies have partially addressed these gaps by including clinically relevant evaluation metrics, such as joint angle errors [18,29-32], testing HPE models on clinically relevant movements [29,31,32], or evaluating the impact of different setups and errors [18,33]. The closest systematic efforts, Rode et al [34], who benchmarked 11 monocular estimators against marker-based capture, and Wade et al [35], who reviewed markerless methods for clinical gait, either compare accuracy under controlled conditions or survey limitations qualitatively. None combine simulated clinical setup errors, per-condition robustness reporting, and cross-dataset joint harmonization in a single reproducible pipeline [5].

Datasets for HPE in Rehabilitation

Deep neural networks, like HPE models, must be trained on large, diverse, and precisely annotated datasets to provide accurate and robust joint locations across various settings and populations. Therefore, selecting appropriate datasets for training and evaluation is critical [36].

Over the past decade, several well-known datasets, such as Human3.6M [37], MPII [38], and COCO [39], have been released. However, these datasets focus on everyday activities

XSL•FO
RenderX

within controlled settings and do not include rehabilitation-specific movements. In contrast, more recent clinically relevant datasets incorporate domain-specific movements and participant groups, thereby better capturing the complexities encountered in rehabilitation settings [5,40-42]. However, these clinical datasets are suboptimal for domain-specific HPE training; they either lack full-body joint coverage despite high spatial accuracy (eg, UCO [40]) or rely on lower-precision baselines limited to specific pathologies (eg, KIMORE [41] and FineRehab [42]), failing to provide a comprehensive, high-accuracy standard for general motor rehabilitation.

Furthermore, HPE training datasets often exhibit inconsistencies in joint center definitions due to varying annotation protocols, such as manual labeling versus marker-based derivation, variations in surface marker placement, and the specific biomechanical models applied. This variation stems from the differing methods used to establish ground-truth joint center locations, such as manual annotations or biomechanical models [37-39]. As a result, these discrepancies lead to inconsistencies in the estimated keypoints when training or evaluating on alternative datasets [23,36]. To address this, several strategies have been developed to align joint center locations across datasets. Rapczyński et al [23] applied linear transformations and scale normalization, while Wandt et al [43] introduced a 2D skeleton-morphing network that aligns keypoints using residual blocks. More recently, Sárándi et al [36] used an affine-combining autoencoder to compress 3D joint keypoints from different datasets into a smaller set of latent keypoints, preserving the geometric relationships across datasets. Thus, these alignment methods enable consistent integration of heterogeneous datasets, improving HPE model comparability.

Research Objectives

This study aims to develop and validate a framework for evaluating the applicability of real-time 3D HPE algorithms in simulated clinical rehabilitation settings, benchmarking an exemplary model against a marker-based gold standard after aligning keypoint definitions. We address 2 research questions (RQs):

- RQ1: How can a comprehensive framework be designed to evaluate the spatial and spatiotemporal accuracy of a real-time 3D HPE algorithm for clinical motor rehabilitation?
- RQ2: How does the proposed framework assess the robustness of HPE algorithms against varied camera angles, common setup errors, and clinically relevant movements?

Contributions

This work makes the following contributions: (1) a modular benchmarking framework that evaluates real-time 3D HPE models under simulated clinical conditions, integrating environmental augmentations, and a skeleton morphing module for cross-dataset comparability; (2) a set of clinically motivated spatial and spatiotemporal metrics reported per joint and body segment; (3) an empirical benchmark of a widely used, accessible HPE model (MediaPipe) across camera configurations, movements, and simulated setup errors; and (4) a validation of the morphing module against controlled synthetic offsets, showing that it corrects systematic joint-definition differences without suppressing tracking error.

Methods

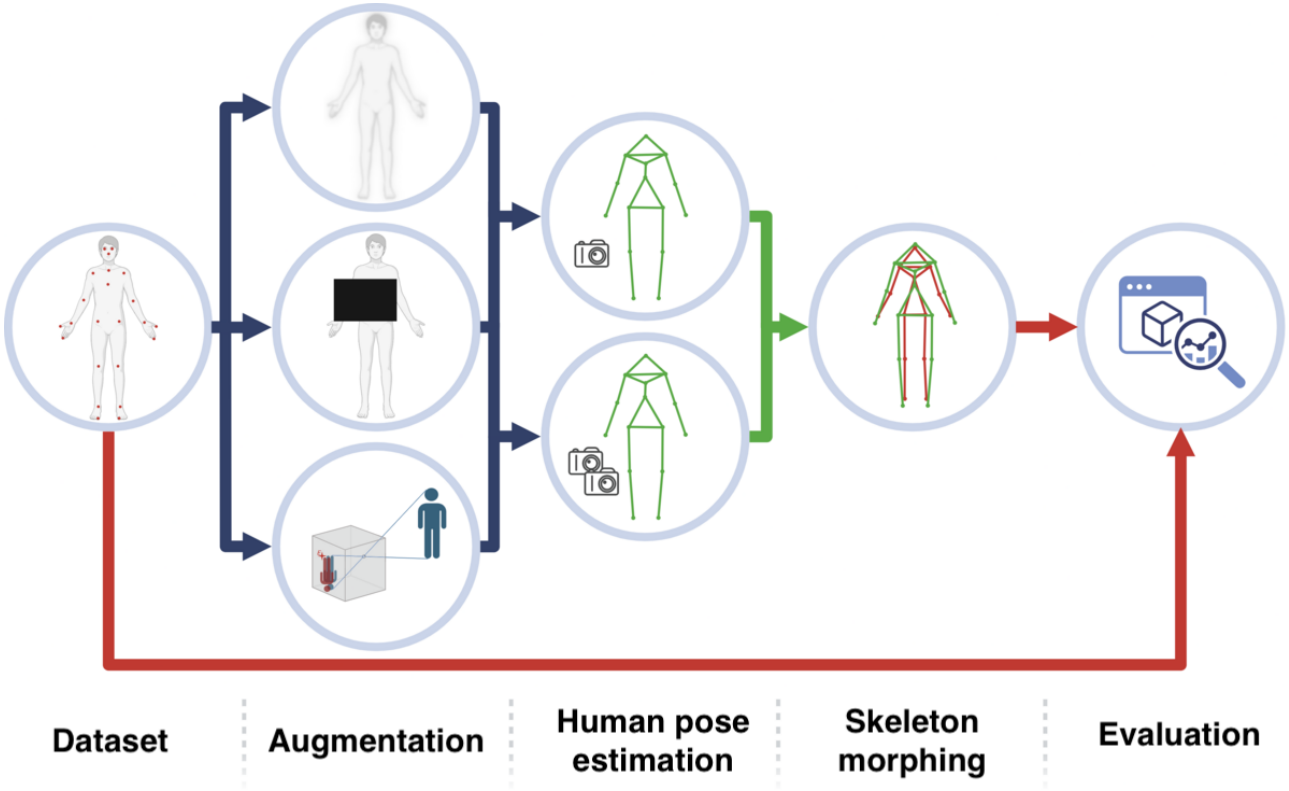
Benchmarking Framework Architecture

The benchmarking framework (illustrated in Figure 1) for real-time 3D HPE models augments the video input according to common errors possible in clinical settings. A modular design was chosen to address the gap between technical engineering benchmarks and actual clinical needs, while allowing for future adaptations. By integrating an augmentation module to simulate real-world setup errors and a skeleton morphing module to align conflicting dataset definitions, the framework solves the problems of a lack of environmental testing and interdataset incomparability identified in the Introduction section. This specific structure ensures that the evaluation measures true clinical robustness rather than just theoretical accuracy, helping developers to choose the most applicable HPE model and ultimately increasing trust in the technology.

It evaluates the spatial and spatiotemporal accuracy and real-time capability of both mono- and multiocular HPE models across different clinically relevant movements and body segments. The pipeline consists of four steps: (1) data augmentation, (2) pose estimation, (3) skeleton morphing, and (4) evaluation.

XSL•FO
RenderX

Figure 1. Modular benchmarking framework for the clinical validation of 3D human pose estimation. The pipeline evaluates human pose estimation performance by processing ground-truth data through 4 sequential modules, which are (1) frame augmentation: simulating clinical environmental setup errors, including defocus blur, body segment occlusion, and extrinsic camera calibration errors; (2) human pose estimation: extraction of 3D coordinates using monocular or multiocular camera configurations; (3) skeleton morphing: a dataset-specific residual neural network that harmonizes joint definitions between datasets; and (4) evaluation: calculation of domain-specific spatial and spatiotemporal metrics. Visual coding indicates data types: blue arrows represent video streams, green arrows represent keypoints in the human pose estimation-native format, and red arrows represent keypoints in the gold-standard dataset format. Image created in BioRender by EAL [44] and licensed under CC BY 4.0 [45].



Augmentation Module

This module simulates the variability of clinical environments by augmenting video inputs and calibration data. It applies image perturbations (eg, blur, occlusion, and lighting changes) and adds noise to camera calibration and temporal synchronization. This allows us to evaluate the algorithm's robustness under suboptimal, yet realistic, conditions.

HPE Module

This module provides a template to incorporate any standard 3D HPE model. It supports both monocular models (inferring 3D pose from single frames) and multiocular setups (triangulating 2D keypoints from multiple views). This step generates a raw skeletal prediction based on the augmented input frames.

Skeleton Morphing Module

As available HPE models are trained on different, at times not publicly available datasets, their definition of the anatomical location of joint center positions varies. To address the discrepancy between the surface-based keypoints of the HPE model and the internal anatomical joint centers of the marker-based gold standard, we implemented a skeleton morphing module. The skeleton morphing module is a residual neural network, based on the morphing model by Wandt et al [43], that maps the HPE keypoints onto the gold-standard (Plug-in Gait) joint-center definition. It takes the 16 predicted joints (48 values) as input, projects them to a latent representation through a fully connected layer, refines them through 2 residual blocks (fully connected layers with leaky rectified linear unit [ReLU] activation and dropout), and outputs a per-joint correction added to the input pose through a residual connection. It is trained per dataset to minimize the Procrustes-aligned mean squared error between the morphed and gold-standard skeletons, so that the correction captures the systematic offset between anatomical protocols rather than pose- or subject-specific variation. We furthermore validate the skeleton morphing module by adding linear and nonlinear offsets and noise to both the ground-truth marker-based data and MediaPipe's pose output rather than testing a second real HPE model, because in real data the joint-definition offset is confounded with tracking error and cannot be isolated. The synthetic protocol lets us quantify offset recovery and noise preservation separately (full architecture, training, and validation in Multimedia Appendix 1).

Evaluation Module

The final module assesses the model's applicability for clinical rehabilitation by calculating spatial and spatiotemporal metrics (defined in Table 1) and monitoring real-time inference speed. In clinical settings, key metrics include joint positions and their trajectories, such as movement velocity. Thus, evaluating the model's accuracy across relevant domains, such as spatial and temporal, is essential. To evaluate the intrinsic morphological correctness of the estimated pose, we aligned the estimated pose

XSL•FO
RenderX

with the ground truth using Procrustes alignment [46]. While rigid coordinate system alignment would preserve the absolute trajectory information necessary, for example, for global gait parameters, Procrustes alignment isolates the algorithm's ability to reconstruct the internal body geometry by normalizing for translation, rotation, and scale. This allows for a focused assessment of the algorithm's relative joint angle accuracy, crucial for a wide variety of stationary clinical assessments.

Table . Metrics for evaluating the human pose estimation model's spatial accuracy, spatiotemporal accuracy, and inference time. The metrics were chosen to facilitate clinical evaluation across relevant domains for rehabilitation applications. Their respective equations can be found in Multimedia Appendix 1.

Metric	Description	Unit
Spatial accuracy		
MPJPE[a]	The MPJPE calculates the Euclidean distance between the predicted keypoints and their respective ground truth over all frames and joints to evaluate the spatial accuracy. Lower is better.	mm
MPJAE[b]	The MPJAE calculates the average deviation between predicted and the ground truth joint angles across all axes and joints. The respective angles are calculated as the deviation from the T-pose [47]. Lower is better.	degrees
Spatiotemporal accuracy		
MPJAVE[c]	The MPJAVE calculates the average deviation between the smoothed (Savitzky-Golay filter, windowlength=21T+1samples, polynomialorder=3) predicted joint angle and ground truth over a time window, based on the temporal derivative of the joint angle. Lower is better.	°/s
Pearson *r*	The mean-per-joint Pearson correlation coefficient (*r*) measures spatiotemporal similarity between the predicted joint trajectories and ground truth. Closer to 1 is better.	1
Real-time capability		
Inference time	Inference time reflects the model's real-time applicability averaged over all frames. Lower is better.	ms

[a]MPJPE: mean-per-joint position error.

[b]MPJAE: mean-per-joint angular error.

[c]MPJAVE: mean-per-joint angular velocity error.

It should be noted that although the coefficient of multiple correlation has been recommended to assess the similarity of waveforms because it accounts for baseline offset and scaling differences [48], this study uses the Pearson correlation coefficient (*r*). Since the applied morphing model compensates for intersensor and dataset discrepancies in baseline and scale, their effects are negligible.

Proof-of-Concept Implementation

To demonstrate the applicability of the proposed framework, we conducted a proof-of-concept validation study using a popular, accessible real-time HPE algorithm (MediaPipe) and a domain-specific dataset.

Dataset and Experimental Protocol

The dataset, which served as the study's ground truth, was acquired using a marker‐based motion capture system (Qualisys AB) hardware, synchronized via transistor-transistor logic (TTL) with RGB video recordings from 6 camera angles (FLIR Systems Inc.). Nineteen healthy participants (mean age 30, SD 8.2 years; mean BMI 25, SD 3.7 kg/m²; 9 females) were recorded. Fifteen participants were used to train and test the skeleton morphing model and 4 for the subsequent benchmarking framework itself. Detailed acquisition, synchronization, and processing protocols are provided in Multimedia Appendix 1.

The movement protocol was designed in collaboration with subject matter experts in physiotherapy and movement science to represent common motor rehabilitation tasks. The protocol included 10 repetitions of left-sided planar upper- and lower-body movements and whole-body functional tasks (eg, sit-to-stand). A detailed list of movement sequences is provided in Multimedia Appendix 1.

Skeleton Morphing Model Training

The model was trained using leave-one-out cross-validation and tested on a subset of the VisLab dataset, consisting of 15 participants (6 females), with a mean age of 32 (SD 11) years, a mean height of 168 (SD 10) cm, and a mean weight of 75 (SD 13) kg. Leave-one-out cross-validation was used to maximize training data usage while ensuring the model learned a

XSL•FO
RenderX

generalized mapping of the joints instead of memorizing subject-specific anthropometry. These participants were excluded from the subsequent benchmarking framework. Detailed information on the model architecture and training is provided in Multimedia Appendix 1.

HPE Model Selection

To demonstrate the functionality of our framework, the commonly used real-time 3D HPE algorithm MediaPipe (Pose Landmarker Full) was chosen. As a lightweight, single-stage model optimized for accessibility and real-time use rather than maximal accuracy, MediaPipe is not the most accurate HPE model available [49]. We selected it as a representative, readily deployable baseline, and the framework itself is model agnostic. MediaPipe's 3D world coordinates were used for monocular pose estimation, while for multiocular pose estimation, MediaPipe's 2D image coordinates were triangulated via direct linear transformation (DLT) [50]. The inference was performed on an AMD Ryzen 9 5950X CPU with 64 GB RAM and an NVIDIA GeForce RTX 3090 GPU.

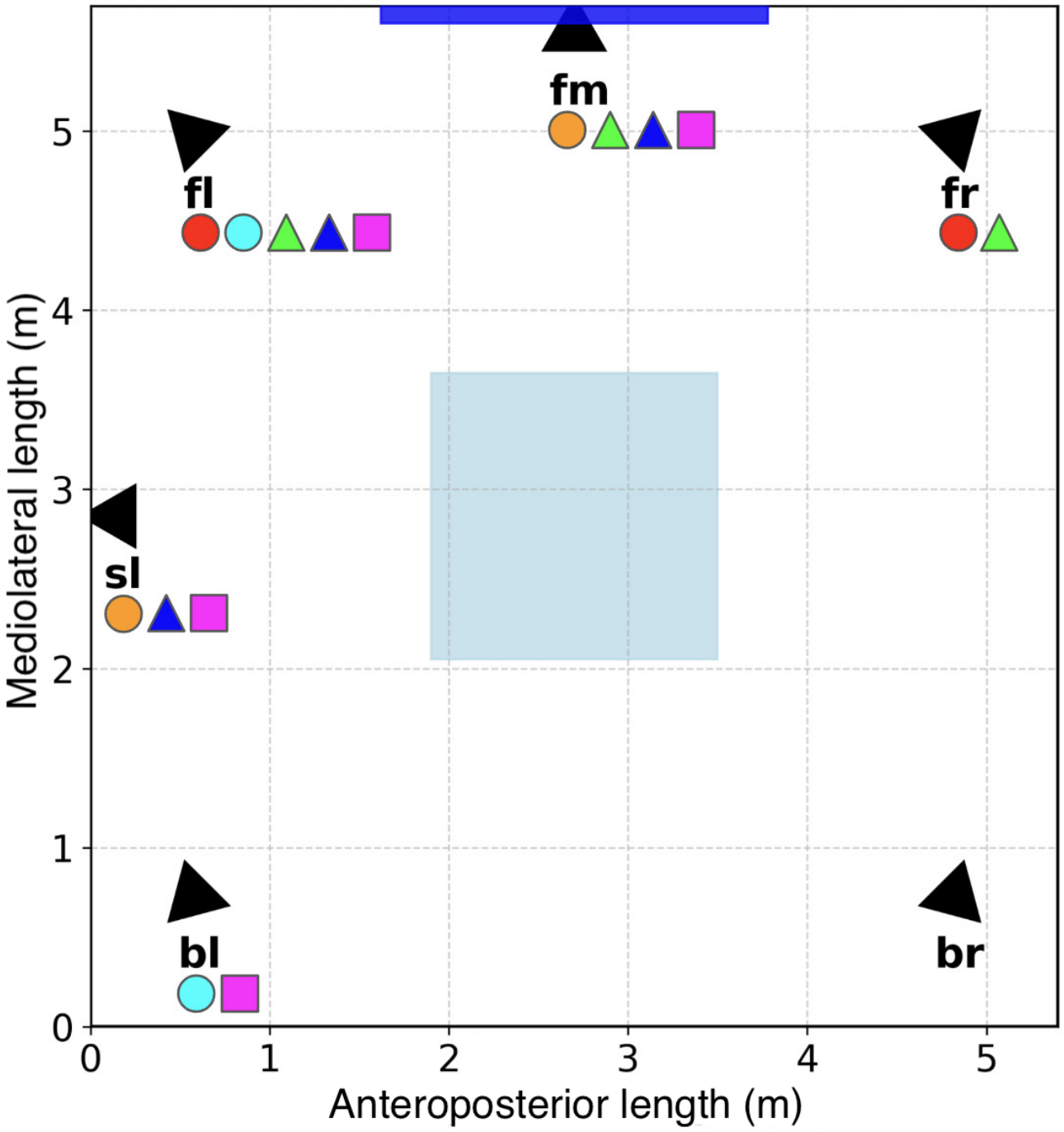
Simulation of Setup and Environmental Errors

For the simulation of common errors in clinical environments, the image frames were augmented using the Python library imgaug [51] and pixlib [52]. Unfocused camera setups were simulated using strong Gaussian blur, while dim ambient lighting was replicated through exposure reduction, achieved by blending frames with a black background at 15% opacity. To simulate environmental obstructions, random occlusions covering 10% to 25% of the image size were superimposed onto the central region, and the uniform laboratory background was replaced with synthetic noise. Furthermore, for multiocular setups, custom scripts were used to perturb the extrinsic calibration parameters by adding Gaussian noise to the projection matrix (resulting in an approximate 2.5-pixel reprojection error) and to introduce temporal desynchronization between camera streams by randomly offsetting frames by 0 to 1 frame samples. A visualization of these augmentations can be found in the Multimedia Appendix 1.

In addition to the named error sources, the effect of different camera placements was tested. The cameras' respective placements are shown in Figure 2, and the camera combinations used for multiocular setups were front left (fl), front right (fr); fl, back left (bl); front middle (fm), side left (sl); fl, fm, fr; fl, fm, sl; and fl, fm, bl, sl. The specific camera positions and their height (1m) were selected to represent typical spatial constraints in clinical environments. For instance, the "fm" and "sl" positions simulate common single-tripod setups, while corner-mounted configurations (eg, fl, fr) reflect permanent installations in rehabilitation gyms where floor space must remain unobstructed.

XSL•FO
RenderX

Figure 2. Top-down schematic of the VisLab laboratory setup, illustrating red-green-blue video camera positions and the movement area. The setup was used to acquire the gold-standard motion capture data from 19 healthy participants. Colored symbols below camera IDs indicate the 2-camera (dots), 3-camera (triangles), and 4-camera (rectangles) multiocular setups tested. Bl: back left; br: back right; fl: front left; fm: front middle; fr: front right; sl: side left.



Statistical Analysis of Evaluation

The evaluation pipeline results are based on a subset of the VisLab dataset, which includes 4 participants (2 females) with a mean age of 29 (SD 4) years, a mean height of 170 (SD 12) cm, and a mean weight of 73 (SD 13) kg.

Nonnormality in mean-per-joint angular error (MPJAE), the main evaluation metric, was confirmed using Q-Q plots. The Wilcoxon signed-rank test was used to compare camera positions and augmentations across all movement conditions. To control for false positives due to multiple comparisons, the Benjamini-Hochberg procedure was applied [53]. However, due to the high sample number (N=127,828), statistical significance tests are highly sensitive, yielding $P<.001$ even for clinically negligible differences. Therefore, the interpretation of results prioritizes effect sizes (rank-biserial correlation [RBC]) and bias over P values. To keep the main tables concise, they report the primary metric (MPJAE) together with its bias and effect size. Corresponding MPJPE, mean-per-joint angular velocity error (MPJAVE), and Pearson r values, along with tables of statistical significance, are provided in Multimedia Appendix 1. The magnitude of these differences was further assessed using the RBC as effect size. Differences between movement types were assessed using pairwise post hoc

XSL•FO
RenderX

comparisons using the Games-Howell test with Benjamini-Hochberg correction to control for false positives [53,54].

Ethical Considerations

The study was approved by the Data Protection Services of the Norwegian Agency for Shared Services in Education and Research (reference number: 878754) and is in accordance with the Declaration of Helsinki. Multimedia Appendix 1 provides additional information on the VisLab dataset.

Results

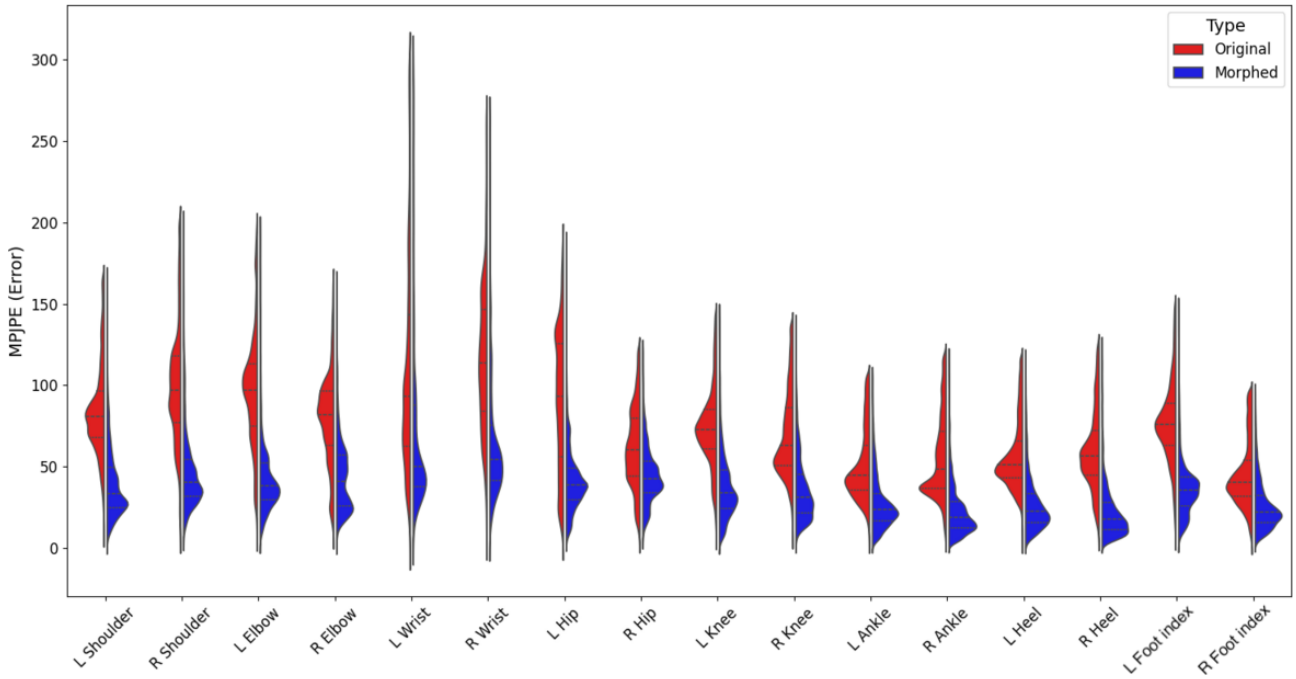
Impact of Keypoint Morphing on Baseline Accuracy

To ensure the framework allows for accurate interdataset comparison, we evaluated the skeleton morphing module. The morphing model decreases the overall Procrustes-aligned mean per joint position error (PA-MPJPE) by 55.29 mm (57.2%) and its SD by 24.69 mm (n=79,423, 50.7%). We observed a significant decrease in the MPJPE error across all tracked keypoints after applying the morphing model (Wilcoxon signed‐rank test, W=1783; P<.001), alongside a significant reduction in variability (Levene test, $F_{1,\ 158,844}$=13,008.1; P<.001; Figure 3).

Joint‐wise analyses revealed significant improvements across all keypoints, with a large effect (RBC mean 0.86, SD 0.10). Notably, while the elbow, wrist, and hip joints exhibited some side‐dependent differences in effect size, the mean effect size difference between the left and right sides was minimal ($RBC_{left\text{-}right}$=0.02).

To confirm that this improvement reflects a correction of the anatomical offset rather than a suppression of genuine tracking error, we additionally validated the morphing model against controlled synthetic offsets with known ground truth (Multimedia Appendix 1, validation method of morphing model). On held-out participants, the model achieved a mean offset recovery rate of 65.0% for movement-dependent (nonlinear) offsets, compared to 21.0% for a per-joint linear correction, while maintaining a noise-preservation ratio of 0.97-1.00, confirming that it harmonizes joint definitions without masking tracking error.

Figure 3. Mean-per-joint position error before and after morphing MediaPipe keypoints to the Plug-in Gait joint center positions. The morphing model was trained on 15 healthy participants to improve cross-dataset comparison validity. L: left; MPJPE: mean-per-joint position error; R: right.



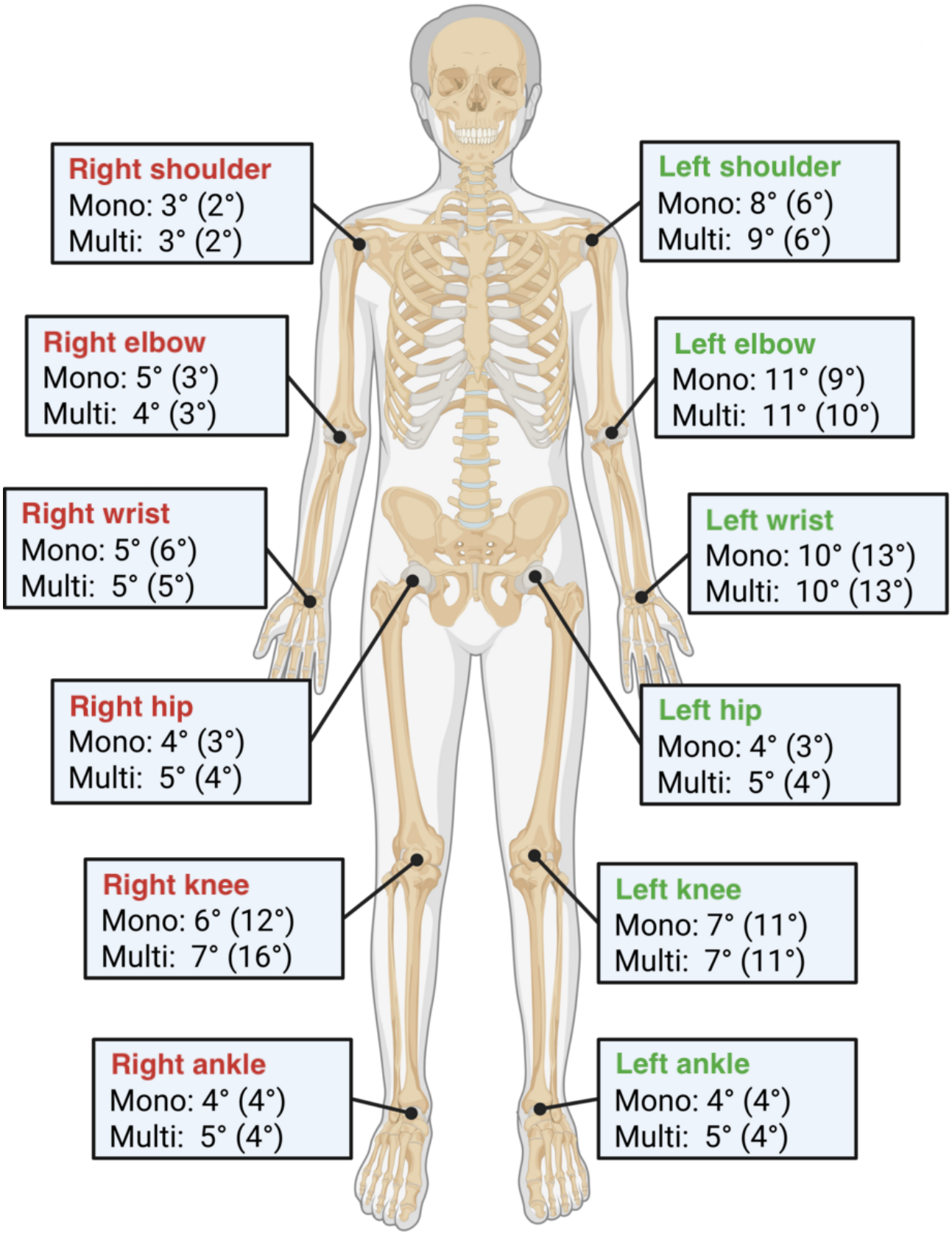
Joint-Specific Sensitivity Analysis

The framework's joint-wise resolution identified anatomical inconsistencies of the HPE model. By evaluating AEs for each joint, we identified which joints are most prone to inaccuracies. The results are based on nonaugmented inputs for the default camera positions (mono: fm, multi: fl-fr) and all movements.

Figure 4 illustrates that the MPJAEs generally tend to be slightly higher in multiocular setups for the default camera placements. Distal upper body joints, such as the wrists and elbows, exhibit higher AEs than proximal or lower body joints in both setups. The knee and left upper body joints show a relatively high variability.

XSL•FO
RenderX

Figure 4. Mean-per-joint angular error and SD for each joint in monocular and multiocular human pose estimation setups. Results are based on nonaugmented inputs for the default camera positions on 4 healthy participants (N=127,828 frames). Image created in BioRender by EAL [55] and licensed under CC BY 4.0 [45]. Mono: monocular; multi: multiocular.



XSL•FO
RenderX

Benchmarking Robustness to Simulated Clinical Setup Errors

The impact of common setup errors on measurement accuracy was assessed for the default camera positions (mono: fm; multi: fl, fr). Table 2 displays the MPJAE, its bias, and effect size across simulated setup errors, averaged over all joints and movements; corresponding MPJPE and MPJAVE distributions are shown in Figure 5.

Table . Mean (SD) values of mean-per-joint angular error across simulated setup errors for monocular and multiocular human pose estimation setups[a].

Setup errors	MPJAE[b] (degrees), mean (SD)		MPJAE effect size		MPJAE bias (degrees), n (%)	
	Mono	Multi	Mono	Multi	Mono	Multi
Background	6 (8)	7 (10)	0.14	0.14	0.3 (5.1)	1.0 (15.3)
Defocus	6 (8)	6 (9)	0.02	0.02	0.1 (1.3)	0.1 (1.3)
Occlusion	7 (9)	8 (10)	0.14	0.24	0.7 (12.7)	1.5 (24.3)
Underexposure	6 (8)	6 (8)	−0.02	0.02	−0.1 (−1.1)	0.0 (−0.2)
Decalibrated	—[c]	14 (14)	—	0.95	—	7.8 (123.4)
Desynchronized	—	6 (8)	—	0.06	—	0.0 (0)

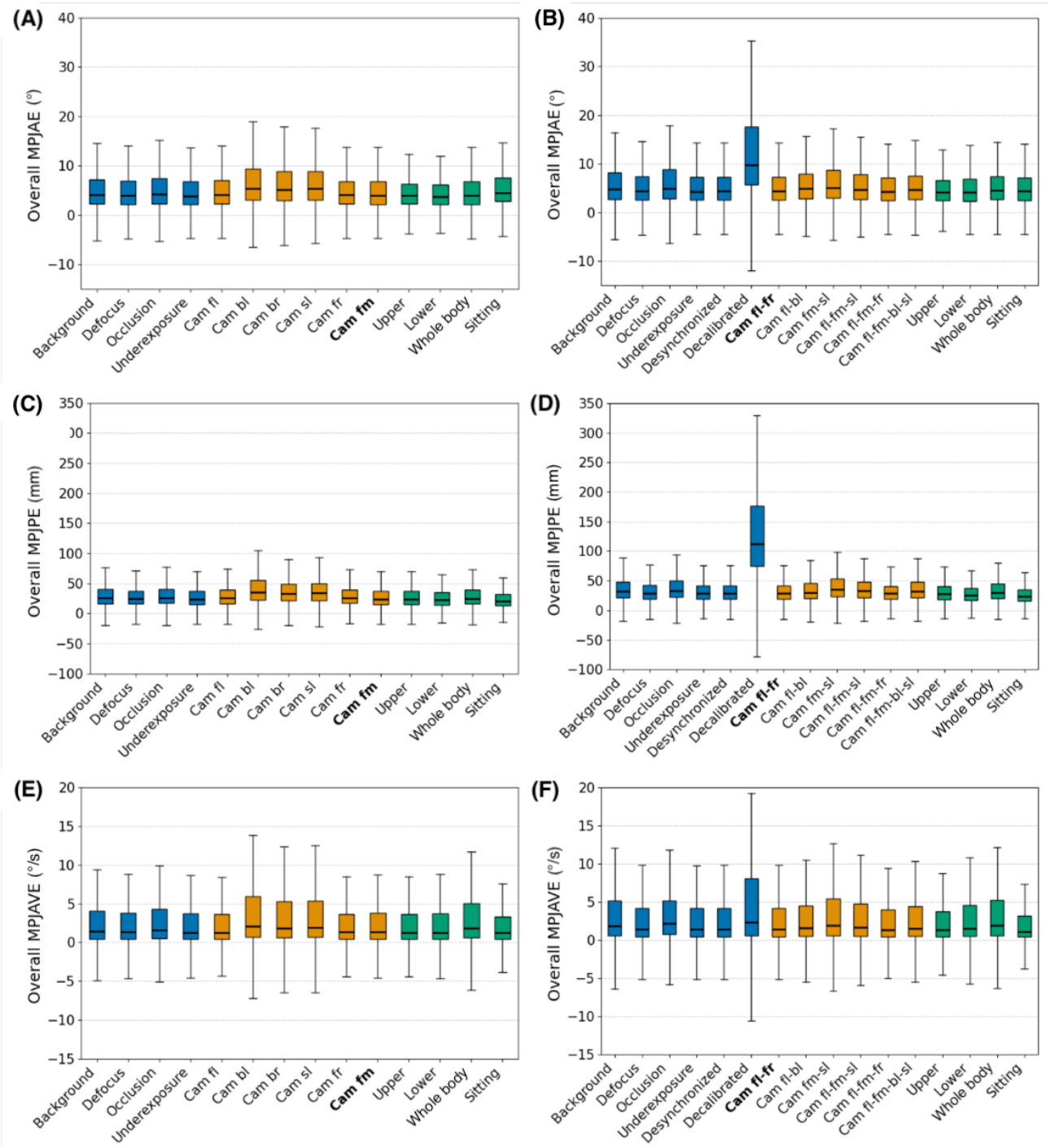
[a]The rank-biserial correlation effect size and bias are given in comparison to the baseline (no augmentation). Results are averaged over all joints and movements using default camera positions (mono: front-middle; multi: front-left and front-right) on 4 healthy participants (N=127,828 total frames).

[b]Mean-per-joint angular error.

[c]Not applicable.

XSL•FO
RenderX

Figure 5. Boxplots of mean joint metrics comparing monocular and multiocular human pose estimation setups. The metrics displayed are (A and B) mean-per-joint angular error, (C and D) mean-per-joint position error, and (E and F) mean-per-joint angular velocity error. The left column (A, C, and E) represents results from the monocular setup, while the right column (B, D, and F) represents the multiocular setup. Data are grouped by setup errors, camera positions, and movement categories, summarizing the evaluation of 4 healthy participants across all simulated scenarios (N=127,828 frames). Bold markings indicate default camera positions. bl: back left; br: back right; Cam: camera; fl: front left; fm: front middle; fr: front right; MPJAE: mean-per-joint angular error; MPJAVE: mean-per-joint angular velocity error; MPJPE: mean-per-joint position error; sl: side left.



The framework verified that defocused and underexposed frames had negligible effects on measurement accuracy in both monocular and multiocular setups ($RBC_{defocus}$=0.02 and $RBC_{underexposure}$ between –0.02 and 0.02) and negligible biases ($Bias_{defocus}$=0.1 degrees and $Bias_{underexposure}$ between –0.1 degrees and 0.0 degrees) compared to the baseline condition.

Increased background noise had only a small effect (RBC=0.14) in both configurations. In the multiocular setup, the framework captured an MPJAE increase of 1 degree (15% relative increase), rising from a baseline of 6 (SD 8) degrees to 7 (SD 10) degrees and MPJPE to 45 (SD 56) mm. The MPJAVE also increased to 5 (SD 11)°/s, while the Pearson *r* decreased to 0.90. The mono - ocular setup also remained close to the baseline, with an MPJAE of 6 (SD 8) degrees and an MPJPE of 33 (SD 27) mm.

Occlusion in the frame presented a challenge for both monocular and multiocular systems, with a small effect (RBC_{mono}=0.14

XSL•FO
RenderX

and RBC_{multi}=0.24). However, the metrics captured a deterioration in the multiocular setup, quantifying an MPJAE increase of 2 degrees (24.3%) to 8 (SD 10) degrees. Meanwhile, MPJPE rose to 45 (SD 53) mm, and MPJAVE reached 5 (SD 9)°/s. Temporal agreement was also affected, with the Pearson *r* dropping to 0.89.

The framework indicates a high sensitivity to extrinsic calibration of multiocular setups. By introducing a calibration error, the framework registered a large degradation (RBC=0.95), with MPJAE increasing significantly by 8 degrees (123% relative increase), rising from a baseline of 6 (SD 8) degrees to 14 (SD 14) degrees. Alongside, the MPJPE increased to 137 (SD 92) mm, and MPJAVE reached 8 (SD 16)°/s. Spatiotemporal agreement decreased considerably, with the Pearson *r* dropping to 0.65.

Finally, camera desynchronization had a negligible impact on the multiocular system (RBC=0.06 and bias=0 degrees).

Benchmarking Robustness Across Camera Placements

The framework provided a detailed map of spatial reliability across different camera placements, allowing for the identification of optimal sensor placement. Table 3 summarizes MPJAE, its bias, and effect size for each camera placement relative to the default position (mono: fm, multi: fl-fr); corresponding MPJPE and MPJAVE distributions are shown in Figure 5.

Table . Mean (SD) of mean-per-joint angular error for all camera placements across all movements[a].

Placement	MPJAE[b] (degrees), mean (SD)	Effect size	Bias (degrees), n (%)
fl[c]	6 (8)	−0.01	0.0 (−0.5)
fm[d]	6 (8)	—[e]	—
fr[f]	6 (8)	0.03	0.1 (1.6)
br[g]	8 (10)	0.46	1.8 (30.2)
bl[h]	7 (9)	0.41	1.6 (28.1)
sl[i]	8 (10)	0.49	2.2 (37.1)
fl, fr	6 (8)	—	—
fl, bl	7 (8)	0.12	0.4 (6)
fm, sl	7 (9)	0.33	1.1 (16.8)
fl, fm, fr	6 (8)	−0.06	−0.2 (−3.1)
fl, fm, sl	7 (9)	0.14	0.3 (5.1)
fl, fm, bl, sl	6 (9)	0.04	0.1 (2.3)

[a]The rank-biserial correlation effect size and bias are given in comparison to the baseline (fm and fl,fr). Results are shown for monocular and multiocular setups on nonaugmented inputs from 4 healthy participants (N=127,828 total frames).

[b]Mean-per-joint angular error.

[c]fl: front left.

[d]fm: front middle.

[e]Not applicable.

[f]fr: front right.

[g]br: back right.

[h]bl: back left.

[i]sl: side left.

The evaluation pipeline verified that changes in camera position had only a minimal impact on pose estimation accuracy (Table 3). Across monocular setups, the metrics indicated that MPJAE remained approximately 6 (SD 8) degrees to 8 (SD 10) degrees, MPJPE ranged between 30 (SD 22) mm and 47 (SD 46) mm, and MPJAVE remained low at 4 (SD 8)°/s to 6 (SD 12)°/s, with high temporal agreement (Pearson *r*=0.92 - 0.96). The framework confirmed similar stability in multiocular setups (MPJAE=6 (SD 8) degrees to 7 (SD 9) degrees, MPJPE=33 (SD 22) mm to 45 (SD 39) mm, MPJAVE=4 (SD 9)°/s to 5 (SD 11)°/s; Pearson *r*=0.93 - 0.95), while quantifying that multiocular configurations were more consistent across placements compared to monocular setups.

Further, the framework indicated that differences in camera placement were statistically significant (*P*<.001). In monocular setups, the pipeline confirmed high reliability in frontal placements (fl and fr), registering negligible effect sizes (RBC=–0.01 to 0.03). In contrast, back - facing (br and bl) and side - facing (sl) placements yielded medium effect sizes (RBC between 0.41 and 0.49) with biases reaching up to 2.2 degrees (37.1%). For multiocular setups, most camera combinations showed negligible changes (eg, [fl, fm, fr] with RBC=–0.06 and [fl, fm, bl, sl] with RBC=0.04). Notably, stereoscopic setups with one side- or back-facing camera ([fm, sl] and [fl, bl]) exhibited a moderate effect ($RBC_{fm,\ sl}$=0.33, $bias_{fm,\ sl}$=1.1

XSL•FO
RenderX

degrees or 16.8%) to a small effect ($RBC_{fl, bl}$=0.12, $bias_{fl, bl}$=0.4 degrees or 6%).

Influence of Movement Type on Accuracy

We evaluated the influence of various movement types—upper body, lower body, and whole-body standing and sitting movements—on measurement accuracy by analyzing their effect on the chosen metrics (Table 4 and Figure 5). The bias and effect size of the MPJAE are presented in Multimedia Appendix 1. The analysis uses only nonaugmented inputs from the default camera position.

Table . Mean (SD) of mean-per-joint angular error across different movement categories[a].

Movement type	MPJAE[b] (degrees), mean (SD)		Number of samples
Upper	6 (7)[c]	6 (8)[d]	30,521
Lower	6 (7)[c]	6 (7)[d]	29,818
Standing	6 (8)[c]	6 (8)[d]	37,967
Sitting	6 (9)[c]	6 (8)[d]	29,522

[a] Results are based on nonaugmented inputs from the default camera positions in 4 healthy participants, demonstrating the influence of movement type on pose estimation accuracy.

[b]Mean-per-joint angular error.

[c]Mono.

[d]Multi.

Table 4 shows that movement type only slightly influenced pose estimation accuracy. Across upper, lower, and whole-body standing and sitting movements, the metrics remained stable with an MPJAE close to 6 (SD 7-9) degrees for both monocular and multiocular setups. MPJPE varied between 25 (SD 17) mm for sitting and 36 (SD 23) mm for whole-body standing movements, while MPJAVE was consistently low, ranging from 3 (SD 8)°/s for sitting movements to 4 (SD 10)°/s across upper-, lower-, and standing-body movements. Pearson *r* (0.89 - 0.95) values indicated strong temporal agreement across all movement categories. For both monocular and multiocular setups, sitting movements demonstrated the lowest errors and best temporal agreement compared to whole-body standing movements, which showed the highest spatial errors and temporal variability. Although post hoc comparisons (Multimedia Appendix 1) reveal statistically significant differences (P<.001) among certain movement types, the effect sizes remained largely negligible. In monocular setups, the largest observed effect size (Hedges g=0.24 for lower vs sitting) corresponded to only about a 1 degree difference in MPJAE. Similarly, in multiocular conditions, the largest effect size (Hedges g=0.16 for standing vs upper) also represented a negligible practical difference.

Benchmarking of Real-time Capability

Last, the model's real-time capability was evaluated by comparing processing latency across camera configurations. This analysis considers various camera counts, providing insights into each model's performance and feasibility for real-time applications. The results are based on nonaugmented inputs for the default camera position. Processing times per frame increased incrementally with the addition of camera inputs. A single-camera setup achieved an average processing time of 23.5 (SD 0.2) ms, corresponding to a frame rate of approximately 43 fps. As complexity increased, the processing time rose to 26.3 (SD 0.7) ms for 2 cameras and 28.2 (SD 0.9) ms for 3 cameras. The 4-camera configuration resulted in the highest latency at 30.6 (SD 1) ms (~33 fps). Across all configurations (N=127,828 frames), the SD remained low, indicating highly consistent processing times regardless of the setup complexity.

Discussion

Principal Findings

Applying the framework to MediaPipe, we found the evaluated model robust to simulated underexposure, defocus, and background variation, but markedly degraded by body-segment occlusion, insufficient multiocular calibration, and side- or back-facing camera placements. The skeleton morphing module reduced the joint-definition offset between the HPE output and the marker-based reference (PA-MPJPE mean reduction of 55.29, SD 24.69 mm; N=79,423 frames, −57.2% reduction) while, on synthetic data, leaving injected noise intact, confirming that it harmonizes joint definitions without masking genuine tracking error. Beyond these specific results, the study's main contribution is the evaluation pipeline itself. Because its augmentation, pose estimation, skeleton morphing, and evaluation modules are interchangeable, present or future HPE models can be assessed against the requirements of a given rehabilitation workflow. The source code for the framework and morphing model is available in GitHub [56].

Augmentation: Validating Robustness and Defining Setup Guidelines

By simulating plausible configuration errors, the framework identifies the operational limits of evaluated HPE models and informs evidence-based setup guidelines. In this proof-of-concept, the framework indicated that underexposure, defocus, and simulated clinical backgrounds had negligible effects on the given HPE model's performance, likely due to its training on partially augmented datasets. This is consistent with previous studies [18,21,23]. In summary, this suggests that lightweight, real-time HPE models such as the one evaluated here can tolerate the lighting and background variation tested

XSL•FO
RenderX

here. This is an encouraging but preliminary indication for less controlled clinical or home environments that require confirmation on patient populations.

In contrast, occlusion remains problematic for HPE systems despite strategies such as per-point visibility classifiers [49] or integrating temporal information (eg, transformer-based models or kinematic constraints) [24]. Therefore, clinical setups must ensure that occlusions are minimized—environmental as well as self-occlusion.

Furthermore, the framework highlighted the sensitivity of multiocular setups to setup errors. While multiocular camera stream desynchronization did not degrade accuracy for slow movements, insufficient camera calibration significantly did. This highlights the need for precise calibration in multiocular setups, which is a well-known problem [57,58]. Thus, particularly for clinical applications, this resource-intensive and error-prone manual process could benefit from novel methods that facilitate automatic calibration [57] or do not require calibration [58].

Regarding camera placement, the framework can be used to indicate preferable setup configurations. In our proof-of-concept, the significant error increase in side-facing setups (>37% bias) indicates that clinical setups should prioritize frontal or corner-mounted views to minimize self-occlusion (consistent with Aguilar-Ortega et al [40]). View-dependent accuracy loss in markerless capture is well documented, particularly for sagittal and nonfrontal perspectives [35], which our per-placement results quantify directly. The framework also indicated that complex multiocular setups, like our multiocular example (DLT-MediaPipe), might not improve accuracy or occlusion robustness relative to modern monocular setups. While Neupane et al [21] report that multiocular systems can improve accuracy, Nogueira et al [50] also noted their sensitivity to occlusion. More recent multiocular HPE models that achieve a significantly increased spatial accuracy (MPJPE: 7.9 mm [58]) estimate the keypoints directly in 3D space rather than reconstructing 3D poses from 2D keypoints using, for example, DLT [50]. However, this often comes at the cost of real-time performance. Consequently, the added complexity and costs of multiocular setups may not justify modest gains in accuracy, for example, in telerehabilitation applications [21,50].

Skeleton Morphing: Ensuring Validity in Interdataset Comparisons

A key framework component is the morphing model that is designed to align keypoint annotations across datasets and which we validate on the MediaPipe-to-Plug-in Gait mapping using controlled synthetic offsets. While its application enhances the validity of comparative analyses and benchmarking, due to dataset-dependent keypoint variations, prior studies did not apply this step. However, similar keypoint morphing methods have been developed, focusing primarily on training models across multiple datasets with different keypoint definitions [23,36,43]. Rapczyński et al [23] achieved a~37% improvement in MPJPE using linear methods to adjust for keypoint offsets and normalize scale across datasets before training. Although demonstrated to be effective, this linear approach cannot adequately capture nonlinear relationships between different keypoint annotation methods. However, Rapczyński et al [23] approach does not require a large dataset for the initial morphing model training, akin to that of Wandt et al [43], who used only one participant. We did not investigate how different training dataset sizes affect our morphing model's performance, but a smaller subset may suffice, given the similarity of our model's architecture to that of Wandt et al [43].

Evaluation: Proposed Metrics and Thresholds in Clinical Settings

Beyond the MPJPE common in engineering benchmarks, clinical use demands domain-specific criteria across 3 dimensions: spatial accuracy, spatiotemporal consistency to capture movement dynamics, and latency for real-time feedback. We consider each in turn.

Spatial Accuracy Requirements

As expected, AE (MPJAE) was less sensitive to global positional inaccuracies than MPJPE because joint angles depend on relative segment orientations rather than absolute keypoint positions.

However, while the approach used to calculate the joint angles as a relative difference from a participant-specific T-pose is simple and consistent, it may not replicate the physiotherapists' by-eye judgment method. Thus, Mercadal-Baudart et al [59] proposed projecting the respective body segments onto a principal anatomical plane before calculating joint angles. This better simulates clinical assessments but only allows single-plane joint angle analysis.

To be able to place the observed MPJAE of 6 (SD 8) degrees into perspective, previous studies suggest that errors above 5 degrees might be misleading for clinical applications [19]. On the other hand, physiotherapists can only discern joint angles of >12 degrees in the visual assessment of single-plane movements [60]. The marker-based ground-truth is itself imperfect, with reported AEs from marker placement (<10 degrees) [61], soft-tissue artifacts (<3 degrees) [12], and joint-center estimation (<3 degrees) [62]. As HPE training data are often marker-derived, these errors may propagate into HPE estimates.

Finally, the framework's spatial accuracy results were comparable with previous studies evaluating MediaPipe [59,63] and fall within the range recently reported for newer monocular estimators benchmarked against marker-based capture on clinical exercises, where per-joint position errors spanned roughly 72 - 122 mm in the image plane and knee-flexion errors 9 - 22 degrees [34]. This indicates that the limits of our framework are not specific to the exemplar model but characteristic of the current generation of markerless estimators.

Spatiotemporal Consistency Requirements

Accurately estimating movement velocity and acceleration based solely on single-sensor position data is, in general, challenging. Inherent sensor noise and variability propagate and amplify during velocity and acceleration calculation, reducing the validity of such measurements [64]. Smoothing the position data beforehand may mitigate this, but its effectiveness depends on the chosen smoothing method and may hide relevant sensor noise or spatiotemporal inaccuracies. Thus, such metrics (like

XSL•FO
RenderX

the MPJAVE) should be rather interpreted as a measure of spatiotemporal consistency or a measure of sensor jitter.

The sensitivity to sensor jitter of MPJAVE is evident in challenging scenarios, such as when occlusion is present, which leads to a higher frequency of insufficiently detected poses. The resulting noise from discontinuities (jumps) in the movement trajectories increases the overall MPJAVE slightly, underlining the necessity of smoothing velocity curves [64,65]. However, the observed MPJAVE of 4 (SD 8)°/s is acceptable, considering low velocities of reaching movements (75 - 103°/s) [66] and during slow walking (85 - 200°/s) [67].

The Pearson *r* supports this observation and shows a high spatiotemporal agreement of the movement trajectories between the gold-standard and HPE models. This agreement is consistent with previous studies [68].

Latency Requirements for Real-Time Feedback

While for sports applications, motion capture systems typically sample between 50 and 250 Hz [69], slower rates (~25 - 40 Hz) can be permissible for motor rehabilitation applications due to slower gross movements [70,71]. In our experiment, MediaPipe (full model) maintained a stable rate of ~40 Hz. However, the inference speed of HPE models depends significantly on the architecture of the model and the hardware used. Lighter models, which are at times faster, often compromise spatial accuracy [72].

Proposed Benchmarking Thresholds

To provide a clinical assessment from these separate metrics, we propose an application-dependent interpretation of the metrics. An algorithm should satisfy spatial accuracy (MPJAE<5 degrees), followed by spatiotemporal accuracy (Pearson r>0.90). Finally, robustness is determined by the RBC effect size being lower than 0.3 across setup errors. This approach prevents a "good average score" from masking critical failures in specific clinical requirements and provides enough flexibility for different applications with different emphases.

Framework Validity Across Rehabilitation Domains

Finally, the evaluation module assessed the framework's validity across diverse movement tasks. Contrary to previous research emphasizing the necessity of training data including domain-specific movements [5,40], our observations suggest that, at least for this class of model, HPE trained on large and diverse general datasets can generalize well across diverse clinically relevant movement types, a finding that should be confirmed on additional models. While whole-body standing movements showed the highest spatial errors and sitting movements the lowest, the differences between movement types and monocular and multiocular setups remain negligible.

The joint-wise analysis exhibited decreased spatial accuracy of the left body side. As the simple movements were predominantly performed with the left body side, this naturally showed higher spatial errors than the static right side. Consequently, when assessing joint-wise spatial errors, the specific movement task and lateral asymmetries need to be taken into consideration. Additionally, distal joints were more vulnerable to spatial errors due to their increased flexibility and range of motion [73].

Practical Considerations

Clinical deployment also requires attention to computational scalability. The evaluated model runs in real time on a standard workstation, and individual components can be swapped as HPE models evolve, so throughput depends on the chosen model rather than the framework itself. Patient-population validation and workflow integration are addressed under the Limitations section.

Limitations

A few limitations of this study must be acknowledged. First, while our dataset covers a wide range of relevant movements for a relatively diverse population, it does not include pathological movements or anthropometric deviations. Although our results indicate that HPE models accurately estimate joint center locations across various movements, pathological movement patterns may affect accuracy, especially when mobility aids are used (eg, crutches, orthotics, and prosthetics) [74].

Second, this proof-of-concept applies the framework to a single exemplar model (MediaPipe). Because the framework is model-agnostic, benchmarking additional architectures, including those with different joint topologies, is a direct application of the pipeline rather than a change to it and is planned for future work. We note that the morphing architecture follows methods (Wandt et al [43], Rapczyński et al [23], and Sárándi et al [36]) already shown to generalize across datasets with differing skeleton formats.

Finally, regarding robustness, the current framework uses standard augmentations that simulate environmental noise. While sufficient for this proof-of-concept, they do not fully replicate complex semantic occlusions found in clinics. In addition, the framework's modular design facilitates future integration of realistic 3D augmentations, such as inserting virtual objects (eg, wheelchairs) or persons (eg, therapists), to test scenario-specific distractions.

Conclusions

This study establishes a technical foundation for the standardized evaluation of markerless tracking in clinical rehabilitation. By validating a modular pipeline, comprising augmentation, pose estimation, skeleton morphing, and evaluation modules, we have provided a transparent method for stakeholders in health care to assess the feasibility of real-time 3D HPE algorithms for their specific applications.

Applying this framework to a popular, accessible real-time algorithm indicated that the evaluated model was robust to environmental lighting variation and performed consistently across the tested movement set in healthy participants. However, the framework effectively identified critical operational limitations: specifically, occlusion, poor multiocular calibration, and side-facing camera placements were found to degrade tracking accuracy significantly. These findings highlight that while current algorithms are viable for general monitoring, strictly controlled setups are required to minimize line-of-sight obstructions.

XSL•FO
RenderX

Future research should extend this validation to pathological populations, integrate additional augmentations, and extend the set of postprocessing strategies and metrics (eg, Bland-Altman analysis) to further evaluate these systems. By providing a reproducible and domain-specific benchmarking approach, this framework facilitates effective integration of computer vision into motor rehabilitation workflows.

Acknowledgments

The authors thank Elise K. Vonstad for her technical assistance with processing marker-based motion capture data, Mihailo Petrovic for assisting with data collection, and Olav F. P. Larsen, William G. Tresselt, and Tord Natlandsmyr for their technical support in data cleaning. We also thank NTNU's VisLab for providing access to equipment and resources. We acknowledge the use of BioRender.com for the creation of Figure 1 and Figure 4. Grammarly was used during manuscript preparation to refine spelling, grammar, and text readability. Generative AI (Gemini) was used for software development support and initial title ideation, specifically, in assisting with data visualization scripting and debugging across the data processing and skeleton morphing training pipelines. Prompt transcripts were not archived, as they were limited to iterative debugging and code optimization. The authors have verified all AI-assisted code and assume full accountability for the technical accuracy of this work.

Funding

This research protocol was funded by the Department of Computer Science at NTNU through the Strategic Research Areas funding scheme.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Methods and results, including the VizLab dataset description (marker set, red-green-blue camera and synchronization setup, experimental procedure, and data processing), metric equations, the skeleton morphing model architecture and training, the synthetic offset-recovery validation, and additional statistical tables (mean-per-joint position error, mean-per-joint angular velocity error, Pearson *r*, *P* values, and CIs).
[PDF File, 1551 KB - rehab_v13i1e91360_app1.pdf]

References

1. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chronic Dis 1985;38(1):27-36. [doi: 10.1016/0021-9681(85)90005-0] [Medline: 3972947]
2. Bai J, Li G, Lu X, Wen X. Automatic rehabilitation assessment method of upper limb motor function based on posture and distribution force. Front Neurosci 2024;18:1362495. [doi: 10.3389/fnins.2024.1362495] [Medline: 38440394]
3. Poole JL, Whitney SL. Assessments of motor function post stroke. Phys Occup Ther Geriatr 2001 Dec 1;19(2):1-22. [doi: 10.1300/J148v19n02_01]
4. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J NeuroEngineering Rehabil 2012 Dec;9(1):21. [doi: 10.1186/1743-0003-9-21]
5. Stenum J, Cherry-Allen KM, Pyles CO, Reetzke RD, Vignos MF, Roemmich RT. Applications of pose estimation in human health and performance across the lifespan. Sensors (Basel) 2021 Nov 3;21(21):7315. [doi: 10.3390/s21217315] [Medline: 34770620]
6. Dos Santos RB, Fiedler A, Badwal A, et al. Standardized tools for assessing balance and mobility in stroke clinical practice guidelines worldwide: a scoping review. Front Rehabil Sci 2023;4:1084085. [doi: 10.3389/fresc.2023.1084085] [Medline: 36896249]
7. Maura RM, Rueda Parra S, Stevens RE, Weeks DL, Wolbrecht ET, Perry JC. Literature review of stroke assessment for upper-extremity physical function via EEG, EMG, kinematic, and kinetic measurements and their reliability. J Neuroeng Rehabil 2023 Feb 15;20(1):21. [doi: 10.1186/s12984-023-01142-7]
8. Vonstad EK, Vereijken B, Bach K, Su X, Nilsen JH. Assessment of machine learning models for classification of movement patterns during a weight-shifting exergame. IEEE Trans Hum-Mach Syst 2021 Jun;51(3):242-252. [doi: 10.1109/THMS.2021.3059716]
9. Cho KH, Lee HJ, Lee WH. Test-retest reliability of the GAITRite walkway system for the spatio-temporal gait parameters while dual-tasking in post-stroke patients. Disabil Rehabil 2015;37(6):512-516. [doi: 10.3109/09638288.2014.932445] [Medline: 24957081]
10. van Meulen FB, Klaassen B, Held J, et al. Objective evaluation of the quality of movement in daily life after stroke. Front Bioeng Biotechnol 2016 Jan 13;3:210. [doi: 10.3389/fbioe.2015.00210] [Medline: 26793705]

XSL•FO
RenderX

11. Otaka E, Otaka Y, Kasuga S, et al. Clinical usefulness and validity of robotic measures of reaching movement in hemiparetic stroke patients. J Neuroeng Rehabil 2015 Dec;12(1):66. [doi: 10.1186/s12984-015-0059-8]
12. Benoit DL, Damsgaard M, Andersen MS. Surface marker cluster translation, rotation, scaling and deformation: their contribution to soft tissue artefact and impact on knee joint kinematics. J Biomech 2015 Jul 16;48(10):2124-2129. [doi: 10.1016/j.jbiomech.2015.02.050] [Medline: 25935684]
13. Hindle BR, Keogh JWL, Lorimer AV. Inertial-based human motion capture: a technical summary of current processing methodologies for spatiotemporal and kinematic measures. Appl Bionics Biomech 2021;2021:6628320. [doi: 10.1155/2021/6628320] [Medline: 33859720]
14. Zhu W, Ma X, Liu Z, Liu L, Wu W, Wang Y. MotionBERT: a unified perspective on learning human motion representations. Presented at: 2023 IEEE/CVF International Conference on Computer Vision (ICCV); Oct 1-6, 2023; Paris, France p. 15039-15053. [doi: 10.1109/ICCV51070.2023.01385]
15. Jiang T, Lu P, Zhang L, et al. RTMPose: real-time multi-person pose estimation based on MMPose. arXiv. Preprint posted online on Jul 3, 2023. [doi: 10.48550/arXiv.2303.07399]
16. Maji D, Nagori S, Mathew M, Poddar D. YOLO-pose: enhancing YOLO for multi person pose estimation using object keypoint similarity loss. 2022 Presented at: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); Jun 19-20, 2022; New Orleans, LA p. 2636-2645. [doi: 10.1109/CVPRW56347.2022.00297]
17. Menychtas D, Petrou N, Kansizoglou I, et al. Gait analysis comparison between manual marking, 2D pose estimation algorithms, and 3D marker-based system. Front Rehabil Sci 2023;4:1238134. [doi: 10.3389/fresc.2023.1238134] [Medline: 37744429]
18. Pagnon D, Domalain M, Reveret L. Pose2Sim: an end-to-end workflow for 3D markerless sports kinematics-part 1: robustness. Sensors (Basel) 2021 Sep 30;21(19):6530. [doi: 10.3390/s21196530] [Medline: 34640862]
19. McGinley JL, Baker R, Wolfe R, Morris ME. The reliability of three-dimensional kinematic gait measurements: a systematic review. Gait Posture 2009 Apr;29(3):360-369. [doi: 10.1016/j.gaitpost.2008.09.003] [Medline: 19013070]
20. Badiola-Bengoa A, Mendez-Zorrilla A. A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise. Sensors (Basel) 2021 Sep 7;21(18):5996. [doi: 10.3390/s21185996] [Medline: 34577204]
21. Neupane RB, Li K, Boka TF. A survey on deep 3D human pose estimation. Artif Intell Rev 2024;58(1):24. [doi: 10.1007/s10462-024-11019-3]
22. Fu Y, Huang C, Li J, et al. HDiffTG: a lightweight hybrid diffusion-transformer-GCN architecture for 3D human pose estimation. Presented at: 2025 International Joint Conference on Neural Networks (IJCNN); Jun 30 to Jul 5, 2025; Rome, Italy p. 1-8. [doi: 10.1109/IJCNN64981.2025.11227571]
23. Rapczyński M, Werner P, Handrich S, Al-Hamadi A. A baseline for cross-database 3D human pose estimation. Sensors (Basel) 2021 May 28;21(11):3769. [doi: 10.3390/s21113769] [Medline: 34071704]
24. Liu W, Bao Q, Sun Y, Mei T. Recent advances of monocular 2D and 3D human pose estimation: a deep learning perspective. ACM Comput Surv 2023 Apr 30;55(4):1-41. [doi: 10.1145/3524497]
25. Mitchell J, Shirota C, Clanchy K. Factors that influence the adoption of rehabilitation technologies: a multi-disciplinary qualitative exploration. J Neuroeng Rehabil 2023;20(1):80. [doi: 10.1186/s12984-023-01194-9]
26. open-mmlab/mmpose. GitHub. 2020. URL: https://github.com/open-mmlab/mmpose [accessed 2024-10-16]
27. Andriluka M, Iqbal U, Insafutdinov E, et al. PoseTrack: a benchmark for human pose estimation and tracking. Presented at: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 18, 2018 to Jun 23, 2023; Salt Lake City, UT p. 5167-5176. [doi: 10.1109/CVPR.2018.00542]
28. Dubey S, Dixit M. A comprehensive survey on human pose estimation approaches. Multimed Syst 2023 Feb;29(1):167-195. [doi: 10.1007/s00530-022-00980-0]
29. Hamilton RI, Glavcheva-Laleva Z, Haque Milon MI, et al. Comparison of computational pose estimation models for joint angles with 3D motion capture. J Bodyw Mov Ther 2024 Oct;40:315-319. [doi: 10.1016/j.jbmt.2024.04.033] [Medline: 39593603]
30. Needham L, Evans M, Cosker DP, et al. The accuracy of several pose estimation methods for 3D joint centre localisation. Sci Rep 2021 Oct 19;11(1):20673. [doi: 10.1038/s41598-021-00212-x] [Medline: 34667207]
31. Kusunose M, Inui A, Nishimoto H, et al. Measurement of shoulder abduction angle with posture estimation artificial intelligence model. Sensors (Basel) 2023 Jul 16;23(14):6445. [doi: 10.3390/s23146445] [Medline: 37514738]
32. van den Hoorn W, Lavaill M, Cutbush K, Gupta A, Kerr G. Comparison of shoulder range of motion quantified with mobile phone video-based skeletal tracking and 3D motion capture-preliminary study. Sensors (Basel) 2024 Jan 15;24(2):534. [doi: 10.3390/s24020534] [Medline: 38257626]
33. Chen K, Gabriel P, Alasfour A, et al. Patient-specific pose estimation in clinical environments. IEEE J Transl Eng Health Med 2018;6:2101111. [doi: 10.1109/JTEHM.2018.2875464] [Medline: 30483453]
34. Rode D, Dunkel A, Willi R, Wolf P, Xiloyannis M, Riener R. Assessment of monocular human pose estimation models for clinical movement analysis. Sci Rep 2025 Nov 5;15(1):38767. [doi: 10.1038/s41598-025-22626-7] [Medline: 41193590]
35. Wade L, Needham L, McGuigan P, Bilzon J. Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. PeerJ 2022 Feb 25;10:e12995. [doi: 10.7717/peerj.12995] [Medline: 35237469]

XSL•FO
RenderX

36. Sarandi I, Hermans A, Leibe B. Learning 3D human pose estimation from dozens of datasets using a geometry-aware autoencoder to bridge between skeleton formats. Presented at: IEEE Workshop on Applications of Computer Vision (WACV); Jan 2-7, 2023; Waikoloa, HI p. 2955-2965. [doi: 10.1109/WACV56688.2023.00297]
37. Ionescu C, Papava D, Olaru V, Sminchisescu C. Human3.6M: large scale datasets and predictive methods for 3D human sensing in natural environments. IEEE Trans Pattern Anal Mach Intell 2014 Jul;36(7):1325-1339. [doi: 10.1109/TPAMI.2013.248] [Medline: 26353306]
38. Andriluka M, Pishchulin L, Gehler P, Schiele B. 2D human pose estimation: new benchmark and state of the art analysis. 2014 Presented at: 2014 IEEE Conference on Computer Vision and Pattern Recognition; Jun 23-28, 2014; Columbus, OH p. 3686-3693. [doi: 10.1109/CVPR.2014.471]
39. Jin S, Xu L, Xu J, et al. Whole-body human pose estimation in the wild. In: Vedaldi A, Bischof H, Brox T, Frahm JM, editors. Computer Vision – ECCV 2020: 16th European Conference 2020:196-214. [doi: 10.1007/978-3-030-58545-7_12]
40. Aguilar-Ortega R, Berral-Soler R, Jiménez-Velasco I, et al. UCO Physical rehabilitation: new dataset and study of human pose estimation methods on physical rehabilitation exercises. Sensors (Basel) 2023 Oct 31;23(21):8862. [doi: 10.3390/s23218862] [Medline: 37960561]
41. Capecci M, Ceravolo MG, Ferracuti F, et al. The KIMORE dataset: kinematic assessment of movement and clinical scores for remote monitoring of physical rehabilitation. IEEE Trans Neural Syst Rehabil Eng 2019 Jul;27(7):1436-1448. [doi: 10.1109/TNSRE.2019.2923060] [Medline: 31217121]
42. Li J, Xue J, Cao R, et al. FineRehab: a multi-modality and multi-task dataset for rehabilitation analysis. 2024 Presented at: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); Jun 17-18, 2024; Seattle, WA p. 3184-3193. [doi: 10.1109/CVPRW63382.2024.00324]
43. Wandt B, Rudolph M, Zell P, Rhodin H, Rosenhahn B. CanonPose: self-supervised monocular 3D human pose estimation in the wild. Presented at: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 19-25, 2021; Nashville, TN (virtual event) p. 13289-13299. [doi: 10.1109/CVPR46437.2021.01309]
44. Modular benchmarking framework for the clinical validation of 3D human pose estimation (HPE). Bio Render. URL: https://app.biorender.com/citation/6942a5635a4830e52ade8db0 [accessed 2026-10-05]
45. CC BY 4.0: attribution 4.0 international deed. creative commons. URL: https://creativecommons.org/licenses/by/4.0/deed.en [accessed 2026-10-26]
46. Goodall C. Procrustes methods in the statistical analysis of shape. J R Stat Soc Series B Stat Methodol 1991 Jan 1;53(2):285-321. [doi: 10.1111/j.2517-6161.1991.tb01825.x]
47. Batpurev T. TemugeB/joint_angles_calculate. GitHub. 2021. URL: https://github.com/TemugeB/joint_angles_calculate [accessed 2026-09-30]
48. Ferrari A, Cutti AG, Cappello A. A new formulation of the coefficient of multiple correlation to assess the similarity of waveforms measured synchronously by different motion analysis protocols. Gait Posture 2010 Apr;31(4):540-542. [doi: 10.1016/j.gaitpost.2010.02.009] [Medline: 20303272]
49. Lugaresi C, Tang J, Nash H, et al. MediaPipe: a framework for building perception pipelines. arXiv. Preprint posted online on Jun 14, 2019. [doi: 10.48550/arXiv.1906.08172]
50. Nogueira AFR, Oliveira HP, Teixeira LF. Markerless multi-view 3D human pose estimation: a survey. Image Vis Comput 2025 Mar;155:105437. [doi: 10.1016/j.imavis.2025.105437]
51. Wada K, Crall J, Tanaka S, et al. aleju/imgaug. GitHub. 2020. URL: https://github.com/aleju/imgaug [accessed 2026-09-30]
52. Olafenwa A. ayoolaolafenwa/pixellib. GitHub. 2022. URL: https://github.com/ayoolaolafenwa/PixelLib [accessed 2026-09-30]
53. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
54. Games PA, Howell JF. Pairwise multiple comparison procedures with unequal N's and/or variances: a Monte Carlo study. J Educ Stat 1976 Jun;1(2):113-125. [doi: 10.3102/10769986001002113]
55. Per-joint mean-per-joint position error (MPJPE) before and after morphing Mediapipe keypoints to the Plugin-Gait joint center positions. Bio Render. URL: https://app.biorender.com/citation/67a5ebf0c8ec10acf848355f [accessed 2026-10-05]
56. Dermanu/HPEClinicalEvaluationFramework. GitHub. 2025. URL: https://github.com/dermanu/HPEClinicalEvaluationFramework [accessed 2026-09-30]
57. Lee SE, Shibata K, Nonaka S, Nobuhara S, Nishino K. Extrinsic camera calibration from a moving person. IEEE Robot Autom Lett 2022;7(4):10344-10351. [doi: 10.1109/LRA.2022.3192629] [Medline: 36713301]
58. Cai Y, Zhang W, Wu Y, Jin C. FusionFormer: a concise unified feature fusion transformer for 3D pose estimation. In: Wooldridge M, Dy J, Natarajan S, editors. AAAI'24/IAAI'24/EAAI'24: Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence 2024:900-908. [doi: 10.1609/aaai.v38i2.27849]
59. Mercadal-Baudart C, Liu CJ, Farrell G, et al. Exercise quantification from single camera view markerless 3D pose estimation. Heliyon 2024 Mar 30;10(6):e27596. [doi: 10.1016/j.heliyon.2024.e27596] [Medline: 38510055]
60. Abbott E, Campbell A, Wise E, Tidman SJ, Lay BS, Kent P. Physiotherapists could detect changes of 12 degrees or more in single-plane movement when observing forward bending, squat or hand-over-head: a cross-sectional experiment. Musculoskelet Sci Pract 2022 Oct;61:102594. [doi: 10.1016/j.msksp.2022.102594] [Medline: 35667319]

XSL•FO
RenderX

61. della Croce U, Cappozzo A, Kerrigan DC. Pelvis and lower limb anatomical landmark calibration precision and its propagation to bone geometry and joint angles. Med Biol Eng Comput 1999 Mar;37(2):155-161. [doi: 10.1007/BF02513282] [Medline: 10396818]
62. Leboeuf F, Reay J, Jones R, Sangeux M. The effect on conventional gait model kinematics and kinetics of hip joint centre equations in adult healthy gait. J Biomech 2019 Apr 18;87:167-171. [doi: 10.1016/j.jbiomech.2019.02.010] [Medline: 30808499]
63. Kim JW, Choi JY, Ha EJ, Choi JH. Human pose estimation using MediaPipe pose and optimization method based on a humanoid model. Appl Sci 2023;13(4):2700. [doi: 10.3390/app13042700] [Medline: 39449731]
64. Winter DA. Biomechanics and Motor Control of Human Movement, 4th edition: John Wiley & Sons, Inc; 2009. URL: https://onlinelibrary.wiley.com/doi/10.1002/9780470549148 [accessed 2025-01-27]
65. Karashchuk P, Rupp KL, Dickinson ES, et al. Anipose: a toolkit for robust markerless 3D pose estimation. Cell Rep 2021 Sep 28;36(13):109730. [doi: 10.1016/j.celrep.2021.109730] [Medline: 34592148]
66. Thrane G, Sunnerhagen KS, Murphy MA. Upper limb kinematics during the first year after stroke: the stroke arm longitudinal study at the University of Gothenburg (SALGOT). J Neuroeng Rehabil 2020 Dec;17(1):76. [doi: 10.1186/s12984-020-00705-2]
67. Williams G, Mentiplay BF, Kahn MB, Clark RA, Banky M. The effect of walking speed on lower limb angular velocity. Arch Phys Med Rehabil 2018 Oct;99(10):e12-e13. [doi: 10.1016/j.apmr.2018.07.041]
68. Mroz S, Baddour N, McGuirk C, et al. Comparing the quality of human pose estimation with BlazePose or OpenPose. 2021 Dec 8 Presented at: 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART); Dec 8-10, 2021; Paris, France p. 1-4. [doi: 10.1109/BioSMART54244.2021.9677850]
69. van der Kruk E, Reijne MM. Accuracy of human motion capture systems for sport applications; state-of-the-art review. Eur J Sport Sci 2018 Jul;18(6):806-819. [doi: 10.1080/17461391.2018.1463397] [Medline: 29741985]
70. Castro-Luna G, Jiménez-Rodríguez D. Relative and absolute reliability of a motor assessment system using KINECT® camera. Int J Environ Res Public Health 2020 Aug 11;17(16):5807. [doi: 10.3390/ijerph17165807] [Medline: 32796619]
71. Shih MC, Wang RY, Cheng SJ, Yang YR. Effects of a balance-based exergaming intervention using the Kinect sensor on posture stability in individuals with Parkinson's disease: a single-blinded randomized controlled trial. J Neuroeng Rehabil 2016 Dec;13(1):78. [doi: 10.1186/s12984-016-0185-y]
72. Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, Grundmann M. BlazePose: on-device real-time body pose tracking. arXiv. Preprint posted online on Jun 17, 2020. [doi: 10.48550/arXiv.2006.10204]
73. Xing H, Yang J. WPL-Based constraint for 3D human pose estimation from a single depth image. Sensors (Basel) 2022 Nov 22;22(23):9040. [doi: 10.3390/s22239040] [Medline: 36501742]
74. Cimorelli A, Patel A, Karakostas T, Cotton RJ. Validation of portable in-clinic video-based gait analysis for prosthesis users. Sci Rep 2024 Feb 15;14(1):3840. [doi: 10.1038/s41598-024-53217-7] [Medline: 38360820]

Abbreviations

AE: angular error
bl: back left
br: back right
DLT: direct linear transformation
fl: front left
fm: front middle
fr: front right
HPE: human pose estimation
MPJAE: mean-per-joint angular error
MPJAVE: mean-per-joint angular velocity error
MPJPE: mean-per-joint position error
PA-MPJPE: Procrustes aligned mean-per-joint position error
RBC: rank-biserial correlation
ReLU: rectified linear unit
RGB: red, green, blue-depth
RQ: research question
sl: side left
TTL: transistor-transistor logic

XSL•FO
RenderX

Edited by S Munce; submitted 15.Jan.2026; peer-reviewed by D Hussain, D Menychtas; revised version received 15.Jul.2026; accepted 08.Sep.2026; published 09.Nov.2026.

Please cite as:
Lorenz EA, Vereijken B, Bach K, Su X
A Benchmarking Framework for Real-Time 3D Pose Estimation in Motor Rehabilitation: Development and Proof-of-Concept Study
JMIR Rehabil Assist Technol 2026;13:e91360
URL: https://rehab.jmir.org/2026/1/e91360
doi:10.2196/91360

© Emanuel Alexander Lorenz, Beatrix Vereijken, Kerstin Bach, Xiaomeng Su. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 9.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Requirements of the Integrated Home-Based Rehabilitation Tool, Neurorehabilitation Ecosystem for Sustained Therapy: Multicenter Focus Group Study With Stroke Survivors, Caregivers, and Health Care Professionals

Merle Matijsen[1,2], MSc; Cheriel Hofstad[1], MSc; Susana Rodriguez Gonzalez[3,4], MD; Xavier Buxó[3,4], MD; Vera Stara[5], PhD; Sara Leonzi[5], MSc; Margherita Rampioni[5], PhD; Anna Mura[6], PhD; Noël Keijsers[1,2,7], Prof Dr

[1]Department of Research, Sint Maartenskliniek, Hengstdal 3, Nijmegen, The Netherlands

[2]Department of Sensorimotor Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen, Nijmegen, The Netherlands

[3]Department of Physical Medicine and Rehabilitation (PMR), Neurorehabilitation Unit, Vall d'Hebron Hospital Universitari, Barcelona, Spain

[4]PMR Research Group, Vall d'Hebron Institut de Recerca, Barcelona, Spain

[5]Centre for Innovative Models for Aging Care and Technology, National Institute of Health and Science on Aging, IRCCS INRCA, Ancona, Italy

[6]Institute of Neurosciences, University Miguel Hernández-CSIC, Alicante, Spain

[7]Department of Sensorimotor Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands

Corresponding Author:
Merle Matijsen, MSc
Department of Research, Sint Maartenskliniek, Hengstdal 3, Nijmegen, The Netherlands

Abstract

Background: Due to the high burden on health care, home-based rehabilitation (HBR) has gained increasing interest. A new HBR program for stroke survivors, containing a gaming app with upper-limb exercises, monitoring system, and virtual coach was being developed.

Objective: This study aimed to assess the user requirements of an HBR tool including a gaming app, monitoring system, and virtual coach, and to examine potential differences between end users and countries.

Methods: Thirteen stroke survivors, 12 caregivers, and 15 health care professionals from centers in the Netherlands, Italy, and Spain, participated in focus groups or interviews. Each center used the same interview guide with open questions about each component. An inductive thematic analysis was conducted separately at each center, and results were combined during a physical meeting.

Results: User requirements were categorized into three main themes: (1) customization: aligning with individual preferences and capabilities; (2) motivational elements: these included reminders, a variety of levels and games, and ease of use; and (3) feedback elements: maintaining interactions with therapists. These themes apply to both home-based exercises as well as daily-life activities during HBR. There were minor differences between end users or centers.

Conclusions: All end users across the participating countries emphasized the importance of integrating gamified exercises, monitoring, and virtual coaching into an HBR system. The user requirements for such a system can be categorized into three key areas: customization, motivational elements, and feedback elements.

(JMIR Rehabil Assist Technol 2026;13:e79382) doi:10.2196/79382

KEYWORDS

home–based rehabilitation; user requirements; qualitative research; stroke; healthcare professionals; multi–country

Introduction

The economic burden of stroke on health care systems is substantial due to large numbers of people affected by stroke and the high costs associated with acute treatment and long-term care. This treatment is necessary due to the diverse range of poststroke impairments, including reduced upper-limb function [1]. In 2019, over 12.2 million new stroke cases were reported worldwide, marking a 70% increase in stroke incidence since 1990 [2,3]. Due to population aging, this upward trend is expected to continue in the coming years [4], leading to even higher costs [5]. Furthermore, a shortage of medical personnel can arise due to the expected increase in stroke cases. To reduce

XSL•FO
RenderX

costs and personnel, a shift in focus toward home-based rehabilitation (HBR) has recently emerged [6]. HBR offers additional benefits including reduced travel time for stroke survivors, the possibility to increase therapy intensity, increased accessibility allowing users to perform the activities throughout the day, and longer therapy duration [7].

Since higher training intensity leads to greater improvements in most patients [8,9], a critical element for the success of HBR is adherence. However, it is challenging with HBR to achieve adherence consistently, especially given that in general, treatment adherence tends to decrease over extended periods [10]. This decrease is often the result of low acceptability and usability, the inability to stay engaged, or integrate the practice into daily life [10]. Nonetheless, adherence to treatment is crucial for patients with chronic stroke, since behavioral changes in arm use require a substantial training dosage schedule [11]. Adherence rates in HBR for stroke vary significantly, ranging between 13% and 140% [10]. Several factors positively influence the adherence rate in HBR, including rehabilitation programs with game-like training exercises with multiplayer modes, sufficient variability in games, and adjustable difficulty levels. Moreover, the availability of monitoring and feedback options to track users' progress and motivate users also increases adherence [10,12]. Technical support for issues, such as malfunction of the system, and the possibility to customize the system to meet individual users' needs are also important considerations [10,12]. Implementing these factors in HBR can increase adherence and thus improve its overall success.

When considering HBR, multiple end users can be identified, such as stroke survivors, health care professionals, and caregivers. All these end users have a different relation to individuals with stroke. These different relations could result in different perspectives; for example, a health care professional might be more objective, whereas the caregivers and stroke survivors are more subjective. However, Demir et al [13] found that stroke survivors, health care professionals, and caregivers generally share the same priorities during stroke rehabilitation, focusing on improving the health condition post stroke, arm and leg mobility, and increasing independence. Nonetheless, HBR might require other demands compared to usual care. In home settings, caregivers and stroke survivors might prioritize ease of use, whereas health care professionals might focus primarily on health outcomes. As a result, health care professionals might have specific requirements for outcome visualization and monitoring, whereas caregivers are more concerned with the practical challenges of implementing rehabilitation therapy in a home setting. These potential differences underscore the importance of a flexible and adaptive system design that meets the needs of diverse users and should be carefully considered in developing a successful HBR system. However, the different viewpoints of these end users on user requirements for HBR is still unknown.

One advantage of HBR is its potential for scalability and implementation across multiple countries, making it accessible to a diverse range of populations. However, differences in health care systems and cultural practices across Europe must be considered. Across Europe, there are several clusters with comparable health care systems, with Spain, Italy, and the Netherlands belonging to the same cluster [14]. However, Italy and Spain present a higher overlap with each other compared to the Netherlands [14]. Additionally, health care systems in Italy and the Netherlands rely more on inpatient care, while Spain has a stronger outpatient care tradition compared to Italy and the Netherlands [15], suggesting that Spain may have greater experience with home-based exercises. There are also some cultural differences. For example, Italy and Spain are more collectivistic, while the Netherlands is more individualistic [16]. Thus, the Netherlands might be more receptive to HBR, since this is a more individualistic therapy form. Studies have even shown that countries across Europe use different cognitive emotion regulation strategies [17]. These differences, both in health care systems and cultural practices, can potentially influence the requirements of an HBR tool. To accommodate these differences in system design, a more prominent role for the coaches in countries less familiar with HBR and enhance group interaction features where social engagement is valued could be integrated. However, there is currently no information available on these possible differences.

Recently, within an international project involving partners from Italy, Spain, and the Netherlands, a new HBR solution was defined, called the Neurorehabilitation Ecosystem for Sustained Therapy (NEST). As of January 2021, this solution is still in the design phase and has not been developed. This innovative rehabilitation system offers an integrated rehabilitation solution for stroke survivors, featuring the following components: (1) the RGS (Rehabilitation Gaming System) app for smartphones, which offers gamified exercises using augmented reality for upper-limb training, focusing on finger, wrist, elbow and shoulder movements while holding a smartphone; (2) the RGS-wear, a smartwatch-based system for monitoring upper-limb movements during daily activities; and (3) a virtual coach, the agent for well-being assistance (AWA), integrated into both the RGS app and RGS-wear, providing motivational feedback on training performance and daily arm use as well as assistance in the use of the gamified exercises. The combination of these three components, which were found relevant in previous literature [10,18,19], makes NEST a unique HBR tool. However, whether specific user requirements for NEST exist and if the system fulfills these needs remains unclear.

In summary, no specific user requirements are known for a NEST-based HBR system, including optimization of gamified training exercises, monitoring of arm use, and virtual coaching. More insight into user requirements specific to NEST could help to further optimize and develop NEST. In addition, there is a lack of knowledge about its implementation across different end users and multiple countries. It is essential to conduct research with multiple end users and across different countries to gain more insight into user requirements. Therefore, the first aim of this study was to define the user requirements for a newly developed upper-limb training system that includes the three aspects of NEST for the end users: stroke survivors, health care professionals, and caregivers. Our second aim was to identify possible differences in user requirements among these end user groups and across three European countries (The Netherlands, Italy, and Spain).

XSL•FO
RenderX

Methods

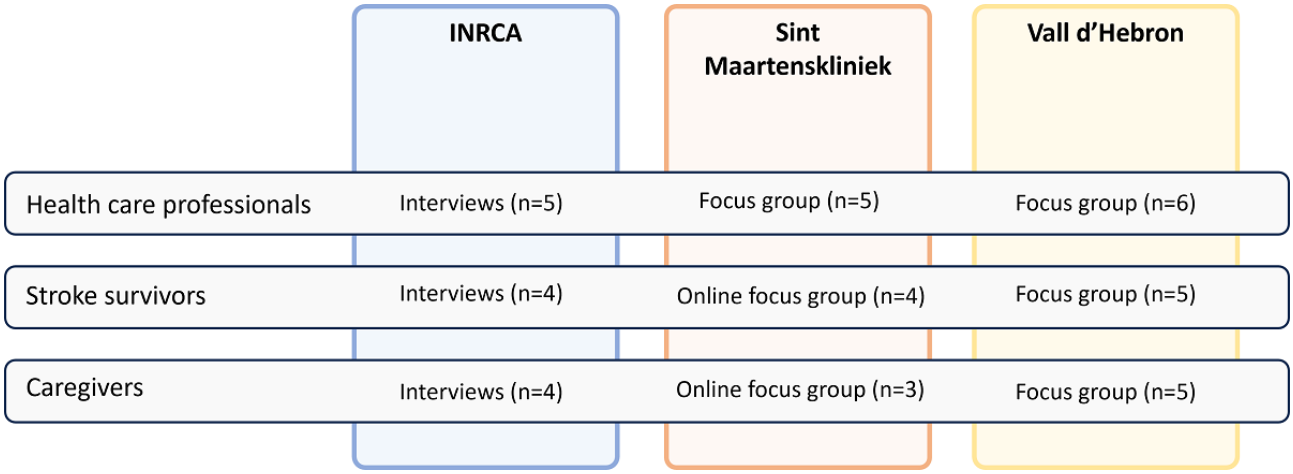
Research Context

The NEST project is a collaborative effort between hospitals and industry partners aimed at developing effective rehabilitation technologies to enhance the recovery of function in patients with stroke. The industrial partners, Eodyne en Hankamp Rehab, are specialized in the development and exploitation of rehabilitation technologies. The clinical partners include Sint Maartenskliniek in the Netherlands, Vall d'Hebron University Hospital in Spain, and Instituto Nazionale di Ricovero e Cura per Anziani in Italy. All clinical centers have significant experience in stroke rehabilitation.

Study Design

A multicenter, qualitative study was conducted at the three clinical sites planned to include three on-site focus groups, segregated by end user. The participants presented the three primary users of the NEST solution: (1) health care professionals, (2) caregivers, and (3) stroke survivors. The end user groups were identified based on their likelihood of being future system users. The planned group size for each group was 5 participants. However, recruiting took place in the beginning of 2022; therefore, difficulties arose due to COVID-19 restrictions. In addition, not all the clinical sites could perform on-site focus groups because of COVID-19 limitations. The Dutch partner performed online focus group sessions with caregivers and stroke survivors and an on-site focus group session with health care professionals. At the Italian clinical rehabilitation center, focus groups were not feasible; hence, all participants were individually interviewed. The Spanish partner conducted all the focus groups on site. An overview of the study design can be found in Figure 1.

Figure 1. Overview of the focus groups, type of focus group and interviews in each center with the number of participants. INRCA: Instituto Nazionale di Ricovero e Cura per Anziani.



Ethical Considerations

All participants gave written consent before the start of the focus group or interview. Medical–ethical approval for this study was not required in Spain and Italy. In the Netherlands, the focus group protocol was submitted to Commissie Mensgebonden Onderzoek Arnhem/Nijmegen, which determined that a full ethical review was not required and granted approval to conduct the study (2022 - 13564). In addition, this study was conducted in accordance with the Declaration of Helsinki of the World Medical Association.

Participants

Health Care Professionals

Health care professionals, such as physiatrists, physical therapists, or occupational therapists, with a minimum experience of 5 years in the rehabilitation of stroke survivors were included.

Stroke Survivors

Potential stroke survivors were identified and contacted by clinicians at the three centers using electronic medical records. Interested participants were then assessed for eligibility based on inclusion and exclusion criteria.

The inclusion criterion was being an adult stroke survivors (18 years or older) with a paretic arm. Exclusion criteria were individuals with aphasia and/or cognitive impairment (Mini Mental Examination score of <21), those with minimal smartphone proficiency, and those affected by other neurological disorders.

Caregivers

Informal caregivers were recruited via the stroke survivors who participated in the study. In addition, social media, such as Facebook, was used to recruit caregivers. Across all three countries, informal caregivers were consistently defined as individuals who provide unpaid assistance and support to stroke survivors who cannot fully care for themselves independently. Caregivers could participate if they had been or currently were taking care of a stroke survivor with a poststroke impaired arm. They were excluded if they had minimal smartphone proficiency. If caregivers showed interest, they were further assessed to determine if they met the inclusion and exclusion criteria.

Structure of the Focus Groups and Data Collection

Focus groups and interviews were conducted by a native language–speaking interviewer accompanied by an assistant

XSL•FO
RenderX

who took notes. Both were experienced researchers at the respective centers. Each focus group lasted between 1 and 2 hours, with breaks available for stroke survivors if needed. Interviews lasted approximately 30 minutes. All three clinical centers used the same English interview guide, which was translated into the local language for consistency across focus groups and interviews. Questions were tailored to the perspective of the participant group. The detailed English interview guide for stroke survivors can be found in Multimedia Appendix 1, with additional details about the interview guide for caregivers and health care professionals. The interviewer started with an introduction of the NEST system, including its three components (the RGS app, RGS-wear, and the AWA coach), followed by brief videos illustrating how NEST operates. In addition, the researcher explained that the videos showed a prototype version of NEST and that the final components were not physically available yet. They also mentioned that the user requirements obtained with the study are necessary for further advanced development of NEST. Participants were then asked for permission to audio record the session. Participants were questioned on three central topics: HBR in general, participants' perspective on the three components of NEST, and their general opinion about NEST. While user requirements for system design were explored, the focus groups also allowed discussion of participants' preferences, expectations, and general opinions. At the start of each topic, the interviewer posed an open question, for example, "What do you need in order to be able to use the RGS–mobile app in daily practice?" to initiate an open discussion. To stimulate further discussion, more specific questions were asked, such as, "What do you think of the different exercises and games?" If both the participants and researcher felt all points were discussed, the next topic was introduced. After the focus group or interview, participants received a summary to check if they had any additional remarks.

Data Analysis

Each clinical site independently transcribed all recorded audio tapes of the interviews and focus groups in full. Manual transcription was conducted at each site, and analysis was performed using Atlas.ti software (Lumivero). An inductive content analysis approach was carried out, in which an analytical framework is created in multiple steps: familiarization, coding, identifying themes, reviewing themes, and defining themes [20,21]. First, each transcript was read in full by the researcher to become familiar with the content. Second, the researcher assigned codes that closely reflected the participants' wording, called "open coding" [22]. Third, the initial codes were refined and organized into categories (eg, multiplayer mode and technical skills) [23]. Lastly, two researchers at each site discussed categories to create concept and initial themes until consensus on the themes was reached. In instances where consensus was not achieved, a third researcher was consulted to make the final decision.

After the three separate inductive thematic analyses at each clinical center, two researchers from the Spanish site, one from the Dutch site, and two from the Italian site met in person to integrate their findings. The objective of this meeting was to identify user requirements across sites. The method of Van der Elst et al [24] was used for combining results. First, each site summarized and presented their site-specific factors to the others. Clarifying questions were asked and additional contextual information was provided where needed to ensure a shared understanding of the findings. Second, all sites prepared sticky notes with their analytic categories, each site using a different color. During an in-person collaborative mapping meeting, comparable categories from different sites were grouped to identify similarities and differences, forming the initial synthesis of site-level findings into broader cross-site patterns. Third, all researchers engaged in a group discussion to interpret the similarities and to develop higher-order themes reflecting shared user requirements. Differences between sites were discussed explicitly and examined in relation to contextual factors, rather than being treated as discrepancies to be resolved. Finally, themes were refined and prioritized based on their relevance to the study objectives and their significance across sites. When divergent interpretations occurred, these were retained as country-specific findings.

The sticky notes also contained information on which end user group discussed each category, allowing for the identification of similarities and differences between end user groups. All steps were repeated for each topic: general HBR, the RGS app, RGS-wear, and the AWA coach. Lastly, all factors were categorized into essential requirements, desirable features, and contextual factors by the first author; this classification was reviewed and approved by the other authors.

Results

Overview

In total 13 stroke survivors, 12 caregivers, and 15 health care professionals participated in the study. Participants were distributed into six focus groups and 13 interviews. Basic characteristics of the participants are shown in Table 1. There are no differences in characteristics between the participants of the three clinical sites, except for a longer time since stroke in stroke survivors at the Dutch site.

An overview of the essential requirements, desirable features and contextual factors can be found in Table 2. All participants expressed enthusiasm for NEST, highlighting the relevance of the three components, and indicated a potential interest in testing it.

XSL•FO
RenderX

Table . Characteristics of participants in the focus groups and interviews.

	INRCA			VH			SM		
	SS[a]	CG[b]	HP[c]	SS	CG	HP	SS	CG	HP
Participants, n	4	4	5	5	5	5	4	3	5
Sex (male), n	2	1	1	4	1	2	4	1	2
Age (years), median (range)	67 (61-71)	53 (45-68)	35 (31-55)	63 (58-67)	47 (25-63)	43 (28-56)	57 (44-58)	61 (29-80)	45 (35-60)
Time since stroke (months), median (range)	2 (1-3)	2 (1-3)[a]	—[d]	3 (1-5)	3 (1-5)[a]	—	72 (8-173)	6 (5-78)[a]	—
Experience (years), median (range)	—	—	10 (4-30)	—	—	8 (3-29)	—	—	19 (8-28)

[a]SS: stroke survivor.

[b]CG: caregiver.

[c]HP: health care professional.

[d]Not applicable.

XSL•FO
RenderX

Table . Essential requirements, desirable requirements, and contextual factors of NEST[a].

Components	Factors
Essential requirements	
RGS[b] app	• Variety of levels and games
RGS app and RGS-wear	• Simple and intuitive • Technical support • Privacy concerns • RGS app without RGS-wear • Initial training
RGS-wear and AWA[c] coach	• Quality of movements
AWA coach	• Goal-oriented coach • Positive attitude • Reminders[d] • Feedback[e]
Desirable features	
RGS app	• Multiplayer mode • Adding more functionalities • Adjust to preferences
RGS-wear	• Circumstance stroke survivor
AWA coach	• Feedback by stroke survivor • Emotional slider • Frequency reminders • Type of signal
Contextual factors	
General	• Additional to on-site therapy • Timing home-based training • Role caregiver[f]
RGS app	• Level of lesion
RGS app and RGS-wear	• Monitoring[g] • Basic technical skills
RGS-wear	• Self-responsibility

[a]NEST: Neurorehabilitation Ecosystem for Sustained Therapy.

[b]RGS: Rehabilitation Gaming System.

[c]AWA: agent for well-being assistance.

[d]On both games and daily tasks.

[e]During and after gameplay and from the virtual coach and therapist.

[f]Either supportive or limited.

[g]Involving the (local) therapist.

Similarities in User Requirements Across End User Groups and Countries

The RGS App

Several factors for the RGS app were collectively identified by the three clinical sites, categorized into four main themes: game requirements, technical requirements, role of the therapist, and customization. Game requirements included a need for a variety of gamified exercises and difficulty levels within games. A multiplayer mode was also discussed, which could potentially boost motivation but might also lead to demotivation. To prevent unfair competition resulting from differences in upper-limb function, customization is essential in multiplayer mode, which should be designed to ensure fairness and minimize user burden. This concern is shown in the following quote:

> *A collaboration is bothersome if something is dependent on you … how would that work? If you do something together and you fail, and the whole team*

XSL•FO
RenderX

> *fails. I would not like that.* [Stroke survivor, the Netherlands]

The technical requirements underscored that the RGS app must be simple and intuitive to use in order for them to stay motivated. However, although a basic level of technical skills is still necessary for stroke survivors. Additionally, technical support must be available to address any issues that arise. The importance of the therapist's role during the use of NEST was emphasized. The therapist should provide clear instructions before patients take the RGS app home. Initial sessions in the outpatient clinic introducing the system were suggested so that the therapist could define both the goals and the treatment pathway best suited to the patient. In addition, NEST should provide continuous supervision and monitoring during use. In this way, the therapist would ensure support, motivation, and movement corrections. Such monitoring is particularly important in the early stages of HBR and may gradually decrease over time. The following was said during an interview:

> *The therapist should motivate the patients and control that the exercises are done in the most appropriate way.* [Stroke survivor, Italy]

The factor customization is related to the personalization of the RGS app to suit individual stroke survivors. Depending on the location of the lesion and capabilities of the stroke survivor, specific games should be selectable or not. Furthermore, the ability to adjust games to personal preferences, such as displaying the digital hand on the paretic side, modifying difficulty levels, and tailoring games to the users' hobby was highlighted. Also, the possibility to add more functionalities, such as balance games, were mentioned.

RGS-Wear

The factors concerning RGS-wear were categorized into three main topics: monitor requirements, technical requirements, and customization. With monitoring, the smartwatch should be capable of measuring both the quality and quantity of movements to gain insight into whether compensation strategies were used during both games and daily life. For example, if a stroke survivor focuses on reaching movements, the RGS-wear should be able to register a reaching movement. These measurements should be saved and analyzed periodically by the therapist. Analysis of arm use in daily life also enables the therapist to further specify and tailor interventions, which was also described by a therapist:

> *Upper limb movement restoration during stroke recovery is a particularly important pillar of rehabilitation practice.* [Health care professional, Spain]

The technical requirements of RGS-wear align with those of the RGS app: it needs to be simple and intuitive to use, although a basic level of technical skills is still necessary. Participants indicated that if the smartwatch did not function properly, it would lead to demotivation in using the system. Additionally, a clear instruction on how to use the smartwatch must be provided. The desire to use the RGS app independently of RGS-wear was also discussed.

> *It needs to be user friendly, I think. Especially since people with a stroke are often a bit older, because the smartwatches I know are for example quite small.* [Caregiver, the Netherlands]

Customization options for RGS-wear should take into consideration the user's specific circumstances, such as use of a sling and esthetics. Esthetics refer to aspects such as the size, design, and color of the smartwatch. Some participants, for example, preferred smaller, more traditional designs over large, modern-looking smartwatches.

AWA Coach

For the AWA coach, the factors were categorized into three main themes: reminder requirements, feedback requirements, and customization. Reminders should not only prompt the stroke survivor to start the games but also encourage the use of the paretic arm during daily tasks. These reminders should be presented with a positive attitude. The importance of reminders was highlighted:

> *Coach reminders could become an important part of the rehabilitation in some of the post–stroke patients in order to improve their motivation and adherence to treatment. These reminders should also be optional and tailored to user's specific needs.* [Health care professional, Spain]

No specific interest in an avatar was expressed across end user groups. Feedback on upper-limb performance should be provided during gameplay, after gameplay, and during daily activities for which RGS-wear data can be used. Personalized feedback from the therapist would also be very helpful and motivating. In addition, also some patients would like to send feedback to the AWA coach in case of necessity. For example, a health care professional mentioned:

> *Games offered in home rehabilitation should include a feedback system for the patient to check his progress (eg, game score, growing tree, graph etc). There should be weekly feedback showing daily activities done, use of the arm, etc. A notification should be sent to the therapist if the patient is playing or exercising less than expected.* [Health care professional, Italy]

The frequency of reminders and feedback on daily life use of the paretic arm should be tailored to each user's preferences. Cognitive ability could play a role, since some participants may simply forget to perform the exercises, making reminders more important. Feedback should be goal-oriented and therefore tailored to the user. The type of signal for reminders should be customized to the stroke survivor. However, no definitive consensus on the preferred signal type emerged during the focus groups and interviews. In general, verbal and written reminders seemed to be fine.

Differences Between End User Groups

An overview of similarities and differences across end user groups and countries for each topic is shown in Table 3. Overall, most user requirements focused on the needs of stroke survivors. Also, caregivers and health care professionals primarily considered requirements from the perspective of stroke

XSL•FO
RenderX

survivors, placing less emphasis on their own requirements for system design. Overall, the end user groups were largely in consensus regarding the user requirements of NEST. However, a few topics were not addressed by all groups, with caregivers not addressing all topics mentioned by stroke survivors and health care professionals. First, caregivers did not share opinions about the early introduction of the system, possibly already in the clinic. They also did not address the ability of the smartwatch to measure the quality of movements. Lastly, caregivers did not express a preference for goal-oriented feedback and the method of delivering feedback. In contrast, caregivers showed interest in having access to the data collected by RGS-wear, a point not mentioned by stroke survivors and health care professionals. Stroke survivors mentioned the possible negative impact of caregivers on motivation if caregivers push the stroke survivors too much at home. Therapists focused on the upper-limb function of stroke survivors and which patients should be able to use the system.

XSL•FO
RenderX

Table . Overview of similar and different factors among the 3 sites. The dashed line means that the topic was mentioned by health care professionals and stroke survivors. The solid line means that the topic was only mentioned by health care professionals. The absence of a line means that the topic was mentioned by all end users and therefore there was consensus.

Topics	Italy	Netherlands	Spain
General home–based			
Integrating requirements			
HBR[a] as additional intervention			
Getting familiar			
Monitoring by therapist			
Involve local therapist			
Motivational requirements			
Motivational components			
Timing home–based training			
Supportive caregiver			
Limited role caregiver			
RGS[b] app			
Game requirements			
Variety of levels and games			
Multiplayer mode			
Adding more functionalities			
Technical requirements			
Simple and intuitive			
Basic technical skills			
Technical support			
Privacy concerns			
Role of the therapist			
Supervision and monitoring			
Initial sessions			
Customization			
Adept to level of lesion			

XSL•FO
RenderX

Topics			Italy	Netherlands	Spain
		Adjust to preferences			
RGS-wear					
	Monitor requirements				
		Quality of movements			
		Analyzed by therapist			
		Self–responsibility stroke survivor			
	Technical requirements				
		Simple and intuitive			
		Basic technical skills			
		Clear instructions			
		Privacy concerns			
	Customization				
		RGS app without RGS–wear			
		Circumstances stroke survivor			
AWA[c] coach					
	Reminder requirements				
		Positive attitude			
		Reminder for doing games			
		Reminder for daily tasks			
		No interest in avatar			
	Feedback requirements				
		Feedback during game			
		Feedback after game			
		Feedback during daily tasks			
		Feedback from therapist			
		Send own feedback to coach			
		Emotional slider			

XSL•FO
RenderX

Topics		Italy	Netherlands	Spain
Customization				
	Frequency reminders			
	Type of signal			
	Goal–oriented coach			

[a]HBR: home-based rehabilitation.

[b]RGS: Rehabilitation Gaming System.

[c]AWA: agent for well-being assistance.

Although caregivers and health care professionals primarily prioritized the needs of stroke survivors, certain user-specific requirements emerged. Caregivers emphasized that NEST should function without their help, as they should not be responsible for monitoring or motivating its use. Given their caregiving responsibilities, additional tasks could increase their burden and shift their role from a supportive partner to that of a therapist or even nurser. Instead, they preferred the option to track progress when interested or when they considered it was necessary to intervene. Additionally, caregivers expressed interest in playing games together with stroke survivors, enabling active participation rather than solely in a motivational capacity role.

Health care professionals considered whether NEST would reduce or potentially increase their workload. They noted that stroke survivors often require assistance with exercises, and if therapists still need to provide coaching and feedback on their training, their burden could even rise. Collaborative management of NEST with local therapists could reduce the burden. However, they emphasized that NEST could support stroke survivors' vitality, which remains the most important. Additionally, therapists highlighted that smartwatch data could offer valuable insights into motivation strategies, which they can use during on-site therapy, resulting in increased quality of the therapy.

Differences Between Countries

All countries showed high overlap in topics discussed. Some differences between sites were observed, which may be related to differences in patient characteristics across countries, such as the higher proportion of patients with chronic stroke at the Dutch site. However, a few differences are worth noting. At the Dutch site, therapists and stroke survivors highlighted the beneficial effect of involving local therapists in NEST. Additionally, Dutch stroke survivors mentioned that caregivers could negatively impact motivation due to the already high burden of caregiving tasks. Privacy concerns regarding the use of the RGS app and smartwatch were raised both at the Spanish and Dutch sites. Both sites also mentioned the importance of making the AWA coach goal-oriented and enabling stroke survivors to provide feedback on their own performance. The Italian site shared similar views, highlighting the importance of awareness and self-responsibility among stroke survivors. They noted that the smartwatch's tracking of daily arm use could increase the stroke survivors' awareness of their functionality, thereby empowering them. Unique to the Spanish site was the suggestion to expand the RGS app to include mobility and balance exercises.

General Feedback on HBR

In addition to the specific requirements for NEST, this study identified broader user requirements for HBR. The factors discussed about general HBR were categorized into two main themes: integration requirements and motivational requirements. Integration requirements refer to incorporating HBR into the overall rehabilitation program as an additional intervention rather than a replacement for on-site rehabilitation. Key aspects include providing adequate instructions for using the HBR system, getting familiar with the HBR system during rehabilitation before starting the intervention at home, and ensuring ongoing monitoring by the therapist during home use. In addition, the timing for initiating the HBR training period should be personalized for each stroke survivor. The motivational requirements emphasize the importance of incorporating elements that enhance motivation with the HBR system. The caregiver must have a supportive role throughout the training period, helping to sustain engagement and adherence to the program.

Discussion

The first objective of this study was to identify the user's needs of a newly developed upper-limb training system that includes gamified exercises, monitoring, and a virtual coach. This was investigated with focus groups and interviews with stroke survivors, caregivers, and health care professionals in clinical centers in Italy, Spain, and the Netherlands. Participants across all user groups and countries highlighted the relevance of each component of NEST, the RGS app, RGS-wear, and the AWA coach. The participants recognized the potential of NEST, including these elements, as an HBR tool. In order for this combination to be successful, motivational elements such as gaming, feedback, and reminders are important. In addition, during the focus group sessions and interviews, several improvements were suggested, mainly concerning adding motivational components and enhancing system customization. If these suggestions are integrated, the system has the potential to maintain high adherence and effectively train upper-limb function. In addition, these aspects may help to make HBR more

XSL•FO
RenderX

goal-oriented, potentially enhancing stroke survivors' awareness of arm use in daily life.

In line with previous studies, this study highlights the need for HBR systems to have three demands in particular: (1) customization to emphasize the need to tailor HBR technology to individual patients [10,19,25-27]; (2) motivational elements, including features such as reminders [19], feedback [10,12,18,19,25-28], a variety of games or levels [7,10,12,19,26,29,30], and the ease of use [7,10,19]; and (3) active engagement of the therapist to ensure proper execution, progress monitoring, and provision of feedback [7,10,19,25,28-30]. These general recommendations for HBR should apply not only during exercises but also in daily life. HBR should be customized to support daily life goals, motivate patients to use the affected upper limb in daily activities, and provide insights into daily-life patterns that are useful for therapists. Since the primary goal of upper-limb rehabilitation is to promote active use of the affected limb in everyday life, these requirements and factors help facilitate the transfer of rehabilitation gains to daily life.

Several studies have underscored the importance of active engagement of the therapist during HBR (ie, active monitoring) to enhance patient adherence and progress [7,10,28-30]. Additionally, HBR can highlight specific goals and therefore provide useful insights for therapy [31]. This study investigated the presence of a virtual coach in detail, highlighting the importance of its interaction with the patient. It can help the patient with daily structure by using reminders and feedback [32]. On a more detailed level, this study found that measuring both the quantity and quality of movements during daily life and exercises and providing goal-based feedback are crucial, as this may help to control compensatory movements and provide input for therapists to further tailor rehabilitation. Only two studies investigated user requirements for home monitoring in detail [19,33]. However, Langerak et al [19] investigated home monitoring using sensors and Cavuoto et al [33] used a smartphone, whereas this study focused on the use of a smartwatch. Still, all three studies revealed similar results. This indicates the need for careful consideration of monitoring methods and the type of feedback provided to stroke survivors.

This was also the first study that took the perspectives of various end users on the user requirements of NEST into account. Previous studies mostly focused on stroke survivors or held focus group sessions or interviews with the stroke survivor and caregiver together, indicating the need for insights from therapists and caregivers independently, free from influence of the stroke survivor [10,25]. Studies that have included stroke survivors, caregivers, and health care professionals have generally found minimal differences in user requirements between these groups [7,34]. However, despite the significant overlap in user requirements among end users, some minor differences were noted. This suggests the need to pay attention to these differences when further developing the NEST system, as tailoring the systems to accommodate these nuances could enhance its effectiveness and user satisfaction. Caregivers did not express opinions on the early introduction of the system, quality measurement, or goal-oriented feedback. Instead, they emphasized the importance of maintaining their role as a partner rather than a therapist. The absence of these therapeutic considerations from the caregivers' answers suggests a more social perspective of their role, most likely because they are not directly involved in the therapy process. Consequently, caregivers may view themselves primarily as supporters rather than as system users. Previous reviews have showed that social support positively influences both the quality of life and physical recovery of the stroke survivor, underscoring the need to consider caregivers in HBR [35,36]. Interestingly, the user requirements mentioned by stroke survivors were also mentioned by therapists and caregivers, suggesting that these groups have a good understanding of the stroke survivors' needs.

In line with the high overlap between end users, there was a strong consensus on user requirements across countries. A previous review showed an agreement across 15 countries on the importance of place, relationships, and therapy [25], though it only focused on HBR without incorporating technology. In addition, a study comparing HBR with technology between Bangladesh and the United Kingdom found similar user requirements but noted more sociotechnical challenges in Bangladesh [37]. In this study, only subtle differences in addressed topics were noted, including privacy concerns raised at the Spanish site, negative caregiver influence at the Dutch site, and self-awareness at the Italian site. The minor differences may reflect the inclusion of sites exclusively from Western countries, where major cultural variation is less likely. Although only small, these differences should be taken into account when further developing the NEST system. Specific social aspects, such as social stigma and relation dynamics among users, were not explicitly addressed nor raised by end users during the focus groups in any country. However, these topics might be more country-specific and could contribute to national differences. The variations in this study may stem from the distinct participant characteristics between the countries. Dutch participants in the chronic phase of stroke recovery likely offered more comprehensive insights into experiences with caregivers at home, unlike the acute stroke survivors at the Italian and Spanish sites. The literature supports that stroke over time affects the experience of being at home, for example, a change of feeling in bodily routines and relationships [38]. In addition, the high motivation observed among Italian stroke survivors may have highlighted the self-awareness. These findings indicate that clinical characteristics, home environments, and coping mechanisms influence the perceptions of HBR by stroke survivors.

One limitation of this study was the predominance of male stroke survivors and the age imbalance among participants across different clinical sites. This could have influenced the needs of the stroke survivors for NEST, since age and sex can influence the severity of stroke [39,40]. Also, women more often live alone after hospital discharge [39]. Additionally, the inclusion criteria for this study, including minimal cognitive impairments, no aphasia, and smartphone usage, likely do not represent the entire stroke survivor population. These inclusion criteria may have contributed to a more positive attitude toward NEST. However, despite the imbalance, our results are aligned with those of previous research, and there was a high degree of overlap in findings across all countries. Another limitation was

XSL•FO
RenderX

the participants' difficulty in conceptualizing NEST, which was still in development. To address this, videos of the RGS app were shown, and when participants asked for more clarity, they were asked to describe their ideal scenario. This approach often led to open discussions about potential features and improvements. The study also faced limitations due to the use of varied interview formats (focus groups, online focus groups, and interviews) imposed by the constraints of the COVID-19 pandemic. Interviews may generate a broader range of items, potentially enhancing depth, whereas focus groups may elicit more personal information due to a peer environment [41]. Furthermore, the online focus groups may have had constrained discussion due to the less natural setting. Nevertheless, previous research showed that content generated in either interviews or focus groups have a high overlap [41], and the same applies to physical or online focus groups [42-44]. Furthermore, the use of identical interview guides ensured consistency in discussion topics. The high degree of overlap in results suggests that the same questions were effectively addressed across the countries. Therefore, this aspect was not considered further in the data analysis. Lastly, no specific questions on between-country or end user differences, such as culture or care pathways, were included, as the focus remained on the primary aim, including user requirements for system development. These topics could have highlighted potential differences; yet, they were not raised by participants, suggesting limited perceived relevance. Therefore, more research is needed on this topic.

In conclusion, all end users across the participating countries emphasized the importance of integrating gamified exercises, monitoring, and a virtual coaching component into an HBR system. The user requirements for such a system can be categorized into three key areas, customization, motivational elements (including ease of use), and feedback elements, which are essential for maximizing user engagement and effectiveness. These areas extend beyond HBR exercises and are also important for daily-life activities. Users' needs found in this study are largely similar; however, subtle differences between end users (eg, social perspective caregivers) and western countries (caregiver burden, privacy concerns, and self-awareness) should be considered when further developing the NEST system.

Acknowledgments

We would like to thank all stroke survivors, caregivers, and health care professionals who participated in the study. An additional thanks to Marco Galli for his help with data analysis from the Spanish site.

Funding

This work was supported by the Active Assisted Living program (AAL–2020–7–227–CP) (https://www.aal-europe.eu/projects/nest/). NK is supported by the following grants: ZonMw–TopZorg (TZO,10070022010004), Interreg NWE–RE:HOME, Eurostars–REACTION, interreg Deutschland–Nederland [Lifehelper], and EFRO OP Oost Myostride.

Data Availability

The data that support the findings of this study are available from the corresponding author, MM, upon reasonable request.

Authors' Contributions

Conceptualization: AM, CH, MM, MR, NK, SL, SRG, VS, XB
Formal analysis: MM, CH, SRG, VS, SL, MR
Investigation: MM, CH, SRG, VS, SL, MR
Writing–original draft: MM
Writing–review and editing: AM, CH, MR, NK, SL, SRG, VS, XB

Conflicts of Interest

AM is a scientific advisor at Eodyne Systems S.L., dedicated to bringing science-grounded health solutions to society.

Multimedia Appendix 1
Interview guide used during the focus group sessions and interviews.
[PDF File, 201 KB - rehab_v13i1e79382_app1.pdf]

References

1. Raghavan P. Upper limb motor impairment after stroke. Phys Med Rehabil Clin N Am 2015 Nov;26(4):599-610. [doi: 10.1016/j.pmr.2015.06.008] [Medline: 26522900]
2. Béjot Y, Bailly H, Graber M, et al. Impact of the ageing population on the burden of stroke: The Dijon Stroke Registry. Neuroepidemiology 2019;52(1-2):78-85. [doi: 10.1159/000492820] [Medline: 30602168]
3. Feigin VL, Brainin M, Norrving B, et al. World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. Int J Stroke 2022 Jan;17(1):18-29. [doi: 10.1177/17474930211065917] [Medline: 34986727]

XSL•FO
RenderX

4. Wafa HA, Wolfe CDA, Emmett E, Roth GA, Johnson CO, Wang Y. Burden of stroke in Europe: Thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years. Stroke 2020 Aug;51(8):2418-2427. [doi: 10.1161/STROKEAHA.120.029606] [Medline: 32646325]
5. Luengo-Fernandez R, Violato M, Candio P, Leal J. Economic burden of stroke across Europe: A population-based cost analysis. Eur Stroke J 2020 Mar;5(1):17-25. [doi: 10.1177/2396987319883160] [Medline: 32232166]
6. Zheng J, Hou M, Liu L, Wang X. Knowledge structure and emerging trends of telerehabilitation in recent 20 years: A bibliometric analysis via CiteSpace. Front Public Health 2022;10:904855. [doi: 10.3389/fpubh.2022.904855] [Medline: 35795695]
7. Brouns B, Meesters JJL, Wentink MM, et al. Why the uptake of eRehabilitation programs in stroke care is so difficult-a focus group study in the Netherlands. Implement Sci 2018 Oct 29;13(1):133. [doi: 10.1186/s13012-018-0827-5] [Medline: 30373611]
8. Lang CE, Lohse KR, Birkenmeier RL. Dose and timing in neurorehabilitation: prescribing motor therapy after stroke. Curr Opin Neurol 2015 Dec;28(6):549-555. [doi: 10.1097/WCO.0000000000000256] [Medline: 26402404]
9. Lohse KR, Lang CE, Boyd LA. Is more better? Using metadata to explore dose-response relationships in stroke rehabilitation. Stroke 2014 Jul;45(7):2053-2058. [doi: 10.1161/STROKEAHA.114.004695] [Medline: 24867924]
10. Neibling BA, Jackson SM, Hayward KS, Barker RN. Perseverance with technology-facilitated home-based upper limb practice after stroke: A systematic mixed studies review. J Neuroeng Rehabil 2021 Feb 24;18(1):43. [doi: 10.1186/s12984-021-00819-1] [Medline: 33627126]
11. Ward SA, Parikh S, Workman B. Health perspectives: International epidemiology of ageing. Best Pract Res Clin Anaesthesiol 2011 Sep;25(3):305-317. [doi: 10.1016/j.bpa.2011.05.002] [Medline: 21925398]
12. Chen Y, Abel KT, Janecek JT, Chen Y, Zheng K, Cramer SC. Home-based technologies for stroke rehabilitation: A systematic review. Int J Med Inform 2019 Mar;123:11-22. [doi: 10.1016/j.ijmedinf.2018.12.001] [Medline: 30654899]
13. Demir YP, Balci N, Ünlüer N, et al. Three different points of view in stroke rehabilitation: Patient, caregiver, and physiotherapist. Top Stroke Rehabil 2015 Oct;22(5):377-385. [doi: 10.1179/1074935714Z.0000000042] [Medline: 25823910]
14. Reibling N, Ariaans M, Wendt C. Worlds of healthcare: A healthcare system typology of OECD countries. Health Policy 2019 Jul;123(7):611-620. [doi: 10.1016/j.healthpol.2019.05.001] [Medline: 31133444]
15. Wendt C. Changing healthcare system types. Soc Policy Adm 2014 Dec;48(7):864-882. [doi: 10.1111/spol.12061]
16. De Cristofaro V, Pellegrini V, Van Zomeren M. The relation between system justification and collective action in individualistic versus collectivistic European countries. Test Psychom Methodol Appl Psychol 2022;29(1):55-70. [doi: 10.4473/TPM29.1.5]
17. Potthoff S, Garnefski N, Miklósi M, et al. Cognitive emotion regulation and psychopathology across cultures: A comparison between six European countries. Pers Individ Dif 2016 Aug;98:218-224. [doi: 10.1016/j.paid.2016.04.022]
18. Emmerson KB, Harding KE, Lockwood KJ, Taylor NF. Home exercise programs supported by video and automated reminders for patients with stroke: A qualitative analysis. Aust Occup Ther J 2018 Jun;65(3):187-197. [doi: 10.1111/1440-1630.12461] [Medline: 29527686]
19. Langerak AJ, Regterschot GRH, Selles RW, et al. Requirements for home-based upper extremity rehabilitation using wearable motion sensors for stroke patients: A user-centred approach. Disabil Rehabil Assist Technol 2024 May;19(4):1392-1404. [doi: 10.1080/17483107.2023.2183993] [Medline: 36905631]
20. Busetto L, Wick W, Gumbinger C. How to use and assess qualitative research methods. Neurol Res Pract 2020;2(14):14. [doi: 10.1186/s42466-020-00059-z] [Medline: 33324920]
21. Ritchie J, Lewis J, McNaughton Nicholls C, Ormston R. Qualitative Research Practice: A Guide for Social Science Students and Researchers, 2nd edition: SAGE Publications; 2014.
22. Punch K. Introduction to Social Research: Quantitative and Qualitative Approaches, 3rd edition: SAGE Publications; 2014.
23. Fossey E, Harvey C, McDermott F, Davidson L. Understanding and evaluating qualitative research. Aust N Z J Psychiatry 2002 Dec;36(6):717-732. [doi: 10.1046/j.1440-1614.2002.01100.x] [Medline: 12406114]
24. Van der Elst K, Bremander A, De Groef A, et al. European Qualitative research project on Patient-preferred outcomes in Early Rheumatoid Arthritis (EQPERA): Rationale, design and methods of a multinational, multicentre, multilingual, longitudinal qualitative study. BMJ Open 2019 Mar 27;9(3):e023606. [doi: 10.1136/bmjopen-2018-023606] [Medline: 30918031]
25. Lavis H, van Vliet P, Tavener M. Stroke survivor, caregiver and therapist experiences of home-based stroke rehabilitation: A thematic synthesis of qualitative studies. Phys Ther Rev 2023 Mar 4;28(2):157-173. [doi: 10.1080/10833196.2023.2180710]
26. Sivan M, Gallagher J, Holt R, Weightman A, O'Connor R, Levesley M. Employing the International Classification of Functioning, Disability and Health framework to capture user feedback in the design and testing stage of development of home-based arm rehabilitation technology. Assist Technol 2016;28(3):175-182. [doi: 10.1080/10400435.2016.1140689] [Medline: 26852630]
27. Standen PJ, Threapleton K, Connell L, et al. Patients' use of a home-based virtual reality system to provide rehabilitation of the upper limb following stroke. Phys Ther 2015 Mar;95(3):350-359. [doi: 10.2522/ptj.20130564] [Medline: 25212521]

XSL•FO
RenderX

28. Wingham J, Adie K, Turner D, Schofield C, Pritchard C. Participant and caregiver experience of the Nintendo Wii Sports™ after stroke: qualitative study of the trial of Wii™ in stroke (TWIST). Clin Rehabil 2015 Mar;29(3):295-305. [doi: 10.1177/0269215514542638] [Medline: 25125442]
29. Palmcrantz S, Borg J, Sommerfeld D, et al. An interactive distance solution for stroke rehabilitation in the home setting - a feasibility study. Inform Health Soc Care 2017 Sep;42(3):303-320. [doi: 10.1080/17538157.2016.1253015] [Medline: 27918220]
30. Szturm T, Imran Z, Pooyania S, Kanitkar A, Mahana B. Evaluation of a game based tele rehabilitation platform for in-home therapy of hand-arm function post stroke: Feasibility study. PM R 2021 Jan;13(1):45-54. [doi: 10.1002/pmrj.12354] [Medline: 32107868]
31. Reunanen MAT, Järvikoski A, Talvitie U, Pyöriä O, Härkäpää K. Individualised home-based rehabilitation after stroke in eastern Finland--the client's perspective. Health Soc Care Community 2016 Jan;24(1):77-85. [doi: 10.1111/hsc.12190] [Medline: 25676161]
32. Kropf J, Hungerländer NA, Gand K, Schlieter H. Transforming rehabilitation to virtually supported care - the vCare project. Presented at: ICMI '20 Companion: Companion Publication of the 2020 International Conference on Multimodal Interaction; Oct 25-29, 2020; Virtual (Netherlands) p. 261-264. [doi: 10.1145/3395035.3425313]
33. Cavuoto LA, Subryan H, Stafford M, et al. Understanding user requirements for the design of a home-based stroke rehabilitation system. Proc Hum Factors Ergon Soc Annu Meet 2018 Sep;62(1):1037-1041. [doi: 10.1177/1541931218621239] [Medline: 35465216]
34. Sauerzopf L, Luft A, Maeusli V, Klamroth-Marganska V, Sy M, Spiess MR. Technology use for home-based stroke rehabilitation in switzerland from the perspectives of persons living with stroke, informal caregivers, and therapists: Qualitative interview and focus group study. JMIR Rehabil Assist Technol 2024 Jul 18;11:e59781. [doi: 10.2196/59781] [Medline: 39024576]
35. Kruithof WJ, van Mierlo ML, Visser-Meily JMA, van Heugten CM, Post MWM. Associations between social support and stroke survivors' health-related quality of life--a systematic review. Patient Educ Couns 2013 Nov;93(2):169-176. [doi: 10.1016/j.pec.2013.06.003] [Medline: 23870177]
36. Northcott S, Moss B, Harrison K, Hilari K. A systematic review of the impact of stroke on social support and social networks: Associated factors and patterns of change. Clin Rehabil 2016 Aug;30(8):811-831. [doi: 10.1177/0269215515602136] [Medline: 26330297]
37. Rony RJ, Amir S, Ahmed N, et al. Understanding the sociocultural challenges and opportunities for affordable wearables to support poststroke upper-limb rehabilitation: Qualitative study. JMIR Rehabil Assist Technol 2024 May 28;11:e54699. [doi: 10.2196/54699] [Medline: 38807327]
38. Meijering L, Nanninga CS, Lettinga AT. Home-making after stroke. A qualitative study among Dutch stroke survivors. Health Place 2016 Jan;37:35-42. [doi: 10.1016/j.healthplace.2015.11.006] [Medline: 26702961]
39. Appelros P, Åsberg S. Sex differences in stroke. Handb Clin Neurol 2020;175(299-312):299-312. [doi: 10.1016/B978-0-444-64123-6.00021-7] [Medline: 33008533]
40. Ohya Y, Matsuo R, Sato N, et al. Modification of the effects of age on clinical outcomes through management of lifestyle-related factors in patients with acute ischemic stroke. J Neurol Sci 2023 Mar 15;446:120589. [doi: 10.1016/j.jns.2023.120589] [Medline: 36807976]
41. Guest G, Namey E, Taylor J, Eley N, McKenna K. Comparing focus groups and individual interviews: Findings from a randomized study. Int J Soc Res Methodol 2017 Nov 2;20(6):693-708. [doi: 10.1080/13645579.2017.1281601]
42. Reid DJ, Reid FJM. Online focus groups: An in-depth comparison of computer-mediated and conventional focus group discussions. Int J Mark Res 2005 Mar;47(2):131-162. [doi: 10.1177/147078530504700204]
43. Stewart DW, Shamdasani P. Online focus groups. J Advert 2017 Jan 2;46(1):48-60. [doi: 10.1080/00913367.2016.1252288]
44. Woodyatt CR, Finneran CA, Stephenson R. In-person versus online focus group discussions: A comparative analysis of data quality. Qual Health Res 2016 May;26(6):741-749. [doi: 10.1177/1049732316631510] [Medline: 26935719]

Abbreviations

AWA: agent for well-being assistance
HBR: home-based rehabilitation
NEST: Neurorehabilitation Ecosystem for Sustained Therapy
RGS: Rehabilitation Gaming System
SM: Sint Maartenskliniek
VH: Vall d'Hebron University Hospital

XSL•FO
RenderX

Edited by S Munce; submitted 20.Jun.2025; peer-reviewed by D Bethere; revised version received 02.Mar.2026; accepted 23.Mar.2026; published 29.Apr.2026.

Please cite as:
Matijsen M, Hofstad C, Rodriguez Gonzalez S, Buxó X, Stara V, Leonzi S, Rampioni M, Mura A, Keijsers N
User Requirements of the Integrated Home-Based Rehabilitation Tool, Neurorehabilitation Ecosystem for Sustained Therapy: Multicenter Focus Group Study With Stroke Survivors, Caregivers, and Health Care Professionals
JMIR Rehabil Assist Technol 2026;13:e79382
URL: https://rehab.jmir.org/2026/1/e79382
doi:10.2196/79382

© Merle Matijsen, Cheriel Hofstad, Susana Rodriguez Gonzalez, Xavier Buxó, Vera Stara, Sara Leonzi, Margherita Rampioni, Anna Mura, Noël Keijsers. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 29.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of Mobile Software "SRCardioCare" Prototype for Implementing Home-Based Exercise Program Among Patients After Adult Cardiac Surgical Revascularization: Qualitative Feasibility Study

Ajith Kumar Pichai[1], MPT; A Sathya[1], ME, PhD; Thillaigovindarajan SenthilKumar[1], MPhil, MPT, PhD; Sadhanandham Shanmugasundaram[1], MBBS, DNB (Med), DNB (Cardiology); R Karthik[2], PhD

[1]Department of Cardiopulmonary Physiotherapy, Sri Ramachandra Faculty of Physiotherapy, Sri Ramachandra Institute of Higher Education and Research (DU), No.1, Sri Ramachandra Nagar, Chennai, India

[2]Centre for Cyber Physical Systems, Vellore Institute of Technology, Chennai, India

Corresponding Author:
Thillaigovindarajan SenthilKumar, MPhil, MPT, PhD
Department of Cardiopulmonary Physiotherapy, Sri Ramachandra Faculty of Physiotherapy, Sri Ramachandra Institute of Higher Education and Research (DU), No.1, Sri Ramachandra Nagar, Chennai, India

Abstract

Background: Noncommunicable diseases are a global concern with high mortality. Among these, cardiovascular disease requires more attention due to recurrence with altered physical activity. "SRCardioCare" (Sri Ramachandra Cardio Care) is an integrated mobile software that was developed to engage patients with effective communication and e-media support. We intend to explore the development of mobile software and its perceived impacts among health care professionals.

Objective: This study aimed to develop a more economical and feasible platform for cardiac rehabilitation following conservatively managed coronary arterial disease, heart failure, postoperative adult cardiac surgical revascularization, and other cardiac surgeries, and to develop software that facilitates effective communication among participants and health care professionals.

Methods: The software application was developed based on the experts' interviews. The core components that are included in the software were assessed for their usefulness and applicability among people with cardiac disease using standardized questions. Physicians, nurse practitioners, and physiotherapists' opinions were obtained. The developed app features providing e-media content for patients and pre- or postphysical activity response, including vitals and feedback from patients at set regular intervals, which were updated to the software.

Results: Opinions obtained from practicing physicians (cardiologists), nurse practitioners, and physiotherapists were hopeful for the development and future implementation of "SRCardioCare" among patients. "SRCardioCare" is designed for the effective implementation of physical activity among patients after conservatively managed coronary arterial disease, heart failure, postoperative adult cardiac surgical revascularization, and other cardiac surgeries. An integrated communication medium and regular postphysical activity feedback of vitals may offer safety in implementing physical activity among the vulnerable in a remote setup.

Conclusions: Remote rehabilitation is an essential and unexplored forum of practice in the field of rehabilitation, yet it requires wearable technology for remote monitoring and virtual reality and mixed reality for enhancing the adherence of the participants. To incorporate the telehealth forum effectively, especially in settings like India, the design must include an economically feasible and convenient model.

Trial Registration: Clinical Trials Registry – India CTRI/2025/03/082747; https://tinyurl.com/2zesstdf

(JMIR Rehabil Assist Technol 2026;13:e69197) doi:10.2196/69197

KEYWORDS

e-media; mobile software; home-based cardiac rehabilitation; cardiac rehabilitation; coronary artery bypass grafting

Introduction

Technical support in the field of health care is an emerging trend fostering the development of telecommunication between specialists and other medical professionals with patients [1,2]. Owing to advancements in telemonitoring and telerehabilitation, various settings are being explored for diverse clinical conditions. Mortality rate due to noncommunicable diseases is represented at 72.3% globally, requiring closer attention toward

XSL•FO
RenderX

the enhancement of prevention among the vulnerable population [3]. An important measure to be taken is to educate the patient on self-care and self-management strategies to prevent serious illness and its secondary complications. Satisfaction with doing physical activity among the patient population during the COVID-19 pandemic lockdown was high, and the telerehabilitation enhanced the patient-reported outcomes but was not limited to patient adherence [4,5].

Mobile software used in health care promotes inner strength and resilience among people with advanced age, which would be considered the core components of an individual's health and quality of life [6]. Commercially available mobile-based software relies on behavioral and motivational theory, but very little literature exists for its applicability and success [7,8]. This encouraged health care professionals to implement physical activity with mobile health apps in enhancing physical health outcomes. Any intervention model shall satisfy the patient's psychological well-being and meaningful choice, leading to sustained engagement in strategies implemented [9]. Data procurement using the software may require a hardware medium to source the most reliable and precise values during physical activities of an individual, which might act as a prime source for gaining confidence for one to actively use the software in remote settings [10,11].

Vital monitoring and supervised exercise program at remote settings would be a needed feature for prescribing the physical activity for an adult with cardiovascular disease, especially during postsurgical revascularization. Components with multiple domain-based software have appeared to be highly effective in engagement and the level of participation [12]. We aim to develop a mobile software that enables effective communication and feasible applicability, with safety measures incorporated for postsurgical revascularization patients, with provision for e-media–based exercise prescriptions. Patient adherence rate and rehospitalization, functional capacity, quality of life, and patient satisfaction would be analyzed after the effective implementation of the software application during hybrid cardiac rehabilitation.

Methods

This was a qualitative feasibility trial, which was a part of a pilot trial before the main trial. An interface was developed for engaging cardiovascular disease patients and health care professionals with an effective mobile software–based technology, which was analyzed for its usefulness and applicability. Purposive sampling was performed among participants to reflect the exposure to cardiac rehabilitation enabled with technology, including physicians (n=5), nurses, cardiac intensive care nurses (n=10), and physiotherapists (n=15) of a total of 25 (Table 1). The systematic face-to-face interview was conducted, and the study was checked for its quality using Consolidated Criteria for reporting qualitative studies (COREQ) [13,14]. The responses received from the health care professionals are stated using the LIKERT scale. Software application would include medication adherence, wound care education, nutritional education, exercise counseling and training, 24/7 communication with the health care team, and associated telemonitoring with postexercise feedback.

The procedure of the experiment trial includes that the intervention group and control group would be trained for physical activity as per routine hospital protocol. Following traditional care, the interventional group would be introduced to a software interface for training, and the control group would receive training based on the routine phase II center-based rehabilitation (Table 1). The intervention group would be trained specifically for regular practice of "physical activity at home" with a safer and effective mode using prescribed dosage. Training would be provided to the participants for recording their vitals, including heart rate, oxygen saturation, blood pressure, and rating of perceived exertion, and they would be trained for documenting these in the software interface.

Table . Baseline characteristics of interview participants

Characteristic	Doctors	Nurses	Physiotherapists
Sex			
Male, n (%)	4 (80)	2 (20)	9 (60)
Female, n (%)	1 (20)	8 (80)	6 (40)
Age (years), mean (SD)	42 (5)	37 (12)	37 (8)
Experience, mean (SD)	16.8 (10)	14.9 (13)	15 (8)
Occupation	• Cardiologist (n=2) • Physicians (n=3)	• Assistant nurse (n=6) • Clinical in-charge (n=2) • Senior nurse (n=2)	• Clinical therapist (n=8) • Senior therapist (n=7)

Education

The management of the postoperative recovery period after a cardiovascular incident would be challenging, requiring an advanced approach for management due to various postoperative complications (Table 2). SRCardioCare enables patients to be effectively educated, where information about conditions, risk factors, and secondary prevention as well as exercise counseling and training is provided through video presentations. Telemonitoring data were collected from the patients, along with postexercise feedback and regular reminders of medications, including communication with the patient.

XSL•FO
RenderX

Table . An outline implementation of software to the participants with a follow-up for 12 weeks.

Module	Week 1 - 2	Week 3 - 4	Week 5 - 6	Week 7 - 8	Week 9 - 10	Week 11 - 12
Introduction and orientation	✓					
Training to self-evaluate vitals	✓					
Evaluation for risk stratification	✓					
Counseling - Medication		✓				
Counseling - Nutrition		✓				
Counseling - Functional capacity		✓				
Counseling - Setting a goal		✓				
Counseling - Physical activity			✓			
Re-evaluation - Functional capacity			✓			
Re-evaluation – Recovery vitals				✓		
Counseling - Energy conservation				✓		
Counseling - Relaxation and coping strategies				✓		
Counseling - Secondary prevention strategies					✓	
Counseling - Behavior changing techniques					✓	✓
Counseling - Leading a quality and comfortable life						✓

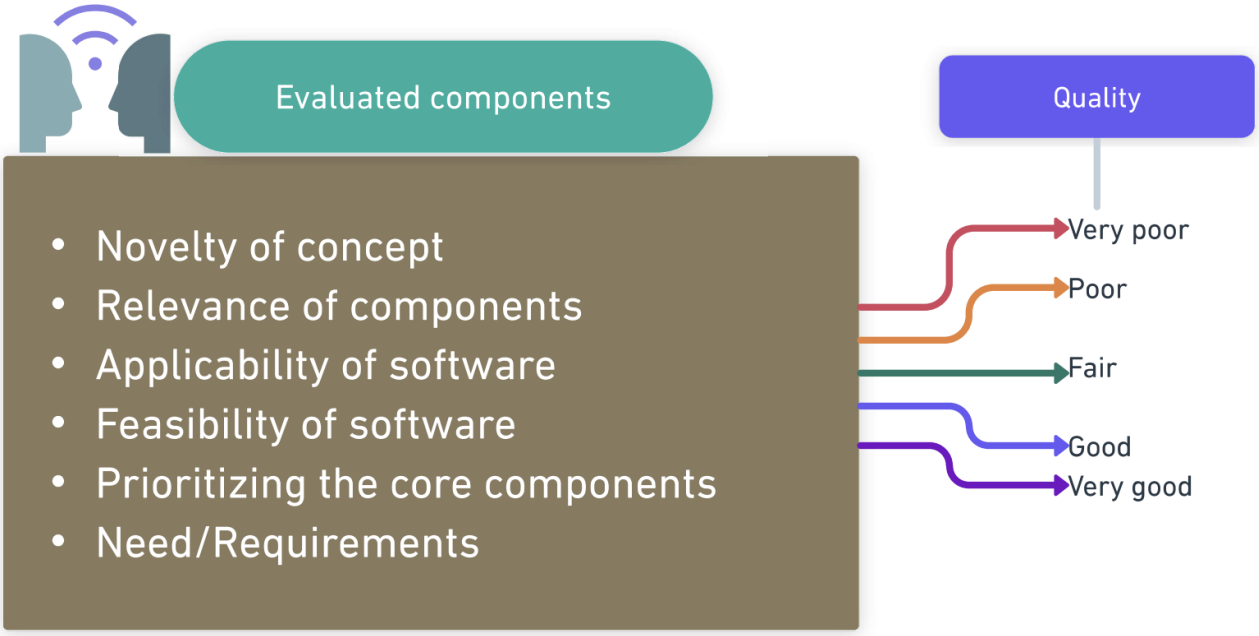
Narrative View of Feasibility Checking Among Health Care Professionals

A systematic interview was conducted among various health care professionals using the following domain (Figure 1), which was stratified from the interview guide. Data were collected from the participants using descriptive field notes and interviews, which were repeated once after the required corrections were made based on feedback provided by the participants in the software. Literature supports that there have been deprived scenarios for implementing the physical activity following the cardiac events, including postoperative cardiac revascularization surgeries and other cardiac diseases. Since understanding the trends falling toward the implementation of newer interventions in health care, most of the health care workers agreed to develop the software. The following feedback was added to the prototype design as suggested by participants:

1. Components of psycho-social counseling
2. Nutritional education components
3. Behavioral change techniques
4. Ensuring safety for implementing physical activity in remote settings

XSL•FO
RenderX

Figure 1. Survey among health care professionals for assessing the quality of software.



System Implementation

To build this mobile app for the rehabilitation of patients, Flutter was used to build the frontend and user interfaces, Node and Express for the backend, and MySQL as the primary database.

Technology Stack

Front-End Development With Flutter

The frontend of the SRCardioCare app has been entirely made using Flutter—a UI framework developed by Google to create cross-platform apps. Some of the advantages of Flutter include the following:

1. Cross-platform support: Flutter can compile to both iOS and Android, our primary targets for this app, from a single code base, maintaining and updating new features with a small development team.
2. Developer experience: Flutter is paired with Dart, a language convenience that was high with building and debugging code, especially on physical devices. The hot reload allows us to observe changes live and assist in preventing the reinstallation of the app frequently.
3. Ecosystem: the ecosystem of packages and libraries available in Flutter. Some of those libraries include: ‘fl_chart’ for visualizations and ‘chewie’ for video playback.

Interface Design

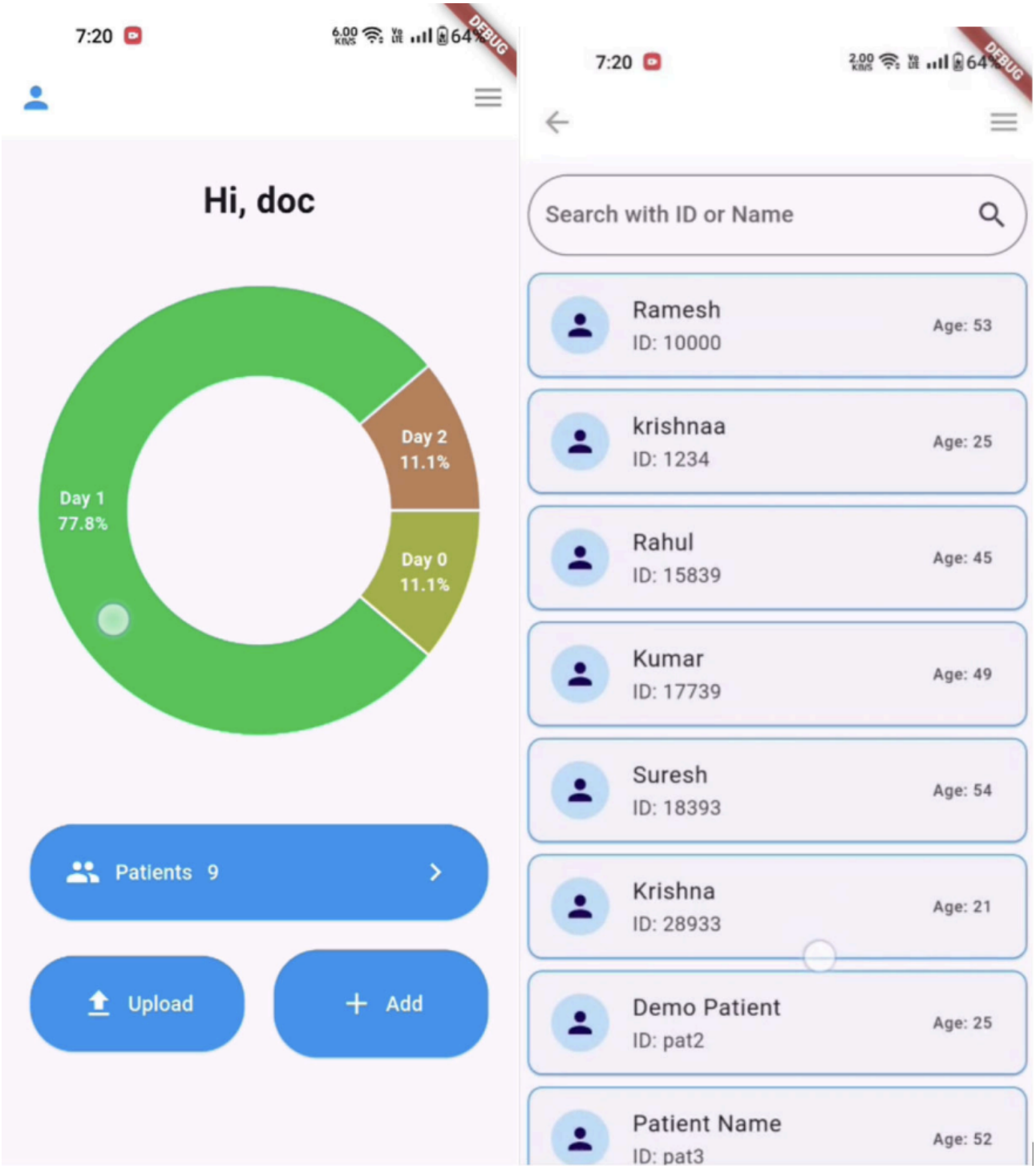
The user interface (UI) of SRCardioCare focuses on simplicity and accessibility, ensuring that patients with varying levels of tech literacy can navigate the platform with ease. The UI comprises several key features:

Dashboard

A personalized dashboard is given for every user who logs into the app, for both patients and physicians (Figure 2).

XSL•FO
RenderX

Figure 2. Developed software design home page for doctors with uploading videos options and patient progress.



- Patients: The dashboard shows the average of their vital parameters during their course of therapy. Links to all important actions such as viewing the current day's exercises, help, and contact information of their dedicated physician are featured.
- Physicians: The dashboard shows a visualization representing the different days different patients are in; it provides them with a broad overview of their patients' progress and offers them the ability to add a new patient.

Video Playback

The Flutter `chewie` package on top of a video controller was used to present the video playback of the exercise videos uploaded by the Physician (Figure 3). Each patient can only view the video prescribed to them, in order, and only on the specific day selected by the Physician. This ensures patients cannot share sensitive information, and their physician can always update the exercise based on the feedback from the patient.

XSL•FO
RenderX

Figure 3. Uploaded videography of the exercises prescribed for the patient.

XSL•FO
RenderX

Feedback

After the patients complete the day's exercises, they are provided with a form to input their vital parameters, such as heart rate, rating of perceived exertion, oxygen saturation, total body weight, and any remarks or doubts (Figure 4). This allows us to create a valuable record of their parameters throughout their journey and helps the physician tailor the exercise program further.

XSL•FO
RenderX

Figure 4. Documentation of post-physical activity vitals for patients.

XSL•FO
RenderX

Technical Challenges

The primary challenge faced when developing this system was integrating both the frontend and backend together with APIs. We used a simple REST (Representational State Transfer) setup, but for safety, we implemented the interfaces of the models and data both in TypeScript and Flutter independently. The difficulty level is high due to Flutter's rigid typing system and inability to convert JSON directly to appropriate types in Flutter. The manually written parsers for each model were used.

Backend

The backend was developed with Express; it runs using a bleeding-edge runtime called Bun. We preferred Bun instead of Node because of the following advantages: The typescript support directly runs a typescript file without having to transpile it to JavaScript. In HTTP/ Web APIs, the performance of Bun appears hastened compared to Node. Bun offers a great standard library, providing useful functions like hashing, which was used to verify our passwords on authentication.

Example- Bun hashing:

let isValid=await Bun.password.verify (password, user[0].password)

Data Storage, Privacy, and Management

Currently, SRCardioCare uses MinIO, a S3-compatible block storage service that is self-hosted. This was used to store the videos uploaded by the physicians. It is a drop-in replacement for AWS S3; therefore, no code changes are required when moving the data. According to the Healthcare Data Protection Regulations by HIPAA (Health Insurance Portability and Accountability Act) and GDPR (General Data Protection Regulation), a clear consent would be obtained from the participants, data usage would be strictly limited and restricted, and participants would be given full rights to remove and access their data as required.

Advantages are that the files are isolated in different buckets, and the access control is strictly granted based on the rules from the database; access is restricted to patients and physicians for maintaining the integrity of data.

The access is limited based on the mechanism of presigning URLs, ensuring no patient ever has permanent access to the videos and preventing them from sharing the exercises without a physician's prescription. Timely backups of the data for redundancy using external scripts such as *cron* and volume backups provided by the Cloud Provider could be scheduled.

System Integration and Module Development

The emerging growth of coronary artery disease and Acute Coronary Syndrome needs an efficient rehabilitation technique that satisfies the desire of affected patients. Traditional methods have more limitations, including monitoring the patient's risk when they are away and the need for continuous personalized care. This design addresses the above-mentioned limitation by supporting a personalized rehabilitation module and assisting them in a fast recovery from the disease. The proposed design methodology is a real-time cloud-based design to continuously monitor the patient and collect all the vital data of the patient.

The main goal of this proposed mobile app is to analyze and manage risk factors to reduce disability, improve functional capacity, lessen activity-related adverse effects, educate patients, and improve the quality of life.

To support the above-mentioned goal, the following design principles are adhered to: designed specifically for individual patients, collect vitals, processed offline, and prepared into a report specific to patients; development of a mobile application integrating technological advancements and cloud computing into health care, and continuous tracking of the patient based on the exercise.

Ethical Considerations

Ethical clearance was obtained from the Sri Ramachandra Ethical Committee IEC/24/AUG/187/26. The registration number for this trial is CTRI/2025/03/082747. This study is a part of a randomized controlled trial, in which the developed SRCardioCare app would be implemented to compare the effects of home-based cardiac rehabilitation with center-based cardiac rehabilitation for adult surgical revascularization patients (CTRI/2025/03/082747). Written informed consent was obtained from all participants prior to enrollment. Participants were informed about the purpose of the study, the voluntary nature of participation, and their rights to decline to answer any questions or withdraw from the study at any time without any consequences. All interview data were deidentified prior to analysis. Personal identities were removed and confidentiality was maintained through secure, password-protected data storage accessible only to the research team. Participants did not receive any financial or material compensation for their participation in the study.

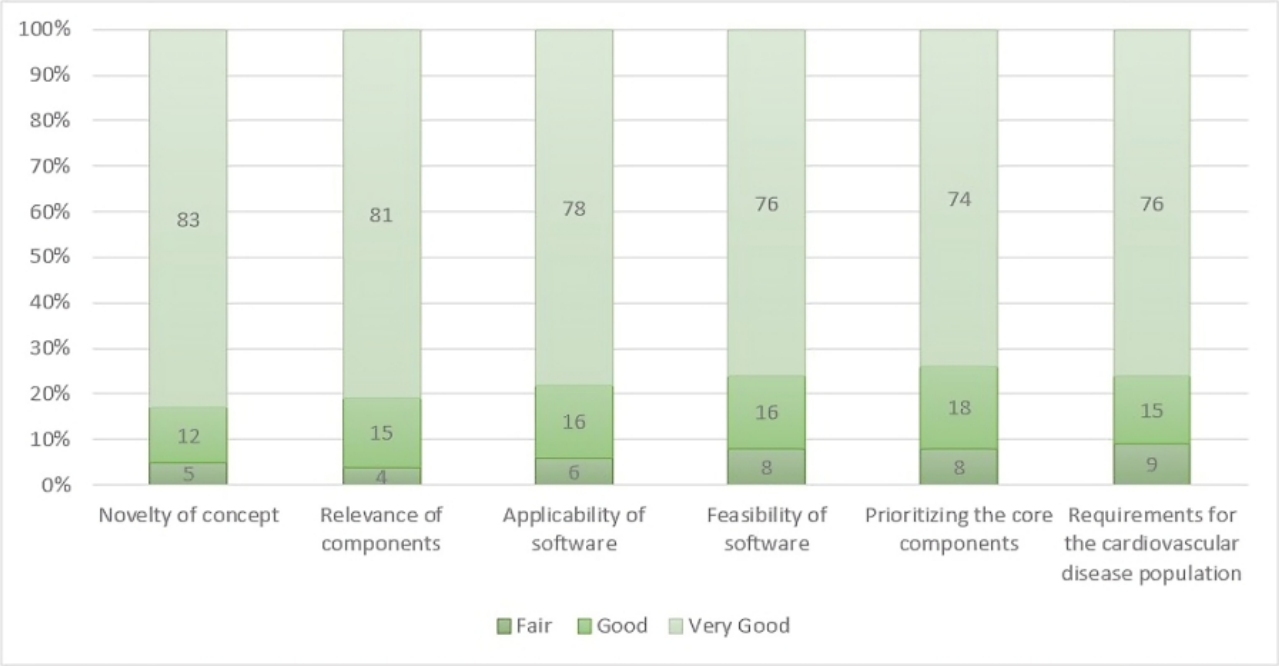
Results

Principal Results

After the presentation of the developed mobile software (app) SRCardioCare, with various expert health care professionals, a total of 25 participants were interviewed for their opinion about the developed software. The acceptability of developed software from the professionals was high due to its application of multidisciplinary approaches, along with high relevance and clinical need (Figure 5). UI-enabled ease toward navigation to their dashboards and exercise videos without difficulties. Respondents provided positive feedback about the app and highlighted their satisfaction with features such as comments after exercises and vitals documentation. But vital documentation emerged as a poor efficiency due to varying levels of digital literacy among participants. Workflow issues addressed by the clinician would be time constraints in managing and replying to the vital details and comments provided by the participant's end.

XSL•FO
RenderX

Figure 5. Responses from health care professionals: the built software was introduced to health care professionals in Sri Ramachandra Hospital and Medical Center, including 5 physicians, 10 nurse practitioners, and 15 physiotherapists. Responses were obtained using survey questions.



- **Acceptability:** Health care professionals described the app-based platform as highly acceptable due to its integration with technology advancements and applicability in cardiac rehabilitation. The novelty of such innovation with tele-rehabilitation contributed to its acceptance.

"The app is found useful for cardiac rehabilitation, especially for the heart failure population…"

- **Usability:** Participants perceived the app as easy to navigate with its simplified dashboards and order of sequence for training without hurdles. Order of week-wise module with training videos supported hassle-free provision for participants.

"The *instruction provided in the app makes easy navigation in the app, providing good follow-ups…"*

- **Satisfaction:** Health care professionals expressed a completeness in app features that include a multidisciplinary counseling forum and improved patient engagement. The essential feature considered mandatory would be the acquisition of the post-vital monitoring and communication module by health care professionals.

"I personally felt safe suggesting the patient get enrolled in cardiac rehabilitation using this software…"

- **Barriers:** The manual entry of vital signs by the participants could affect the accuracy and compliance, as per the current trial protocol. Unless a monitoring device is prepared, the defects might be considered noncompliant. Technical challenges in linking the frontend and backend components were described.

"Preloaded video components were slow occasionally and could reduce the medical terminologies used…,

The app could also be used for heart failure, where more precautionary measures need to be taken..."

- **Facilitators:** The components, which included educational and personalized training, motivated health care professionals to prescribe the application. The visual dashboards and consistent communication between patients and providers engage patients.

"It is good that the app displays the number of days the participants have to work on… and the exercises provided were kept changing every week, making participants more involved…"

- **Workflow Issues:** Data flow and coordination between modules were concerns of workflow-related issues, and manual data entry by the participants would be considered a difficulty, indeed an issue. The current software is limited to such a feature; the second phase of the current trial would be an innovation of a real-time monitoring device that would gather data automatically from the participants' continuing exercise training. The integration of real-time monitoring devices was recommended to enhance workflow efficiency.

"The alerting features and regular reminders would enhance the participation of patients effectively if done properly…"

"Rather than doing exercises independently, this app helps participants to perform without missing the sessions…"

Experts supported the SRCardioCare mobile application, emphasizing the necessity of the innovation, stating that telehealth adoption is rapidly increasing. Content analysis of health care professionals' suggestions and feedback revealed that they valued safety as a major requirement, stressing the necessity of continuous monitoring and postexercise feedback by the enrolled participants toward cardiac rehabilitation.

XSL•FO
RenderX

App Outputs

Personalized Treatment

After the presentation of the developed mobile software (app), SRCardioCare, with various expert health care professionals, a total of 25 participants were interviewed for their opinion about the developed software (Figure 1). Based on the vital measurements, the physician will prescribe personalized exercises suited for each patient (Figure 3). The exercises are uploaded to their login and assigned to that patient based on the patient ID. The physician will upload a sequence of six nondownloadable videos that comprise the exercises to be done and will change after four days. The patient can log into the app but can only view the first exercise, which is assigned to their ID. These exercises may vary after four days. The same progress can also be viewed from the portal.

Baseline Assessment of Patient

This module is basic for the Remote Rehabilitation system. Before using this remote system (Mobile App), orientation for the patient is given on this app and the need of this mobile app and the postoperative self-evaluation, importance of their continuous monitoring of their vital parameters of the patient such as heart rate, oxygen saturation, rating of perceived exertion, total body weight, and provided ways of managing the risk in postoperative periods are explained (Figures 4 and 6).

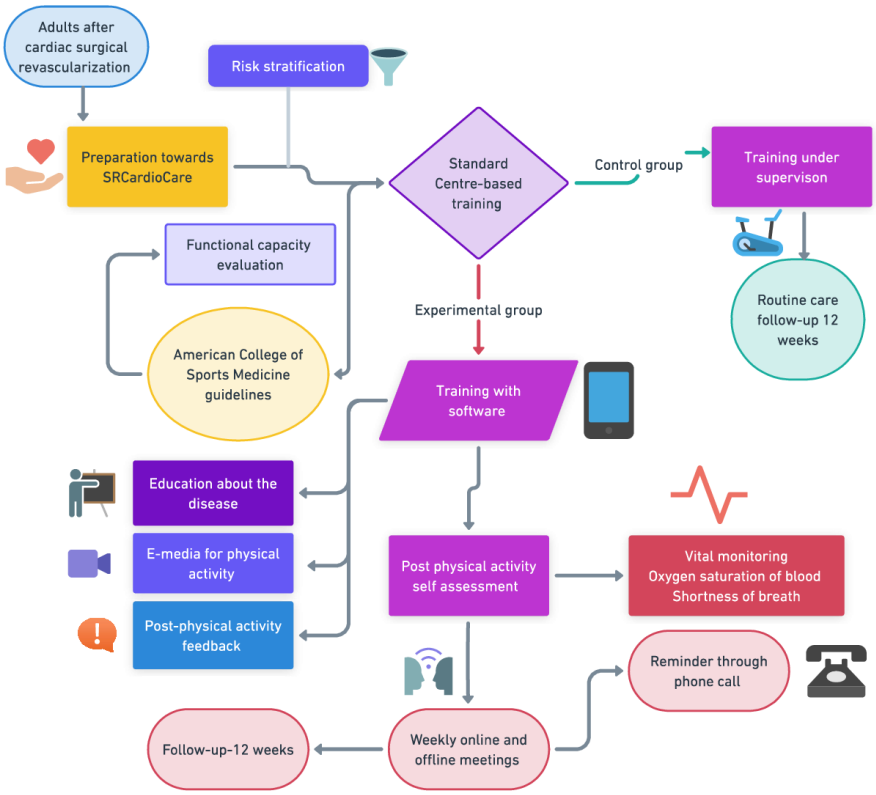
XSL•FO
RenderX

Figure 6. The screenshot shows the history of patients in the physician's login, along with the vital parameters measured.

XSL•FO
RenderX

To support the patients for the above-mentioned program, a Mobile App was designed. The mobile app is facilitated with role-based access, such as physician and patient. A physician will be assigned to many patients. During the first two weeks of this program physician will record the vitals of the patient before they are discharged from the hospital, and based on the vital signs, a personalized rehabilitation program will be designed for the patient. When the patient returned home, they needed constant treatment with remote monitoring (Figure 7). The patient can log into the system and view the personalized exercises loaded to them by their physicians (Figure 3).

Figure 7. SRCardioCare Trial protocol for participants after introducing the software. The chart was developed using the Whimsical tool.



Risk Factor Analysis

The main objective of this module is to assess the patient and provide continuous monitoring of their vital parameters. If needed, changes in the exercises can also be made. In the current app, the patient has to manually enter the vital parameter in the mobile app. As a future enhancement, it is planned to incorporate telemonitoring tools to record vital parameters. The vital parameters of the patients will be tracked and monitored by the physician. If any abnormality arises in the vital parameter, alert the caretaker or speed up the recovery process. This tracking prevents or reduces the adverse effects on cardiac activity.

This data is used to identify patterns and correlations between physical activity and vital signs, helping us assess the effectiveness of the rehabilitation plan and make informed adjustments. If a risk factor is detected (eg, prolonged elevated heart rate), the system sends alerts to both the patient and their health care provider, prompting immediate review and intervention. It also helps to predict the patient's health before the risk arises.

Data Analysis and Feedback Module

This mobile app establishes a communication between the physician and the patient at a remote setup. Through the app, the patient can readily present queries to the physician. They can record their vitals and post them to the physician. Physicians will resolve the queries raised by the patient on a priority basis. Physicians can analyze the patient's vital data stored in the cloud and get detailed insights about the patient's health, and they can predict the potential risk and avoid or reduce the adverse effects.

Discussion

Principal Findings

Cardiovascular disease remains the first among the major causes for years of life lost globally, ischemic heart disease being one of the causes if grouped globally [3,15]. The global risk of noncommunicable diseases is rising and requires immediate intervention to prevent complications [16]. Face-to-face implementation of exercises for chronic disabilities such as low back pain did not prove to be more effective than an exercise-based mobile app for managing the patients' issues, according to a recent randomized controlled trial [17]. Decisions

XSL•FO
RenderX

toward the development of a newer approach to deliver effective telemedicine and telerehabilitation are highly warranted [18]. A trial shows that intervention using a mobile app resulted in promising outcomes, stating that enhanced vitals, exercise capacity, and satisfaction toward exercise participation [19]. Device-dependent cardiac rehabilitation program for patients enhances the effective implementation of the protocol, as well as participants' interest toward participation is also promisingly enabled [20].

Behavioral change techniques would be an emerging intervention that would initiate the hopeful scope toward physical fitness in society. The awareness should be obtained at the guardian level and the care provider level to ensure the implementation of physical activity among the vulnerable population [21]. A systematic review reveals that the mobile app–based intervention would enhance diet, physical activity, and sedentary behavior when applied modestly [12]. Another systematic review suggests that despite strong evidence, people fail to follow the suggested physical activity and guidance from clinicians to lead a healthy life [22].

Cardiac rehabilitation has been prescribed to patients after cardiac events, but effective implementations are becoming poor due to poor participation. Which might require an alternative method of implementation, including technology-enabled implementation of cardiac rehabilitation at remote setups as well as in-center setups [23-25]. Effective implementation of a telehealth program depends on the various domains that are to be incorporated to gain effective clinical outcomes for the participants under safer practice. That might include telemonitoring, medication adherence, education, self-awareness about their functional capacity, understanding, and mastery of the telehealth model for self-implementation by the patients [26,27].

Patients would be thoroughly evaluated by a health care professional before introducing the software app, and the assessed comorbid conditions would be considered for exercise prescriptions. Prior training would be provided based on the suggestions provided by physicians. The concepts providing the e-supported rehabilitation criteria would include implementing cardiac rehabilitation through an economically cheaper method. Henceforth, the app usage would be based on subscription, which depends on the requirements of the duration of rehabilitation sessions suggested by physicians.

Conclusion

Globally, the literature supports that remote rehabilitation for cardiac diseases is effective. A mobile app offers trustworthy standards for the long-term adherence of patients with noncommunicable diseases. The developed software may help us to build a protocol for implementing remote rehabilitation following cardiac surgeries, heart failure, and conservatively managed acute coronary syndromes. Further upgradation may include telemonitoring support using the hardware gadget that would be designed to provide precise data of vitals during and post-exercise sessions. This leads the researchers to focus more on areas of telemonitoring devices that are more feasible and economical; deficient procurement of telemonitoring would be a limitation of this study.

Acknowledgments

The module was developed with the help of our co-author from Sri Ramachandra Engineering Technology; hence, no developmental cost was included. Maintenance of the software would be taken up by the principal investigator and the co-author team until the software is used for the complete implementation. Whimsical AI tool was used to draft Figure 6 (for design purposes only) [28].

Data Availability

Datasets generated during or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Gallagher R, Roach K, Sadler L, et al. Mobile technology use across age groups in patients eligible for cardiac rehabilitation: survey study. JMIR Mhealth Uhealth 2017 Oct 24;5(10):e161. [doi: 10.2196/mhealth.8352] [Medline: 29066425]
2. Oclaman JM, Murray ML, Grandis DJ, Beatty AL. The association between mobile app use and change in functional capacity among cardiac rehabilitation participants: cohort study. JMIR Cardio 2023 May 15;7:e44433. [doi: 10.2196/44433] [Medline: 37184917]
3. Naghavi M, Abajobir AA, Abbafati C, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet 2017 Sep;390(10100):1151-1210. [doi: 10.1016/S0140-6736(17)32152-9]
4. Krzowski B, Boszko M, Peller M, et al. Mobile application and digital system for patients after myocardial infarction: early results from a randomized trial. Pol Arch Intern Med 2023 Sep 29;133(9):16452. [doi: 10.20452/pamw.16452] [Medline: 36876854]

XSL•FO
RenderX

5. Bordas-Martinez J, Matéu Gómez L, Cámara Menoyo D, et al. Patient-reported outcomes measures (PROMs) and patient-reported experience measures (PREMs) of COVID-19 telerehabilitation: prospective pilot program. Medicine (Baltimore) 2022 Aug 5;101(31):e29639. [doi: 10.1097/MD.0000000000029639] [Medline: 35945781]
6. Salarvand S, Farzanpour F, Gharaei HA. The effect of personalized mobile health (mHealth) in cardiac rehabilitation for discharged elderly patients after acute myocardial infarction on their inner strength and resilience. BMC Cardiovasc Disord 2024 Feb 19;24(1):116. [doi: 10.1186/s12872-024-03791-5] [Medline: 38373888]
7. Li X, Zhao L, Xu T, et al. Cardiac telerehabilitation under 5G internet of things monitoring: a randomized pilot study. Sci Rep 2023;13(1). [doi: 10.1038/s41598-023-46175-z]
8. Senanayake S, Halahakone U, Abell B, et al. Hybrid cardiac telerehabilitation for coronary artery disease in Australia: a cost-effectiveness analysis. BMC Health Serv Res 2023 May 20;23(1):512. [doi: 10.1186/s12913-023-09546-w] [Medline: 37208666]
9. Najafi F, Nalini M. Hospital-based versus hybrid cardiac rehabilitation program in coronary bypass surgery patients in western Iran: effects on exercise capacity, risk factors, psychological factors, and quality of life. J Cardiopulm Rehabil Prev 2015;35(1):29-36. [doi: 10.1097/HCR.0000000000000087] [Medline: 25402170]
10. Batalik L, Dosbaba F, Hartman M, Batalikova K, Spinar J, Palazón-Bru A. Benefits and effectiveness of using a wrist heart rate monitor as a telerehabilitation device in cardiac patients: a randomized controlled trial. Medicine (Baltimore) 2020 Mar;99(11):e19556. [doi: 10.1097/MD.0000000000019556] [Medline: 32176113]
11. Dalli Peydró E, Sanz Sevilla N, Tuzón Segarra MT, Miró Palau V, Sánchez Torrijos J, Cosín Sales J. A randomized controlled clinical trial of cardiac telerehabilitation with a prolonged mobile care monitoring strategy after an acute coronary syndrome. Clin Cardiol 2022 Jan;45(1):31-41. [doi: 10.1002/clc.23757] [Medline: 34952989]
12. Schoeppe S, Alley S, Van Lippevelde W, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 7;13(1):127. [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
13. Buus N, Perron A. The quality of quality criteria: Replicating the development of the Consolidated Criteria for Reporting Qualitative Research (COREQ). Int J Nurs Stud 2020 Feb;102:103452. [doi: 10.1016/j.ijnurstu.2019.103452]
14. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care ;19(6):349-357. [doi: 10.1093/intqhc/mzm042]
15. Dorje T, Zhao G, Tso K, et al. Smartphone and social media-based cardiac rehabilitation and secondary prevention in China (SMART-CR/SP): a parallel-group, single-blind, randomised controlled trial. Lancet Digit Health 2019 Nov;1(7):e363-e374. [doi: 10.1016/S2589-7500(19)30151-7] [Medline: 33323210]
16. Fullman N, Yearwood J, Abay SM, et al. Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. The Lancet 2018 Jun;391(10136):2236-2271. [doi: 10.1016/S0140-6736(18)30994-2]
17. López-Marcos JJ, Díaz-Arribas MJ, Valera-Calero JA, et al. The added value of face-to-face supervision to a therapeutic exercise-based app in the management of patients with chronic low back pain: a randomized clinical trial. Sensors (Basel) 2024 Jan 16;24(2):567. [doi: 10.3390/s24020567] [Medline: 38257659]
18. Beatty AL, Beckie TM, Dodson J, et al. A new era in cardiac rehabilitation delivery: research gaps, questions, strategies, and priorities. Circulation 2023 Jan 17;147(3):254-266. [doi: 10.1161/CIRCULATIONAHA.122.061046] [Medline: 36649394]
19. Snoek JA, Prescott EI, van der Velde AE, et al. Effectiveness of home-based mobile guided cardiac rehabilitation as alternative strategy for nonparticipation in clinic-based cardiac rehabilitation among elderly patients in europe: a randomized clinical trial. JAMA Cardiol 2021 Apr 1;6(4):463-468. [doi: 10.1001/jamacardio.2020.5218] [Medline: 33112363]
20. Ruano-Ravina A, Pena-Gil C, Abu-Assi E, et al. Participation and adherence to cardiac rehabilitation programs. A systematic review. Int J Cardiol 2016 Nov 15;223:436-443. [doi: 10.1016/j.ijcard.2016.08.120] [Medline: 27557484]
21. Babu AS, Turk-Adawi K, Supervia M, Jimenez FL, Contractor A, Grace SL. Cardiac rehabilitation in India: results from the International Council of Cardiovascular Prevention and Rehabilitation's Global Audit of Cardiac Rehabilitation. Glob Heart 2020 Apr 3;15(1):28. [doi: 10.5334/gh.783] [Medline: 32489801]
22. Satyamurthy A, Prabhu N, Padmakumar R, Babu AS. Feasibility of an exercise-based cardiac rehabilitation algorithm in patients following percutaneous coronary intervention for acute coronary syndrome. Indian Heart J Teach Ser 2020 Jul;72(4):289-292. [doi: 10.1016/j.ihj.2020.07.011]
23. McGillion M, Yost J, Turner A, et al. Technology-Enabled remote monitoring and self-management - vision for patient empowerment following cardiac and vascular surgery: user testing and randomized controlled trial protocol. JMIR Res Protoc 2016 Aug 1;5(3):e149. [doi: 10.2196/resprot.5763] [Medline: 27480247]
24. Patterson K, Davey R, Keegan R, et al. Testing the Effect of a smartphone app on hospital admissions and sedentary behavior in cardiac rehabilitation participants: ToDo-CR randomized controlled trial. JMIR Mhealth Uhealth 2023 Oct 3;11:e48229. [doi: 10.2196/48229] [Medline: 37788043]
25. Zou H, Chair SY, Feng B, et al. A social media-based mindfulness psycho-behavioral intervention (MCARE) for patients with acute coronary syndrome: randomized controlled trial. J Med Internet Res 2024 Feb 20;26(26):e48557. [doi: 10.2196/48557] [Medline: 38376899]

XSL•FO
RenderX

26. Chang HY, Wu HW, Hung CS, et al. Costs and cardiovascular benefits of a fourth-generation synchronous telehealth program on mortality and cardiovascular outcomes for patients with atrial fibrillation: retrospective cohort study. J Med Internet Res 2024 Jan 8;26(1):e48748. [doi: 10.2196/48748] [Medline: 38190237]
27. Patterson K, Davey R, Keegan R, et al. A smartphone app for sedentary behaviour change in cardiac rehabilitation and the effect on hospital admissions: the ToDo-CR randomised controlled trial study protocol. BMJ Open 2020 Dec 15;10(12):e040479. [doi: 10.1136/bmjopen-2020-040479] [Medline: 33323435]
28. Whimsical AI. URL: https://whimsical.com/protocol-NeQDxTXwXKwR1sTbcPUzc [accessed 2026-02-03]

Abbreviations

ACS: acute coronary syndrome
CAD: coronary artery disease
NCD: noncommunicable diseases
UI: User Interface

Edited by J Sarvestan; submitted 26.Nov.2024; peer-reviewed by A Suresh, D Singhal, ST Kulnik; revised version received 28.Jul.2025; accepted 15.Sep.2025; published 23.Feb.2026.

Please cite as:
Kumar Pichai A, Sathya A, SenthilKumar T, Shanmugasundaram S, Karthik R
Development of Mobile Software "SRCardioCare" Prototype for Implementing Home-Based Exercise Program Among Patients After Adult Cardiac Surgical Revascularization: Qualitative Feasibility Study
JMIR Rehabil Assist Technol 2026;13:e69197
URL: https://rehab.jmir.org/2026/1/e69197
doi:10.2196/69197

© Ajith Kumar Pichai, A Sathya, Thillaigovindarajan SenthilKumar, Sadhanandham Shanmugasundaram, R Karthik. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging Large Language Models for Early Detection of Anomaly Work Injury Cases: Data-Driven Approach to Rehabilitation Efficiency

Peter Q Chen[1], MSc; Hayley Y W Gu[1], MSc; Heidi K Y Lo[2], MBChB, MRCPsych; Wing Chung Chang[2], MBChB, MD; Cameron J M Lai[3], MBA; Sun H S Lai[3], BSc, PhD; Andy S K Cheng[4], BSc, PhD; Peter H F Ng[1], MSc, PhD

[1]Department of Computing, Hong Kong Polytechnic University, PQ710, Mong Man Wai Building, Hong Kong, China (Hong Kong)

[2]Department of Psychiatry, University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Total Rehabilitation Management (HK) Limited, Hong Kong, China (Hong Kong)

[4]Department of Health and Physical Education, The Education University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Peter H F Ng, MSc, PhD
Department of Computing, Hong Kong Polytechnic University, PQ710, Mong Man Wai Building, Hong Kong, China (Hong Kong)

Abstract

Background: Large language models (LLMs) have demonstrated potential in automating the analysis of unstructured clinical data, yet their application in rehabilitation therapy for work injury cases remains underexplored.

Objective: We aimed to evaluate the performance of an LLM-assisted approach for the rapid identification of anomalous rehabilitation cases related to work injuries to enhance scalability and precision in case management.

Methods: We retrospectively analyzed 110,346 deidentified work injury cases between 2001 and 2024 from a leading rehabilitation coordination company in Hong Kong, representing approximately 20% of all work injury incidents in the region. LLMs were used to estimate the expected duration of recovery based on free-text injury descriptions. The cases in which the actual number of medically certified sick leave days exceeded the LLM-predicted maximum were classified as anomalies.

Results: The LLM-assisted method achieved high accuracy, with GPT-4o achieving over 73% accuracy in nonanomalous classification and 79% accuracy in all dataset detection, outperforming comparator models. The model maintained high accuracy across subgroups and demonstrated the reliable extraction of information from free-text notes.

Conclusions: The proposed method demonstrated robustness when evaluated on a large-scale dataset with a bimodal age distribution. This study highlights the potential of LLMs to transform rehabilitation workflows by automating anomaly detection at scale. The method also shows promise in tailoring rehabilitation strategies to age-specific needs and leveraging LLM tools for efficient case management. However, a key limitation is that the dataset includes only injury cases from a single geographic region, potentially limiting the generalizability of the findings to other populations or health care systems.

(JMIR Rehabil Assist Technol 2026;13:e80607) doi:10.2196/80607

KEYWORDS

anomaly detection; large language model; work injury; rehabilitation prediction; machine learning

Introduction

Efficient and targeted rehabilitation management is essential to ensure that individuals with severe conditions receive timely and appropriate care [1]. However, current work injury management often lacks the precision needed to allocate resources optimally, resulting in delays and inefficiencies in addressing high-priority cases [2]. A core challenge lies in the misallocation of attention and services, where relatively minor injuries with predictable recovery trajectories are sometimes treated with the same urgency as more complex cases. This misdirection not only burdens the health care system but also diverts valuable clinical time and expertise away from anomalies, cases involving severe injuries or irregular patterns of recovery that require specialized evaluation or intervention [3]. Without a robust mechanism to distinguish these cases early in the workflow, rehabilitation systems risk compromised patient outcomes, inefficient use of limited resources, and waste critical resources. In Hong Kong, the term “anomalies” refers to cases where the injured worker’s records indicate a severe injury requiring special or intensive care [4]. These cases may also suggest potential discrepancies, such as claims that have been overstated for additional compensation or legal benefits, indicating that the incident may not follow typical recovery

XSL•FO
RenderX

patterns according to industrial practice. In the Asia-Pacific context, including Singapore's employment practices guidance and Australian public sector leave management, cases that exceed expected or allowable sick leave provisions are treated as requiring attention, which aligns with our operational use of sick leave exceedance to flag potential anomalies [5-7]. Nonanomalous data refer to cases in which the injured worker experiences a light injury expected to heal in the usual course, with a standard recovery process leading to a timely return to work. These records represent the standard outcomes, without complications or indications of potential fraud. Some anomalies may also signal potential inconsistencies, such as exaggerated claims made for extended compensation or legal advantage. In industrial practice, comprehensive annotations about why a case is "special" are typically unavailable. The only consistent, objective post hoc indicator of atypical recovery trajectory is the realized count of medically certified sick leave days. Consequently, this study operationalizes cases requiring attention via a fast filter that flags cases whose realized sick leave exceeds a large language model (LLM)−estimated expected range. We emphasize that this is a pragmatic triage proxy, not a clinical determination of pathophysiology or fraud.

Addressing this challenge requires innovative methods for quickly and accurately identifying severe cases to optimize the distribution of resources. With the advancement of artificial intelligence (AI) techniques, clinical decision support systems have been increasingly employed across various domains to assist therapists in decision-making [8-11]. In the context of workplace injury, recent research has integrated machine learning methodologies, such as the variational autoencoder, for predicting sick leave outcomes and establishing a high alertness cliff [12]. Nevertheless, the prediction process still partially depends on the initial judgment of work injury case managers, who serve as the primary decision-makers in these cases. Senior work injury case managers consistently achieve higher accuracy compared to AI-based predictions [13]. Even with the assistance of neural networks, AI cannot rapidly achieve an acceptable level of performance without proper data preprocessing and customization. The research gap lies in the lack of efficient, data-driven methods to proactively identify anomalies in work injury management workflows. Current practices predominantly rely on random case assignment and retrospective corrections, resulting in wasted time and resources.

Recently, LLMs have demonstrated exceptional capabilities in processing language-related data, even passing the United States Medical Licensing Examination [14,15]. It has also been studied in several medical fields [16-18]. Numerous studies and surveys have investigated LLMs' ability to assume specific roles based on provided profiles, with results indicating that LLMs can effectively simulate profiled characters [19-21]. Simply prompting LLMs with a data description can generate responses in less than 1 minute without requiring additional model training. However, a critical research gap remains in determining how to constrain the outputs of LLMs and how to effectively design methods that leverage LLMs to detect anomalies in incoming injury cases efficiently.

Unlike traditional rehabilitation workflows, where senior work injury case managers must spend considerable time manually identifying anomalous cases, we developed an LLM-assisted method to streamline and accelerate this process. By leveraging prompt engineering techniques, we structured the input and constrained the output format to support accurate and efficient initial screening. The method is grounded in clinical reasoning: each injured worker is expected to follow a typical recovery trajectory, reaching a work-ready state within a medically anticipated time frame based on the nature and severity of the injury. LLMs, trained on extensive digital corpora that include medical and occupational content, are well positioned to infer such expectations and assist in detecting deviations from normative recovery patterns [22].

The scenario mirrors current practices in work injury management, where cases are often assigned randomly to junior or senior work injury case managers, only to discover later that certain cases would have benefited from senior-level expertise from the outset. This misallocation frequently results in delays and inefficient use of resources.

To this end, we proposed a novel approach leveraging LLMs to detect anomalies in occupational rehabilitation in the context of work injury management. Our method offered a potentially fast, scalable, and highly accurate solution for identifying severe cases based on data from work injury cases. Furthermore, this research collected over 110,000 work injury cases from a local company in Hong Kong, which handles nearly 20% of the total work injury cases yearly [23]. The data were used to validate our method, and several pilot studies were conducted for feasibility assessment, including model selection. Meanwhile, this research aims to uncover the key factors that characterize anomalies in this dataset, providing deeper insights into the decision-making process and facilitating a more informed allocation of resources. Our objectives are 2-fold: (1) developing a robust LLM-based method for anomaly detection in work injury management and (2) utilizing exploratory data analysis to uncover potential age and work-injury patterns in these anomaly cases.

Methods

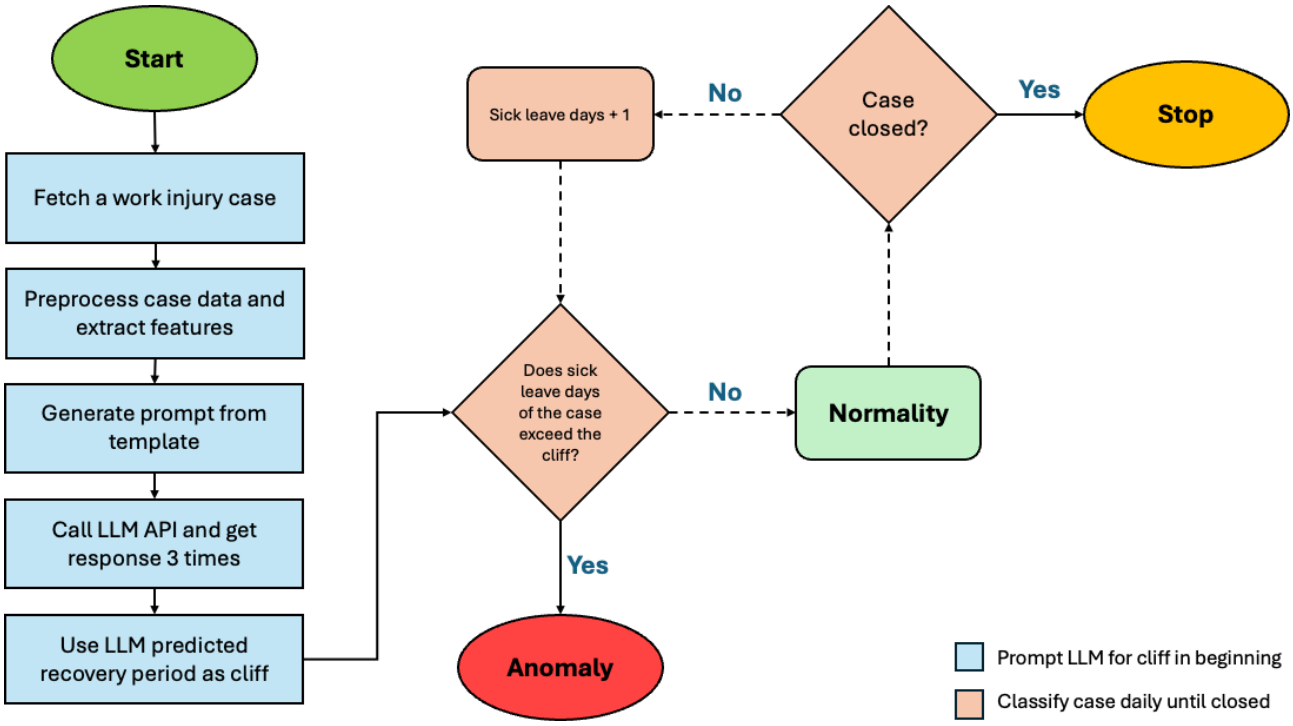
LLM-Assisted Anomaly Fast Detection Method

This method uses a fast and reliable alertness cliff to classify cases more effectively. If the total number of medically certified sick leave days for an injured worker exceeds this cliff, the case is flagged as potentially anomalous and prioritized for review by senior-level personnel or a detailed evaluation. An LLM predicts the expected duration of sick leave for each case from injury and accident descriptions. By comparing realized sick leave with the model-predicted recovery days, cases exceeding the cliff are classified as anomalies, and those within the expected range are considered normal. To enhance precision, 3 aggregation rules are used to define the decision cliff, referred to as the cliff: the maximum, the average, and the median of 3 independent LLM−generated recovery estimates. A case whose realized sick leave exceeds this cliff is classified as an anomaly. To mitigate variability and improve reliability, the LLM is queried 3 times per case using the same prompt structure, and the final decision is derived from the aggregated predictions to produce a robust, data-driven anomaly detection process [24].

XSL•FO
RenderX

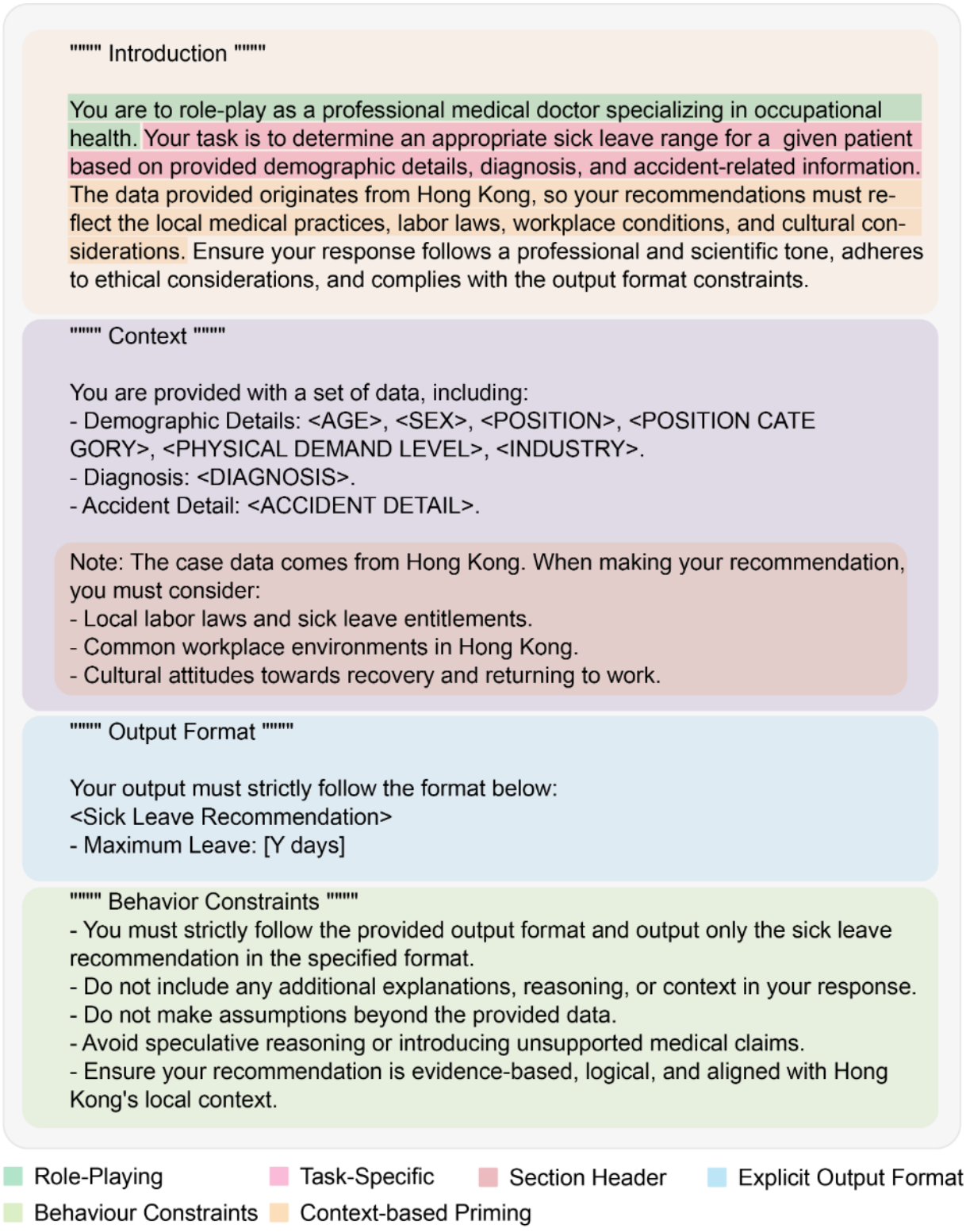
As illustrated in Figure 1, the workflow begins when a new work injury case is received. The "Query LLM 3 times and aggregate results" step rapidly determines the cliff using an LLM. The procedure begins by preparing case information, followed by preprocessing to retrieve demographics and extract key details, such as accident and diagnosis information. This content is embedded into a prompt based on the template shown in Figure 2. For each case, the LLM application programming interface (API) is queried 3 times to obtain predicted recovery periods, which serve as the cliff indicator for anomaly classification. A case is classified as nonanomalous if its sick leave has not yet exceeded the predicted cliff. For ongoing cases, sick leave days are incremented and reassessed against the cliff until the case is closed. Once the sick leave surpasses the cliff, the case is classified as an anomaly and referred to a senior work injury case manager for intervention.

Figure 1. Large language model (LLM)−assisted anomaly fast detection method. API: application programming interface.



XSL•FO
RenderX

Figure 2. Concrete prompt used to acquire information from a large language model.



The primary outcome is the accuracy of anomaly and nonanomalous classification in work injury cases, assessed on a dataset from a leading local work injury management company with expert-verified labels. The systematic use of LLM predictions enhances detection sensitivity while reducing manual workload, enabling case managers to focus on genuinely complex cases that require expert attention.

In routine operations, rich clinical detail is often unavailable at intake, so triage relies on minimal text and basic demographics. To enable low-latency and low-cost prioritization, a case-specific cliff is defined as the LLM-estimated expected duration of sick leave, serving as a data-driven prior on typical recovery given available notes. As the case progresses, if the running tally of medically certified sick leave exceeds this prior, particularly early in the timeline, the case is automatically queued for senior review. This mechanism functions as a workload triage heuristic rather than a diagnostic judgment, triggering timely escalation in cases of information scarcity,

XSL•FO
RenderX

improving allocation precision, and deferring definitive clinical determinations to expert assessment.

The method was validated using a dataset of 110,346 real-world work injury cases provided by a leading work injury management company in Hong Kong.

Prompt Template

Figure 2 demonstrates the detailed prompt engineering template, which utilizes multiple prompting techniques to enhance LLM performance. For clarity, the section header technique, such as "Introduction," "Context," "Output Format," and "Behavior Constraints," is used in the prompt template [25]. In the first part of the prompt template, the role-playing technique enhances contextual understanding, adaptability, and response accuracy by simulating specific personas, perspectives, or expertise in a given scenario [26]. To improve response relevance and coherence, the template specifies that all cases occurring in Hong Kong should utilize the context-based priming technique [27]. In the "Context" section of the prompt template, the input data for the injury case, including demographic details (eg, age, occupation), diagnosis details (textual description), and accident details (textual description), were included. In the "Output Format" part, the Explicit Output Format avoids undesired reasoning steps (eg, Chain-of-Thought) or other deviations, ensuring the LLM generates responses strictly in the intended structure without adding irrelevant content [28]. The "Behavior Constraints" part also serves a similar purpose, ensuring that the model's responses remain factual, precise, and contextually appropriate. Prompts are explicitly contextualized to Hong Kong's legal and clinical environment to enhance ecological validity, with strict output schemas for parsability and consistency [29]. Explicit instructional constraints are embedded to reduce hallucination and enforce adherence to the analytical task.

Pilot Study

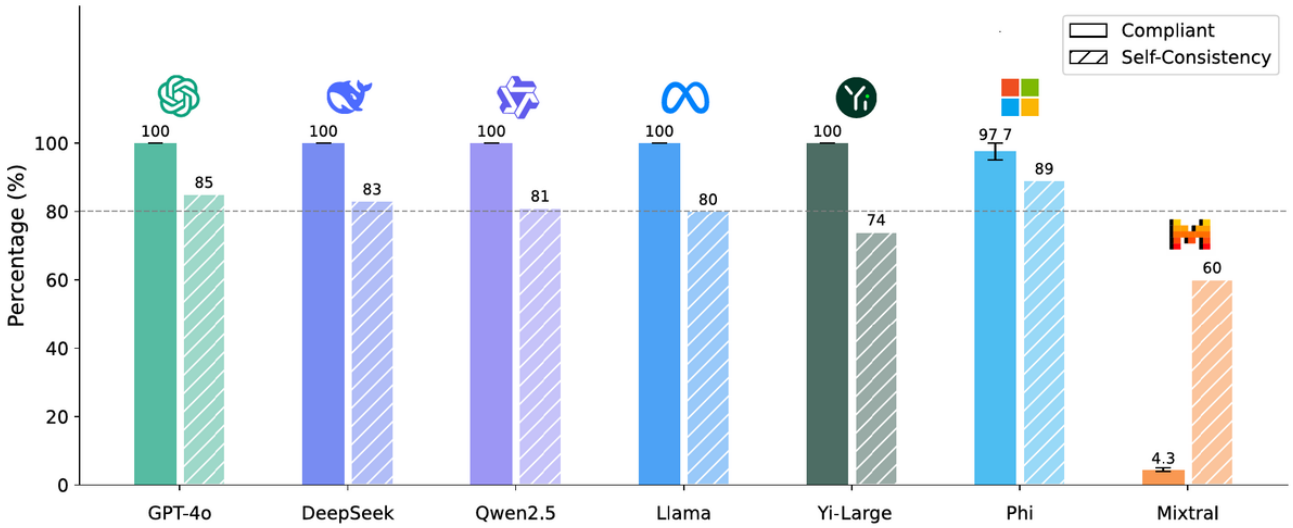
In our pilot study, we examined the varying strengths of different LLMs (eg, mathematical reasoning) and recognized that model size and architecture significantly influence performance [25,30-32]. To ensure a robust evaluation, we selected the largest and most widely recognized models from diverse LLM families, encompassing a broad range of architectures and capabilities.

Our primary objective was to determine whether these models could interpret a predefined prompt template and generate outputs that conformed to the required structure. Rather than examining their reasoning or predictive abilities, we focused on consistency, introducing 2 metrics: Compliance, which measures adherence to guidelines for producing the desired content, and Self-Consistency, which assesses whether the same response is generated across 3 repeated trials. We randomly sampled 100 cases from the dataset, prompted each model 3 times per case, and recorded the number of outputs that followed the required format. We focus on per-case anomaly triage using an upper bound on expected sick leave, so the prior work's center-focused reproducibility framework and large-run benchmarking do not fit our objective, evaluation needs, or deployable low-stochasticity protocol [32].

Figure 3 presents the results of this pilot study. Compliance represents the proportion of responses that satisfied our prescribed guidelines, while Self-Consistency quantifies each model's consistency across repeated prompts. Among 7 leading commercial LLMs, ChatGPT-4o, DeepSeek-V3, Qwen2.5-72B Instruct, Llama-3.1-405B Instruct, and Yi-Large achieved perfect Compliance; other models failed to avoid generating undesired content. We also examined internal consistency by comparing outputs across 3 prompts, categorizing them as identical across all trials, identical in at least 2, or unique each time. ChatGPT-4o exhibited the highest Self-Consistency, and several other models met our chosen 80% cliff for the subsequent experiments.

However, certain models with high Self-Consistency struggled with Compliance. This discrepancy is especially concerning, given our objective of providing a fast and reliable anomaly detection system for work injury management companies. Strict adherence to instructions is critical: any deviation can introduce erroneous or fabricated data, ultimately undermining the detection process. Ensuring compliance is thus essential to maintain the integrity of anomaly detection.

Figure 3. Compliance and self-consistency of different large language models.



XSL•FO
RenderX

Work Pipeline

Based on the results of the pilot study, we selected DeepSeek-V3, ChatGPT-4o, LLaMA-3.1-405B-Instruct, and Qwen2.5-72B to serve as LLM agents in our framework. All selected models achieved 100% compliance with the prompt template and demonstrated over 80% self-consistency across repeated responses.

A request-response framework was implemented using the FireWorks API in Python. Decoding parameters were configured for low-variance outputs using a temperature of 0.2 and a top-p of 1.0, which empirically reduced variability while maintaining robust adherence to the required output schema. A temperature of 0.0 was considered for full determinism; however, some models exhibited occasional output truncation or schema noncompliance at strictly deterministic settings during preliminary checks.

The demographic information, incident accident details, and clinical diagnoses were embedded into a standardized prompt template, as shown in Figure 2. Each prompt was submitted to the LLM via an API. A Python script extracted the responses to generate a predicted maximum duration of medically certified sick leave for each case. The data were then visualized to assess the accuracy of the LLM-based method.

Data Sources

Our dataset originates from a leading work injury management company in Hong Kong, which manages approximately 20% of work injury cases annually, covering records from 2001 to 2024. The dataset comprises 110,346 cases, with a gender distribution of 41.3% female and 53.6% male. The study population is predominantly Chinese, with individuals ranging in age from 18 to 83 years. This broad demographic coverage provides a robust basis for analyzing patterns across different age groups and genders within a relatively homogeneous ethnic context. Input leakage was not possible in this study. All records are confidential medical data, fully classified, and never publicly released or exposed to the models beyond the controlled evaluation pipeline, thereby preventing any external contamination that could bias LLM outputs or compromise validity.

Within this dataset, 15,575 cases were recorded as having zero sick leave days, which were treated as potential data entry errors. We exclude zero-day legitimate cases as outliers because they are immediately escalated to senior rehabilitation coordinators on day zero and handled outside our fast detection system, which targets anomalies only after predicted sick leave durations are exceeded. An additional 9230 cases had nonzero sick leave durations but contained missing values. For the primary predictive analysis, we used 85,541 cases that reported a nonzero number of medically certified sick leave days and had no missing data. These were inputs for the LLMs to predict the expected duration of normal recovery. Although excluded from the prediction task, the remaining data groups were also analyzed to extract relevant insights, given their substantial size.

Data Preprocessing

The dataset underwent staged preprocessing to ensure consistency and analytical suitability. Records outside the target time window were removed, implausible values were constrained within reasonable bounds, and entries with nulls in critical analytical fields were excluded. Noncritical descriptive fields with missing values were imputed using a neutral placeholder to preserve coverage while signaling incompleteness.

Categorical features were standardized through controlled vocabulary mapping, consolidation of multivalued entries into explicit multicategory indicators, and aggregation of low-frequency categories to mitigate sparsity. Text fields were sanitized by removing noninformative placeholders, and duplicates were eliminated based on content equivalence. A focused set of salient variables was retained for downstream analysis.

Data Analysis

In this study, we utilize a comprehensive set of metrics to rigorously evaluate the performance of the LLM-assisted anomaly detection method, encompassing classification accuracy, error magnitude, and model reliability. Classification accuracy, a core metric, is calculated based on 3 cliff-based methods: method 1 (maximum of 3 LLM predictions), method 2 (average of 3 predictions), and method 3 (median of 3 predictions). To further assess prediction deviations between realized sick leave days and LLM-predicted cliffs, we compute mean absolute error (MAE), mean squared error, root mean squared error (RMSE), mean absolute percentage error (MAPE), and mean percentage error (MPE). Additionally, Compliance (adherence to structured output formats) and Self-Consistency (reproducibility of outputs across repeated prompts) are quantified to ensure model reliability. We evaluate misclassification deviation, summarized through percentiles (eg, 50th and 75th percentiles), to analyze error distribution. Furthermore, exploratory data analysis is conducted to provide insights into the dataset, including descriptive statistics such as injury frequency, demographic distributions (eg, age, gender, occupation), anomaly prevalence, and misclassification patterns across key variables like body part and industry type. Together, these metrics and calculations form a robust framework for assessing the precision, reliability, and operational effectiveness of the proposed LLM-based anomaly detection system.

We evaluated triage performance using cliff-based classification derived from LLM-predicted “cliffs” of expected sick leave duration. For each case, the LLM was queried 3 times with the same prompt. The per-case decision cliff was then defined by 1 of the 3 aggregation rules: method 1, method 2, and method 3. A case was classified as an anomaly if it realized medically certified sick leave days exceeded the chosen cliff; otherwise, it was classified as nonanomalous. The primary performance metric was accuracy, computed as the proportion of correctly classified cases over the evaluation set, and reported overall and stratified by anomaly and nonanomalous subsets to characterize trade-offs across decision rules and models.

XSL•FO
RenderX

To assess reliability and error magnitude, we further quantified misclassification deviation, defined for errors as the absolute difference between realized sick leave days and the decision cliff, summarized via percentiles (eg, 50th and 75th).

To gain a deeper understanding of the phenomenon in the dataset, salary is treated as a composite proxy that captures differences across job types, seniority, contract structures, and work experience. Given the absence of granular role-level pay scales, salary should not be interpreted as a pure measure of experience but as an indicator shaped by occupational category and tenure.

Ethical Considerations

This study introduced an innovative anomaly detection method for work injury rehabilitation, validated using real-world cases from Hong Kong. The project has been approved by the PolyU Institutional Review Board (reference HSEARS20250406002). A pilot study was first conducted to evaluate the performance of several well-known commercial LLMs, which informed the selection of the most effective models for the subsequent experiments.

The dataset was provided by one of the largest work injury management companies in Hong Kong, which manages nearly 20% of the government-reported work injury rehabilitation cases. The data were shared exclusively for research purposes under a strict confidentiality agreement. Prior to transfer, all records were anonymized by the provider to ensure the protection of personal information.

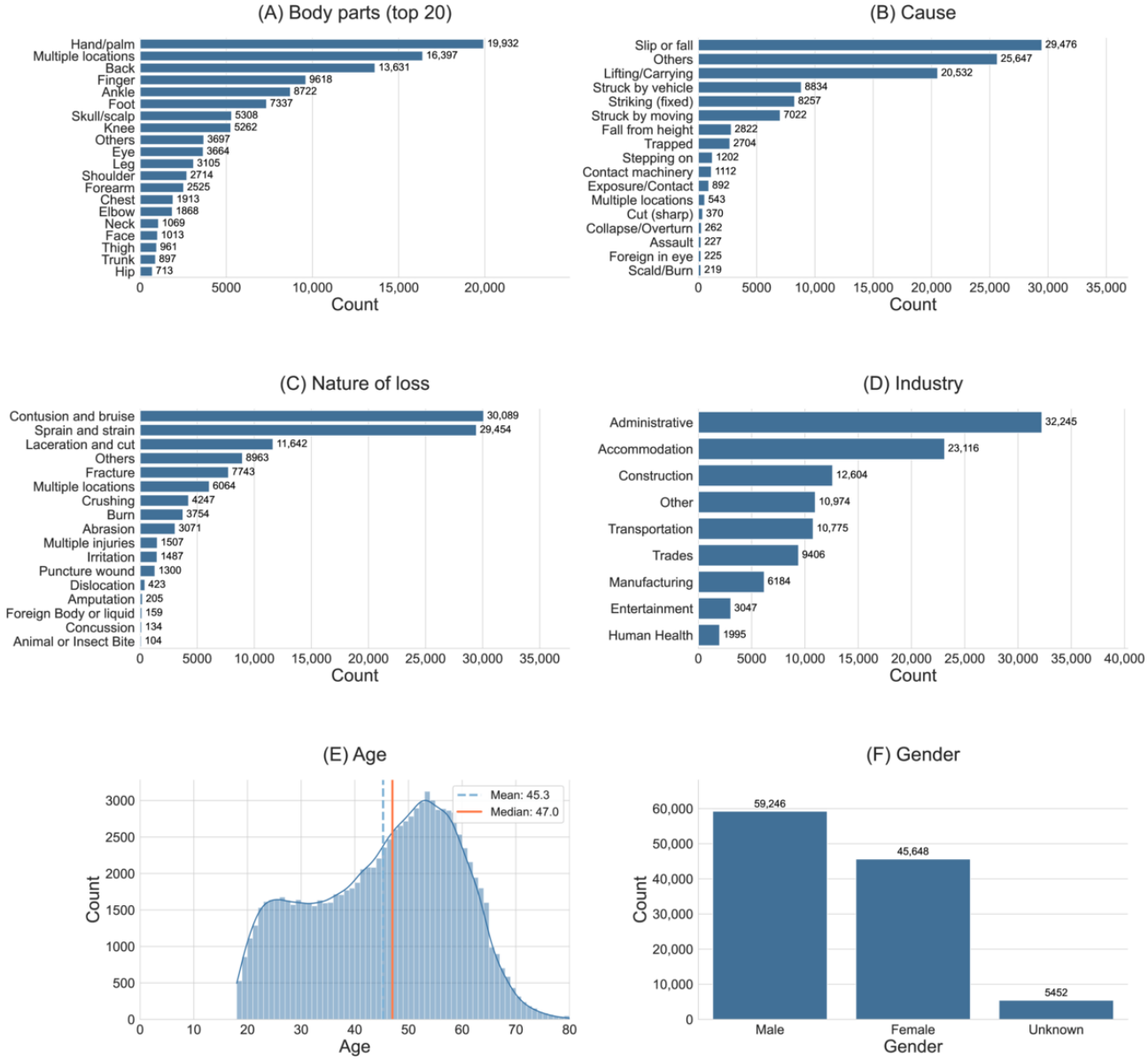
Results

Work Injury Dataset

Figure 4 presents a comprehensive analysis of 110,346 deidentified occupational injury cases managed by a leading rehabilitation coordination company in Hong Kong between 2001 and 2024. Injuries predominantly involved peripheral anatomical regions, with fingers (n=16,397) and backs (n=13,631) collectively accounting for nearly one-quarter of all incidents, followed by hand or palm injuries (n=9618) and ankle injuries (n=8722). Consistent with these anatomical findings, the most common types of loss were contusions or bruises (n=30,089) and sprains or strains (n=29,454), whereas open wounds, such as lacerations and cuts (n=11,642), and fractures (n=7743), occurred less frequently. Industry-specific data indicate a substantial burden arising from labor-intensive service sectors, notably “Administration and support services” (n=32,245) and “Accommodation and food service activities” (n=23,116), collectively accounting for more than half of all cases and surpassing the construction sector (n=12,604) in this dataset. Precipitating events were predominantly same-level slips, trips, and falls (n=29,476) and manual lifting or carrying tasks (n=20,532). Demographically, male workers represented a modest majority (53% of claims); nevertheless, female workers comprised a substantial proportion (42%). The age distribution was right-skewed, with a mean age of 45.3 (SD 13.3) years and a median of 47.0 (IQR 21.0) years.

XSL•FO
RenderX

Figure 4. Demographic data statistics in the dataset. (A) The top 20 most common categories of injured body parts. (B) Most common categories in the industry. (C) Most common categories in the nature of loss. (D) Most common categories in the cause of injury. (E) Gender distribution in the dataset. (F) Age distribution in the dataset.



Performance Assessment of LLMs

Table 1 shows the LLM classification accuracy across the selected models for different data categories. All LLMs mentioned in Figure 3 have been tested. For the entire dataset, all selected LLMs achieved more than 70% accuracy in the maximax, expected maximum, and median cliff criteria methods. Among the selected LLMs, GPT-4o achieved the best performance in all cliff criteria, while DeepSeek-V3 had the lowest performance. In the separate anomaly dataset containing only anomalies, the prediction accuracy of all selected models exceeds 95%. The best performance for the nonanomalous dataset comes from GPT-4o as well, achieving more than 76% under maximax cliff criteria, over 73% accuracy under expected maximum cliff criteria, and a median cliff criterion.

XSL•FO
RenderX

Table . Case classification accuracy across models[a].

	Method 1 (%)	Method 2 (%)	Method 3 (%)
Qwen	78.71	78.26	78.19
DeepSeek	76.31	75.75	75.66
Llama	78.15	76.93	76.76
GPT-4o	81.77	79.95	79.51
Anomaly			
Qwen	97.72	97.79	97.79
DeepSeek	97.59	97.81	97.80
Llama	97.69	98.02	97.97
GPT-4o	96.32	97.16	97.21
Nonanomalous			
Qwen	71.71	71.07	70.98
DeepSeek	68.48	67.62	67.51
Llama	70.96	69.16	68.95
GPT-4o	76.40	73.61	72.99

[a]Method 1 uses the maximum value among the 3 large language models (LLMs) predictions as a cliff to classify anomalies and nonanomalous. Method 2 uses the average of the 3 LLM predictions as a cliff for classification. Method 3 uses the median of the 3 LLM predictions as a cliff for classification.

Table 1 further highlights the trade-offs in classification accuracy when different criteria are applied. When using the expected maximum as the classification criterion, the model achieves higher accuracy in anomaly detection but at the cost of reduced accuracy in nonanomalous classification. However, misclassifying nonanomalous cases is relatively less consequential, as such cases typically resolve quickly, with injured workers returning to work in a short period. In contrast, misclassifying anomalies can have significant financial and operational implications. Suppose an anomaly is incorrectly classified as a normal case. In that case, the company may need to allocate additional resources to reassign a senior work injury case manager later in the process, leading to prolonged recovery times and potentially missed rehabilitation windows. From an anomaly detection perspective, Llama demonstrates the highest detection rate. However, when considering overall performance across both anomaly and nonanomalous classification, GPT-4o outperforms other models, making it the most balanced and practical choice for real-world deployment. These findings highlight the importance of selecting an LLM that optimally balances accuracy, adherence to instruction, and overall classification performance.

Table 2 shows that across the 3 aggregation strategies, absolute and squared errors remain high: MAE is approximately 72 days, and RMSE is around 158 days for all methods, indicating substantial pointwise deviations and volatility in predicting the maximum sick leave duration. Relative errors are also large: MAPE ranges from 169.86% to 195.48%, and MPE exceeds 100% for all methods, evidencing pronounced systematic overestimation. Among the alternatives, the median-based cliff (method 3) yields the lowest relative error (MAPE=169.86%, MPE=106.77%) and slightly lower dispersion, whereas using the maximum prediction as the cliff (method 1) amplifies both bias and variance; the mean (method 2) lies in between. Despite these differences, the small gaps in MAE or RMSE across methods suggest that aggregation choice alone does not resolve the core error magnitude, and bias calibration or robustness enhancements are warranted. Directly using LLMs to predict sick leave duration from demographics yields large absolute and relative errors (≈72-day MAE, ≈158-day RMSE, MAPE >169% with systematic overestimation), revealing unstable and biased point forecasts, which motivates our shift to a fast exceedance−based detection method rather than relying on raw predictions.

XSL•FO
RenderX

Table . Standard metrics for 3 methods[a].

	MAE[b]	MSE[c]	RMSE[d]	MAPE[e] (%)	MPE[f] (%)
Method 1	72.37	24,741.36	157.29	195.48	136.39
Method 2	72.42	25,173.20	158.66	172.39	110.32
Method 3	72.54	25,266.41	158.95	169.86	106.77

[a]Method 1 uses the maximum value among the 3 large language models (LLMs) predictions as a cliff to classify anomalies and nonanomalous. Method 2 uses the average of the 3 LLM predictions as a cliff for classification. Method 3 uses the median of the 3 LLM predictions as a cliff for classification.

[b]MAE: mean absolute error.

[c]MSE: mean squared error.

[d]RMSE: root mean squared error.

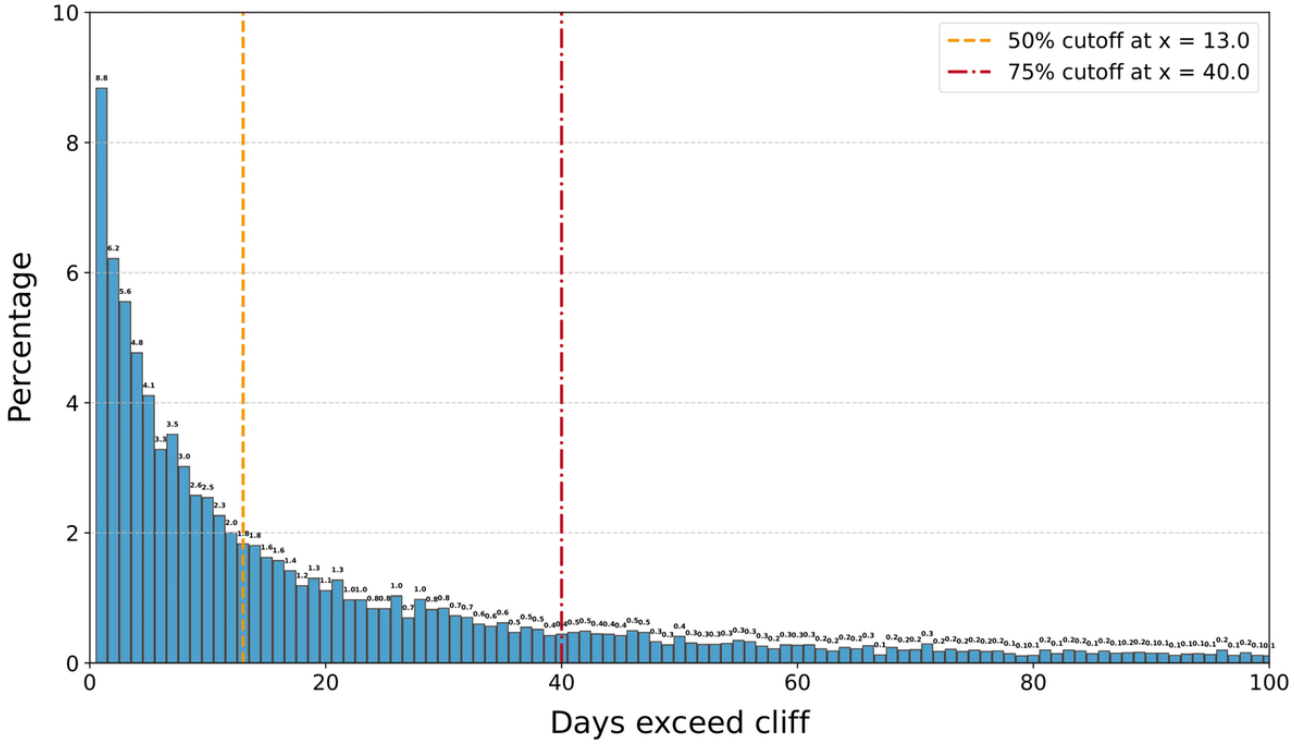
[e]MAPE: mean absolute percentage error.

[f]MPE: mean percentage error.

LLM Misclassifications Case Study

Since method 1 achieved the highest overall accuracy, this section primarily focuses on analyzing the predictions of the LLM based on method 1, as well as those generated by ChatGPT-4o, which demonstrated the highest average accuracy among all the selected models. Figure 5 illustrates the distribution of erroneous predictions that exceed the specified cliff when method 1 is applied. The yellow dashed line represents the 50th percentile of the cumulative distribution, corresponding to a deviation of 13 days. In contrast, the 75th percentile of the cumulative distribution corresponds to a deviation of 40 days. These results indicate that half of the erroneous predictions deviate from the ground truth by no more than 13 days, further underscoring the robustness of the proposed method. The previous work on similar datasets shows that the traditional variational autoencoder could only achieve an average error of 107.447 days, and they failed to predict a cliff for classifying the anomaly and nonanomalous cases [12,13]. Table 2 shows the absolute errors of direct LLM predictions, demonstrating substantial bias, while Figure 5 highlights erroneous cases flagged by our fast detection method, underscoring that our approach is intended to assist rather than replace rehabilitation coordinators.

Figure 5. Distribution of erroneous predictions beyond the cliff.



To gain a deeper insight into the LLM's prediction performance, we examined the distribution of GPT-4o outputs using method 1 (as described in the previous section). Specifically, 14,751 nonanomalous cases were misclassified as anomalies, and 847 anomalies were misclassified as normal. Figure 6 illustrates the distribution of the key variables within these misclassified nonanomalous cases. Finger has the highest proportion of misclassifications, at approximately 23%, whereas other body parts each account for around 10%. In the "Nature of Loss" variable, "Sprain & Strain" accounts for over 28% of the misclassifications, followed by "Contusion & Bruise" and "Laceration & Cut," both of which exceed 15%; the remaining

XSL•FO
RenderX

categories each fall below 10%. Regarding the Industry, the Administrative and Support Service sector and the Accommodation and Food Service Activities sector exhibit notably higher proportions of misclassifications (over 20%) relative to others. Finally, in the "Position Category" variable, most misclassifications occur under the "unknown" category.

Figure 6. Distribution of large language model misclassifications nonanomalous across the key variables (top 10).

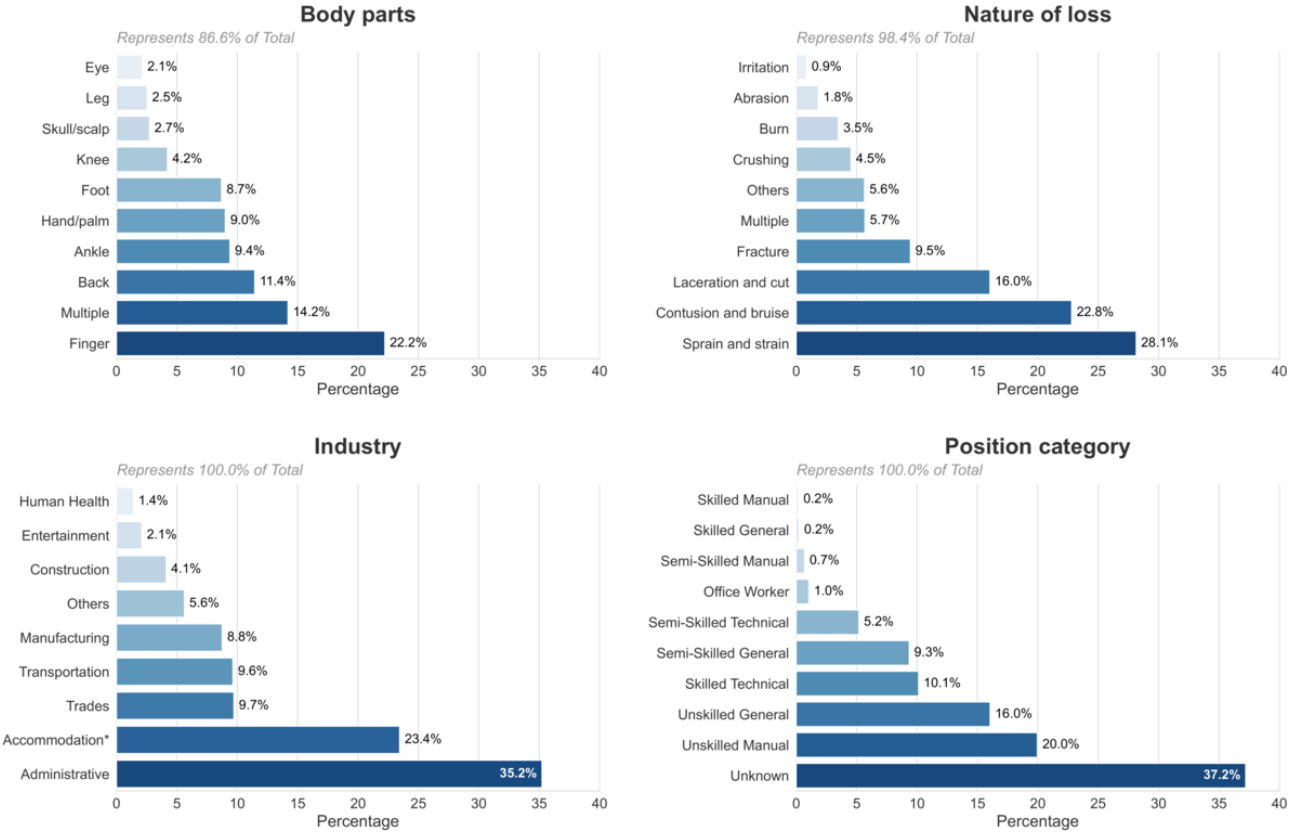


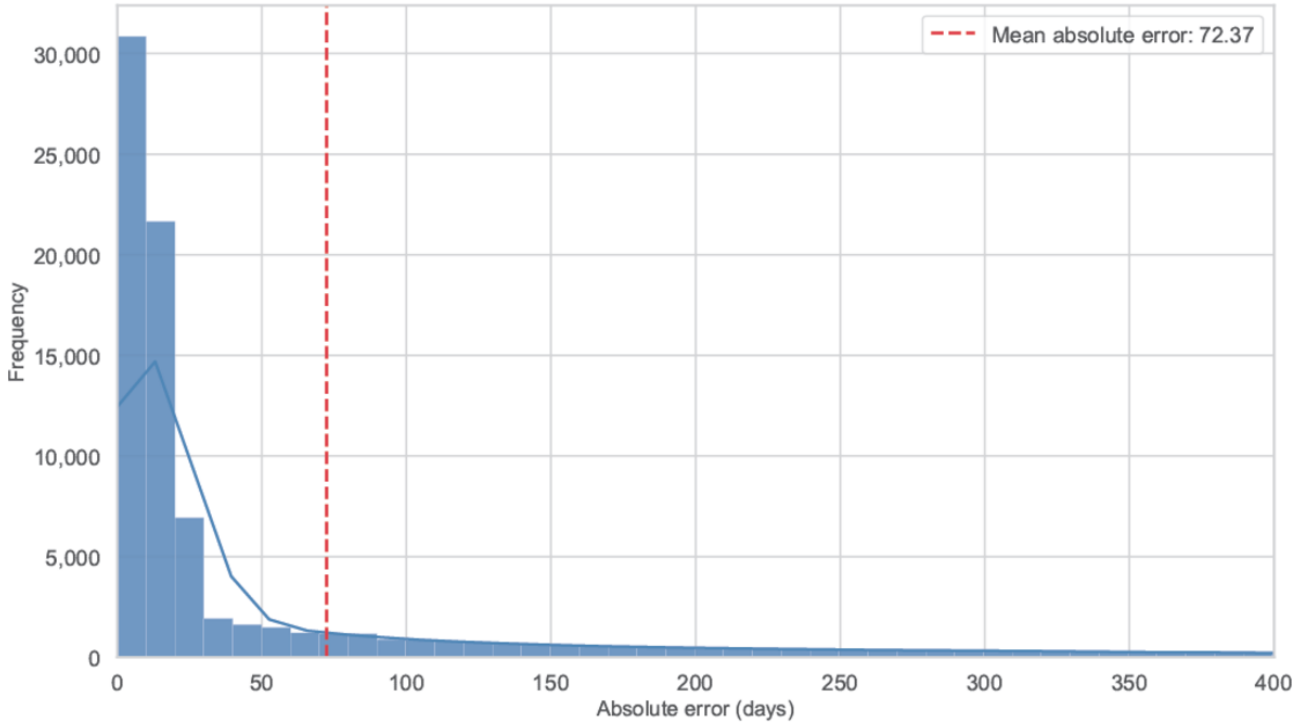
Figure 6 also suggests that certain misclassifications may arise from incomplete data, such as the absence of a position category, which impairs the LLM's ability to predict the cliff accurately. Without this critical information, the model must rely on its intrinsic knowledge, leading to increased variability and uncertainty. Additionally, for variables such as the nature of loss, injured body parts, and industry type, human judgment is also prone to significant bias, particularly in those variables with high percentages of misclassification [13].

LLM Direct Prediction Error Distribution

Figure 7 shows the distribution of absolute errors for raw LLM day predictions; the mean error is 72 days, which indicates that direct prediction yields deviations too large for practical use. Therefore, we use the LLM as a fast anomaly filter that flags cases whose realized sick leave exceeds an estimated upper bound, rather than predicting sick leave days directly.

XSL•FO
RenderX

Figure 7. Distribution of the large language model raw prediction absolute error.



Case Study of Our Approach

In this case study, we analyze a 51-year-old female worker who sustained a back injury while lifting a heavy basket of packaged bread. The injury was classified under Body Parts: Back, Nature of Loss: Sprain & Strain, and Cause: Injured whilst lifting or carrying, with the worker employed as a packer in the Accommodation and Food Service Activities industry. The actual sick leave duration was 6 days, while the LLM predictions for recovery duration were consistent across all methods: result 1=14 days, result 2=14 days, and result 3=14 days, leading to cliff values of max_max=14 days.

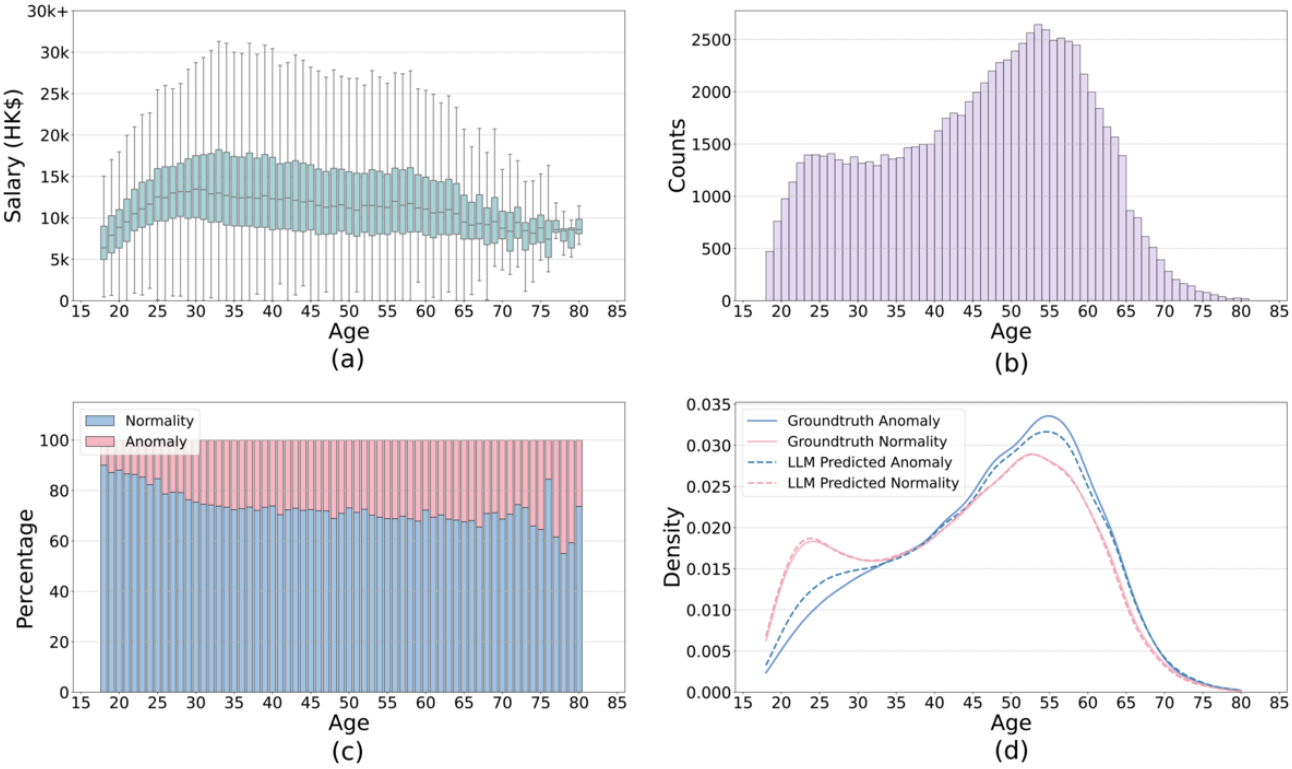
The difference between the actual sick leave and the maximum predicted cliff sick leave duration max_max was –8 days, indicating that the actual recovery period was well within the predicted range and did not exceed the anomaly cliff. This demonstrates the effectiveness of the LLM-assisted method in identifying cases that conform to expected recovery patterns, ensuring that resources are not misallocated to cases that align closely with typical recovery trajectories.

Exploratory Insights From the Dataset

To examine anomalies, we present a salary box plot against age in Figure 8. The top-left panel of the figure presents a box plot of salary against age, where the median salary for each age group is extracted to reveal a clearer trend. The median salary initially increases, peaking between ages 35 and 40, before gradually declining. The top-right panel illustrates the distribution of age counts from 18 to 80 years, showing an initial peak in the early 20s, followed by a decline until the 40s, after which the number of cases rises again, reaching its highest peak around the age of 55 years. The bottom-left panel displays the percentage of anomalies across different age groups, revealing a steady increase in anomaly prevalence with age. Finally, the bottom-right panel illustrates the normalized age distributions of both anomalies and nonanomalous, alongside the LLM's predicted anomalies and nonanomalous. The red line (representing nonanomalous) closely mirrors the overall age distribution, while the blue line (representing anomalies) steadily increases and peaks near the age of 60 years. Notably, the LLM's predicted distributions align closely with the ground-truth trends.

XSL•FO
RenderX

Figure 8. Overview of salary and age distributions with anomaly analysis. (A) Salary boxplot across ages. (B) Age distribution of all data. (C) Percentage of anomaly across ages. (D) Age distribution of anomaly or normality data.



Discussion

Principal Findings

The performance of the proposed LLM-assisted anomaly fast detection method demonstrates promising results, 110,346 deidentified work-injury cases between 2001 and 2024 from a leading work injury management company in Hong Kong, representing approximately 20% of all work injury incidents in the region. Compared to previous research, which primarily assessed AI prediction accuracy by comparing it to work injury case managers' judgments, our study provides a more comprehensive evaluation [12]. Prior studies demonstrated that AI predictions could surpass human work injury case managers in both nonanomalous and anomalous cases. However, while they also attempted to predict anomaly cliffs, their findings did not explicitly report the accuracy of such predictions. In contrast, our approach ensures that the accuracy of anomaly cliff prediction is systematically analyzed, contributing to a more reliable anomaly detection framework.

In Figure 7, we illustrate the trend of salary changes, considering salary as a key indicator of work experience. It is well known that physical ability peaks in the 20s, remains stable or slightly declines until around 30 - 35, and then drops significantly thereafter [33,34]. Comparing this physical ability trend with the salary trend, we observe that workers in their early 20s possess peak physical strength but lack experience. As a result, they may be more prone to injuries; however, their quick recovery often prevents these cases from becoming anomalies. This explains the initial peak in injury cases around the age of 20 years.

By the time workers reach their 30s, they have gained significant experience (as indicated by higher salaries), while their physical ability has not yet declined substantially. Consequently, the number of work injury cases decreases between the ages of 30 and 40 years. However, after 40, salaries may begin to decline slightly as workers are unable to maintain the same working hours as before, while their physical ability drops sharply. This results in a second peak in work-related injury cases. At this stage, injuries are more severe, and recovery is less likely, contributing to an increase in both the number and proportion of anomalies.

Limitations and Future Directions

This study tests LLMs for spotting unusual rehabilitation cases. The models read injury diagnoses and accident notes, estimate a normal recovery time, and flag cases that fall outside that range. We judge the approach by how well it separates nonanomalous from anomalies, not by exact prediction accuracy. However, the limitations are that no comparison is made with human rehab coordinators, only general-purpose LLMs without medical models, and data from a single region; therefore, the findings may not be generalizable. As LLMs improve, they could streamline the rehabilitation triangle and resource planning. This research can be readily transferred to other rehabilitation systems within Hong Kong, but adaptation to other regions would be domain specific due to significant differences in work injury frameworks, legal and policy requirements, and cultural practices.

Deterministic decoding with temperature 0.0 could further enhance reproducibility. As a robustness check, future or supplementary analyses can compare performance and schema compliance at 0.0 versus 0.2 to quantify any trade-off between

XSL•FO
RenderX

determinism and adherence to output constraints. If comparable, deployment would favor 0.0 to maximize reproducibility; if not, a small nonzero temperature remains justified to preserve formatting reliability under operational constraints.

Conclusions

We present an LLM-based approach that estimates expected recovery time from injury records and flags deviations as anomalies, streamlining rehabilitation triage. Tested on more than 110,000 Hong Kong work-injury cases, the method improved classification efficiency; GPT-4o delivered the most balanced accuracy, with DeepSeek-V3 and Qwen2·5-72B Instruct close behind. Demographic analysis reveals that injuries are more frequent yet milder in younger workers, whereas those aged 40 and above experience more anomalies, reflecting reduced resilience. The approach advances data-driven rehabilitation coordination and optimizes resource allocation.

Funding

We acknowledge the support of the Innovation and Technology Fund (ITP/039/24LP) of the Government of the Hong Kong Special Administrative Region for providing the support.

Data Availability

Medical notes and the full list of responses from large language models will not be shared to protect the privacy of the patients. The responses of the large language models or derived data are available from the corresponding author.

Authors' Contributions

PQC and PHFN conceived the study, oversaw its administration, and provided supervision. PQC, HYWG, CJML, and SHSL accessed, verified, and curated the data; performed the formal analysis and methodology; developed the software; validated the results; and produced the visualizations. PQC prepared the original draft, while HYWG, HKYL, WCC, ASKC, and PHFN reviewed and edited the manuscript. PHFN additionally secured funding and supplied resources. All authors had full access to all study data and approved the final version for publication.

Conflicts of Interest

None declared.

References

1. Szeto GPY, Cheng ASK, Lee EWC, Schonstein E, Gross DP. Implementing the work disability prevention paradigm among therapists in Hong Kong: facilitators and barriers. J Occup Rehabil 2011 Mar;21(1):76-83. [doi: 10.1007/s10926-010-9256-2] [Medline: 20652377]
2. Remus L, Grope M, Lemke S, Bethge M. An innovative case management intervention for people at high risk of permanent work disability to improve rehabilitation coverage and coordination of health services: a randomized controlled trial (AktiFAME, DRKS00024648). BMC Health Serv Res 2022 Mar 15;22(1):342. [doi: 10.1186/s12913-022-07482-9] [Medline: 35292005]
3. Brownlee S, Chalkidou K, Doust J, et al. Evidence for overuse of medical services around the world. Lancet 2017 Jul 8;390(10090):156-168. [doi: 10.1016/S0140-6736(16)32585-5] [Medline: 28077234]
4. Cooper DA, Slaughter JE, Gilliland SW. Reducing injuries, malingering, and workers' compensation costs by implementing overt integrity testing. J Bus Psychol 2021 Jun;36(3):495-512. [doi: 10.1007/s10869-020-09681-9]
5. Paid sick and carer's leave. Fair Work Ombudsman. URL: https://www.fairwork.gov.au/leave/sick-and-carers-leave/paid-sick-and-carers-leave [accessed 2025-12-11]
6. Management of staff leave in the Australian public service. : Australian National Audit Office (ANAO); 2022 URL: https://www.anao.gov.au/work/performance-audit/management-staff-leave-the-australian-public-service [accessed 2025-12-11]
7. Tripartite standards: frequently asked questions (FAQs). TAFEP. 2024. URL: https://www.tal.sg/tafep/-/media/tal/tafep/getting-started/files/tripartite-standards-faqs [accessed 2025-11-12]
8. Lutz W, Deisenhofer AK, Rubel J, et al. Prospective evaluation of a clinical decision support system in psychological therapy. J Consult Clin Psychol 2022 Jan;90(1):90-106. [doi: 10.1037/ccp0000642] [Medline: 34166000]
9. Antoniadi AM, Du Y, Guendouz Y, et al. Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review. Appl Sci 2021;11(11):5088. [doi: 10.3390/app11115088]
10. Musen MA, Middleton B, Greenes RA. Clinical decision-support systems. In: Biomedical Informatics: Computer Applications in Health Care and Biomedicine: Springer; 2021:795-840. [doi: 10.1007/978-3-030-58721-5_24]
11. Borges do Nascimento IJ, Abdulazeem HM, Vasanthan LT, et al. The global effect of digital health technologies on health workers' competencies and health workplace: an umbrella review of systematic reviews and lexical-based and sentence-based meta-analysis. Lancet Digit Health 2023 Aug;5(8):e534-e544. [doi: 10.1016/S2589-7500(23)00092-4] [Medline: 37507197]

XSL•FO
RenderX

12. Ng PHF, Chen PQ, Sin ZPT, Lai SHS, Cheng ASK. Smart Work Injury Management (SWIM) system: a machine learning approach for the prediction of sick leave and rehabilitation plan. Bioengineering (Basel) 2023 Jan 28;10(2):172. [doi: 10.3390/bioengineering10020172] [Medline: 36829666]
13. Yeung YYK, Chen PQ, Ng PHF, Cheng ASK. Evaluation of the accuracy of the Smart Work Injury Management (SWIM) system to assist case managers in predicting the work disability of injured workers. J Occup Rehabil 2025 Jun;35(2):320-332. [doi: 10.1007/s10926-024-10199-7] [Medline: 38874680]
14. Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023 Feb 8;9(1):e45312. [doi: 10.2196/45312] [Medline: 36753318]
15. Malgaroli M, Schultebraucks K, Myrick KJ, et al. Large language models for the mental health community: framework for translating code to care. Lancet Digit Health 2025 Apr;7(4):e282-e285. [doi: 10.1016/S2589-7500(24)00255-3] [Medline: 39779452]
16. Deiner MS, Deiner RY, Fathy C, et al. Use of large language models to classify epidemiological characteristics in synthetic and real-world social media posts about conjunctivitis outbreaks: infodemiology study. J Med Internet Res 2025 Jul 2;27:e65226. [doi: 10.2196/65226] [Medline: 40601927]
17. Huang W, Wei W, He X, et al. ChatGPT-assisted deep learning models for influenza-like illness prediction in mainland China: time series analysis. J Med Internet Res 2025 Jun 27;27:e74423. [doi: 10.2196/74423] [Medline: 40577658]
18. Huang J, Lai H, Zhao W, et al. Large language model–assisted risk-of-bias assessment in randomized controlled trials using the revised risk-of-bias tool: evaluation study. J Med Internet Res 2025 Jun 24;27:e70450. [doi: 10.2196/70450] [Medline: 40554779]
19. Li N, Gao C, Li M, Li Y, Liao Q. EconAgent: large language model-empowered agents for simulating macroeconomic activities. 2024 Presented at: Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); Aug 11-16, 2024; Bangkok, Thailand p. 15523-15536 URL: https://aclanthology.org/2024.acl-long [accessed 2025-12-12]
20. Zhang Z, Zhang-Li D, Yu J, et al. Simulating classroom education with LLM-empowered agents. Presented at: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics; Apr 29 to May 4, 2025; Albuquerque, NM p. 10364-10379 URL: https://aclanthology.org/2025.naacl-long [accessed 2025-12-12] [doi: 10.18653/v1/2025.naacl-long.520]
21. Wang L, Ma C, Feng X, et al. A survey on large language model based autonomous agents. Front Comput Sci 2024 Dec;18(6):186345. [doi: 10.1007/s11704-024-40231-1]
22. Longwell JB, Hirsch I, Binder F, et al. Performance of large language models on medical oncology examination questions. JAMA Netw Open 2024 Jun 3;7(6):e2417641. [doi: 10.1001/jamanetworkopen.2024.17641] [Medline: 38888919]
23. Occupational safety and health statistics. Labour Department. URL: https://www.labour.gov.hk/tc/osh/content10.htm [accessed 2025-12-11]
24. Wang X, Wei J, Schuurmans D, et al. Self-consistency improves chain of thought reasoning in language models. arXiv. Preprint posted online on Mar 21, 2022. [doi: 10.48550/arXiv.2203.11171]
25. Brown TB, Mann B, Ryder N, et al. Language models are few-shot learners. 2020 Presented at: 34th Conference on Neural Information Processing Systems (NeurIPS 2020); Dec 6-12, 2020; Vancouver, Canada p. 1877-1901 URL: https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf [accessed 2025-12-11]
26. Shanahan M, McDonell K, Reynolds L. Role play with large language models. Nature 2023 Nov;623(7987):493-498. [doi: 10.1038/s41586-023-06647-8] [Medline: 37938776]
27. Dong Q, Li L, Dai D, et al. A survey on in-context learning. arXiv. Preprint posted online on Dec 31, 2022. [doi: 10.48550/arXiv.2301.00234]
28. Kojima T, Gu SS, Reid M, Matsuo Y, Iwasawa Y. Large language models are zero-shot reasoners. 2022 Presented at: 36th Conference on Neural Information Processing Systems (NeurIPS 2022); Nov 28 to Dec 9, 2022; New Orleans, LA p. 22199-22213 URL: https://openreview.net/pdf?id=e2TBb5y0yFf [accessed 2025-12-12]
29. Yu F, Quartey L, Schilder F. Exploring the effectiveness of prompt engineering for legal reasoning tasks. Presented at: Findings of the Association for Computational Linguistics; Jul 9-14, 2023; Toronto, Canada p. 13582-13596 URL: https://pdfs.semanticscholar.org/9661/dda80241923420976 47c37423eb7fccee298.pdf [accessed 2025-12-12]
30. Achiam J, Adler S, Agarwal S, et al. GPT-4 technical report. arXiv. Preprint posted online on Mar 15, 2023. [doi: 10.48550/arXiv.2303.08774]
31. Hoffmann J, Borgeaud S, Mensch A, et al. Training compute-optimal large language models. arXiv. Preprint posted online on Mar 29, 2022. [doi: 10.48550/arXiv.2203.15556]
32. Wang JJ, Wang VX. Assessing consistency and reproducibility in the outputs of large language models: evidence across diverse finance and accounting tasks. SSRN. Preprint posted online on Mar 21, 2025. [doi: 10.2139/ssrn.5189069]
33. Burr DB. Muscle strength, bone mass, and age-related bone loss. J Bone Miner Res 1997 Oct;12(10):1547-1551. [doi: 10.1359/jbmr.1997.12.10.1547] [Medline: 9333114]
34. Westerterp KR. Changes in physical activity over the lifespan: impact on body composition and sarcopenic obesity. Obes Rev 2018 Dec;19 Suppl 1:8-13. [doi: 10.1111/obr.12781] [Medline: 30511504]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
API: application programming interface
LLM: large language model
MAE: mean absolute error
MAPE: mean absolute percentage error
MPE: mean percentage error
RMSE: root mean squared error

Edited by S Munce; submitted 14.Jul.2025; peer-reviewed by AA Barr, E Fu, VX Wang; accepted 19.Nov.2025; published 30.Jan.2026.

Please cite as:
Chen PQ, Gu HYW, Lo HKY, Chang WC, Lai CJM, Lai SHS, Cheng ASK, Ng PHF
Leveraging Large Language Models for Early Detection of Anomaly Work Injury Cases: Data-Driven Approach to Rehabilitation Efficiency
JMIR Rehabil Assist Technol 2026;13:e80607
URL: https://rehab.jmir.org/2026/1/e80607
doi:10.2196/80607

© Peter Q Chen, Hayley Y W Gu, Heidi K Y Lo, Wing Chung Chang, Cameron J M Lai, Sun H S Lai, Andy S K Cheng, Peter H F Ng. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predicting Geriatric Rehabilitation Stays of ≤4 Weeks After Hip Fracture Surgery: Machine Learning Approach Using Physical Activity and Patient Data

Sanne M Krakers[1,2], MSc; Frank J Wouda[1], PhD; Dieuwke van Dartel[1,3], MSc; Miriam MR Vollenbroek-Hutten[1,4], Prof Dr; Johannes H Hegeman[1,2], MD, Prof Dr; Up&Go After a Hip Fracture Group[5]

[1]Department of Biomedical Signals and Systems, University of Twente, Enschede, The Netherlands

[2]Department of Trauma Surgery, Ziekenhuisgroep Twente, Almelo, The Netherlands

[3]Reggeborgh Research Fellow, ZGT Academy, Ziekenhuisgroep Twente, Almelo, The Netherlands

[4]Board of Directors, Medisch Spectrum Twente, Enschede, The Netherlands

[5]See Acknowledgments

Corresponding Author:
Sanne M Krakers, MSc
Department of Biomedical Signals and Systems
University of Twente
Drienerlolaan 5
Enschede, 7522 NB
The Netherlands
Phone: 31 (0)53 489 9111
Email: s.m.krakers@utwente.nl

Abstract

Background: In 2022, over 18,000 patients aged ≥70 years were hospitalized in the Netherlands for a hip fracture, with 50% requiring geriatric rehabilitation after surgery. Increasing geriatric rehabilitation patient numbers, staff shortages, and rising pressure on health care budgets make adequate care challenging. To make geriatric rehabilitation more future-proof, a stronger focus on home-based rehabilitation is needed. Early identification of patients likely to be discharged soon enables timely discharge planning and coordination of support at home. Early geriatric rehabilitation discharge planning may help organize home-based rehabilitation more effectively by arranging home care services in advance. This can facilitate smoother transitions toward home and prevent discharge delays, which is important to ensure optimal bed occupancy.

Objective: This study aims to develop machine learning (ML) models to predict a geriatric rehabilitation stay of 4 weeks or less in a skilled nursing home for older patients after hip fracture surgery, using continuously monitored physical activity data from the first week of geriatric rehabilitation and patient characteristics.

Methods: This prospective cohort study (January 2019-August 2024) included 100 patients. Patient characteristics and physical activity data from the MOX1 accelerometer (Maastricht Instruments BV) were collected during the first rehabilitation week. Principal component analysis was used to reduce the physical activity features. ML models were developed using Bayesian hyperparameter optimization and refined if necessary. The performance of the single best-performing configuration per remaining ML model type was evaluated, and the most important features for predicting the length of geriatric rehabilitation stay were identified.

Results: Of the 3 ML models evaluated (support vector machine [SVM], ensemble of decision trees, and neural network), the SVM achieved the highest performance, with 19 out of 20 correct predictions (accuracy=0.95, 95% CI 0.85-1.00; precision=0.91, 95% CI 0.71-1.00; recall=1.00, 95% CI 1.00-1.00; F_1-score=0.95238, 95% CI 0.83-1.00; area under the curve [AUC]=0.97, 95% CI 0.83-1.00). The most important features for predicting the length of geriatric rehabilitation stay across the best-performing ML models included the continuously monitored physical activity data, time in the emergency room, functional ambulation category (FAC) at hospital discharge, age, Katz Index of Independence in Activities of Daily Living–6 (Katz-ADL6) at hospital discharge, Montreal Cognitive Assessment (MoCA), availability of nonprofessional help, surgery type, Charlson Comorbidity Index (CCI), gender, and hemoglobin level at hospital admission.

XSL•FO
RenderX

Conclusions: This study developed several ML models to predict a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home for older patients after hip fracture surgery. Among these models, the SVM proved to be highly accurate in its predictions with an accuracy of 0.95 (95% CI 0.85-1.00), precision of 0.91 (95% CI 0.71-1.00), recall of 1.00 (95% CI 1.00-1.00), F_1-score of 0.95 (95% CI 0.83-1.00), and AUC of 0.97 (95% CI 0.88-1.00).

(JMIR Rehabil Assist Technol 2026;13:e79331) doi:10.2196/79331

KEYWORDS

accelerometer; continuous physical activity monitoring; geriatric rehabilitation; hip fracture; length of stay; machine learning; prediction

Introduction

In the Netherlands, more than 18,000 patients aged 70 years and older were hospitalized with a hip fracture in 2022 [1]. After hip fracture surgery, more than 50% of the patients are admitted to a geriatric rehabilitation department at a skilled nursing home [2,3]. Geriatric rehabilitation is defined by the European Consensus group as "a multidimensional approach involving diagnostic and therapeutic interventions. Its purpose is to optimize functional capacity, promote activity, and preserve functional reserve and social participation in older people with disabling impairments" [4]. Based on incidence trends and increasing life expectancy, the number of hip fracture patients is expected to double by 2050 [5]. The expectation is that the group of geriatric rehabilitation patients will also increase. However, the number of professionals able to provide care for these patients will not grow at the same pace. In addition, health care budgets are under increasing pressure [6]. Together, these developments make the provision of geriatric rehabilitation care increasingly challenging.

To make geriatric rehabilitation future-proof, a stronger focus on home-based rehabilitation is essential, especially given the shifting vision of Dutch health care that emphasizes treatment closer to home [7]. Achieving this requires the early identification of patients who are likely to be discharged within 4 weeks of geriatric rehabilitation admission. Such early identification enables timely discharge planning and the coordination of necessary support at home. The 4-week cutoff marks the transition between shorter and longer rehabilitation trajectories, as defined by the Dutch Diagnosis Treatment Combinations. Early geriatric rehabilitation discharge planning may help organize home-based rehabilitation more effectively by arranging home care services in advance. This can facilitate smoother transitions toward home and prevent unnecessary discharge delays, which is important for maintaining optimal bed occupancy.

Early geriatric rehabilitation discharge planning requires prediction models to make objective predictions for the length of geriatric rehabilitation stay. By relying on data rather than assumptions, these models enable more accurate discharge planning. Machine learning (ML) is a subset of artificial intelligence that enables computers to learn from data without being explicitly programmed [8-10]. The use of ML for developing length of stay prediction models in various medical fields has shown promising results [8,10-12]. However, to the best of our knowledge, no ML models exist that predict the length of geriatric rehabilitation stay at a skilled nursing home for older patients following hip fracture surgery. In a recent study, we identified factors that influence the length of geriatric rehabilitation stay [13]. However, this study did not aim to predict the exact length of geriatric rehabilitation stay, and its multivariate model explained only 32% of the variance. This suggests that additional factors influencing the length of stay exist that were not included in the study. Since early mobilization and physical activity enhance functional recovery, continuously monitored physical activity using accelerometers during the first week of geriatric rehabilitation might be a relevant factor to include [14-16]. Another important factor could be cognitive functioning, assessed with a cognitive screening test, as patients with cognitive impairments may have difficulty performing physiotherapy exercises [17].

The aim of this study is to develop ML models to predict a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home for older patients after hip fracture surgery, using continuously monitored physical activity data from the first week of geriatric rehabilitation and patient characteristics.

Methods

Study Design

A prospective cohort study was conducted from the first of January 2019 until the first of August 2024. Older patients aged ≥70 years who underwent hip fracture surgery at the Centre for Geriatric Traumatology at Ziekenhuisgroep Twente (ZGT) and rehabilitated at one of the 3 participating skilled nursing homes (TriviumMeulenbeltZorg, ZorgAccent, and Carintreggeland) were included. Exclusion criteria were pathological or periprosthetic fractures, severe cognitive impairment, total hip replacement, plaster allergy, or contact isolation. Patients were enrolled one day after hip fracture surgery or one day before discharge to one of the collaborating skilled nursing homes.

Data Collection

The primary outcome was the length of geriatric rehabilitation stay at the skilled nursing homes, stratified into 2 lengths of stay groups, including ≤4 and >4 weeks. This division was based on the structure of the Dutch Diagnosis Treatment Combinations, which fund geriatric rehabilitation and classify rehabilitation trajectories into one of 5 length of stay groups (1-2, 2-4, 4-8, 8-13, and 13-17 weeks) [18,19]. The 4-week cutoff corresponds to the transition point between shorter and longer rehabilitation pathways.

For this study, the physical activity of all enrolled patients was continuously monitored during the first week of geriatric

XSL•FO
RenderX

rehabilitation. Patient characteristics up to the first week of rehabilitation were collected from a transmural care pathway database. This care pathway was implemented to synchronize the care processes between the hospital and geriatric rehabilitation departments and provide more insight into the rehabilitation process of older patients after hip fracture surgery. Patients with an incomplete data collection (<5 days of physical activity data of the first week of geriatric rehabilitation and/or missing patient characteristics) were excluded from the data analysis.

The physical activity of all enrolled patients was continuously monitored using an MOX1 accelerometer (Maastricht Instruments BV) attached with a custom-made patch 10 cm proximal of the patella on the ipsilateral side of the operated hip. This small, lightweight, waterproof device (35x35x10 mm, 11 g, IPX8) contains a tri-axial accelerometer sensor, which measures raw acceleration data for the X-, Y-, and Z-axes at a sampling frequency of 25 Hz, storing the data directly in its internal memory (1.5 GB) for up to 7 days. Data analysis was performed offline after the raw acceleration data were uploaded to a computer via a USB connection.

The collected patient characteristics (baseline, in-hospital, hospital discharge, and geriatric rehabilitation variables; Textbox 1) were selected based on previous studies demonstrating their value in predicting rehabilitation outcomes or length of stay [20-31].

Textbox 1. Patient characteristics included in this study.

Baseline variables

- Age
- Gender
- Premorbid living situation
- Living alone versus living together
- Availability of nonprofessional help
- Need for climbing stairs
- Prefracture Mobility Score (PFMS)
- Katz Index of Independence in Activities of Daily Living-6 (Katz-ADL6)
- Charlson Comorbidity Index (CCI)
- American Society of Anesthesiologists physical status classification (ASA) score
- Short Nutritional Assessment Questionnaire (SNAQ)

In-hospital variables

- Hemoglobin level at hospital admission
- Hip fracture type
- Time spent in the emergency room
- Surgical treatment
- Postoperative weight-bearing protocol
- In-hospital complications

Hospital discharge variables

- Functional Ambulation Categories (FAC) at hospital discharge
- Fracture Mobility Score (FMS) at hospital discharge
- Katz-ADL6 at hospital discharge
- Discharge destination (skilled nursing home)
- Length of hospital stay

Geriatric rehabilitation variables

- Montreal Cognitive Assessment (MoCA)

Data Analysis for Accelerometer

The raw acceleration data were analyzed using a MATLAB algorithm (R2024b; The MathWorks, Inc) developed and validated for older hospitalized patients by Van Dijk-Huisman et al [32]. This resulted in the total intensity of physical activity per day by determining the signal magnitude area of the acceleration data; the total time spent sitting or lying, standing,

XSL•FO
RenderX

and walking per day; the position classification (sitting or lying, standing, walking) every 4 seconds; as well as the total measurement time per day. The intensity of physical activity reflects the raw accelerometer output and is expressed as counts per day, not as categorical levels of activity. If the position classification changed between 2 consecutive 4-second time steps, it was defined as a transition. For each patient, the following metrics were calculated for each day:

- The intensity of physical activity (counts per hour)
- Time spent sitting or lying (minutes per hour)
- Time spent standing (minutes per hour)
- Time spent walking (minutes per hour)
- Transitions between sitting or lying and standing (number of transitions per hour)
- Transitions between sitting or lying and walking (number of transitions per hour)
- Transitions between standing and walking (number of transitions per hour)

Different features were extracted from each metric, including the overall mean, overall SD, weekend mean, weekend SD, weekday mean, weekday SD, median, IQR, root mean square, variance, minimum value, maximum value, minimum-maximum range, coefficient of variance, skewness, kurtosis, and characteristics of the third-degree polynomial curve describing the shape of the pattern. The characteristics of the third-degree polynomial curve were defined as coefficients a, b, c, and d, which were derived from the third-order polynomial equation: $y = ax^3 + bx^2 + cx + d$. More detailed information about the different continuously monitored physical activity features can be found in Multimedia Appendix 1.

After data analysis, data within each of the 2 length-of-stay groups were randomly divided into a training (80%) and testing (20%) set, which were then combined to ensure class balance.

Feature Reduction

Principal component (PC) analysis, a common method for feature dimensionality reduction, was applied exclusively to the features extracted from the physical activity data of the training set. As a preprocessing step, the data were standardized to have zero mean and unit variance. Each PC, obtained after applying PC analysis, consists of a weighted combination of the features extracted from the physical activity data. The PCs were sorted in decreasing order of their eigenvalues, indicating how much of the data's variance is captured by its corresponding PC. The PCs contributing to a cumulative explained variance of 95% were selected as input for the ML models.

Feature selection was also applied to the patient characteristics of the training set. Features that did not occur or exhibited an identical distribution between the 2 lengths of stay groups were excluded, as they provided no discriminative value for the ML models. Feature selection was further guided by prior research [13]. When variables were conceptually linked or measured similar aspects of a patient's condition, priority was given to the variable that was more commonly used in clinical practice or provided a more direct assessment of functional status. The remaining patient characteristics were incorporated into the ML models along with the selected PCs.

ML Models

First, the performance of 5 commonly used supervised ML models for classification was considered: support vector machine (SVM), neural network (NN), decision tree, naïve Bayes (NB), and an ensemble of decision trees (EoDT). These models were selected because they represent a diverse set of widely used supervised learning models, allowing for a comprehensive evaluation of which modeling strategy is most suitable for this clinical prediction task. For each model, hyperparameters were optimized using Bayesian optimization with an expected-improvement-plus acquisition function and tenfold cross-validation on the training set. Bayesian optimization was chosen because it enables efficient hyperparameter tuning and allows for the development and comparison of multiple ML models simultaneously.

For the SVM, 4 kernel functions (linear, radial basis function, polynomial, and Gaussian) were evaluated separately, allowing the kernel-related model performance to be systematically compared. For each kernel, Bayesian optimization tuned the remaining hyperparameters. For the NN, decision tree, NB, and EoDT, all model-specific hyperparameters were included in the Bayesian optimization procedure. The optimal hyperparameter values for each ML model were determined by the minimum observed value of the objective function.

Running Bayesian optimization for all models over many iterations can be computationally expensive. Therefore, model performance was evaluated after the first 30 iterations to identify a subset of well-performing models. In this screening phase, multiple configurations of the same model type, for example, the SVM models with different kernels, could be retained if they showed similar performance. The area under the receiver operating characteristic curve was used for initial model comparison, after which the models with the highest area under the curve (AUC) underwent continuous Bayesian optimization for an additional 30 iterations. Subsequently, the final hyperparameters of these ML models were evaluated. If notable deviations were identified, a new ML model was trained with adjusted hyperparameters.

Finally, only the single best-performing configuration per remaining ML model type was evaluated on the test set. Model performance was assessed using the confusion matrix, actual length of stay of the misclassified patients, accuracy score, precision score, recall score, F_1-score, and AUC. CIs were obtained using bootstrap resampling; the test set was resampled with replacements 1000 times while keeping the training set and ML models fixed. The 95% CIs were defined by the 2.5th and 97.5th percentiles. A Shapley Additive Interpretation (SHAP) summary plot based on the test set was used to determine the relationship between the 2 length of stay groups and its main predictors for each of the best-performing ML models [33]. The overall best-performing ML model was retrained and retested using 2 additional dataset configurations, (1) patient characteristics only and (2) physical activity data only, to conduct an ablation study evaluating the contribution of each dataset compared with the combined datasets.

XSL•FO
RenderX

Ethical Considerations

The Medical Ethical Committee of Twente stated that this study did not require an assessment by the Medical Ethical Committee according to Dutch law. The study was approved by the Institutional Review Board of ZGT (ZGT17-40). All patients gave written informed consent to participate. Study data obtained from the enrolled patients were deidentified and managed in accordance with relevant data protection regulations. Patients did not receive any financial or material compensation for their participation.

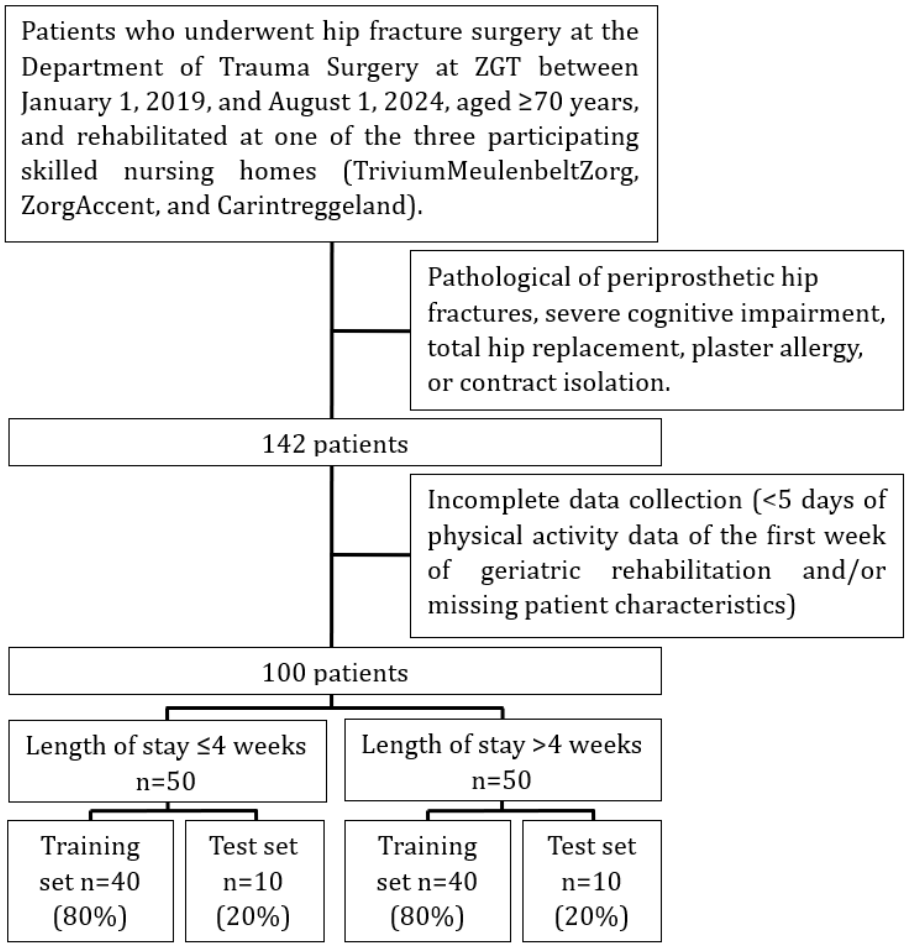
Results

Patient Characteristics and Physical Activity Data

A total of 142 older patients were included in this study (Figure 1). Of these, 42 were excluded due to incomplete data (<5 days of physical activity data of the first week of geriatric rehabilitation and/or missing patient characteristics). The missing physical activity data were mainly caused by sensor-related issues, such as battery depletion and removal of the device for charging and data upload. This resulted in a final sample of 100 patients. Patients were divided into ≤4 weeks (n=50) and >4 weeks (n=50) length of geriatric rehabilitation stay and subsequently split into a training and test set. The distribution of the length of geriatric rehabilitation stay of both sets can be seen in Figure 2. Patient characteristics of the training and test set are shown in Table 1. The mean age in the ≤4 weeks group was 84.5 (SD 5.9) years, while the mean age in the >4 weeks group was 82.8 (SD 6.9) years. The median Montreal Cognitive Assessment (MoCA) score was 21.5 (IQR 18-24) and 19.5 (IQR 17-22.5), respectively. The patient characteristics of the test set showed similar results.

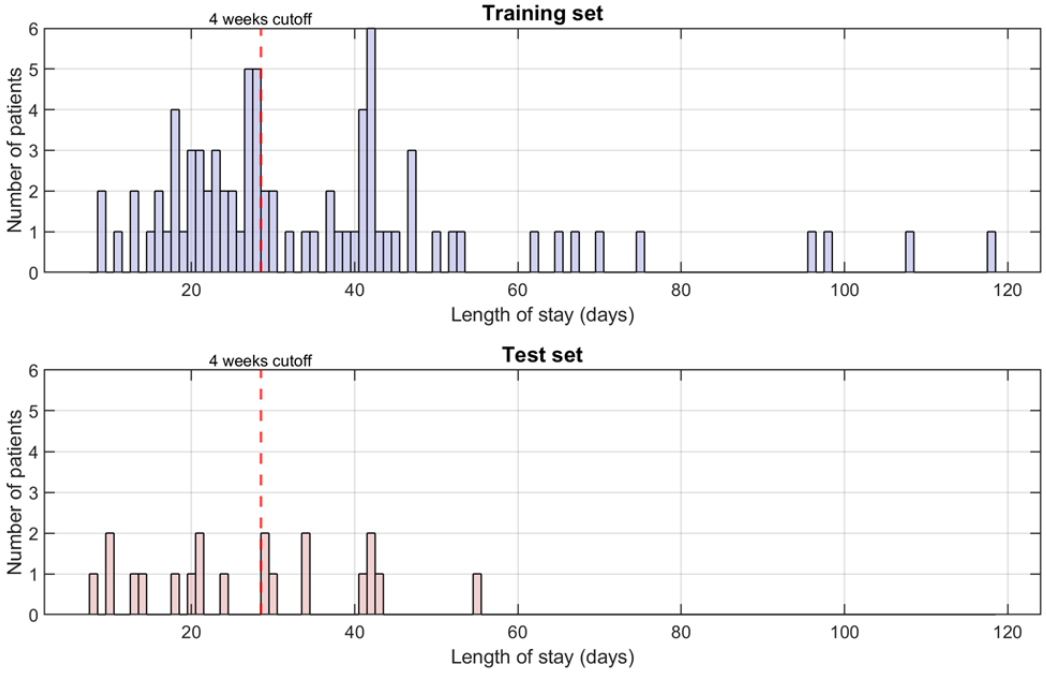
Individual patient trajectories of intensity of physical activity during the first 7 days of geriatric rehabilitation are displayed, along with the mean intensity for patients with a length of stay of 4 weeks or less and for those with more than 4 weeks, including 95% CI (Figure 3).

Figure 1. Flowchart patient recruitment.



XSL•FO
RenderX

Figure 2. Length of geriatric rehabilitation stay (days): training set vs test set.



XSL•FO
RenderX

Table 1. Patient characteristics training and test set.

Patient characteristics training and test set	Training set		Test set	
	Length of geriatric rehabilitation stay		Length of geriatric rehabilitation stay	
	≤4 weeks (n=40)	>4 weeks (n=40)	≤4 weeks (n=10)	>4 weeks (n=10)
Baseline variables				
Age (years)[a], mean (SD)	84.5 (5.9)	82.8 (6.9)	82.8 (6.0)	82.2 (6.9)
Female (gender)[a], n (%)	27 (68)	33 (83)	7 (70)	8 (80)
Premorbid living situation[a], n (%)				
Independent without home care services	30 (75)	26 (65)	8 (80)	9 (90)
Independent with home care services	9 (23)	14 (35)	1 (10)	1 (10)
Residential home	1 (3)	0 (0)	1 (10)	0 (0)
Living together[a], n (%)	12 (30)	14 (35)	2 (20)	7 (70)
Availability nonprofessional help[a], n (%)	31 (78)	38 (95)	6 (60)	7 (70)
Need for climbing stairs[a], n (%)	11 (28)	7 (18)	4 (40)	3 (30)
PFMS[a,b], n (%)				
1-Freely mobile without aids	20 (50)	11 (28)	5 (50)	4 (40)
2-Mobile outdoors with one aid	4 (10)	1 (3)	1 (10)	2 (20)
3-Mobile outdoors with 2 aids or frame	16 (40)	27 (68)	4 (40)	4 (40)
4-Some indoor mobility but never goes outside without help	0 (0)	1 (3)	0 (0)	0 (0)
5-No functional mobility	0 (0)	0 (0)	0 (0)	0 (0)
Premorbid Katz-ADL6[a,c], median (IQR)	0 (0-0)	0 (0-1)	0 (0-0)	0 (0-1)
CCI[a,d], median (IQR)	1 (0-2)	1 (0-2)	0 (0-1)	1 (0-3)
ASA[e]score[a], n (%)				
ASA 1	0 (0)	2 (5)	1 (10)	1 (10)
ASA 2	15 (38)	10 (25)	4 (40)	1 (10)
ASA 3	23 (58)	22 (55)	3 (30)	6 (60)
ASA 4	2 (5)	6 (15)	2 (20)	2 (20)
ASA 5	0 (0)	0 (0)	0 (0)	0 (0)
SNAQ[a,f], median (IQR)	0 (0-0)	0 (0-0)	0 (0-0)	0 (0-0)
In-hospital variables				
Hemoglobin level at hospital admission[a], median (IQR)	7.9 (7.6-8.7)	7.8 (7.5-8.6)	7.9 (7.5-8.1)	7.9 (6.0-8.3)
Hip fracture type, n (%)				
Femoral neck	24 (60)	17 (43)	5 (50)	3 (30)
Subtrochanteric femur	16 (40)	23 (58)	5 (50)	7 (70)
Time in the emergency room in minutes[a], median (IQR)	133.5 (103.5-168.0)	133.0 (106.5-161.0)	104.5 (82.0-155.0)	158.5 (124.0-184.0)
Surgical treatment[a], n (%)				
Dynamic hip screw	3 (8)	3 (8)	2 (20)	1 (10)
Intramedullary nail	15 (38)	25 (63)	5 (50)	6 (60)
Hemiarthroplasty	22 (55)	12 (30)	3 (30)	3 (30)

XSL•FO
RenderX

Patient characteristics training and test set	Training set		Test set	
	Length of geriatric rehabilitation stay		Length of geriatric rehabilitation stay	
	≤4 weeks (n=40)	>4 weeks (n=40)	≤4 weeks (n=10)	>4 weeks (n=10)
Postoperative weight-bearing protocol[a], n (%)				
Full weight-bearing	40 (100)	36 (90)	10 (100)	9 (90)
Partial or non–weight-bearing	0 (0)	4 (10)	0 (0)	1 (10)
In-hospital complications, n (%)				
Anemia[a]	6 (15)	7 (18)	1 (10)	3 (30)
Heart failure[a]	1 (3)	2 (5)	0 (0)	1 (10)
Pressure ulcers[a]	3 (8)	2 (5)	0 (0)	0 (0)
Delirium[a]	1 (3)	5 (13)	0 (0)	0 (0)
Pulmonary embolism[a]	0 (0)	1 (3)	0 (0)	0 (0)
Kidney failure[a]	0 (0)	1 (3)	0 (0)	0 (0)
Pneumonia[a]	3 (8)	4 (10)	0 (0)	0 (0)
Urinary tract infection[a]	2 (5)	1 (3)	1 (10)	0 (0)
Fall incident	0 (0)	0 (0)	0 (0)	0 (0)
Wound infection	0 (0)	0 (0)	0 (0)	0 (0)
Reoperation	0 (0)	0 (0)	0 (0)	0 (0)
Others	8 (20)	8 (20)	3 (30)	0 (0)
Hospital discharge variables				
FAC[g] at hospital discharge[a], n (%)				
0-No functional mobility	0 (0)	5 (13)	0 (0)	1 (10)
1-Dependent in mobility level 2	1 (3)	4 (10)	0 (0)	4 (40)
2-Dependent in mobility level 1	9 (23)	17 (43)	0 (0)	2 (20)
3-Independent mobility under supervision	22 (55)	9 (23)	8 (80)	1 (10)
4-Independent mobility on a flat surface	8 (20)	5 (13)	2 (20)	2 (20)
FMS[h] at hospital discharge, n (%)				
1-Freely mobile without aids	0 (0)	0 (0)	0 (0)	0 (0)
2-Mobile outdoors with one aid	0 (0)	0 (0)	0 (0)	0 (0)
3-Mobile outdoors with 2 aids or frame	3 (8)	2 (5)	1 (10)	0 (0)
4-Some indoor mobility but never goes outside without help	37 (93)	33 (83)	9 (90)	9 (90)
5-No functional mobility	0 (0)	5 (13)	0 (0)	1 (10)
Katz-ADL6 at hospital discharge[a], median (IQR)	4 (4-4)	4 (4-5)	4.5 (4-5)	4 (4-5)
Rehabilitation department skilled nursing home, n (%)				
Skilled nursing home A	22 (55)	23 (58)	9 (90)	5 (50)
Skilled nursing home B	12 (30)	13 (33)	1 (10)	3 (30)
Skilled nursing home C	6 (15)	4 (10)	0 (0)	2 (20)
Length of hospital stay in days[a], median (IQR)	7 (6-9)	8 (6.5-9)	6.5 (6-8)	7 (6-8)
Geriatric rehabilitation variables				

XSL•FO
RenderX

Patient characteristics training and test set	Training set		Test set	
	Length of geriatric rehabilitation stay		Length of geriatric rehabilitation stay	
	≤4 weeks (n=40)	>4 weeks (n=40)	≤4 weeks (n=10)	>4 weeks (n=10)
MoCA[a,i], median (IQR)	21.5 (18-24)	19.5 (17-22.5)	21 (20-25)	23 (18-26)

[a]Patient characteristics included in the ML models.

[b]PFMS: Prefracture Mobility Score.

[c]Katz-ADL6: Katz Index of Independence in Activities of Daily Living-6.

[d]CCI: Charlson Comorbidity Index.

[e]ASA: American Society of Anesthesiologists physical status classification.

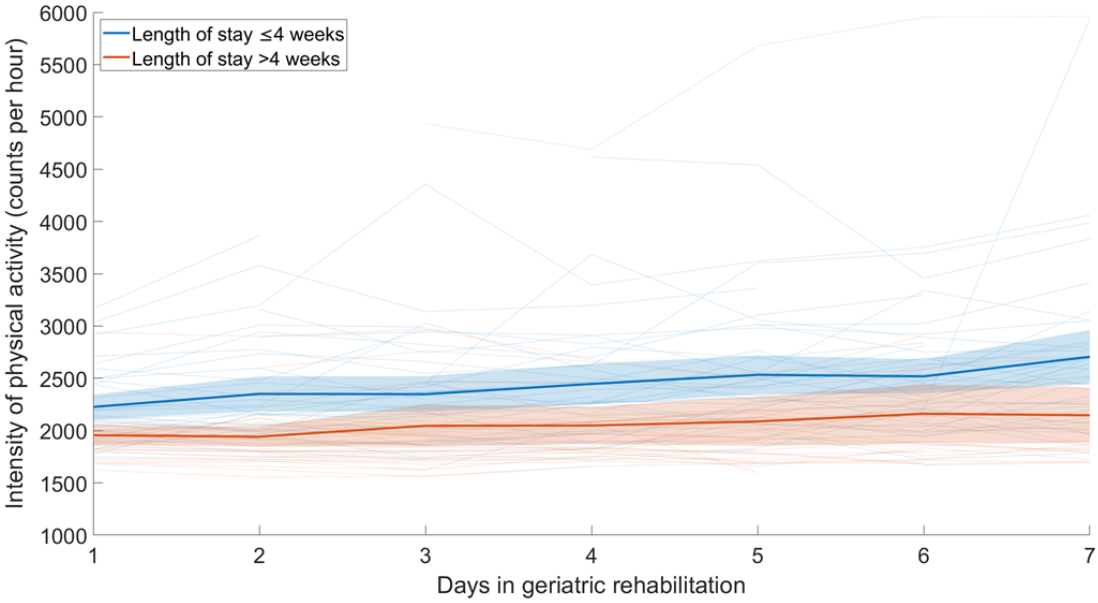
[f]SNAQ: Short Nutritional Assessment Questionnaire.

[g]FAC: Functional Ambulation Categories.

[h]FMS: Fracture Mobility Score.

[i]MoCA: Montreal Cognitive Assessment.

Figure 3. Intensity of physical activity with individual trajectories of the training set and mean physical activity intensity with 95% CI.



Feature Reduction

After applying PC analysis for the feature reduction of the features extracted from the physical activity data, 79 PCs were obtained. The graph of cumulative explained variance versus the number of PCs shows that 19 PCs explained 95% of the variance (Multimedia Appendix 2; Figure S1). These 19 PCs were included in the ML models.

In-hospital fall incidents, in-hospital wound infections, in-hospital reoperations, and other in-hospital complications were excluded because none of these complications occurred or the distributions were identical between the 2 lengths of stay groups. Hip fracture type was excluded from the ML models because it is strongly associated with surgical treatment [13]. Surgical treatment was regarded to be more patient-specific, as it depends on both fracture type and patient condition. FMS at hospital discharge was also excluded due to its strong association with functional ambulation category (FAC) at hospital discharge [13]. The FAC was considered more relevant because it is more widely used in geriatric rehabilitation to assess patient progress. Similarly, admission to which skilled nursing home was excluded due to its association with FAC at hospital discharge [13]. Notably, skilled nursing home A had more patients with lower FAC scores at hospital discharge compared to B and C. This was a coincidence, as the skilled nursing home placement procedure does not depend on the FAC. The FAC at hospital discharge was prioritized for its greater clinical relevance.

ML Models

Based on the AUC of each ML model after the first 30 iterations (Multimedia Appendix 2; Figure S2), the SVM, NN, and EoDT ML models were selected for continuous Bayesian optimization.

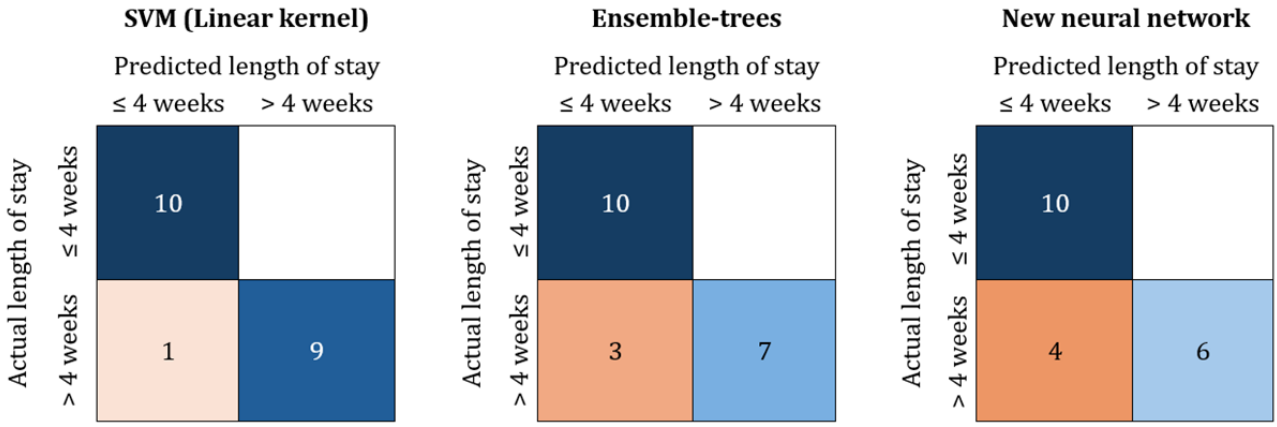
The evaluation of the final hyperparameters after Bayesian optimization resulted in an NN consisting of only one layer with 2 neurons. Since this architecture was considered too simple and likely insufficient for capturing the complexity of the data, a new NN with 2 layers containing 102 and 51 neurons, respectively, was trained for further evaluation.

Figure 4 presents the confusion matrices of the final best-performing configuration of each remaining ML model

XSL•FO
RenderX

type (SVM, EoDT, and NN) for the test set. The confusion matrix evaluates the performance of the ML models by comparing the actual length of stay with the predicted length of stay generated by the models. This provides an overview of the number of correctly classified and misclassified patients.

Figure 4. Confusion matrices of the final best-performing machine learning (ML) models (support vector machine [SVM], ensemble of decision trees [EoDT], and neural network [NN]) tested on the test set (n=20).



The performance of the final best-performing ML models, along with the actual length of stay for patients misclassified as ≤4 weeks, is presented in Table 2. The overall best results were achieved using the SVM model with a linear kernel (accuracy=0.95, 95% CI 0.85-1.00; precision=0.91, 95% CI 0.71-1.00; recall=1.00, 95% CI 1.00-1.00; F_1-score=0.95238, 95% CI 0.83-1.00; AUC=0.97, 95% CI 0.83-1.00). The misclassified ≤4 weeks group included one patient with an actual length of stay of 42 days. The EoDT misclassified 3 patients with lengths of stay of 29, 29, and 42 days, while the new NN misclassified 4 patients with lengths of stay of 29, 30, 41, and 42 days.

Table 2. Evaluation of final best-performing machine learning (ML) models (SVM[a], EoDT[b], and neural network).

Machine learning models	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)	AUC[c] (95% CI)	Actual LOS[d] misclassified ≤4 weeks patients
SVM (linear kernel)	0.95 (0.85-1.00)	0.91 (0.71-1.00)	1.00 (1.00-1.00)	0.95 (0.83-1.00)	0.97 (0.88-1.00)	42 days
Ensemble-trees	0.85 (0.70-1.00)	0.77 (0.50-1.00)	1.00 (1.00-1.00)	0.87 (0.67-1.00)	0.97 (0.88-1.00)	29, 29, and 42 days
New neural network	0.80 (0.60-0.95)	0.71 (0.47-0.93)	1.00 (1.00-1.00)	0.83 (0.64-0.97)	0.92 (0.78-1.00)	29, 30, 41, and 42 days

[a]SVM: support vector machine.

[b]EoDT: ensemble of decision trees.

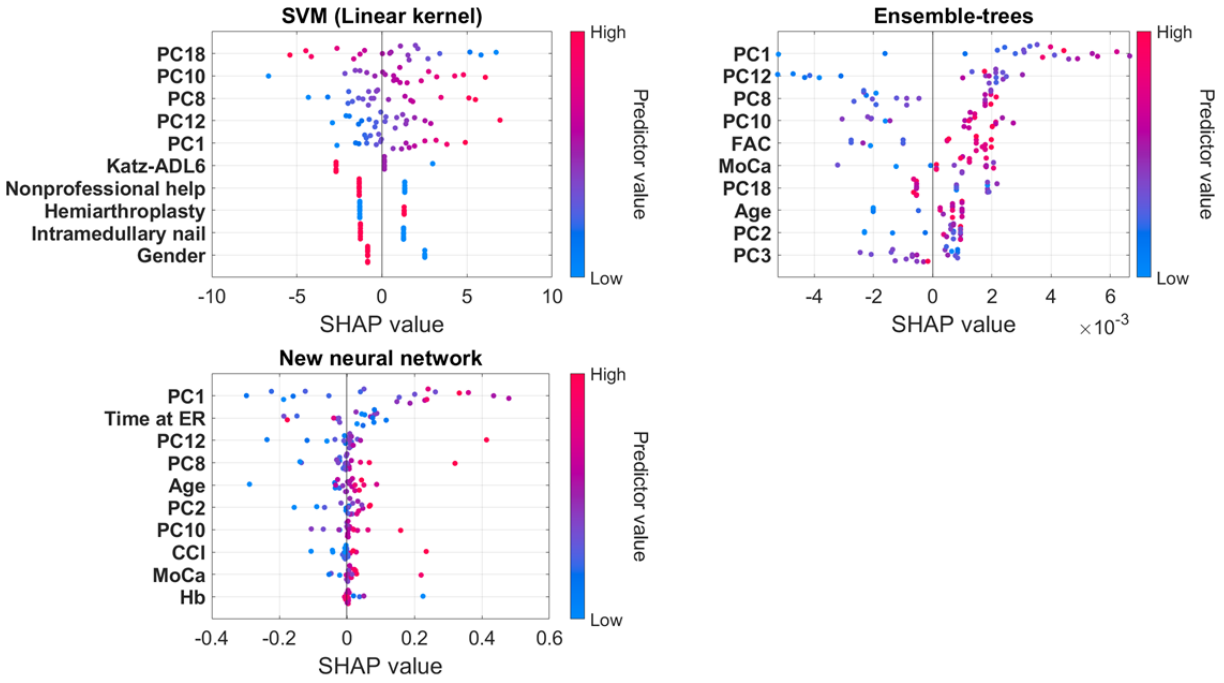
[c]AUC: area under the curve.

[d]LOS: length of stay.

The 10 most important features of the final best-performing ML models, ranked from most to least impactful, and their influence on predictions can be seen in the results of the SHAP analysis (Figure 5). Each dot represents a patient from the test set, with its color indicating the actual feature value (predictor value: red for high values and blue for low values). The SHAP value of each dot represents the contribution of that feature to the prediction for an individual patient. It indicates how much the feature pushes the prediction toward one length of stay group or the other. Negative SHAP values push the prediction toward a length of stay of ≤4 weeks, while positive SHAP values push it toward a length of stay of >4 weeks.

XSL•FO
RenderX

Figure 5. Shapley Additive Interpretation (SHAP) summary plots: the 10 most important features ranked from most to least impactful and their influence on predictions.



For example, in the Shapley summary plot of the SVM (linear kernel) model, the feature availability of nonprofessional help appears blue on the right. This indicates that patients with no availability of nonprofessional help (0=no availability, 1=availability) have positive SHAP values, pushing predictions toward a length of stay of >4 weeks. Based on the SHAP analysis, it can be seen that the most important features for predicting the length of geriatric rehabilitation stay, across the 3 best-performing ML models, were the PCs (weighted combinations of features extracted from the physical activity data); time in the emergency room (ER); FAC at hospital discharge; age; Katz Index of Independence in Activities of Daily Living-6 (Katz-ADL6) at hospital discharge; MoCA; availability of nonprofessional help; surgery type; Charlson Comorbidity Index (CCI); gender; and hemoglobin level at hospital admission.

In the ablation study, the best-performing ML model trained on physical activity data only outperformed the model trained on patient characteristics only. The combined dataset yielded the highest predictive performance. The confusion matrices and detailed performance metrics for the ablation study are provided in Multimedia Appendix 3.

Discussion

Principal Findings

This study developed ML models for the early prediction of a geriatric rehabilitation stay of ≤4 weeks in older patients after hip fracture surgery, using continuously monitored physical activity data and patient characteristics. This resulted in several ML models with varying overall predictive performance. The SVM was the best-performing model, demonstrating excellent predictive capability with an accuracy of 0.95 (95% CI 0.85-1.00), precision of 0.91 (95% CI 0.71-1.00), recall of 1.00 (95% CI 1.00-1.00), F_1-score of 0.95 (95% CI 0.83-1.00), and AUC of 0.97 (95% CI 0.88-1.00). These initial results suggest that early prediction of a geriatric rehabilitation stay of ≤4 weeks is promising.

In our previous study, the FAC at hospital discharge, premorbid living situation, postoperative weight-bearing protocol, surgery type, in-hospital delirium, and in-hospital heart failure were identified as key factors affecting the length of geriatric rehabilitation stay [13]. Among all these previously identified predictors, the current study confirmed that FAC at hospital discharge and surgery type are key features for predicting a geriatric rehabilitation stay of ≤4 weeks. Postoperative weight-bearing protocol, in-hospital delirium, and in-hospital heart failure did not emerge as key predictors in the best-performing models. A possible explanation is that these factors occurred relatively infrequently in our study population, limiting their predictive power in this context. Furthermore, whereas the previous study focused solely on identifying predictors for the length of geriatric rehabilitation stay, the current study developed and evaluated actual prediction models, offering potential for early discharge planning.

A substantial amount of other research also exists on the factors influencing the length of stay after hip fracture surgery. However, many of these studies focus on factors influencing the length of hospital stay [28,30], or the length of stay in other rehabilitation settings (eg, private care rehabilitation, in-hospital rehabilitation, or both the acute hospital and rehabilitation phases) [21-24], making them difficult to compare with this study. These previously identified predictors for the length of geriatric rehabilitation stay include age [22-24]; fracture type [22]; American Society of Anesthesiologists physical status classification score [20,21]; complications (wound infection,

XSL•FO
RenderX

delirium, urinary tract infection, and pneumonia) [22,23]; comorbidities (Parkinson disease, diabetes, and dementia) [22-24]; gait status at hospital discharge [20]; living situation prior to the hip fracture (alone or together) [24,25]; use of mobility aids prior to the hip fracture [21,25]; Barthel score at geriatric rehabilitation admission [25]; Functional Independence Measure score at rehabilitation admission [23]; Abbreviated Mental Test Score (AMTS) [21]; and pain score at rehabilitation admission [24]. Among these previously identified predictors, this study further confirmed that age, comorbidities (CCI), gait status at hospital discharge (FAC), an assessment tool to evaluate a patient's functional ability to perform activities of daily living (Barthel score or Katz-ADL6), and a cognitive score (AMTS or MoCA) are key features for predicting a geriatric rehabilitation stay of ≤4 weeks.

Unlike previous studies, which did not incorporate continuously measured physical activity, this study included this aspect, providing a more detailed insight into the effect of physical activity on the length of geriatric rehabilitation stay. Consequently, the PCs (weighted combinations of features extracted from the physical activity data) of this study emerged as one of the most important features for predicting a geriatric rehabilitation stay of ≤4 weeks. The findings of this study underscore the critical role of physical activity and align with previous studies emphasizing the positive impact of early mobilization and physical activity on the functional recovery of hip fracture patients [14-16,34]. Recent research further supports this by demonstrating that higher levels of physical activity at the time of rehabilitation admission are associated with greater independence in activities of daily living at discharge [35]. Early mobilization may help maintain postural muscle function, leading to improved functional recovery [35]. In contrast, prolonged bed rest, associated with low levels of physical activity, has been linked to a decline in physical function and reduced independence in activities of daily living due to decreased oxygen consumption, metabolic changes, and muscle atrophy [36,37]. Patients with lower physical activity levels at rehabilitation admission may experience slower functional recovery, prolonged dependence in activities of daily living, and consequently a longer geriatric rehabilitation stay.

In addition to physical activity, this study identified other important predictors of the length of geriatric rehabilitation stay that have not been examined before in previous research. Specifically, the availability of nonprofessional help, time in the ER, and hemoglobin level at hospital admission emerged as important predictors. The availability of nonprofessional help was included as it represents a key aspect of functional social support (social relationships that fulfill particular functions in times of need), which has been shown to be associated with rehabilitation outcomes in older adults with hip fractures [27]. Time spent in the ER was included, as recent research has found an association with hospital length of stay [30], although the reason for this association remains unclear. Hemoglobin level at hospital admission was also included, as recent research has found an association with hospital length of stay due to delayed recovery and increased complication risk in patients with low hemoglobin levels [28].

The remaining most important feature across all ML models, gender, also does not align with previously identified predictors of the length of geriatric rehabilitation stay. However, gender has been included as a variable in earlier studies, although it was not found to be a significant predictor. A possible explanation for its role as a key feature in this study, but not in previous ones, may lie in the use of different analytical techniques. In this study, we used ML techniques, whereas previous research relied on traditional statistical analyses.

Overall, while many of the key features for predicting a geriatric rehabilitation stay of ≤4 weeks identified in this study align with previous research, this study also uncovered other predictors, including weighted combinations of features extracted from continuously measured physical activity data, availability of nonprofessional help, and time spent in the ER. These findings highlight the potential influence of factors often overlooked in traditional statistical analyses.

When examining the misclassified ≤4 weeks patients of the 3 best-performing ML models in more detail, the EoDT has 2 patients with a length of stay of 29 days, and the new NN has one patient with a length of stay of 29 days and another with a length of stay of 30 days. However, misclassifying a 29- and 30-day stay as ≤4 weeks may not be a clinically significant issue, as it is close to 28 days. Furthermore, the new NN misclassified a patient with a length of stay of 41 days. However, this patient developed a complication during rehabilitation, which may have contributed to the incorrect prediction. Since complications occurring after the first seven days of geriatric rehabilitation are not taken into account in the models, this may explain the misclassification. In clinical practice, however, treatment plans are adjusted when complications arise, meaning that such a case would likely have been identified as requiring a longer stay regardless. The other 3 misclassified patients across the 3 best-performing ML models had a length of stay of 42 days. No clear explanation was found based on the available data.

Strengths and Limitations of This Study

To the best of our knowledge, this is the first study to develop and compare ML models for predicting a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home at an early rehabilitation stage, using a novel experimental approach with continuously monitored physical activity data and patient characteristics in older patients after hip fracture surgery. Another strength of our study is the inclusion of a diverse range of features from both the physical and social domains, with particular emphasis on the integration of continuously monitored physical activity data. In addition, the MOX1 accelerometer was well tolerated by patients and was easily integrated into the clinical workflow, requiring minimal effort from clinical staff and demonstrating the feasibility of continuous activity monitoring in this population.

Regarding the limitations, a complete dataset was required for ML model development, leading to the exclusion of many patients due to missing a single feature. No data imputation was applied in this study, as the aim was to avoid introducing potential bias from imputed values. This requirement for complete data reduced the sample size. A practical limitation

XSL•FO
RenderX

contributing to missing physical activity data was the inability to check the battery level of the MOX1 accelerometer during use. Although the lights of the accelerometer show whether it is actively measuring, clinical staff did not routinely check this, as the wearable required no daily attention. As a result, wearables occasionally ran out of battery unnoticed, leading to the loss of multiple days of activity data. While model performance was already high, having complete data for all patients could have potentially increased the reliability and robustness of the analyses. Second, given the focus on ML model development, the sample size may be relatively small, making some models prone to overfitting. To mitigate this risk, techniques such as cross-validation and feature reduction were applied. Third, although length of geriatric rehabilitation stay is a continuous variable, this study formulated the problem as a classification task (≤4 weeks vs >4 weeks). This cutoff reflects a clinically meaningful decision point with substantial impact on early discharge planning. For clinicians, knowing after one week whether a patient is likely to be discharged within 4 weeks is far more actionable than attempting to predict the exact number of days. Predicting the exact length of stay in days would introduce unnecessary complexity, require substantially more data, and is not essential for guiding early clinical decisions. Therefore, we consider the current classification framework to be the most suitable and clinically relevant approach for guiding early discharge planning. Future research is recommended to explore appropriate data imputation strategies in a cautious manner, to reduce data loss while ensuring that bias is not introduced. For categorical patient characteristics such as the availability of nonprofessional help, imputation may be challenging and not always be desirable, highlighting the need for careful consideration of which variables to impute. In addition, future work is recommended to validate the models with a larger patient population, refine them if necessary, and integrate the best-performing model into clinical practice.

Clinical Interpretation

The results of this study indicate that early prediction of a geriatric rehabilitation stay of ≤4 weeks or less in a skilled nursing home for older patients after hip fracture surgery is feasible. This marks the first step toward a more future-proof geriatric rehabilitation system. Identifying patients with an expected stay of ≤4 weeks or less after 7 days of rehabilitation enables timely discharge planning and coordination of support at home. Such planning can smooth the transition toward home and prevent discharge delays, which is important to ensure optimal bed occupancy.

Conclusion

Several ML models were developed to predict a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home for older patients after hip fracture surgery, using continuously monitored physical activity data from the first week of geriatric rehabilitation and patient characteristics. Among these models, the SVM proved to be highly accurate in its predictions with an accuracy of 0.95 (95% CI 0.85-1.00), precision of 0.91 (95% CI 0.71-1.00), recall of 1.00 (95% CI 1.00-1.00), F_1-score of 0.95 (95% CI 0.83-1.00), and AUC of 0.97 (95% CI 0.88-1.00).

Acknowledgments

The authors would like to thank the Up&Go after a Hip Fracture Group, including the health care professionals of the rehabilitation departments of the nursing home institutions TriviumMeulenbeltZorg, ZorgAccent, and Carintreggeland for their contribution to the transmural Up&Go after a hip fracture project.

Funding

This research received no external funding.

Data Availability

The datasets analyzed during this study are not publicly available due to the fact that our patients did not give permission to share the data publicly but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: SK, FJW, DvD, MMRV-H, JHH
Data curation: SK
Formal analysis: SK, FJW
Investigation: SK
Methodology: SK, FJW
Supervision: FJW, DvD, MMRV-H, JHH
Validation: SK, FJW
Visualization: SK
Writing – original draft: SK
Writing – review & editing: SK, FJW, DvD, MMRV-H, JHH

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed explanation of the extracted continuously monitored physical activity features.
[DOCX File , 19 KB - rehab_v13i1e79331_app1.docx]

Multimedia Appendix 2
Feature reduction and machine learning model development.
[DOCX File , 312 KB - rehab_v13i1e79331_app2.docx]

Multimedia Appendix 3
Confusion matrices and detailed performance metrics for the ablation study.
[DOCX File , 40 KB - rehab_v13i1e79331_app3.docx]

References

1. CBS. Hospital admissions and patients; diagnostic classification VTV. Statistics Netherlands (CBS). 2024. URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84067NED/table?ts=1570535289637 [accessed 2025-01-28]
2. Mattiazzo GF, Drewes YM, van Eijk M, Achterberg WP. Geriatric rehabilitation care after hip fracture. Eur Geriatr Med 2023 Apr;14(2):295-305 [FREE Full text] [doi: 10.1007/s41999-023-00755-4] [Medline: 36788193]
3. van Dartel D, Vermeer M, Folbert EC, Arends AJ, Vollenbroek-Hutten MMR, Hegeman JH, Dutch Hip Fracture Audit (DHFA) Group. Early Predictors for Discharge to Geriatric Rehabilitation after Hip Fracture Treatment of Older Patients. J Am Med Dir Assoc 2021 Dec;22(12):2454-2460 [FREE Full text] [doi: 10.1016/j.jamda.2021.03.026] [Medline: 33933417]
4. Grund S, Gordon AL, van Balen R, Bachmann S, Cherubini A, Landi F, et al. European consensus on core principles and future priorities for geriatric rehabilitation: consensus statement. Eur Geriatr Med 2020 Apr;11(2):233-238 [FREE Full text] [doi: 10.1007/s41999-019-00274-1] [Medline: 32297191]
5. Sing CW, Lin TC, Bartholomew S, Bell JS, Bennett C, Beyene K, et al. Global Epidemiology of Hip Fractures: Secular Trends in Incidence Rate, Post-Fracture Treatment, and All-Cause Mortality. J Bone Miner Res 2023 Aug;38(8):1064-1075 [FREE Full text] [doi: 10.1002/jbmr.4821] [Medline: 37118993]
6. Minderhout RN, van Ede AF, Voragen L, Verheijen C, Vos HM, Numans ME, et al. Reforming healthcare in the Netherlands: practical population health management and the Plot model: A questionnaire survey and focus group study to assess the willingness and readiness of six regions in the Netherlands. SAGE Open Med 2023;11:20503121231160830 [FREE Full text] [doi: 10.1177/20503121231160830] [Medline: 36949828]
7. Integraal Zorgakkoord: 'samenwerken aan gezonde zorg'. The National Government. For the Netherlands. 2022. URL: https://www.rijksoverheid.nl/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg [accessed 2025-01-28]
8. Lin WT, Wu TY, Chen YJ, Chang YS, Lin CH, Lin YJ. Predicting in-hospital length of stay for very-low-birth-weight preterm infants using machine learning techniques. J Formos Med Assoc 2022 Jun;121(6):1141-1148 [FREE Full text] [doi: 10.1016/j.jfma.2021.09.018] [Medline: 34629242]
9. Deo RC. Machine Learning in Medicine. Circulation 2015 Nov 17;132(20):1920-1930 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.001593] [Medline: 26572668]
10. Tian CW, Chen XX, Shi L, Zhu HY, Dai GC, Chen H, et al. Machine learning applications for the prediction of extended length of stay in geriatric hip fracture patients. World J Orthop 2023 Oct 18;14(10):741-754 [FREE Full text] [doi: 10.5312/wjo.v14.i10.741] [Medline: 37970626]
11. Daghistani TA, Elshawi R, Sakr S, Ahmed AM, Al-Thwayee A, Al-Mallah MH. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. Int J Cardiol 2019 Aug 01;288:140-147. [doi: 10.1016/j.ijcard.2019.01.046] [Medline: 30685103]
12. Muhlestein WE, Akagi DS, Davies JM, Chambless LB. Predicting Inpatient Length of Stay After Brain Tumor Surgery: Developing Machine Learning Ensembles to Improve Predictive Performance. Neurosurgery 2019 Sep 01;85(3):384-393 [FREE Full text] [doi: 10.1093/neuros/nyy343] [Medline: 30113665]
13. Krakers SM, Woudsma S, van Dartel D, Vermeer M, Vollenbroek-Hutten MMR, Hegeman JH. Rehabilitation of Frail Older Adults after Hip Fracture Surgery: Predictors for the Length of Geriatric Rehabilitation Stay at a Skilled Nursing Home. J Clin Med 2024 Aug 03;13(15):4547 [FREE Full text] [doi: 10.3390/jcm13154547] [Medline: 39124813]
14. Riemen AH, Hutchison JD. The multidisciplinary management of hip fractures in older patients. Orthop Trauma 2016 Apr;30(2):117-122 [FREE Full text] [doi: 10.1016/j.mporth.2016.03.006] [Medline: 27418950]
15. Morri M, Forni C, Marchioni M, Bonetti E, Marseglia F, Cotti A. Which factors are independent predictors of early recovery of mobility in the older adults' population after hip fracture? A cohort prognostic study. Arch Orthop Trauma Surg 2018 Jan;138(1):35-41. [doi: 10.1007/s00402-017-2803-y] [Medline: 28956152]

XSL•FO
RenderX

16. van Dartel D, Wang Y, Hegeman JH, Vermeer M, Vollenbroek-Hutten MMR. Patterns of physical activity over time in older patients rehabilitating after hip fracture surgery: a preliminary observational study. BMC Geriatr 2023 Jun 16;23(1):373 [FREE Full text] [doi: 10.1186/s12877-023-04054-2] [Medline: 37328743]
17. Makizako H, Shimada H, Doi T, Tsutsumimoto K, Lee S, Hotta R, et al. Cognitive functioning and walking speed in older adults as predictors of limitations in self-reported instrumental activity of daily living: prospective findings from the Obu Study of Health Promotion for the Elderly. Int J Environ Res Public Health 2015 Mar 11;12(3):3002-3013 [FREE Full text] [doi: 10.3390/ijerph120303002] [Medline: 25768239]
18. van de Haar HJ, Alme MN. Geriatric rehabilitation in the Netherlands and their preparedness for the increase in older adults: a scoping review. Discov Public Health 2025 Jan 13;22(1):13. [doi: 10.1186/s12982-025-00399-8]
19. Frowijn T, Vermeer M, Koop R, Schreuder R. USER draagt bij aan voorspelling revalidatieduur GRZ. Tijdschrift voor Ouderengeneeskunde 2016:4 [FREE Full text]
20. Adrados M, Wang K, Deng Y, Bozzo J, Messina T, Stevens A, et al. A Simple Physical Therapy Algorithm Is Successful in Decreasing Skilled Nursing Facility Length of Stay and Increasing Cost Savings After Hip Fracture With No Increase in Adverse Events. Geriatr Orthop Surg Rehabil 2021;12:2151459321998615 [FREE Full text] [doi: 10.1177/2151459321998615] [Medline: 33815865]
21. Richards T, Glendenning A, Benson D, Alexander S, Thati S. The independent patient factors that affect length of stay following hip fractures. Ann R Coll Surg Engl 2018 Sep;100(7):556-562 [FREE Full text] [doi: 10.1308/rcsann.2018.0068] [Medline: 29692191]
22. Ireland AW, Kelly PJ, Cumming RG. Total hospital stay for hip fracture: measuring the variations due to pre-fracture residence, rehabilitation, complications and comorbidities. BMC Health Serv Res 2015 Jan 22;15(1):17 [FREE Full text] [doi: 10.1186/s12913-015-0697-3] [Medline: 25609030]
23. Thornburgh Z, Samuel D. Factors Influencing Length of Stay and Discharge Destination of Patients with Hip Fracture Rehabilitating in a Private Care Setting. Geriatrics (Basel) 2022 Mar 31;7(2):44 [FREE Full text] [doi: 10.3390/geriatrics7020044] [Medline: 35447847]
24. Daly N, Fortin C, Jaglal S, MacDonald SL. Predictors of Exceeding Target Inpatient Rehabilitation Length of Stay After Hip Fracture. Am J Phys Med Rehabil 2020;99(7):630-635. [doi: 10.1097/phm.0000000000001386]
25. Faut L, van Schieveen C, Schalk BWM. Met een heupfractuur op de Geriatrische Revalidatie: Factoren gerelateerd aan een kortere opnameduur. Tijdschrift voor Ouderengeneeskunde 2021:5 [FREE Full text]
26. Orwig D, Hochberg MC, Gruber-Baldini AL, Resnick B, Miller RR, Hicks GE, et al. Examining Differences in Recovery Outcomes between Male and Female Hip Fracture Patients: Design and Baseline Results of a Prospective Cohort Study from the Baltimore Hip Studies. J Frailty Aging 2018;7(3):162-169 [FREE Full text] [doi: 10.14283/jfa.2018.15] [Medline: 30095146]
27. Zhu Y, Xu BYX, Low SG, Low LL. Association of Social Support with Rehabilitation Outcome among Older Adults with Hip Fracture Surgery: A Prospective Cohort Study at Post-Acute Care Facility in Asia. J Am Med Dir Assoc 2023 Oct;24(10):1490-1496 [FREE Full text] [doi: 10.1016/j.jamda.2023.03.034] [Medline: 37156471]
28. Zhang N, Zhang D, Ren S, Gao Y, Sun W, Yang S. Relationship between preoperative hemoglobin levels and length of stay in elderly patients with hip fractures: A retrospective cohort study. Medicine (Baltimore) 2024 Jun 21;103(25):e38518 [FREE Full text] [doi: 10.1097/MD.0000000000038518] [Medline: 38905374]
29. Millrose M, Schmidt W, Krickl J, Ittermann T, Ruether J, Bail HJ, et al. Influence of Malnutrition on Outcome after Hip Fractures in Older Patients. J Pers Med 2023 Jan 03;13(1):109 [FREE Full text] [doi: 10.3390/jpm13010109] [Medline: 36675770]
30. Clement ND, Farrow L, Chen B, Duffy A, Murthy K, Duckworth AD. Delayed admission of patients with hip fracture from the emergency department is associated with an increased mortality risk and increased length of hospital stay. Emerg Med J 2024 Oct 23;41(11):654-659. [doi: 10.1136/emermed-2023-213085] [Medline: 39379165]
31. Ek S, Meyer AC, Hedström M, Modig K. Hospital Length of Stay After Hip Fracture and It's Association With 4-Month Mortality-Exploring the Role of Patient Characteristics. J Gerontol A Biol Sci Med Sci 2022 Jul 05;77(7):1472-1477 [FREE Full text] [doi: 10.1093/gerona/glab302] [Medline: 34622920]
32. van Dijk-Huisman HC, Bijnens W, Senden R, Essers JMN, Meijer K, Aarts J, et al. Optimization and Validation of a Classification Algorithm for Assessment of Physical Activity in Hospitalized Patients. Sensors (Basel) 2021 Feb 27;21(5):1652 [FREE Full text] [doi: 10.3390/s21051652] [Medline: 33673447]
33. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. Clin Transl Sci 2024 Nov;17(11):e70056. [doi: 10.1111/cts.70056] [Medline: 39463176]
34. Mashimo S, Kubota J, Sato H, Saito A, Gilmour S, Kitamura N. The impact of early mobility on functional recovery after hip fracture surgery. Disabil Rehabil 2023 Dec;45(26):4388-4393. [doi: 10.1080/09638288.2022.2151652] [Medline: 36448297]
35. Shimizu T, Kanai C, Asakawa Y. Relationship between independence in activities of daily living at discharge and physical activity at admission of older postoperative hip fracture rehabilitation inpatients: A retrospective case-control study. Physiother Res Int 2024 Jan;29(1):e2070. [doi: 10.1002/pri.2070] [Medline: 38284469]

XSL•FO
RenderX

36. Graham ZA, Lavin KM, O'Bryan SM, Thalacker-Mercer AE, Buford TW, Ford KM, et al. Mechanisms of exercise as a preventative measure to muscle wasting. Am J Physiol Cell Physiol 2021 Jul 01;321(1):C40-C57 [FREE Full text] [doi: 10.1152/ajpcell.00056.2021] [Medline: 33950699]
37. Creditor MC. Hazards of hospitalization of the elderly. Ann Intern Med 1993 Feb 01;118(3):219-223. [doi: 10.7326/0003-4819-118-3-199302010-00011] [Medline: 8417639]

Abbreviations

AMTS: Abbreviated Mental Test Score
ASA: American Society of Anesthesiologists physical status classification
AUC: area under the curve
CCI: Charlson Comorbidity Index
EoDT: Ensemble of decision trees
ER: emergency room
FAC: functional ambulation categories
Katz-ADL6: Katz Index of Independence in Activities of Daily Living-6
ML: machine learning
MoCA: Montreal Cognitive Assessment
NB: naïve Bayes
NN: neural network
PFMS: Prefracture Mobility Score
PC: principal component
SHAP: Shapley Additive Interpretation
SVM: support vector machine
ZGT: Ziekenhuisgroep Twente

Edited by S Munce; submitted 19.Jun.2025; peer-reviewed by M Santos, H Namba; comments to author 19.Nov.2025; accepted 13.Jan.2026; published 23.Feb.2026.

Please cite as:
Krakers SM, Wouda FJ, van Dartel D, Vollenbroek-Hutten MMR, Hegeman JH, Up&Go After a Hip Fracture Group
Predicting Geriatric Rehabilitation Stays of ≤4 Weeks After Hip Fracture Surgery: Machine Learning Approach Using Physical Activity and Patient Data
JMIR Rehabil Assist Technol 2026;13:e79331
URL: https://rehab.jmir.org/2026/1/e79331
doi:10.2196/79331
PMID:

©Sanne M Krakers, Frank J Wouda, Dieuwke van Dartel, Miriam MR Vollenbroek-Hutten, Johannes H Hegeman, Up&Go After a Hip Fracture Group. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Enhancing Knee Joint Proprioception in Healthy Adults Through Exergame Training With Augmented Feedback: Randomized Controlled Pilot Trial

Yiling Zhang[1], MSc; Luis Felipe García Arias[1], MSc; Hans Timmerman[2], PhD; Ming Cao[1], Prof Dr; Elisabeth Wilhelm[1], PhD

[1]Engineering and Technology Institute, Faculty of Science and Engineering, Universtiy of Groningen, Groningen, The Netherlands

[2]Department of Anesthesiology, Pain Center, University Medical Center Groningen, Groningen, The Netherlands

Corresponding Author:
Elisabeth Wilhelm, PhD
Engineering and Technology Institute
Faculty of Science and Engineering
Universtiy of Groningen
Nijenborgh 4
Groningen, 9747AG
The Netherlands
Phone: 31 0631921169
Email: e.wilhelm@rug.nl

Abstract

Background: Proprioception training is essential for restoring knee function in several medical conditions. Open kinetic chain (OKC) and closed kinetic chain (CKC) exercises are used in active movement interventions to enhance proprioception. Exergames, supported by wearable sensors, offer a solution by providing real-time feedback. Auditory feedback (AF) embedded in the serious game training has shown benefits in upper limb rehabilitation compared to visual feedback (VF) alone. However, the potential of AF in exergames that provide knee training is not known.

Objective: This study presents an exergame platform aimed at enhancing knee joint proprioception through stretching and squatting exercises. The platform allows to provide feedback in 2 modes, namely, VF only and a combination of AF and VF. The AF indicates the joint position by adjusting the loudness of the sound. The VF maps the motion of the lower limb into the game space, where this information is used to control a game object. A randomized controlled trial with 14 participants compared AF to VF only. The hypothesis was that AF would improve knee joint position accuracy, enhancing neuromuscular coordination and lower limb stability.

Methods: A randomized controlled trial was conducted within 14 healthy volunteers to test the exergames for knee joint motor learning using augmented feedback. All participants were required to do a pretest consisting of half and full squats and two tasks, in which participants were asked to reproduce a 45-degree knee bend and to stretch their knee fully. After that, participants played 4 rounds of each of the 2 exergames. Then the tasks of the pretest were repeated. A 1-sided Mann-Whitney *U* test was conducted for answering whether AF has a positive effect on the ability of participants to accurately control the knee joint angle. In addition, we calculated the muscle synergies participants used to complete the exercises. Subjective gaming experience was assessed using the Intrinsic Motivation Inventory and the User Experience Questionnaire.

Results: A total of 14 participants were recruited, including 7 in the experimental group with AF, and 7 in the control group only with VF. The result of the Mann-Whitney *U* test demonstrated that augmented feedback improved knee joint accuracy compared to VF in both the CKC (statistics=41.0; *P*=.04) and OKC (statistics=42.0; *P*=.03) tasks. Additionally, muscle synergy analysis revealed high consistency in different muscle synergy patterns between groups across both game types.

Conclusions: Augmented feedback significantly enhanced knee joint motor learning performance (reflected in the knee joint angle positioning ability) in both CKC and OKC exergame training. Consistent muscle synergy patterns across participants show that the developed exergames are suitable for knee training. Studies in patient populations are needed to establish whether the games could be used in lower limb rehabilitation.

Trial Registration: ClinicalTrials.gov NCT07141290; https://clinicaltrials.gov/study/NCT07141290

XSL•FO
RenderX

(JMIR Rehabil Assist Technol 2026;13:e78525) doi:10.2196/78525

KEYWORDS

home-based exergame; augmented feedback; auditory feedback; knee proprioception; Inertial Measurement Unit

Introduction

Background and Rationale

Proprioceptive training has been reported to be beneficial in restoring knee joint function among clinical populations and athletes. For example, rehabilitation by means of physical therapy can restore this proprioception in patients with knee osteoarthritis [1].

According to Aman et al [2], proprioceptive training can be further divided into 5 types. These are active movement and balance training, passive movement training, somatosensory stimulation training, somatosensory discrimination training, and combined and multiple system training. The meta-analysis does not come to a clear conclusion on which of these training types is superior. It also highlights that one major limitation of the field is that there is no clear definition of the term proprioceptive training. To overcome this, the authors suggest that this term should be reserved for studies that train only based on somatosensory signals while excluding other modalities such as vision [2]. Winter et al [3] suggest a broader definition that includes any intervention aiming to improve proprioception with the ultimate goal to enhance motor function. This definition is the one that applies to most serious games, as they usually involve visual feedback (VF). Winter et al [3] also concluded that active movement interventions are an important aspect of proprioceptive training. Therefore, serious games that contain active movement interventions could be beneficial to improve proprioception.

In physiotherapy, active movements such as open kinetic chain (OKC) and closed kinetic chain (CKC) exercises are used to strengthen the muscles of the knee joint [4]. OKC exercises consist of free movement of the distal joint in space without any loading, such as seated leg extensions, terminal knee extension exercises, hamstring curls, and calf pumps [5]. OKC exercises isolate specific muscle groups, enabling precise, targeted strengthening. Furthermore, OKC exercises can increase the range of motion in patients with anterior cruciate ligament injury [6]. In contrast to OKC exercises, CKC exercises fix or load a distal segment, thereby prohibiting free movement of that segment [5]. Prominent examples of CKC exercises are squats and lunges. CKC exercises promote the cocontraction of muscles to stabilize and control joint movements [6].

In the rehabilitation context, patients ideally would train 3 to 7 days per week. However, due to the lack of physiotherapists, several sessions need to be carried out unsupervised. Exergames are an option for providing feedback, while patients with impairments of the upper and lower limbs, gait disorders, balance disorders, visual-spatial disorder, or cognitive disorder train independently [7]. Especially, video exergames become a promising option for engaging patients in rehabilitative exercises. As reviewed by Gelineau et al [8], for the physical rehabilitation exergames, nonspecific video game systems (eg, Nintendo Wi and Xbox Kinect) need to be combined with specific rehabilitation systems (eg, Rehabilitation Gaming System and virtual gloves) or specific rehabilitation devices (eg, Hand Mentor Pro and Polhemus 3) to achieve satisfying training results in patients with stroke [8].

The rehabilitation systems that can be combined with exergames usually contain sensors to track the participant's motion. Inertial Measurement Units (IMUs), which measure linear acceleration and angular acceleration, can be used to monitor kinetic motion parameters [9]. In addition, some devices use electromyography (EMG) to extract information on the activation of different muscle groups [10,11]. Metrics derived from these different sensor modalities can be used to calculate, analyze, and predict whether a user accomplishes the prescribed exercises. The use of sensorized systems makes it possible to inform exergame users about the knowledge of results (KOR), which can be used as extrinsic feedback. The KOR is defined as information about the measured error between the terminal limb position and the given target position [12]. Traditionally, most video games use a visual representation to inform users about the KOR.

Augmented feedback has been suggested as a strategy to enhance the performance of rehabilitation training with exergames. Among the different feedback modalities that have been suggested for motor learning, auditory feedback (AF), which converts movement-related data into sound, has been successfully used in tasks with low functional task complexity. Sounds can either be used to warn participants (alarms), to make them more aware of their movements (motion sonification), or to amplify errors (error amplification) [13]. The decrease in performance when feedback is removed is smaller in AF as compared to VF, indicating that participants who receive AF do not become dependent on the presence of the feedback [14].

For proprioceptive training in older adults, a combination of a commercially available stepping game (Stomp it) and a game in which participants have to kick a virtual ball (Sport Kinect target kick) or prevent a ball from entering the goal (Sport Kinect goal keeper game) has been investigated. These games mainly focus on OKC tasks. Furthermore, the games that are related to soccer playing require a mix of knee movement and balancing. The feedback in the games is mostly VF. To assess the effect of an 8-week training program with these games, participants were asked to reproduce knee angles when seated. The instructions for the knee angle reproduction assessment were given with a tactile interface. Participants were blind folded during this procedure. Participants of the intervention group reduced the knee proprioception error from 5.2 to 3.5 for a target angle of 30. Similar significant improvements were also reported for other target angles [15]. For patients with knee osteoarthritis, a game that provides VF on CKC exercises was proposed. The exercises in this game were squatting and leg presses against an elastic Theraband. The proprioception was assessed by asking participants to reproduce a knee angle that has previously visually been demonstrated to them. For a target position of 30,

XSL•FO
RenderX

the knee proprioception error improved from 6.25 to 3.5 in a training period of 6 weeks [16]. Another study in patients with osteoarthritis focused on balance games with the Nintendo Wii Fit. The games are, in general, challenging the balance of the user. As the solution the player attempts is not limited, they can contain a mix of open and closed kinematic chain motions. In accordance with its main goal, this study focused on balance-related outcome measures [17]. No study so far has systematically investigated whether the efficacy of AF is related to the type of exercise (OCK vs CKC) in lower limb training.

Objectives

In this study, we developed a video exergame system for enhancing position accuracy through targeted training with incorporated motion tracking and augmented feedback. The system contains 2 different games, one for CKC and one for OKC exercises. The CKC game is called Bunny Game and encourages users to do half squats and deep squats. The OKC game, Fish Game, asks users to perform 4 different knee joint stretching tasks. To provide feedback, the knee joint angle was measured by IMU sensors. This information was used to adjust the position of a game object on the screen (VF) and provide motion sonification (AF).

To verify whether augmented feedback training yields better results than VF, a pilot study was conducted with 14 healthy participants, 7 in the augmented feedback group (experimental group [EG]) and 7 in the VF group (control group [CG]). All participants played both games, one with an OKC task and one with a CKC task in a randomized order. We hypothesized that participants in the EG would significantly improve knee joint position accuracy compared to the CG. In total, 7 EMG signals from the nondominant leg of the participants were collected to analyze the muscle synergies used during training.

Methods

This paper adheres to the CONSORT (Consolidated Standards of Reporting Trials). The CONSORT checklist for this paper can be found in Multimedia Appendix 1.

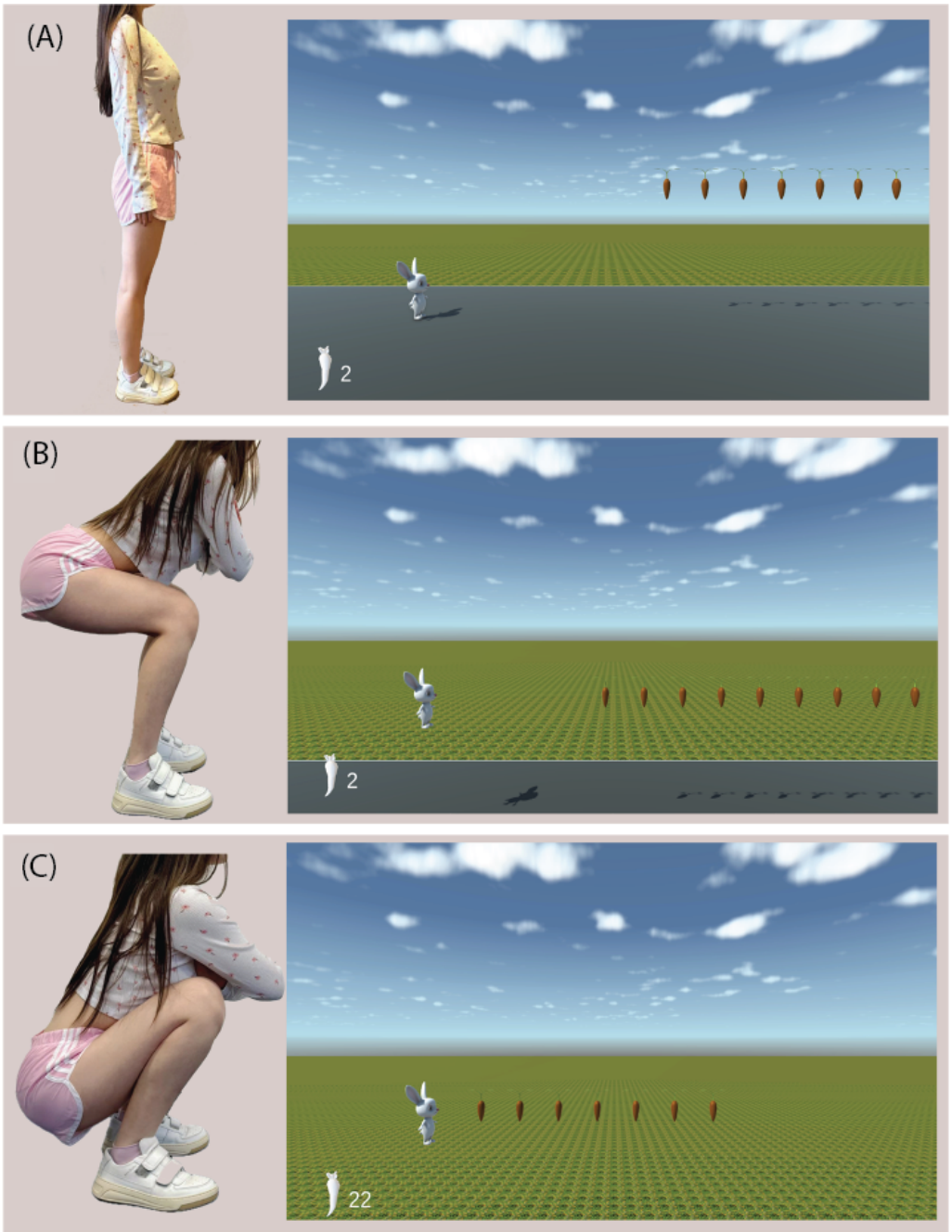
Game Development

CKC Exergame

The user interface (UI) of the CKC exergame, called “Bunny Game,” is shown in Figure 1. In this game, participants performed 2 different types of squats to control the vertical position of the bunny on the screen: the deeper the squat, the higher the bunny jumps. The sonification of the position of the bunny is based on a sequence of brief, percussive tones with varying pitches. The pitch increases as the bunny’s vertical position rises. Carrots were set at 2 heights: one reachable by a half squat, as shown in Figure 1B, and the other by a deep squat, as shown in Figure 1C. When the bunny reaches a carrot, it is collected, and the score increases by 1. Each group consists of 9 carrots, requiring the player to hold the squat position for 2.5 seconds to collect all of them. Each group of carrots is followed by a 20-second gap to allow the player to stand up and rest. One round of the game contains 12 groups of carrots, thereby prompting the player to perform 12 squats per round.

XSL•FO
RenderX

Figure 1. Illustration of the game with closed kinetic chain movements. (A) Preparation in stance gesture. (B) Half squat demonstration and the carrots' height in the user interface. (C) Deep squat demonstration and the height of the carrots that should trigger this motion.

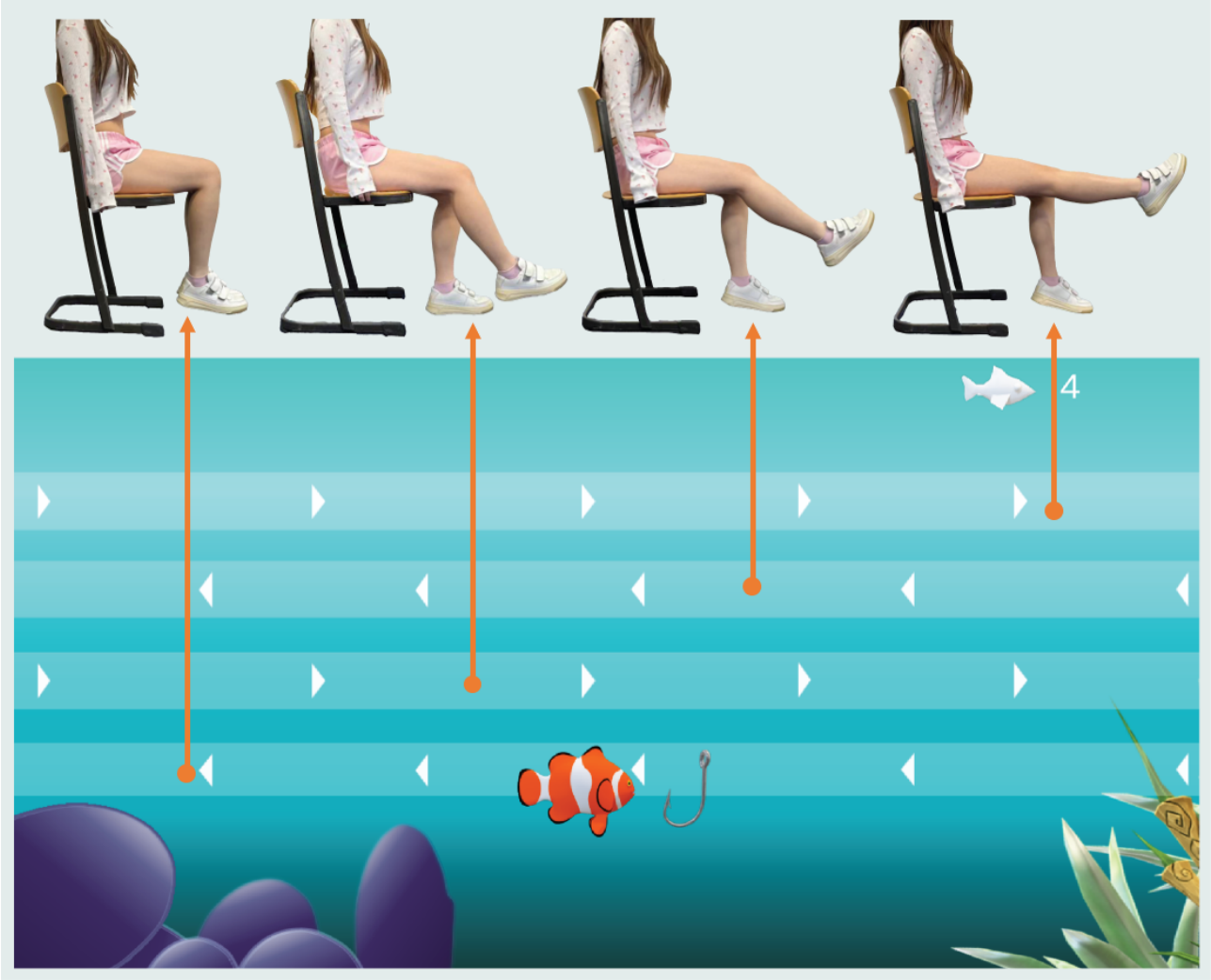


OKC Exergame

The UI of the stretch exergame "Fish Game" is shown in Figure 2. In this game, 12 fish appear randomly along 4 different trails, each corresponding to a specific knee stretch movement. The lowest trail matched the movement of the 90° knee joint, the second lowest trail matched the 120° knee joint bending, the second highest trail matched the 150° knee joint bending, and the highest trail matched the total stretching. The participant rotates the knee joint of the tested leg to move the hook from one trail to the other in a vertical direction. The pitch of the sonification prompt increases as the vertical position of the hook rises. If the hook stays for more than 2 seconds on the trail where a fish appears, the fish is caught, and the score increases by 1. If the fish is not caught within 10 seconds, it reaches the other side of the screen, where it disappears. The word "Miss" is shown on the screen to communicate to the player that the fish escaped.

XSL•FO
RenderX

Figure 2. Illustration of the game with open kinetic chain motion. From left to right are the demonstrations of knee joint angles in 90°, 120°, 150°, and 180°, which correspond to 4 trials.



Auditory Feedback

The sound used in AF mode consisted of a short, percussive tone extracted from a rain sound recording, which gave a crisp, bell-like quality with natural texture. The pitch of the sonification was modulated according to the vertical position of the task object, such that the higher the bunny or hook, the louder the sound. The audio was delivered through the Unity engine, with amplitude scaled linearly from silence to the maximum nonclipping volume. The duration of 1 auditory stimulus was 0.1-0.2 seconds. After this period, the next fragment with a different pitch was played.

Game Integration With Sensors

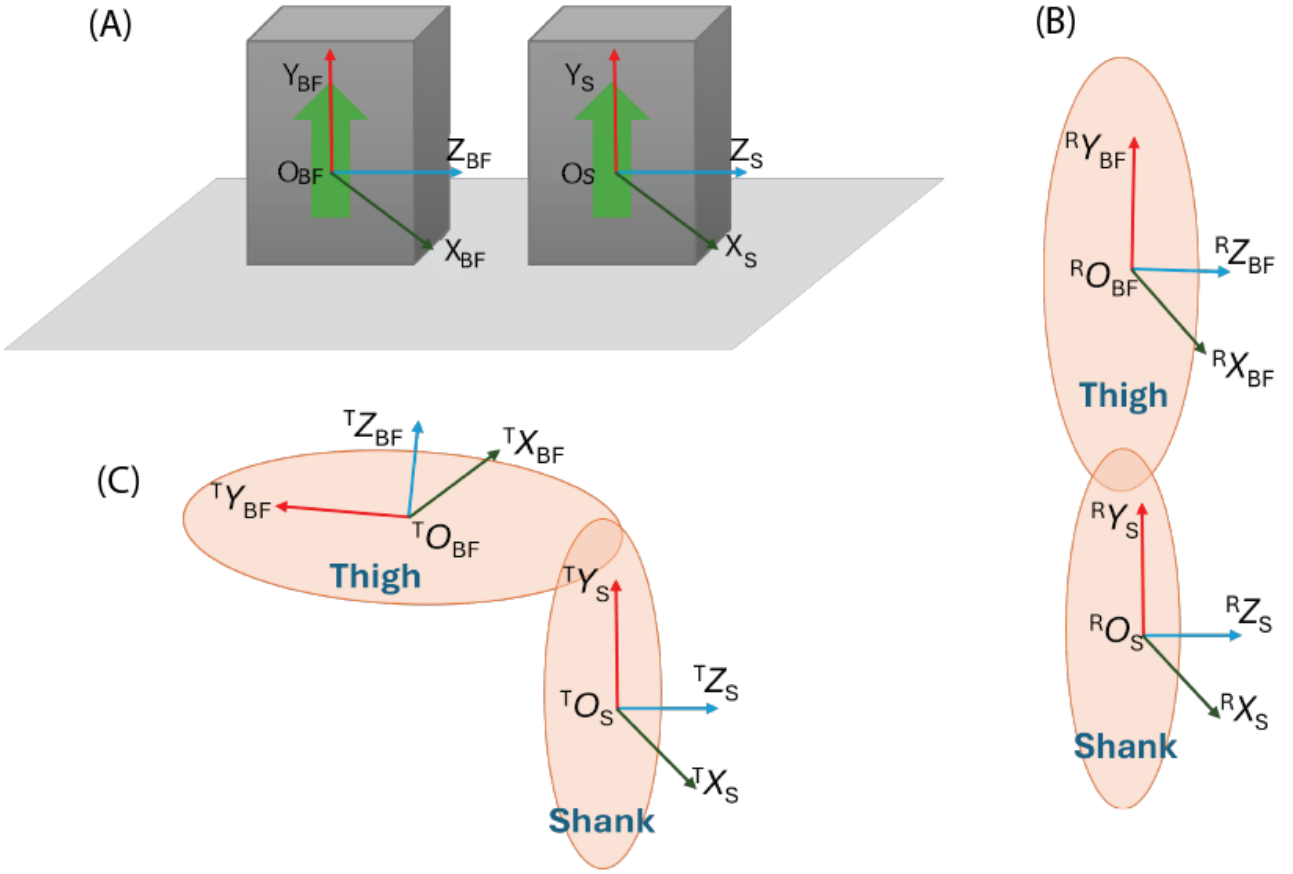
This knee joint training game system was developed using Unity (version 2021.3.0.f1; Unity Technologies). Both kinematic data and EMG data were collected with 9 Avanti sensors (Delsys Europe Ltd). Kinematic data were obtained from 3 Avanti sensors placed on midshank and biceps femoris (BF) muscle on the participant's nondominant leg and the one on the collarbone region of the trunk. Muscular activity of tibialis anterior, gastrocnemius medialis, gluteus medius, gastrocnemius lateralis, peroneus longus, BF, and rectus femoris in the nondominant leg were measured with 6 Avanti sensors. The application programming interface of Delsys was used to stream the data from the IMU sensors in real time into the game environment.

Sensor Calibration

Prior to sensor attachment, the sensors were lying next to each other in a parallel orientation on a flat desk. In this state, the sensors were connected to the game system; when the UI showed that all of the sensors were connected, the initial IMU orientation was recorded to determine the offset of individual sensors. Once all sensors started streaming data, they were attached to the legs of the participant, and the electrode location followed the SENIAM (Surface Electromyography for the Non-Invasive Assessment of Muscle) guidelines [18]. Before testing, the participants were asked to stand in an upright posture to establish a baseline orientation of the IMUs on the user's body. This position offset was used in the kinematic analysis to prevent that difference in body shape affects the game performance. The whole calibration process is shown in Figure 3.

XSL•FO
RenderX

Figure 3. Sensor calibration process illustration: (A) sensors aligned in parallel, (B) sensors attached to the leg in an upright standing position, and (C) sensors in a movement. BF: biceps femoris.



The data obtained during the calibration process were used for the following in-game calculations. To obtain the initial reference quaternion for knee joint angle measurements, the IMUs attached to the BF and the midshank on the nondominant leg were placed parallel to each other on a flat surface. Sensor data were then recorded for 14 seconds at a sampling frequency of 74 Hz, resulting in a buffer of 1000 unit quaternion samples, which was used to compute the reference orientation. Each quaternion in this buffer is represented as $q_i=(w_i, x_i, y_i, z_i)$. The original reference quaternion (q_{avg}) is obtained by averaging all 1000 quaternions component-wise.

To ensure that this average remains a unit quaternion, q_{avg} is normalized by dividing it by its magnitude:



This q_0 serves as the baseline orientation of the sensors.

After this calculation, all of the sensors were attached to the participant by the experimenter, with the midshank sensor placed in the middle of the shank and aligned with the nondominant leg BF sensor. The player was asked to stand upright for 14 seconds, avoiding moving. During this time, the computer calculated the offset (q_{ref}). This information was entered in equation 2 to calculate the rotation q_T between the original quaternion $q_{original}$ and the offset $q_{reference}$ to account for any orientation bias due to the leg shape.



During game training, the real-time sensor orientation is computed by multiplying the measured quaternion (q_R) and the offset (q_T):



The knee angle (q_{knee}) is defined as the angle between the calibrated coordinate system of the midshank  and that of the BF of the nondominant leg . This angle can be obtained from equation 4. The Euler angle of the knee in the Y direction is used as the game controller in games.



Power Calculation

The effect size of AF was estimated to be 0.78, based on the results of Ghai et al [19] and Fujii et al [20]. The main results of this study were evaluated using a repeated-measures, between-factors ANOVA. The power was set to 80%, and an α of .05. G*Power (version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf) was used to calculate the sample size a priori. This calculation indicated that a minimum number of 10 participants, including 5 participants in the EG and 5 participants in the CG, were needed to draw a conclusion. Considering a dropout rate of 30%, the recruitment target was set to 14 participants.

XSL•FO
RenderX

Recruitment

All the participants were recruited among the employees and students of the University of Groningen. To be eligible to participate in this study, participants had to be at least 18 years of age, be healthy based on self-report, and be fluent in either English or Dutch to ensure that they understand the instructions given by the experimenter. Exclusion criteria were pregnancy, orthopedic problems, musculoskeletal disorders, prior surgery on the lower extremities, or abuse of drugs, growth hormones, anabolic steroids, or performance-enhancing substances based on self-report. Furthermore, participants with a known adhesive allergy were excluded to reduce the risk of skin irritation due to sensor attachment. To be able to attach the EMG electrodes, participants were required to wear shorts during the test. Therefore, we also had to exclude participants who were not willing to expose the skin of the legs for any reason.

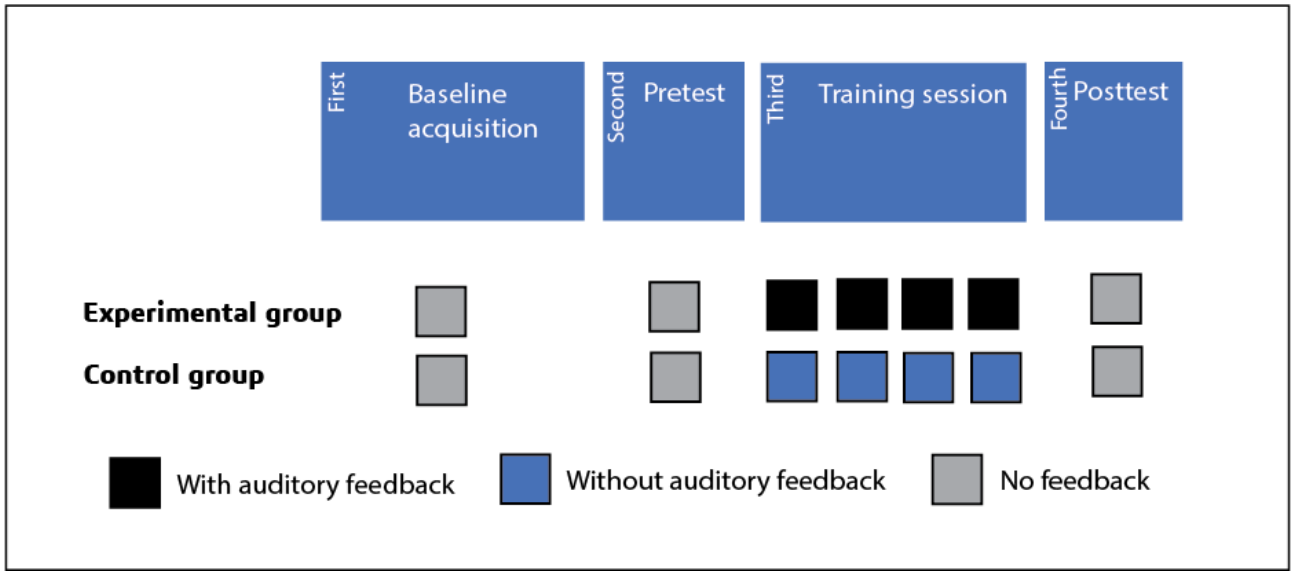
Experiment Design

Overview

The study design of this experiment was a randomized between-group comparison. Participants were randomly assigned to groups using a Microsoft Excel–generated sequence to minimize bias. For comparing the game effectiveness within and without AF, all participants tested both game levels in randomized order. The group the participant belongs to was predetermined via a vector generated with the random function in Excel. This vector was used to assign each participant to 1 group based on the order in which they were recruited. AF was assigned to participants in the EG, but not to the participants in the CG. To rule out confounders, the participants were assigned to the EG and CG at random, apart from the AF; both groups followed the same experimental procedure.

There were 4 blocks in the experiment, as shown in Figure 4.

Figure 4. Flow chart of the study design overview and the main procedures that subjects will undergo in the course of research.



After the sensor calibration, the pretest was performed. During the pretest, each participant performed 10 squats and 10 knee stretches. Participants during the pretest were instructed by written information on the screen on which movement they should perform. Each movement was held for approximately 5 seconds. During the squatting sequence of the pretest, the commands “half squat” or “deep squat” appeared in randomized order. During the stretching sequence of the pretest, the prompts “half stretch (45°)” or “totally stretch (0°)” appeared in random order by the script used for prompt text generation. Participants were seated for this test, and they were previously instructed to obtain a position that is in the middle between totally stretching and having their feet in a 90°. As all participants had a mathematical background, no further explanation of the instructions was necessary. Participants were not blindfolded during the test. Psychometric properties of the test are not available. Participants who were randomly allocated to start with the bunny game started the pretest with the squatting session, while participants who started with the fish game were first prompted with the stretching exercises. To initiate the pretest, the experimenter selected the “Pre-Test” scene in the game menu. This game menu consists of 3 buttons: “Bunny Game,” “Fish Game,” and “Pre-Test,” which allow players to select the different games. In the “Pre-Test” scene, the screen displays 2 buttons: “Stretch Test” and “Squat Test.” The experimenter used these buttons to select the tests in the order indicated on the random allocation list. After completing a test, the exit button returns to the main menu scene, with the Pre-Test button now labeled Post-Test. The Fish Game and Bunny Game options remain available.

After the pretest, each participant completed 4 rounds of training for both games. The training platform for the EG included AF, while for the CG, AF was turned off, leaving only VF.

In the final block, a posttest was conducted, consisting of 10 nondominant knee stretches and 10 squats. The test content and format were identical to those of the pretest. During the pre- and the posttest, participants were asked to perform 10 half squats (target knee angle [TKA]=90°) and full squats (TKA=130°). In addition, participants were asked to sit on a chair and stretch their leg 10 times with a TKA of 135° or 180°, which equals a half and full stretch, respectively.

XSL•FO
RenderX

Study Parameters

The main outcome was the error of position between TKA and actual knee angle (AKA) of the participant along the femoral-tibial axis. The AKA was calculated based on the IMU sensors as described in the Game Development section. The error of position was calculated by subtracting TKA from AKA (equation 5).

$$\text{Error}=\text{AKA}-\text{TKA} \quad \textbf{(5)}$$

According to both the IMU signal and the recorded video during the test, the IMU data are segmented, with each segment representing a test movement. The error of these data segments was calculated. In each test, average error (AE) is classified by the type of movements, with 2 groups (squat test group and stretch test group) of AE in the pretest and 2 groups of AE in the posttest for each participant. The absolute mean value of each group represented the average performance of the participant in the corresponding movement during the test (average mean error [AME]), calculated by equation 6, where *n* denotes the number of times this type of movement appeared in the test, and AE denotes the average error of a movement.



For each participant, the improvement of training was represented by the difference in performance of each group between pretest and posttest. Equation 7 expresses the calculation of the improvement for a participant in 1 movement classification, which is denoted by IE (improvement of error trajectory), AME_{Pre} denotes the absolute mean value of a group of movement errors in the pretest, and AME_{Post} denotes the absolute mean value of the group of movement error in the posttest.

$$IE=AME_{Pre}-AME_{Post} \quad \textbf{(7)}$$

Self-Report Instruments

Participants completed the Task Evaluation Questionnaire of Intrinsic Motivation Inventory (IMI) [21], a multidimensional measurement tool used to assess participants' subjective motivation during game training. The following 4 subscales: interest or enjoyment, perceived competence, perceived choice, and pressure or tension, were assessed based on the Task Evaluation Questionnaire. Each subscale consisted of 5-7 items rated on a 7-point Likert scale (1=not at all true, 4=somewhat true, and 7=very true) [22].

The User Experience Questionnaire (UEQ) [23] was also filled after the experiment. Six dimensions were measured: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. Each dimension was assessed through a set of bipolar adjective pairs (eg, "annoying-enjoyable"), rated on a 7-point scale from –3 to +3. Higher scores indicated a more positive user experience [23].

Muscle Synergy

In this study, muscle synergy analysis was performed to examine the differences in muscle weighting patterns across different synergies during game training, following the methodology described in Hayes et al [24].

EMG signals were collected from 7 muscles (tibialis anterior, gastrocnemius medialis, gluteus medius, gastrocnemius lateralis, peroneus longus, BF, and rectus femoris) on the nondominant leg. The EMG signals were segmented based on task-specific event timestamps. For the Fish Game, segments were extracted from 2 seconds before the timestamp at which a fish was caught up to the event itself. For the Bunny Game, a valid task was defined as the collection of a complete group of 9 consecutive carrots; the corresponding segment spanned from the timestamp of the first carrot to that of the ninth in a single group. Each segment was processed with a fifth-order Butterworth low-pass filter (cutoff: 30 Hz) and high-pass filter (cutoff: 4 Hz). Following filtering, the signals were detrended by subtracting the mean and then rectified by taking the absolute value. Finally, each signal was normalized by dividing it first by its maximum value and then by the SD of the resulting signal to achieve unit variance.

Muscle weights were extracted using nonnegative matrix factorization. To determine the appropriate number of synergies required to reconstruct the original EMG signal, the variance accounted for (VAF) was calculated. For each trial, the number of synergies was identified as the minimum number for which the average VAF dropped below thresholds of 90%, 85%, and 80%, respectively, ensuring a balance between model simplicity and signal reconstruction accuracy.

Statistics

Q-Q plots were used to test for the distribution of the data. As the data were not normally distributed, the nonparametric alternative for a repeated measures ANOVA was used. Therefore, a 1-sided Mann-Whitney *U* test was conducted for answering whether AF has a positive effect on the knee joint training based on the exergame platform. To ensure data quality and reduce the impact of extreme values, outliers were identified and removed using the *z* score method. Specifically, for each variable, *z* scores were calculated, and data points with absolute *z* scores greater than a predefined threshold ($|z|>2.5$) were regarded as extreme data and excluded from the analysis. The threshold was chosen to balance outlier detection while retaining sufficient data for robust statistical analysis.

The null hypothesis (H_0) is that AF does not significantly improve trajectory error compared to without AF during stretching and squatting. The observations are squat test IE in the EG, squat test IE in the CG, stretch test IE in the EG, and stretch test IE in the CG. A significance level of $\alpha=.05$ was used for hypothesis testing.

Ethical Considerations

The study was carried out in April 2025 in the laboratories of the Discrete Technology and Production Automation group of the Engineering and Technology Institute of the University of Groningen in Groningen, The Netherlands. The study was conducted in accordance with the Declaration of Helsinki. The institutional ethics committee (Research Ethics Review Committee [CETO]) of the Faculty of Arts, University of Groningen, reviewed the experiment protocol and had no objection to the proposal (ID 93318312). Before the sessions started, participants had received an information letter describing

XSL•FO
RenderX

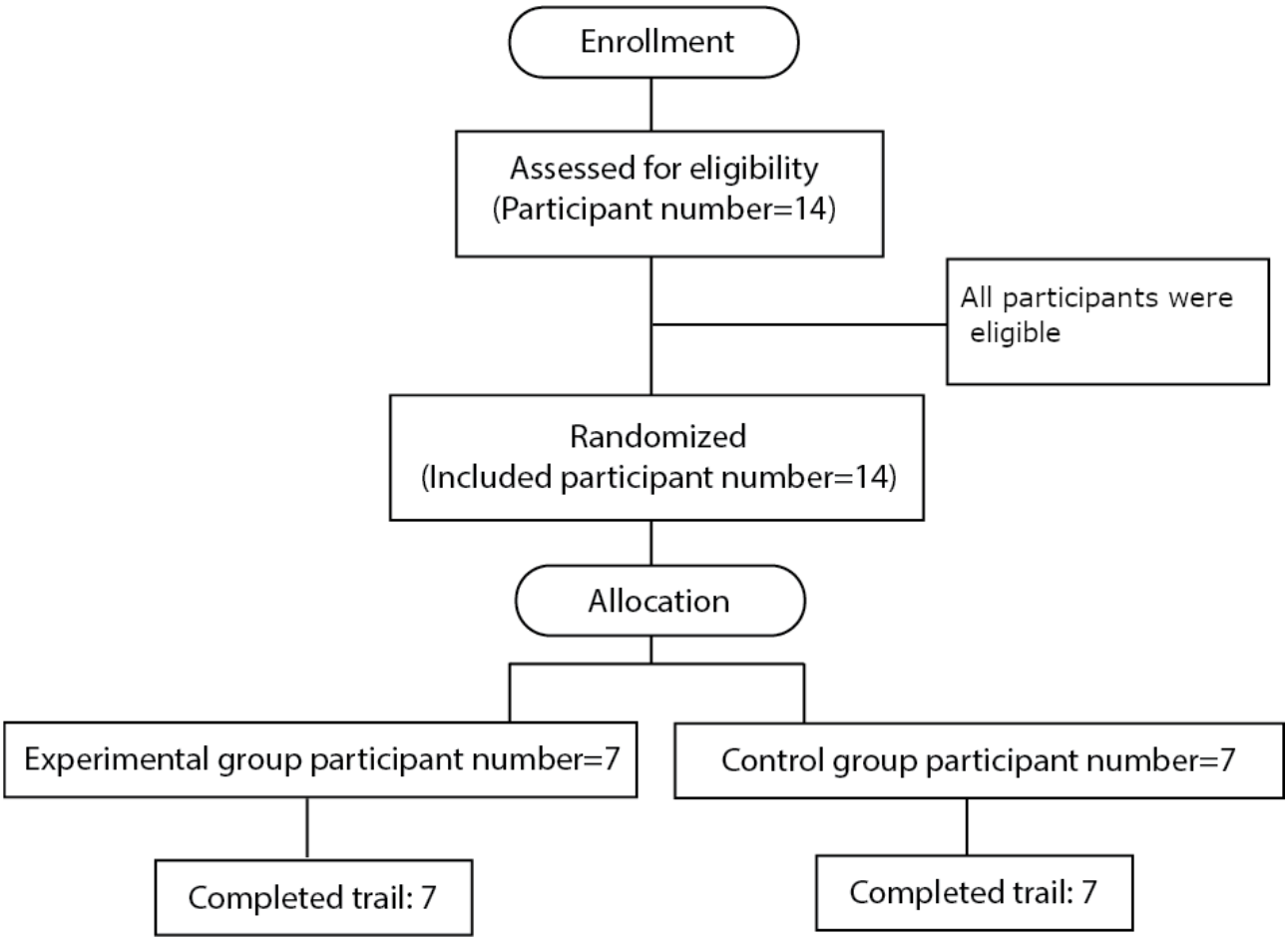
the aim and procedures carried out during the study. They were given the opportunity to ask questions about their participation. Participants, who signed the informed consent form, were required to fill in a short demographic questionnaire and the Waterloo Footedness Questionnaire [25]. Afterward, the experimenter tested their hearing capability by Audiometer 9910 (Amplivox Ltd). Prior to data acquisition, a Data Protection Impact Assessment was carried out to ensure that the data processing is in line with European regulations. As a conclusion of this assessment, the data were stored in a secure virtual research workspace of the University of Groningen. The results are reported as median or average to ensure that individuals cannot be identified. Participants did not receive any compensation for their participation in these experiments.

Results

Participants

A total of 14 participants were recruited for this study: 7 participants were in the EG, and 7 participants were in the CG. All of the participants were right-leg dominant. In the augmented AF group, there were 4 female and 3 male participants; the average age was 27.58 (SD 3.09) years, and the mean BMI was 21.07 (SD 3.04) kg/m^2. The VF group consisted of 3 female and 4 male participants; the average age was 28.69 (SD 4.17) years, and the mean BMI was 21.15 (SD 2.86) kg/m^2. The trial ended when the recruitment target was reached. A flowchart describing the recruitment process is depicted in Figure 5.

Figure 5. Flowchart depicting the recruitment process.



Kinetic Analysis

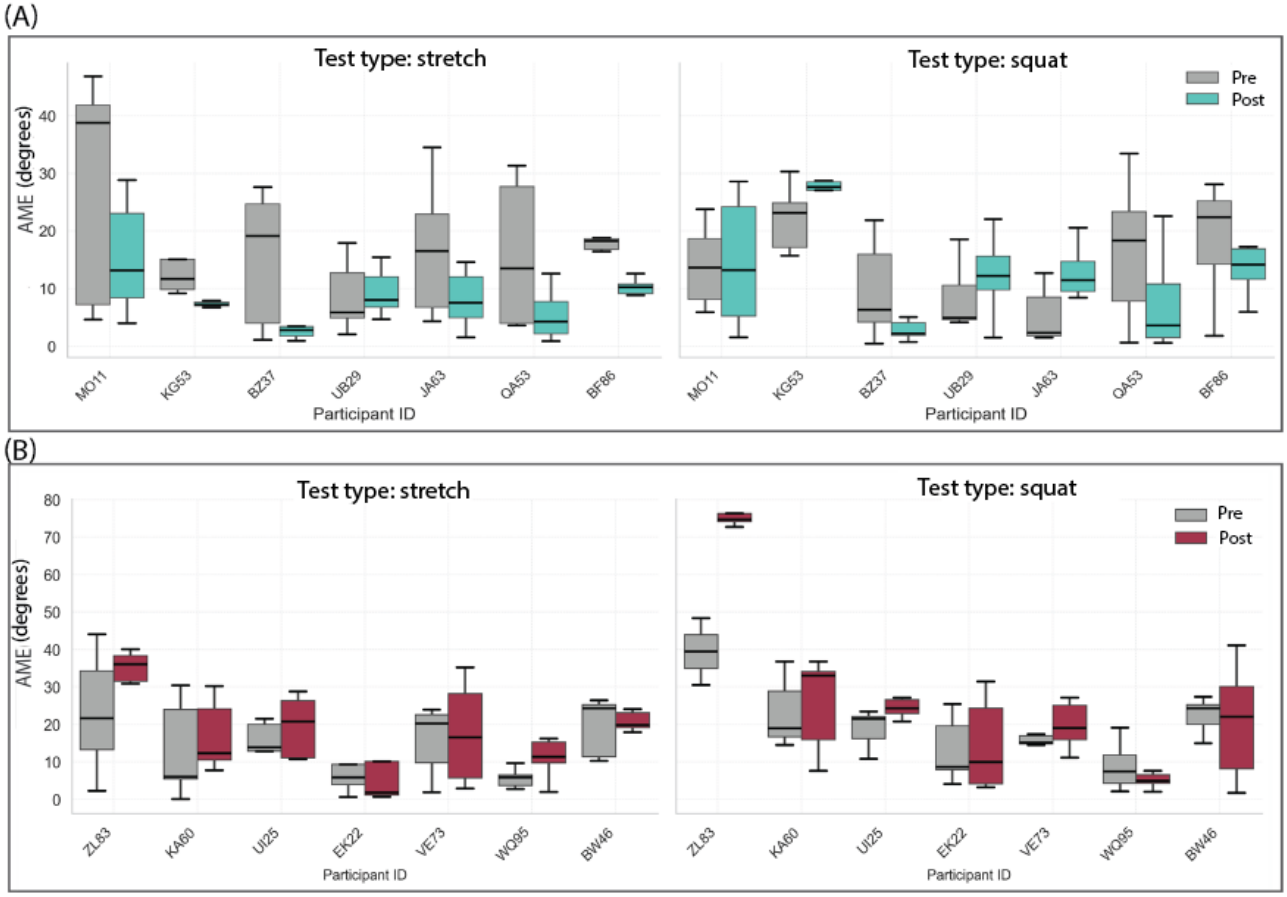
Figure 6 depicts the AME during pre- and posttest per movement category and participant. Figure 6A illustrates the AME of 7 participants in the EG. In the stretch test, all EG participants showed a smaller AME, indicating an improvement in their ability to reproduce the angles. In the squat test, only 3 participants of the EG showed obvious improvement in the posttest. The other 4 participants either had a huge spread in the AME or did not improve in the posttest compared to the pretest.

Figure 6B displays the AME for 7 participants in the CG. During both the stretch and the squat test, the participants of this group showed nearly no improvement in the AME.

The violin plot depicted in Figure 7 illustrates the improvement in knee joint performance for 2 test movements in both the EG and CG. Each violin plot represents the distribution of IE values within 1 category.

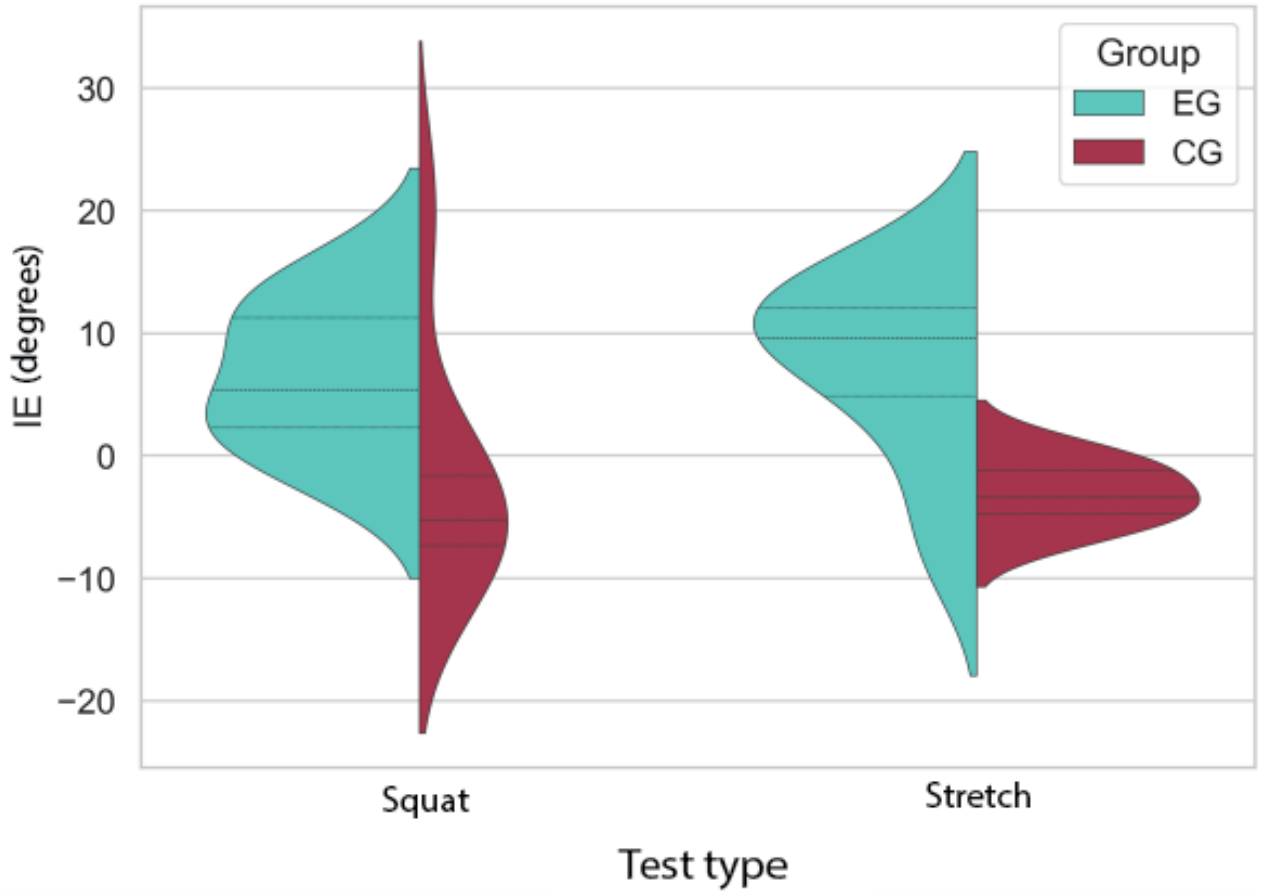
XSL•FO
RenderX

Figure 6. AME in the stretch and squat test in both pre- and posttest for the experimental group (EG) and the control group (CG). (A) AME of 7 participants in the EG during pre- and posttest. (B) AME of 7 participants in the CG during pre- and posttest. AME: average mean error.



XSL•FO
RenderX

Figure 7. Squat and stretch test IE after intervention (with and without augmented feedback) illustration. CG: control group; EG: experimental group; IE: improvement of error trajectory.



For the squat and stretch test groups, the mean value in the EG is higher than in the CG, indicating a notable improvement in squatting performance for the EG compared to the CG. A similar, though slightly smaller, trend is observed between the stretch test in EG and CG. The quartiles (first quartile Q1, median, and third quartile Q3) of each group of tests are depicted in Figure 7. The median improvement of the experimental group was 5.36 (IQR 2.27-11.25) and 10.46 (IQR 5.45-12.07) in the squat and stretch test, respectively. The control group achieved an IE of –5.31 (IQR –7.32 to –1.62) and –3.37 (IQR –5.73 to –1.22) in the squat and stretch test, respectively.

A 1-sided Mann-Whitney *U* test was conducted on the IE datasets, revealing a statistically significant difference between the EG and the CG (statistics=41.0; *P*=.04) for squatting and between the EG and the CG (statistics=42.0; *P*=.03) for stretching.

Participants who received augmented feedback demonstrated greater improvements in both squat and stretch tasks compared to the CG. H_0 was rejected, indicating that integrating augmented AF into the exergame platform positively impacts knee joint motor training.

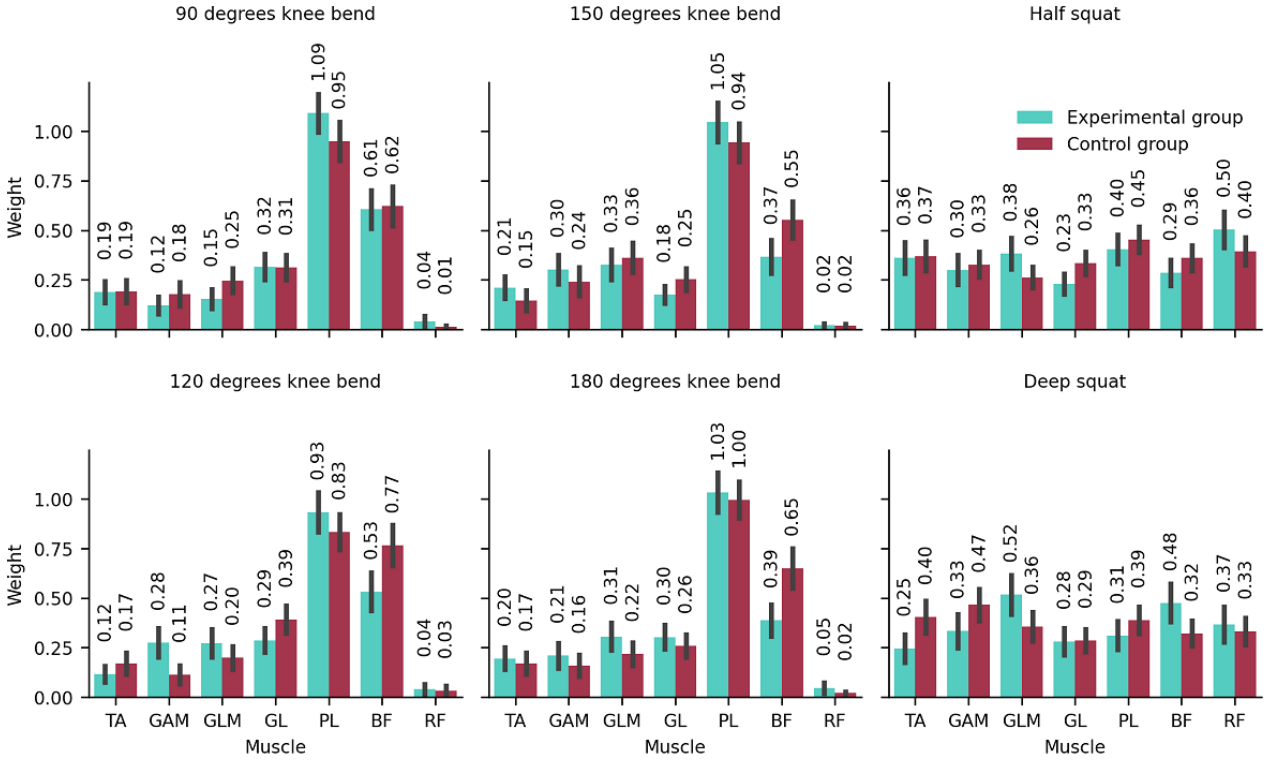
Muscle Synergy

EMG data from all 14 participants (1054 exercise movements in total) were included in the muscle synergy analysis. The number of synergies required to meet the VAF thresholds varied across trials. For the 90% VAF threshold, synergy numbers ranged from 4 to 7; a total of 3 movements had 4 synergies, 98 movements had 5 synergies, 787 movements had 6 synergies, and 166 movements had 7 synergies. For 85% VAF, synergy numbers ranged from 3 to 6; a total of 3 movements had 3 synergies, 29 movements had 4 synergies, 342 movements had 5 synergies, and 682 movements had 6 synergies. For 80% VAF, synergy numbers ranged from 3 to 6; a total of 4 movements had 3 synergies, 115 movements had 4 synergies, 683 movements had 5 synergies, and 252 movements had 6 synergies. The 85% VAF threshold was selected for subsequent analysis to balance signal reconstruction quality and model simplicity, 6 synergies appeared most frequently across participants and movements, suggesting a stable underlying structure in muscle coordination.

One of the extracted muscle synergy weight patterns is presented in Figure 8. The complete results of the synergy analysis can be found in Multimedia Appendix 2. The first and second columns display the 4 tasks from the Fish Game, comparing 2 groups. Overall, the muscle weight distributions between the groups were largely consistent. The third column shows the 2 tasks from the Bunny Game, half squat and deep squat, with similarly consistent muscle weighting patterns observed between the 2 groups.

XSL•FO
RenderX

Figure 8. Muscle synergy analysis including tibialis anterior (TA), gastrocnemius medialis (GAM), gluteus medius (GLM), gastrocnemius lateralis (GL), peroneus longus (PL), biceps femoris (BF), and rectus femoris (RF).



Quantitative Analysis of Self-Report Measures

Table 1 summarizes the IMI subscale scores for EG and CG. The EG reported higher perceived choice, and the CG reported higher interest or enjoyment, perceived competence, and pressure or tension.

The results of UEQ were analyzed from the following 5 subscales: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. In EG, the mean of subscales is 2.738 (SD 0.31), 2.821 (SD 0.14), 2.143 (SD 0.46), 2.500 (SD 0.27), 2.536 (SD 0.43), and 2.000 (SD 1.15), respectively. In CG, the mean (variance) of subscales is 2.357 (SD 0.70), 2.714 (SD 0.03), 2.286 (SD 0.28), 2.286 (SD 0.38), 1.929 (SD 1.31), and 1.679 (SD 2.08), respectively. According to the rating scheme of the questionnaire, these scores indicate that the experience was rated "excellent" [23]. The EG showed higher scores across all dimensions, particularly in attractiveness (mean 2.738, SD 0.31) and novelty (mean 2.000, SD 1.15), indicating a more engaging and novel user experience. Furthermore, to access the overall experience, the subscales were grouped into pragmatic and hedonic quality according to UEQ guidelines [23]. The EG scored higher on both pragmatic quality (mean 2.49) and hedonic quality (mean 2.27) compared to the CG (mean 2.43 and mean 1.80, respectively).

Table 1. Intrinsic Motivation Inventory subscale scores by group.

Subscale	Experimental group, mean (SD)	Control group, mean (SD)
Interest or enjoyment	35 (10)	35 (9)
Perceived competence	29 (7)	29 (7)
Perceived choice	8 (4)	8 (4)
Pressure or tension	–6 (4)	–5 (4)

Discussion

Main Results

The aim of this study was to explore whether augmented AF applied in motor learning during game-based knee joint exercise (involving both CKC and OKC) improves the knee angle proprioception more than VF alone. The results demonstrated that augmented AF exergame training improved the AME of the knee joint angle significantly more than VF alone. After both CKC and OKC game training, participants in the EG demonstrated significantly better knee angle positioning skills. This suggests that the benefit of auditory sonification cues can positively support both CKC and OKC knee joint tasks.

The results of self-report instruments, IMI and UEQ, showed that the experience of the participants was positive in both games. The IMI suggests that the participants felt low pressure while experiencing enjoyment, which is desired in an exergame. EG participants reported a more immersive, attractive, and novel experience compared with CG participants. These findings are consistent with prior research that supports the role of

XSL•FO
RenderX

augmented feedback, especially auditory, in facilitating sensorimotor integration and enhancing motor skill learning [13]. Auditory cues provide immediate, real-time reinforcement, which can guide users' movements, keep users engaged, and increase interactions during motor learning.

As shown in Figure 6, most participants in EG exhibited a reduction in AME during the OKC test following the augmented feedback game training. In addition, the SD of the error decreased from pretest to posttest for each participant, indicating reduced variability and more stable task performance after training. In CG, 6 of 7 participants also had a decreased SD after training. However, the mean AME did not show improvement in the OKC tasks after training. These findings are not in line with previous literature, which reports that improvements can be reached with both CKC [16] and OKC [15] exercises. However, the studies in the literature observe participants over a duration of 6 weeks, while we only investigated a single game session. Furthermore, the participants in the studies reported in the literature were older male individuals and patients with osteoarthritis. Our population was much younger and healthy. Therefore, the baseline might already have been better. Due to differences in the instructions during the assessment, a direct comparison of the absolute values is unfortunately not possible.

For the CKC test (squat test), the performance in EG after training does show smaller improvements than the OKC test result. In total, 4 of 7 participants in the EG decreased the AME after interventions, and 6 participants had a decreased SD after training. Compared to the EG, AME results of the CG show less improvement. Figure 7 demonstrates a more intuitive comparison. In EG, the improvement of knee joint positioning skills is generally greater than the CG in both squat and stretch tests. The OKC game (Fish Game) showed greater improvements in motor learning compared to the CKC game (Bunny Game).

Another observation in this study is the consistency in muscle synergy weight pattern between EG and CG across both game contexts. As shown in Figure 8, the muscle weight distributions are remarkably similar between 2 group participants, suggesting that motor control strategies in the muscle remained stable regardless of group assignment. This trend is also evident in both games. The stability in muscle synergy organization supports the use of game-based interventions as a tool for training repetitive motions.

Limitations

Several limitations should be acknowledged in this study. First, all participants were recruited from the University of Groningen, consisting exclusively of students and researchers. This relatively homogenous sample limits the generalizability of the findings to broader populations. Second, the experiment was conducted within a single day, and the study does not capture long-term retention or progression of motor learning over time. Third, due to the nature of AF, it was not possible to blind the participants with respect to group assignment. Therefore, the subjective measures taken with the questionnaires might be influenced by the fact that participants were aware whether they were in the control or the intervention group.

Fourth, the system has not yet been tested in a clinical population. While the current results are promising, validation in clinical settings is essential to confirm the utility and effectiveness of the developed augmented AF game training system in rehabilitation contexts for the knee joint. Finally, both Fish Game and Bunny Game include 2 distinct types of motor tasks: spatial positioning of the knee joint and knee joint extreme movements. These task types were not separately analyzed in this study.

The most noticeable limitation of our paper is that participants were not blindfolded during the pre- and posttest. Therefore, the extent to which they used vision to compensate for eventual proprioceptive deficits is unknown. Due to this limitation, we can only say that the study had positive effects on motor learning in general. Future studies that test proprioception while occluding vision and apply psychometric testing methods are needed to evaluate the precise contribution of proprioception.

Implication for Research and Future Work

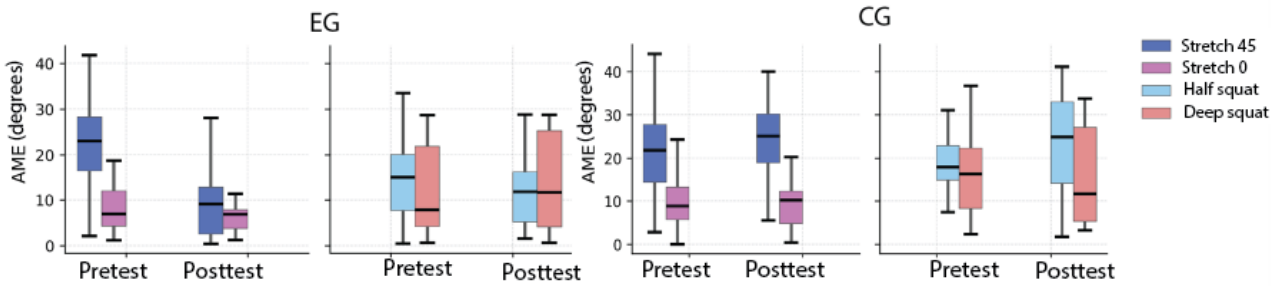
Despite the limitations, this study provides a prototype of an exergame for knee joint motor learning. The result shows the benefit of augmented feedback in interactive, feedback-driven games to support effective motor learning in the lower limbs.

Moving forward, implementing longitudinal study designs would allow for assessment of motor learning retention and the long-term impact of knee joint training based on an augmented feedback exergame. Clinical trials are also essential to explore whether the game can be used in the context of rehabilitation. Finally, further analysis should differentiate between the 2 types of motor tasks, knee joint positioning and extreme joint movement, present in the games. As shown in Figure 9, the distributions of AME between "stretch 45," "stretch 0," "half squat," and "deep squat" are obviously different. For OKC tasks, knee joint positioning shows a bigger variance than the extreme joint movement. For CKC tasks, the knee joint positioning shows a smaller variance than the extreme joint movement. Investigating these categories separately could uncover task-specific motor strategies. Together, these directions can build upon the promising outcomes of this study to inform the design of adaptive, personalized game-based interventions for the knee joint motor skill learning and rehabilitation.

The tasks that we selected in this study were focused on specific exercises that are frequently used in knee rehabilitation. Future studies could also investigate how exergaming can be used for activities of daily living such as walking.

XSL•FO
RenderX

Figure 9. AME distribution comparison among 4 test movements. CG: control group; EG: experimental group.



Conclusions

This randomized controlled trial study explored the effectiveness of CKC and OKC game-based interventions in promoting motor learning with augmented feedback in healthy adults. The results demonstrated that augmented feedback significantly enhanced knee joint motor learning performance in both game contexts, highlighting its value as an intuitive and engaging form of feedback. Additionally, consistent muscle synergy patterns across tasks and groups suggest that the neuromuscular strategies used were stable and repeatable, reinforcing the reliability of augmented feedback game-based training for knee joint motor control. In the future, targeted populations such as clinical groups could be considered to test the effectiveness of rehabilitation.

Acknowledgments

The authors thank Andreia Abud and Lisanne Bakker for their guidance on how to apply the electromyography sensors on the leg muscles.

Funding

This work is supported by China Scholarship Council (202006540022) and the Colombian Ministry of Science, Technology and Innovation (Minciencias).

Authors' Contributions

YZ, EW, MC, and HT designed the study protocol. YZ, LFGA, and EW programmed the game and analyzed the data. YZ drafted the first version of the manuscript. All authors reviewed the manuscript and gave feedback.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1292 KB - rehab_v13i1e78525_app1.pdf]

Multimedia Appendix 2
Complete overview of the results of the muscle synergy analysis.
[DOCX File , 570 KB - rehab_v13i1e78525_app2.docx]

References

1. Wang Y, Wu Z, Chen Z, Ye X, Chen G, Yang J, et al. Proprioceptive training for knee osteoarthritis: a systematic review and meta-analysis of randomized controlled trials. Front Med (Lausanne) 2021;8:699921 [FREE Full text] [doi: 10.3389/fmed.2021.699921] [Medline: 34778281]
2. Aman JE, Elangovan N, Yeh I, Konczak J. The effectiveness of proprioceptive training for improving motor function: a systematic review. Front Hum Neurosci 2015;8:1075 [FREE Full text] [doi: 10.3389/fnhum.2014.01075] [Medline: 25674059]
3. Winter L, Huang Q, Sertic JVL, Konczak J. The effectiveness of proprioceptive training for improving motor performance and motor dysfunction: a systematic review. Front Rehabil Sci 2022;3:830166 [FREE Full text] [doi: 10.3389/fresc.2022.830166] [Medline: 36188962]
4. An J, Son YW, Lee BH. Effect of combined kinematic chain exercise on physical function, balance ability, and gait in patients with total knee arthroplasty: a single-blind randomized controlled trial. Int J Environ Res Public Health 2023;20(4):3524 [FREE Full text] [doi: 10.3390/ijerph20043524] [Medline: 36834218]

XSL•FO
RenderX

5. Ellenbecker TS, Davies GJ. Closed Kinetic Chain Exercise: A Comprehensive Guide to Multiple Joint Exercise. Champaign, IL: Human Kinetics; 2001.
6. Pamboris GM, Pavlou K, Paraskevopoulos E, Mohagheghi AA. Effect of open vs. closed kinetic chain exercises in ACL rehabilitation on knee joint pain, laxity, extensor muscles strength, and function: a systematic review with meta-analysis. Front Sports Act Living 2024;6:1416690 [FREE Full text] [doi: 10.3389/fspor.2024.1416690] [Medline: 38887689]
7. Herren S, Seebacher B, Mildner S, Riederer Y, Pachmann U, Böckler NS, et al. Exergame (ExerG)-based physical-cognitive training for rehabilitation in adults with motor and balance impairments: usability study. JMIR Serious Games 2025;13:e66515 [FREE Full text] [doi: 10.2196/66515] [Medline: 39951650]
8. Gelineau A, Perrochon A, Robin L, Daviet J, Mandigout S. Measured and perceived effects of upper limb home-based exergaming interventions on activity after stroke: a systematic review and meta-analysis. Int J Environ Res Public Health 2022;19(15):9112 [FREE Full text] [doi: 10.3390/ijerph19159112] [Medline: 35897472]
9. Fan B, Li Q, Tan T, Kang P, Shull PB. Effects of IMU sensor-to-segment misalignment and orientation error on 3-D knee joint angle estimation. IEEE Sens J 2022 Feb 1;22(3):2543-2552. [doi: 10.1109/jsen.2021.3137305]
10. Kim JJ, Cho H, Park Y, Jang J, Kim JW, Ryu JS. Biomechanical influences of gait patterns on knee joint: kinematic and EMG analysis. PLoS One 2020;15(5):e0233593 [FREE Full text] [doi: 10.1371/journal.pone.0233593] [Medline: 32470052]
11. Yang J, Lu Z, Chen S, Liu C, Zhao H. Continuous knee joint angle prediction with surface EMG. Biomed Signal Process Control 2024 Sep;95:106354. [doi: 10.1016/j.bspc.2024.106354]
12. Salmoni AW, Schmidt RA, Walter CB. Knowledge of results and motor learning: a review and critical reappraisal. Psychol Bull 1984;95(3):355-386. [doi: 10.1037//0033-2909.95.3.355]
13. Sigrist R, Rauter G, Riener R, Wolf P. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. Psychon Bull Rev 2013 Feb;20(1):21-53 [FREE Full text] [doi: 10.3758/s13423-012-0333-8] [Medline: 23132605]
14. Ronsse R, Puttemans V, Coxon J, Goble DJ, Wagemans J, Wenderoth N, et al. Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. Cereb Cortex 2011 Jun;21(6):1283-1294. [doi: 10.1093/cercor/bhq209] [Medline: 21030486]
15. Sadeghi H, Hakim MN, Hamid TA, Amri SB, Razeghi M, Farazdaghi M, et al. The effect of exergaming on knee proprioception in older men: a randomized controlled trial. Arch Gerontol Geriatr 2017;69:144-150 [FREE Full text] [doi: 10.1016/j.archger.2016.11.009] [Medline: 27923177]
16. Mete E, Sari Z. The efficacy of exergaming in patients with knee osteoarthritis: a randomized controlled clinical trial. Physiother Res Int 2022 Jul;27(3):e1952. [doi: 10.1002/pri.1952] [Medline: 35470534]
17. Manlapaz DG, Sole G, Jayakaran P, Chapple CM. Exergaming to improve balance and decrease the risk of falling in adults with knee osteoarthritis: a mixed-methods feasibility study. Physiother Theory Pract 2022 Nov;38(13):2428-2440. [doi: 10.1080/09593985.2021.1952670] [Medline: 34280069]
18. Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. Development of recommendations for SEMG sensors and sensor placement procedures. J Electromyogr Kinesiol 2000 Oct;10(5):361-374. [doi: 10.1016/s1050-6411(00)00027-4] [Medline: 11018445]
19. Ghai S, Schmitz G, Hwang T, Effenberg AO. Training proprioception with sound: effects of real-time auditory feedback on intermodal learning. Ann N Y Acad Sci 2019 Feb;1438(1):50-61. [doi: 10.1111/nyas.13967] [Medline: 30221775]
20. Fujii S, Lulic T, Chen JL. More Feedback Is Better than Less: Learning a Novel Upper Limb Joint Coordination Pattern with Augmented Auditory Feedback. Front Neurosci 2016;10:251 [FREE Full text] [doi: 10.3389/fnins.2016.00251] [Medline: 27375414]
21. Ryan RM, Mims V, Koestner R. Relation of reward contingency and interpersonal context to intrinsic motivation: a review and test using cognitive evaluation theory. J Pers Soc Psychol 1983 Oct;45(4):736-750. [doi: 10.1037/0022-3514.45.4.736]
22. Alhirsan SM, Capó-Lugo CE, Brown DA. Effects of different types of augmented feedback on intrinsic motivation and walking speed performance in post-stroke: a study protocol. Contemp Clin Trials Commun 2021 Dec;24:100863 [FREE Full text] [doi: 10.1016/j.conctc.2021.100863] [Medline: 34841123]
23. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Holzinger A, editor. HCI and Usability for Education and Work. Berlin, Heidelberg: Springer; 2008:63-76.
24. Hayes HB, Chvatal SA, French MA, Ting LH, Trumbower RD. Neuromuscular constraints on muscle coordination during overground walking in persons with chronic incomplete spinal cord injury. Clin Neurophysiol 2014 Oct;125(10):2024-2035 [FREE Full text] [doi: 10.1016/j.clinph.2014.02.001] [Medline: 24618214]
25. Elias LJ, Bryden M, Bulman-Fleming M. Footedness is a better predictor than is handedness of emotional lateralization. Neuropsychologia 1998 Jan;36(1):37-43. [doi: 10.1016/s0028-3932(97)00107-3] [Medline: 9533385]

Abbreviations

AE: average error
AF: auditory feedback
AKA: actual knee angle
AME: average mean error

XSL•FO
RenderX

BF: biceps femoris
CG: control group
CKC: closed kinetic chain
CONSORT: Consolidated Standards of Reporting Trials
EG: experimental group
EMG: electromyography
IE: improvement of error trajectory
IMI: Intrinsic Motivation Inventory
IMU: Inertial Measurement Unit
KOR: knowledge of result
OKC: open kinetic chain
SENIAM: Surface Electromyography for the Non-Invasive Assessment of Muscle
TKA: target knee angle
UEQ: User Experience Questionnaire
UI: user interface
VAF: variance accounted for
VF: visual feedback

Edited by S Munce; submitted 04.Jun.2025; peer-reviewed by P Marsico, PE Roos; comments to author 15.Oct.2025; revised version received 09.Dec.2025; accepted 23.Dec.2025; published 26.Feb.2026.

Please cite as:
Zhang Y, García Arias LF, Timmerman H, Cao M, Wilhelm E
Enhancing Knee Joint Proprioception in Healthy Adults Through Exergame Training With Augmented Feedback: Randomized Controlled Pilot Trial
JMIR Rehabil Assist Technol 2026;13:e78525
URL: https://rehab.jmir.org/2026/1/e78525
doi:10.2196/78525
PMID:

©Yiling Zhang, Luis Felipe García Arias, Hans Timmerman, Ming Cao, Elisabeth Wilhelm. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of a Stretching Forearm Sleeve for Lateral Epicondylitis: Repeated Measures Study

Adriana M Ríos Rincón[1*], OTR, MSc, PhD; Christine Guptill[2], MS, OTR, PhD; Ann Tran[1], OTR; Rija Kamran[1], PT; Salamah Alshammari[1,3], OTR; Antonio Miguel Cruz[1,4*], MSc, DSC

[1]Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

[2]Department of Occupational Therapy, University of Ottawa, Ottawa, ON, Canada

[3]Occupational Therapy Program, College of Applied Medical Sciences, King Saud bin Abdulaziz University for Health Sciences, Alahsa, Eastern Region, Saudi Arabia

[4]Glenrose Rehabilitation Research, Innovation & Technology, Glenrose Rehabilitation Hospital, Edmonton, AB, Canada

[*]these authors contributed equally

Corresponding Author:
Antonio Miguel Cruz, MSc, DSC
Department of Occupational Therapy
Faculty of Rehabilitation Medicine
University of Alberta
8205 114 St NW
2-64 Corbett Hall
Edmonton, AB, AB T6G 2G4
Canada
Phone: 1 7804926104
Email: miguelcr@ualberta.ca

Abstract

Background: Lateral epicondylitis (LE) is a condition that impairs daily activities due to pain exacerbated by wrist and hand movements. The ArmLock sleeve is a novel, nonsurgical intervention to stretch the wrist extensor muscles by maintaining the elbow in extension, forearm in pronation, and wrist and fingers in flexion.

Objective: This study aimed to assess the effectiveness of sustained tension provided by the ArmLock sleeve on pain and functional outcomes in adults with LE. This novel device supports the forearm by aligning the elbow in extension, the forearm in pronation, and the wrist in flexion, while keeping the metacarpophalangeal and proximal interphalangeal joints of fingers II-V flexed.

Methods: A within-subjects repeated-measures design was used to assess outcomes at baseline, week 6, and week 12. Nineteen participants wore the device at home for 30 minutes daily for 12 weeks. Outcome measures were administered by research assistants and included pain intensity, pain-free grip strength, pressure pain threshold, pain during resisted wrist extension, and composite extensibility of wrist and finger extensors. Repeated-measures 1-way ANOVA and Friedman tests were conducted (α=.05), followed by post hoc comparisons (α=.017, Bonferroni correction).

Results: Significant improvements were observed in 6 of 7 (85.7%) outcome variables, including increased muscle extensibility, enhanced grip strength, and reduced pain intensity. Participants reported decreased pain and functional disability in a self-reported questionnaire.

Conclusions: Wearing the device daily for 12 weeks led to significant improvements in extensibility, grip strength, and pain reduction. Participants also reported decreased pain and disability. These results suggest that the ArmLock sleeve may support symptom relief and functional gains in individuals with LE. Larger, controlled studies are needed to confirm its effectiveness.

Trial Registration: ISRCTN Registry ISRCTN13309889; https://www.isrctn.com/ISRCTN13309889

(JMIR Rehabil Assist Technol 2026;13:e80400) doi:10.2196/80400

KEYWORDS

lateral epicondylitis; musculoskeletal; orthopedic device; rehabilitation; tennis elbow; upper extremity

XSL•FO
RenderX

Introduction

Lateral epicondylitis (LE), known as tennis elbow, is a frequent cause of lateral elbow pain that significantly impacts the performance of work, self-care, and leisure activities, as the pain intensifies with wrist and hand movements [1]. Aggravating movements often include resisted wrist extension, middle finger extension, and forearm supination with a fully extended elbow [2]. Prevalence studies indicate that LE affects approximately 1.3% of the general population [1], with higher rates observed in working populations (2%-14.5%) [3] and among tennis players (14%-41%) [4]. The condition is predominantly seen in the dominant upper extremity of individuals aged 30-64 years, with peak incidence occurring between ages 45 and 54. LE is strongly associated with repetitive and forceful tasks [1]. As an increasing percentage of the aging population remains physically active, the incidence of LE is expected to rise significantly in the coming years [5]. Given the high prevalence of LE, its socioeconomic impact can be substantial. For example, in the United States, nontraumatic epicondylitis accounts for an annual workers' compensation claim incidence rate of 10.6 per 10,000 full-time equivalent employees [6].

The etiology of LE remains a subject of ongoing debate within the medical community [7]. The most widely accepted causative factor is repetitive microtrauma, resulting from overuse of the wrist extensor or supinator muscles. Such overuse commonly arises from prolonged heavy workloads, intensive training in racquet sports, or improper technique during physical activities [7,8]. While LE has historically been described as a tendinosis based on histopathological findings [7], contemporary models conceptualize the condition as a degenerative tendinopathy characterized by maladaptive responses to mechanical loading, altered mechanobiology, and pain mechanisms that extend beyond local tissue pathology [9]. This distinction has important therapeutic implications, supporting interventions that emphasize load modulation, sustained low-load mechanical stimuli, and functional restoration rather than anti-inflammatory approaches alone. As a degenerative tendinopathy, LE is characterized by maladaptive responses to mechanical loading, which may include excessive or uneven tensile forces acting on the tendon and contributing to microtrauma and degenerative changes [9,10]. In degenerative tendons, the typically organized structure of collagen architecture is disrupted, with fibroblasts and atypical vascular tissue creating a tangled, disorganized matrix with reduced tensile capacity [10]. LE involves chronic tendon degeneration at the common origin of the forearm extensor muscles. These muscles include the extensor carpi radialis brevis (ECRB), extensor digitorum, extensor digiti minimi, and extensor carpi ulnaris, with the ECRB tendon most frequently affected at the lateral epicondyle of the humerus [9,10].

While many individuals with LE experience symptom resolution over time, including spontaneous improvement without formal treatment, most also demonstrate a positive response to conservative management. However, a clinically meaningful subset of patients does not achieve symptom resolution. Approximately 5%-10% of individuals develop persistent symptoms, with pain extending beyond several weeks, exceeding what would be expected from spontaneous recovery alone and suggesting progression toward a chronic or recalcitrant presentation. Patients whose symptoms persist despite appropriate rehabilitation interventions, particularly beyond 3-6 months, are commonly classified as having recalcitrant LE. In this subgroup, continued passive observation is unlikely to yield meaningful improvement, underscoring the need for targeted interventions aimed at modifying tendon loading and mechanobiological processes [9].

A wide range of rehabilitation and nonsurgical interventions has been described for LE, including corticosteroid injections, iontophoresis, botulinum toxin A injections, prolotherapy, platelet-rich plasma (PRP) or autologous blood injections, as well as physical and occupational therapy approaches such as bracing and splinting, stretching and strengthening programs, kinesiotaping, extracorporeal shockwave therapy, and laser therapy. Among these, PRP has been identified as a conservative treatment with higher levels of supporting evidence [11,12]. However, current research does not clearly demonstrate that PRP is superior to other nonsurgical approaches [11,12]. Surgical intervention is typically reserved for approximately 3% of patients who fail to respond to nonsurgical treatments after 1 year from symptom onset [13].

Although the evidence remains somewhat ambiguous, rehabilitation interventions are generally recommended as first-line treatment over medications, corticosteroid injections, or surgery [14]. Braces and splints represent a rehabilitation intervention for LE, particularly for individuals with limited access to other rehabilitation services. These devices aim to reduce mechanical load on the extensor tendons by providing external support. Commonly used orthoses include forearm straps, which decrease force transmission at the common extensor origin, and wrist extension splints, which offload the extensor musculature by immobilizing the wrist. However, evidence regarding their effectiveness in alleviating LE symptoms remains inconclusive [1]. Therapeutic exercise, including stretching, is supported for short-term improvements in pain and function; however, there is no clear consensus on an evidence-based multimodal rehabilitation program for LE, nor on optimal exercise dosage parameters, including intensity, duration, frequency, and progression [14].

Stretching exercises targeting the wrist extensor musculature are commonly incorporated as part of multimodal conservative management strategies for LE [14-16]. A frequently described technique involves wrist flexion with ulnar deviation, combined with elbow extension and forearm pronation, with or without concurrent finger flexion [14,17]. Although the direct effects of static stretching on tendon structure remain incompletely understood, mechanical loading is recognized as an important stimulus for tendon cell activity. Experimental and clinical literature suggests that the mechanical load generated by this upper limb position may support tendon healing in LE by encouraging fibroblasts to align new collagen fibers in an orderly manner, improving the tendon's structural integrity [18]. Accordingly, the mechanical load generated during wrist extensor stretching is hypothesized to contribute to these cellular responses, although the magnitude and durability of such effects in degenerative tendinopathy remain variable and require further investigation.

XSL•FO
RenderX

A promising rehabilitation intervention for LE is the ArmLock sleeve, which applies principles used in therapeutic stretching. The device positions the elbow in extension, the forearm in pronation, and the wrist in flexion, while maintaining flexion of the metacarpophalangeal and proximal interphalangeal joints of digits II-V. This configuration reproduces a frequently prescribed stretching posture used in multimodal rehabilitation programs and is intended to apply sustained mechanical loading to the musculotendinous structures implicated in LE. Specifically, this alignment targets the ECRB, extensor digitorum, extensor digiti minimi, extensor carpi ulnaris, and the superficial head of the supinator, which attach near the lateral epicondyle. By providing prolonged, externally supported positioning, the ArmLock sleeve may offer a practical means of delivering consistent mechanical stimulus as part of a rehabilitation approach, particularly when adherence to repeated active stretching is challenging. However, the effectiveness of the ArmLock sleeve in reducing LE symptoms has not yet been determined. This study aimed to assess the effectiveness of sustained tension provided by the ArmLock sleeve on pain and functional outcomes in adults with LE.

Methods

Study Research Design

A within-subjects repeated-measures design was used to address the research objective. Outcomes related to pain and function were assessed at 3 time points: baseline (week 0), midintervention (week 6), and postintervention (week 12). The use of a within-subjects repeated-measures design to evaluate the effects of the ArmLock sleeve on pain and function in adults with LE aligns with contemporary frameworks for generating evidence in rehabilitation technologies and reflects the technology's level of maturity. Specifically, this study was guided by the Framework for Accelerated and Systematic Technology-based Intervention Development and Evaluation Research (FASTER), which cautions against the routine application of pharmaceutical-style randomized controlled trials (RCTs) to early-stage, rapidly evolving technology-based interventions in disability and rehabilitation [19]. Increasingly, such expectations have been questioned due to their high cost and logistical burden, the risk that extended trial timelines render technologies obsolete, and the limited external validity of tightly controlled experimental conditions. FASTER therefore recommends quasi-experimental designs as rigorous and appropriate approaches for phase 2 (progressive usability and feasibility evaluation). Within-subject repeated-measures designs, in particular, are well suited to this phase because they allow participants to serve as their own controls while also enabling group analysis, thereby facilitating the examination of individual change trajectories and effect signals. The effect sizes and response patterns observed in this study provide critical parameters for subsequent sample size estimation, refinement of research protocol, and hypothesis development in future controlled trials. By situating this study within an iterative, technology-specific evidence-generation pathway, consistent with the FASTER framework, our findings contribute to generating foundational evidence, including effect size estimates, necessary to design future RCTs.

Materials

As illustrated in Figure 1, the ArmLock sleeve is a soft upper-limb orthosis designed to provide sustained positioning and postural support of the elbow, forearm, wrist, and hand. The orthosis incorporates an adjustable strapping system within a semistructured textile sleeve to maintain the elbow in extension, the forearm in pronation, and the wrist in flexion, while positioning the metacarpophalangeal and proximal interphalangeal joints of digits II-V in flexion. The orthosis is fabricated from a breathable, elastic textile composite (74% polyester, 20% rayon, and 6% elastic). Its design allows adjustable traction to increase or decrease wrist and finger flexion to the degree required to achieve the intended positioning. The ArmLock sleeve is classified as a class I medical device, is licensed by Health Canada (Medical Device Establishment Licence [MDEL] No. 7429), and is commercially available.

XSL•FO
RenderX

Figure 1. The ArmLock sleeve.



Participants

Participants were eligible for inclusion if they were 18 years or older, able to communicate in English, and had a self-reported diagnosis of LE accompanied by the following symptoms: (1) pain on the lateral side of the elbow elicited by resisted wrist extension, (2) tenderness at the lateral epicondyle, (3) measurable loss in composite muscle flexibility (extensibility), and (4) symptoms persisting for at least 12 weeks.

Exclusion criteria were applied to minimize confounding factors and ensure the safety and appropriateness of participants for this study. Potential participants were excluded if they (1) had a steroid injection for LE within the past 3 months (steroid injections are a treatment for LE and could confound the study results) and (2) had undergone previous surgery for LE (approximately 10% of postsurgical cases remain symptomatic, suggesting severe LE or underlying conditions unrelated to this study) [9].

Exclusion criteria for symptoms likely due to other medical conditions. These conditions could mimic or contribute to LE-like symptoms or reflect different underlying pathologies, making them unsuitable for this study. Such exclusions included those who (1) had been diagnosed with elbow, wrist, or finger arthritis; (2) had sensory and/or motor changes distal to the elbow, such as carpal tunnel syndrome or elbow joint instability; (3) experienced pain associated with radiculopathy, cervical nerve compression, or thoracic outlet syndrome; and (4) practiced high-velocity racquet sports, such as tennis or badminton.

Exclusion criteria for conditions preventing safe or appropriate use of the ArmLock sleeve included those who (1) exhibited a difference of less than 15 degrees in the range of motion (ROM)

XSL•FO
RenderX

of wrist flexion with fingers extended compared to fingers flexed as such cases indicate limited capacity for further mechanical stretching using the ArmLock sleeve; (2) experienced pain during ligament stress tests at the elbow, which may indicate ligament instability, precluding safe use of the device; and (3) had a wound or scarring in the area where the ArmLock sleeve would be applied as these conditions could prevent proper application and pose risks to healing.

Settings

All assessments were performed at Lab 1-45 Corbett Hall, University of Alberta. Participants took the Armlock sleeve home.

Sample Size

We aimed to recruit 31 participants with a power of 0.8, an alpha of .05, and a medium effect size (f=0.30) [20]. The sample size calculation was based on the Patient-Rated Tennis Elbow Evaluation (PRTEE). In previous studies, sample sizes between 28 and 33 participants were enough to detect a change in the PRTEE (P=.002) [21].

Recruitment

The sampling strategy used was nonprobabilistic. Study details were disseminated through various channels, including social media platforms, online classified advertisement services, regional newspapers, and flyers. Flyers were placed in multiple locations, such as bulletin boards at postsecondary institutions, orthopedic supply stores, and medical supply retail stores. Additionally, the research team reached out to local private clinics and professionals using email lists available on professional association websites in Alberta. Potential participants who expressed interest in the study contacted the research team for more information.

Once participants initiated contact, a member of the research team provided a detailed explanation of the study over the phone and conducted screening questions using a standardized telephone script to confirm eligibility. If an individual met the inclusion criteria and expressed interest in participating, the research assistant emailed the information letter for further review and invited them to an in-person session for upper extremity screening. During this session, researchers explained that the purpose of the screening was to confirm eligibility through a physical evaluation. Participants provided signed consent before proceeding with the assessment. The upper extremity screening included (1) isometric resisted muscle testing, (2) sensory examination (light touch), and (3) ROM assessment of the wrist and fingers.

Any asymmetry between sides or a ROM limitation that prevented the participant from tolerating the position required to wear the ArmLock sleeve resulted in exclusion from the study.

Additionally, the following clinical tests were performed:

1. Elbow Varus Instability Stress Test (to rule out elbow joint instability) [22]
2. Spurling Test (to assess cervical nerve compression) [23]
3. Roos Test (to assess thoracic outlet compression) [24]
4. Upper Limb Reflex Testing (to rule out radiculopathy) [25]

If a potential participant tested positive for any of these assessments, they were deemed ineligible for the study, and the research team informed them that the ArmLock sleeve would not be beneficial for their condition. The research team thanked the participants for their time and provided parking compensation for their participation.

Variables

Dependent Variables (Outcome Variables)

A combination of objective and subjective (self-reported) outcome variables was used to assess the effectiveness of the ArmLock sleeve intervention. The following outcome variables were measured:

Composite Wrist and Finger Extensor Muscles

The composite extensibility of the wrist and finger extensor muscles is the muscle tightness of the wrist and finger extensors (objective outcome variable). For this measure, we followed an adaptation of the Mill Test for LE. Participants were positioned with a closed hand, the wrist in dorsiflexion, and the elbow in extension. The examiner then applied wrist flexion, instructing the participant to resist the movement. The test was considered positive and discontinued when the participant reported pain at the lateral epicondyle [26]. No overpressure was applied, and wrist flexion was measured using a 12-inch clear plastic goniometer (Baseline 360°). The value was recorded in degrees.

Pain-Free Grip Strength

Pain-free grip strength (PFG) is the maximal grip strength at which pain appears (objective outcome variable). This is a common measure of pain in LE. In this study, we followed the protocol published in previous studies of tennis elbow [4,27]. First, to minimize the risk of exacerbating injuries during PFG assessment, participants were first introduced to the concept of PFG. Then, participants were asked to stand with the elbow in a complete extension and the shoulder and radioulnar joints in neutral rotation. They were then instructed to slowly squeeze the dynamometer to their maximum strength using the unaffected arm, allowing them to become familiar with the device. Following this, participants repeated the process with the affected arm, maintaining the same gripping rate as the unaffected side but stopping immediately upon experiencing any discomfort. We asked the participants to perform 3 trials with 1-minute rest intervals, as done in previous tennis elbow studies [4,27]. PFG was measured using the Sammons Jamar Plus+ digital hand dynamometer (200-lb capacity; Performance Health, Sammons Preston). The value was recorded in pounds (lb).

Pressure Pain Threshold

The pressure pain threshold (PPT; objective outcome variable) was defined as the minimum amount of pressure required to elicit pain. Participants were seated facing a table with their forearms resting on its surface. The researcher identified the muscle belly of the wrist extensor group by asking the participant to repeatedly extend their middle finger. Testing began on the unaffected side and was then repeated on the affected side, alternating between sides for a total of 3

XSL•FO
RenderX

measurements. A 20-second interval was provided between each measurement. The PPT was measured using the Commander Echo digital algometer with a 1-cm² rubber probe tip (JTECH Medical), applied to the most tender point over the lateral epicondyle, following the method described in a study by Cho et al [4]. The value was recorded in pounds (lb).

Pain During Resisted Wrist Extension

Pain during resisted wrist extension (2 lb) test (subjective outcome variable) is a common symptom of LE. For this assessment (pain with lifting a 2-lb weight), participants were instructed to hold a 2-lb weight while standing with their arm relaxed. They were then asked to slowly lift the weight by flexing the elbow from 0 to 120 degrees and subsequently return to the starting position, following the method described by Cho et al [4]. Pain experienced during the movement was rated by the participant using a Numerical Rating Scale (NRS) from 0 to 10, with 0 = "no pain" and 10 = "worst possible pain."

Pain and Functional Disability

Pain in the affected arm and functional disability were assessed using the PRTEE (subjective [self-reported] outcome variable). The PRTEE is a 15-item questionnaire that evaluates 2 key components: pain (5 items) and functional disability (10 items). The functional disability section is further divided into 2 parts: specific daily activities (eg, turning a doorknob or key, pulling up pants; 6 items) and usual activities (eg, personal care, household tasks, work, and recreational activities; 4 items). The PRTEE is a widely used, validated assessment tool with excellent psychometric properties [28,29]. The questionnaire provides subscale scores for pain and function, as well as a total score, all ranging from 0 to 100. The total score is the sum of the pain and functional disability subscales (maximum score 100). Higher scores indicate greater symptom severity. In this study, we used the pain subscale, the functional disability subscale, and the total score as outcome variables.

Independent Variable

The independent variable in this study was the use of the ArmLock sleeve. Participants were instructed to wear the device for 30 minutes daily for a 12-week period as an intervention to treat LE. The prescribed duration of ArmLock sleeve use was informed by clinical guidance for stretching protocols in lateral elbow tendinopathy and by parameters reported in prior RCTs. Clinical guidelines commonly describe stretching protocols involving multiple repetitions of 30-45 second holds, typically performed twice daily [30]. Across randomized trials, stretching parameters vary substantially, with hold durations ranging from 6 to 45 seconds, delivered over multiple sets (eg, 1-3), multiple repetitions (eg, 6-15), and performed 2 to 3 times per day [15].

Acknowledging this variability, the decision to prescribe 30 minutes of daily ArmLock sleeve use represented an informed estimation of these parameters, corresponding to an approximate cumulative stretch exposure of ~10-33 minutes when considering total time under stretch across repetitions and sessions reported in the literature. For this study evaluating the effectiveness of a novel technology, we intentionally selected a duration toward the upper end of reported parameter ranges to ensure sufficient stretch exposure while also examining participant tolerance to a sustained, cumulative stretching dose that approximates or exceeds conventional protocols. This approach allowed us to balance theoretical therapeutic dosing with feasibility, safety, and tolerability considerations in our evaluation.

Baseline Participant Information

Participants completed a demographic questionnaire that collected information on age, gender, occupation, side affected, current treatments (eg, hand therapy), and symptom duration.

Procedures

Participants who met the inclusion criteria and provided written consent were invited to attend an initial meeting. During this first meeting, participants completed the demographic questionnaire. A research assistant administered the measures of each outcome variable. We started with the ones that were not painful at all and finished with those that were potentially the most painful for participants. The order of the tests was as follows: PRTEE, composite extensibility, PFG, the pain with lifting a 2-lb weight, and PPT. After these assessments, the research assistant provided each participant with an ArmLock sleeve and instructed them on how to properly wear it. Participants wore the device for 15 minutes to check for any signs of skin irritation (eg, red marks or pressure points) or discomfort. The device was adjusted as needed. Participants were instructed to wear the ArmLock sleeve for 30 minutes daily for a 12-week period, at a time of their choosing. Device use was documented using a Daily Log Sheet provided by the research team. For each session, participants recorded the time the sleeve was donned and doffed, the total duration of wear, and any relevant comments (eg, pain, discomfort, or events that prevented adherence to the recommended 30-minute wear time). Participants were asked to return the completed Daily Log Sheet at their subsequent study visit. The research team used the logged information to monitor adherence and document participant-reported experiences with device use.

Participants attended 2 additional sessions: one at week 6 and a final session at week 12. During these visits, the researcher readministered all outcome measures. At the week 12 session, a brief exit interview was conducted to gather participants' feedback on the use and acceptance of the ArmLock sleeve. Additionally, follow-up phone calls were made during weeks 3 and 9 to monitor participants' use of the device and to remind participants to track use in the Daily Log Sheet.

Statistical Analysis

We described the demographics by reporting the frequency and proportion for categorical variables and the mean (SD) for continuous variables. To determine the normality of the data, histogram analysis and the Shapiro-Wilk and Kolmogorov-Smirnov tests were conducted. Repeated-measures one-way ANOVA and Friedman tests were conducted to evaluate differences among the measurement time points (week 0, week 6, and week 12) for normally distributed and non-normally distributed data, respectively. Post hoc comparisons between the measurement time points were evaluated using paired *t* tests and Wilcoxon signed-rank for normally distributed and non-normally distributed data,

XSL•FO
RenderX

respectively. Bonferroni post hoc correction was used to account for the increased risk of type I errors associated with multiple comparisons. A *P* value <0.05 was considered to indicate statistical significance for ANOVA and Friedman tests. The Bonferroni correction level of statistical significance was corrected at a *P* value <.02 (ie, 0.05/3 tests per outcome variable). All the statistical analyses and visual presentations were performed using IBM SPSS (version 29.0.2.0).

Ethical Considerations

The University of Alberta Research Ethics Board reviewed and approved this study (Pro00095330). Potential participants contacted the research team to express interest, and written consent was obtained before participation. All personal data were deidentified prior to data analysis.

Participants received parking compensation for the time spent in each in-person session, including the Upper Extremity Screening for eligibility determination and the assessment sessions at weeks 0, 6, and 12. Additionally, participants received the ArmLock sleeve at no cost and were allowed to keep it after the study was completed.

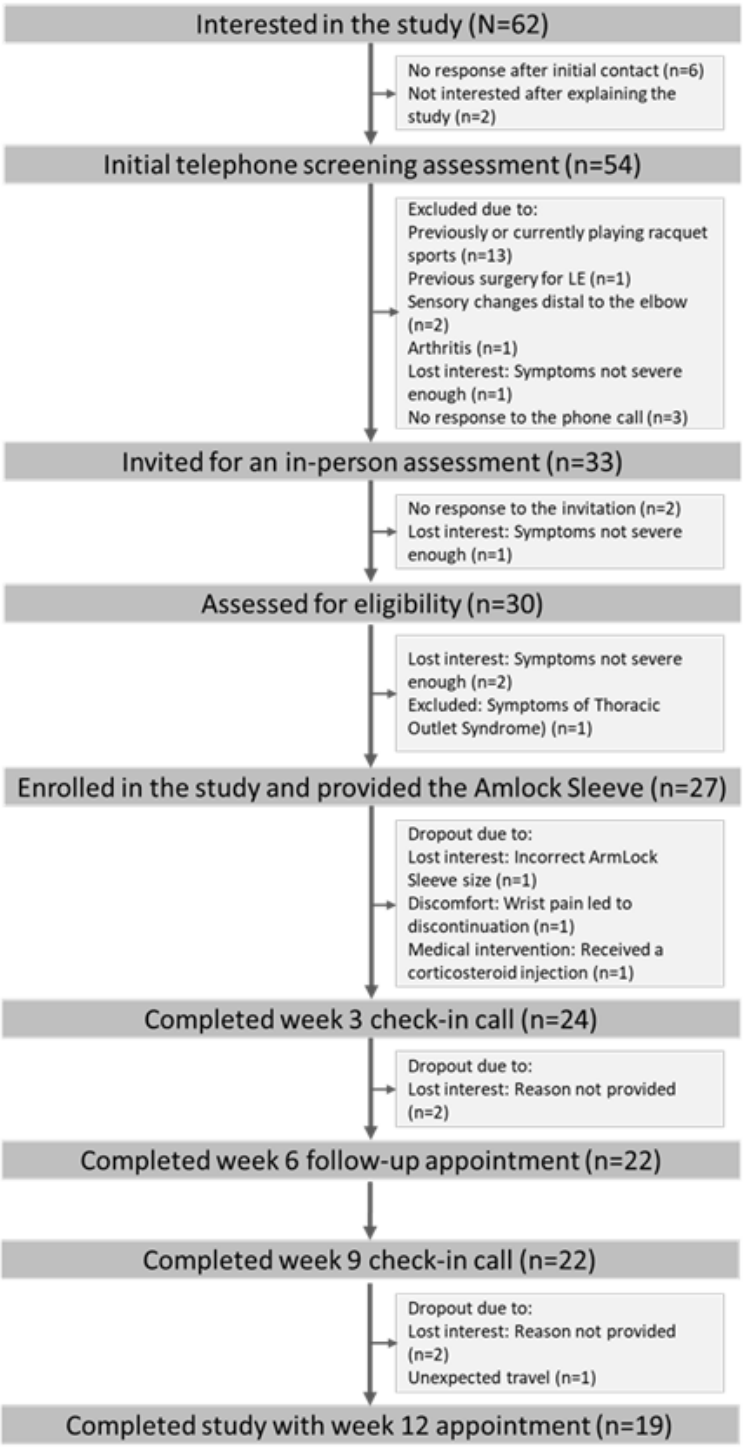
Results

Participant Flow

Figure 2 illustrates the progression of participants throughout the study, between March 23, 2020, and April 30, 2023. A total of 62 individuals from the community contacted the researchers to express interest in participating. Of these, 54 individuals took part in an initial telephone screening assessment, with 21 excluded at this stage. Subsequently, 30 participants proceeded to an in-person eligibility assessment, where 2 declined to participate and 1 was excluded. Ultimately, 27 participants enrolled in the study, completed baseline assessments and received the ArmLock sleeve. All enrolled participants were scheduled to have follow-up assessments conducted at week 6 and week 12. Throughout the study, 8 participants dropped out for various reasons, as detailed in Figure 2. In total, 19 participants completed the study, and their data were included in the final analysis. Based on this participant's flow, the study retention was 70% (19/27). Compliance with the intervention was measured based on the number of days the device was used and the duration of daily use. On average, participants used the device for 62 days, representing 74% of the recommended 84 days. The average daily usage time was 30 (SD 2.62) minutes.

XSL•FO
RenderX

Figure 2. Study flowchart. LE: lateral epicondylitis.



Demographics of Participants

Table 1 presents the demographic characteristics of participants. The majority of participants were middle-aged adults (mean age 44.11, SD 10.57 years), male, and employed. The most commonly reported treatment strategies used in the past to manage tennis elbow included performing ROM exercises and using a brace or elbow strap.

At the time of data collection, nearly half of the participants reported engaging in ROM exercises, approximately one-quarter were taking nonsteroidal anti-inflammatory drugs such as ibuprofen, aspirin, or naproxen, and slightly more than one-quarter reported no nonsurgical concomitant treatments. Additionally, slightly more than half of the participants reported that LE symptoms lasted for 1 hour or more.

XSL•FO
RenderX

Table 1. Demographics of participants (N=19).

Variable	Participants, n (%)
Sex	
Male	13 (68.4)
Female	6 (31.6)
Employment status	
Employed	9 (47.4)
Self-employed	2 (10.5)
Out of work and looking	3 (15.8)
Out of work but not looking	1 (5.3)
Student	1 (5.3)
Retired	2 (10.5)
Unable to work	1 (5.2)
Previous treatment	
Range of motion exercises	13 (68.4)
Brace or elbow strap	13 (68.4)
Nonsteroidal anti-inflammatory drugs	10 (52.6)
Ice	10 (52.6)
Physiotherapy or massage therapy	9 (47.4)
Steroid injections	1 (5.3)
Present treatment	
Range of motion exercises	9 (47.4)
Nonsteroidal anti-inflammatory drugs	5 (26.3)
Ice	3 (15.8)
Physiotherapy or massage therapy	3 (15.8)
Brace or elbow strap	3 (15.8)
Steroid injections	0 (0)
Duration of symptoms	
One hour or more	11 (57.9)
Less than 1 hour	8 (42.1)

Effectiveness of the ArmLock Sleeve on Reducing LE Symptoms

Table 2 presents the descriptive statistics and results of the Friedman tests for the outcome variables across measurement time points. The analysis revealed statistically significant changes in 85.7% (6/7) of the outcome variables, suggesting that the ArmLock sleeve was effective in improving several key measures. Specifically, the intervention resulted in increased composite extensibility of the wrist and finger extensor muscles, enhanced PFG, and reduced pain intensity when lifting a 2-lb weight. Additionally, participants reported decreases in self-reported pain and functional disability, as measured by the PRTEE, with significant improvements observed in the pain subscale, the functional disability subscale, and the total score.

XSL•FO
RenderX

Table 2. Descriptive statistics and Friedman test for the outcome variables across measurement time points (week 0, 6, and 12; N=19)[a].

Outcome variable	Descriptive statistics						Friedman test results		
	Week 0		Week 6		Week 12		Chi-square (*df*)	*P* value	W[b]
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)			
Ext-AA[c] (degrees)	45.0 (8.6)	45.0 (40.0-54.0)	48.6 (5.1)	49.0 (46.0-53.0)	48.0 (8.8)	51.0 (43.0-55.0)	6.08 (2)	*.048*	0.160
PFG-AA[d] (lb)	48.9 (29.9)	43.1 (20.9-76.4)	64.2 (25.5)	59.0 (42.7-88.3)	73.2 (22.9)	76.0 (65.6-88.9)	27.26 (2)	*<.001*	0.717
PPT-AA[e] (lb)	4.1 (1.1)	3.9 (3.2-5.1)	4.3 (1.1)	4.9 (3.4-4.6)	4.7 (1.7)	4.2 (3.5-6.0)	4.89 (2)	.09	0.129
PWL 2lb-AA[f]	3.16 (2.2)	3.00 (1.0-4.0)	1.7 (2.5)	1.0 (0.0-3.0)	0.84 (1.8)	0.0 (0.0.-1.0)	24.00 (2)	*<.001*	0.632
PAINS_SUB[g]	22.0 (8.2)	20.00 (16.0-30.0)	12.5 (9.0)	11.00 (5.0-17.0)	7.9 (9.5)	5.0 (2.0-8.0)	27.26 (2)	*<.001*	0.717
FUNC_SUB[h]	19.6 (10.8)	24.0 (8.5-30.5)	9.7 (9.9)	5.0 (3.0-15.0)	6.0 (9.5)	3.0 (1.0-5.5)	29.95 (2)	*<.001*	0.788
TOTAL_W[i]	41.6 (17.7)	46.0 (27.0-59.0)	22.2 (18.6)	9.0 (9.0-32.5)	14.0 (18.7)	8.0 (4.0-13.0)	28.74 (2)	*<.001*	0.756

[a]Italicized *P* values indicate statistical significance at *P*<.05.

[b]W: Kendall's coefficient of concordance (considered a measure of effect size in the context of the Friedman Test). [c]Ext-AA: composite extensibility of wrist and finger extensor muscles–affected arm.

[d]PFG-AA: pain-free grip strength–affected arm.

[e]PPT-AA: pressure pain threshold–affected arm.

[f]PWL 2lb-AA: pain with lifting a 2-lb weight–affected arm.

[g]PAINS_SUB: Patient-Rated Tennis Elbow Evaluation pain subscale.

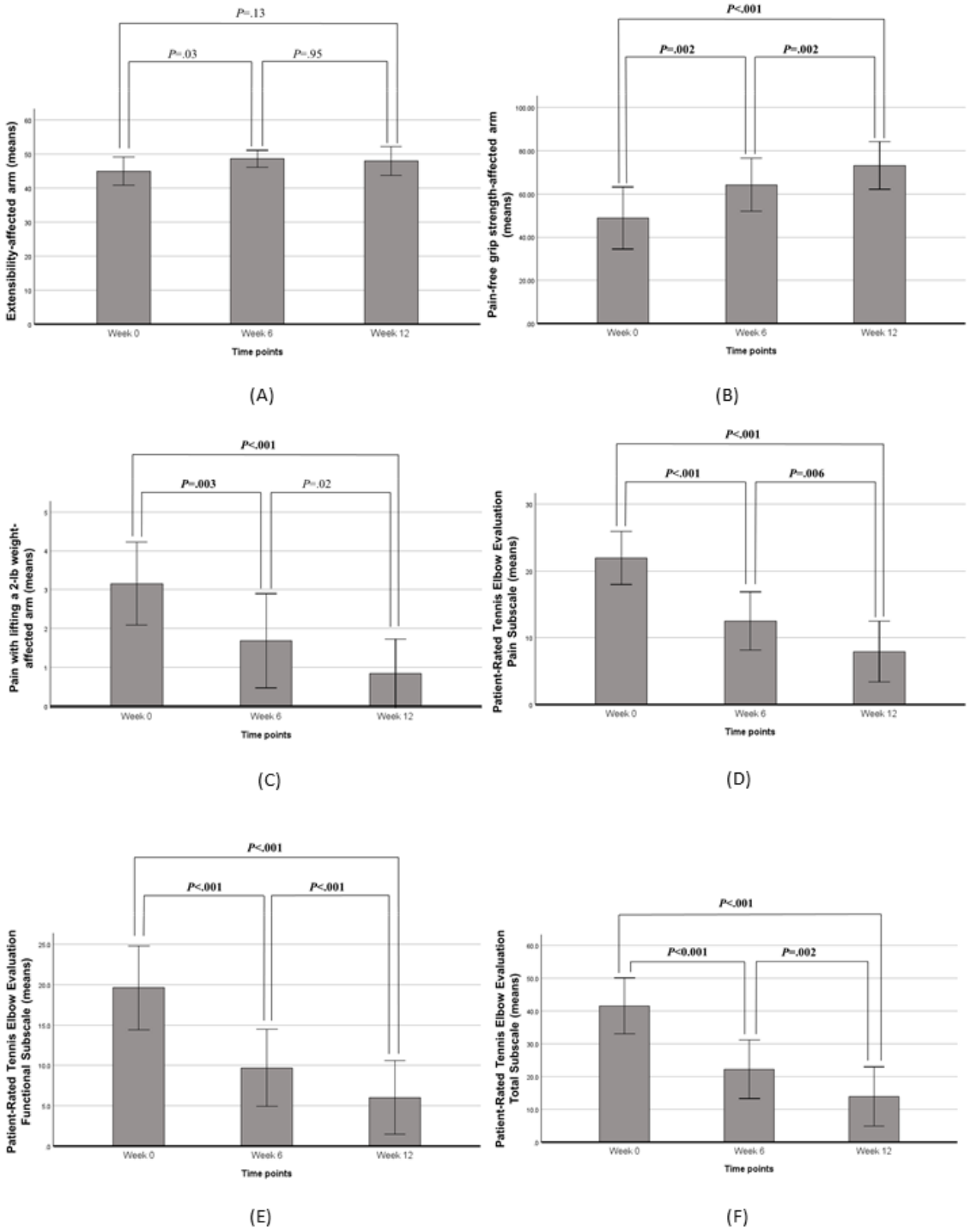
[h]FUNC_SUB: Patient-Rated Tennis Elbow Evaluation functional disability subscale.

[i]TOTAL_W: Patient-Rated Tennis Elbow Evaluation total scale.

Figures 3A-3F illustrates the mean values of each outcome variable for the affected arm across the 3 time points (week 0, week 6, and week 12), along with post hoc comparisons between time points using Bonferroni correction. The pain pressure threshold was excluded from the post hoc comparisons, as the Friedman test did not indicate statistically significant changes for this variable. Error bars represent the SEM at each time point.

XSL•FO
RenderX

Figure 3. Post hoc comparisons between time points. The *P* values in the figure are obtained from the Wilcoxon signed-rank tests.



In Figure 3, Bonferroni correction: $\alpha_{adjusted} = .017$. The $\alpha_{adjusted}$ was calculated as follows: $\alpha_{adjusted} = /m$, where m is the number of tests per outcome variable (ie, $m=3$) and =.05. All the *P* values in bold are statistically significant using the new $\alpha_{adjusted}$.

For composite extensibility of the wrist and finger extensor muscles, the results suggest an initial improvement between week 0 and week 6, followed by stabilization from week 6 to week 12 (Figure 3A). However, despite the observed improvement at week 6, the changes from week 0 to week 6 and week 0 to week 12 were not statistically significant.

PFG showed a progressive improvement over time (Figure 3B), with statistically significant changes at each stage. These findings reinforce the effectiveness of the ArmLock sleeve in enabling participants to achieve a stronger, pain-free grip. The decreasing *P* values indicate that the cumulative improvement from week 0 to week 12 was greater than the improvement observed over shorter intervals.

XSL•FO
RenderX

Pain during resisted wrist extension, measured by the pain with lifting a 2-lb weight test, significantly decreased at each consecutive time point, suggesting a consistent reduction in pain throughout the 12-week period (Figure 3C). The lowest *P* value ($P<.001$) for the comparison between week 0 and week 12 confirms that the overall reduction in pain across the study period was statistically significant.

The subjective assessment of tennis elbow aligned with the objective measures. As shown in Figure 3E and F, both the pain subscale and functional disability subscale of the PRTEE decreased over time, indicating that the ArmLock sleeve had a positive impact on pain reduction and functional limitations in daily activities. Consequently, the total PRTEE score also declined over time, with all observed changes being statistically significant.

Because participants were not restricted from engaging in nonsurgical concomitant care, potential confounding effects were further examined analytically by including concomitant treatment status (yes or no) as a between-subjects factor in a repeated-measures ANOVA. This analysis evaluated whether changes in pain and functional outcomes over the intervention period differed between participants who did (n=14) and did not (n=5) receive concomitant treatments. The results indicated that pain pressure threshold was the only outcome significantly influenced by concomitant treatment status ($\beta=17.073$; $P=0.03$; partial $\eta^2=0.250$). Notably, pain pressure threshold was also the only outcome that did not demonstrate a significant change over time in response to the ArmLock sleeve intervention. Taken together, these findings suggest that concomitant treatments did not meaningfully confound the observed intervention-related improvements in the remaining outcomes.

Discussion

Principal Findings

This study aimed to evaluate the effects of sustained tension provided by the ArmLock sleeve on biomechanical and functional outcomes in adults diagnosed with LE. The results of this within-subjects repeated-measures design study provide evidence of the potential clinical benefits of the ArmLock sleeve in reducing pain and improving function. Despite the small sample size, both subjective and objective measures of pain and functional disability outcomes demonstrated statistically significant reductions, with large effect sizes ($W\geq0.5$). Improvements were observed in composite extensibility of the wrist and finger extensor muscles, PFG, pain with lifting a 2-lb weight, and the pain and functional disability subscales of the PRTEE. These findings suggest that the sustained muscle stretching provided by the ArmLock sleeve contributed to a substantial reduction in symptoms and enhanced functional performance in individuals with LE.

PFG refers to the maximum force an individual can exert during a gripping action without experiencing pain. It is a widely recommended measure for assessing function and treatment progress in individuals with LE [31]. PFG has also been found to be highly correlated with self-reported pain and has demonstrated high accuracy in detecting clinically important changes over time in individuals with LE [32]. Therefore, the significant increase in PFG observed in our study for the 12-week period is supported by a valid and reliable measure capable of detecting meaningful changes in patients with LE.

Hill et al [33] identified a minimal detectable change for PFG in individuals with LE of 9.2 kg force, which is approximately 20.28 lb. In our study, the observed increase in PFG for 12 weeks was 24.3 lb, exceeding the minimal detectable change threshold of 20.28 lb. This suggests that the measured improvement is clinically meaningful and unlikely to be attributed to measurement error.

The results from objective measures align with subjective assessments, reinforcing the intervention's effectiveness. Findings from the PRTEE indicate that participants experienced notable improvements, including reduced pain and discomfort and enhanced arm functionality in daily activities. During the exit interview, participants reported resuming various activities due to the ArmLock sleeve intervention, such as playing hockey and golf, opening jars, typing, picking up a cup of coffee, putting on pants, opening doors, snow shoveling, paddling a canoe, climbing, and carrying heavy bags or groceries. However, 2 (11%) participants reported no perceived improvements with device use. Variations in the device's effectiveness may be influenced by lifestyle factors such as smoking and alcohol consumption, which have been associated with nonsurgical treatment failure [34].

The majority of participants found the ArmLock sleeve easy to put on and remove, indicating good usability. However, some participants faced challenges in adjusting the optimal tightness and correctly using the straps. Among those who completed the study, 3 (16%) participants reported numbness, tingling, or soreness while wearing the device. One participant noted that these symptoms resolved after loosening the wrist strap. Additionally, 1 participant withdrew from the study due to wrist pain. These findings highlight the need for future research to explore potential adverse effects associated with wearing the device and to optimize its design and usage guidelines.

The high retention rate and compliance with the intervention indicate that participants accepted the ArmLock device and tolerated the recommended daily wearing time of 30 minutes. Participants provided feedback on how they adjusted their schedules to accommodate device use, and many appreciated the ability to wear it privately at home. Several participants expressed satisfaction with the outcomes of using the device daily. However, not all were able to do so consistently due to various challenges, including time constraints, forgetfulness, evening fatigue, or a diminished perception of the stretching effect after several weeks of use. These insights highlight the importance of considering individual adherence factors in future implementations of the device.

Finally, the ArmLock sleeve controlled and gentle sustained stretch may promote collagen fiber alignment and increase blood flow, thereby facilitating tendon healing. This mechanism is consistent with the observed reductions in pain and improvements in function demonstrated in this study.

XSL•FO
RenderX

Study Limitations

This study has some limitations. First, the small sample size was influenced by recruitment challenges, particularly in a community setting where potential participants were not referred by health care professionals. Additionally, 15 individuals who initially expressed interest did not respond to follow-up calls or invitations. Future studies should consider recruiting from hospitals and clinics to improve participant retention and sample size. The COVID-19 pandemic further impacted recruitment, delaying the study for 1 year due to restrictions on in-person screening. Second, the lack of a control group limits the ability to attribute improvements solely to the ArmLock sleeve intervention. A placebo, sham device, or standard treatment group would help distinguish the intervention's effects from natural recovery or placebo effects. Third, self-reported compliance may be affected by recall bias or reporting inaccuracies, as participants tracked device usage in a diary. Future studies could incorporate wear-time sensors for more reliable adherence data. Fourth, lifestyle factors such as smoking and alcohol consumption were not measured in our study, highlighting a limitation. Future research on this device should consider assessing these variables to better understand their potential impact on outcomes. Fifth, although LE typically presents with an equal sex distribution [2], the number of male participants in our sample was twice that of female participants. Future research should consider refining recruitment strategies and materials to better engage and include female participants. Finally, the lack of long-term follow-up means it is unclear whether the observed benefits were sustained beyond 12 weeks. Building on the effect estimates generated by this study, future research with larger sample sizes should incorporate extended follow-up assessments at 3 to 6 months postintervention to evaluate the durability of intervention effects.

Conclusion

This within-subjects repeated-measures study provides preliminary evidence supporting the ArmLock sleeve as a promising nonsurgical treatment for LE. Wearing the device daily for 30 minutes across 12 weeks led to statistically significant improvements in more than 80% of the outcome variables. Specifically, the intervention resulted in increased composite extensibility of the wrist and finger extensor muscles, enhanced PFG, and reduced pain intensity when lifting a 2-lb weight. Additionally, participants reported lower self-reported pain and functional disability. These findings suggest that the sustained muscle stretching provided by the ArmLock sleeve contributes to symptom reduction and improved functional performance in individuals with LE. However, future studies with a control group and a larger sample size are needed to provide more conclusive evidence regarding the device's effectiveness.

Acknowledgments

We would like to extend our special thanks to Mykaela Belter, Takeisha Wang, and Carly De Moissac for their valuable assistance with data collection.

Funding

This work was supported by the Mathematics of Information Technology and Complex Systems (Mitacs) Accelerate Program (grant IT15917).

Conflicts of Interest

None declared.

References

1. Sims SEG, Miller K, Elfar JC, Hammert WC. Non-surgical treatment of lateral epicondylitis: a systematic review of randomized controlled trials. Hand (N Y) 2014;9(4):419-446 [FREE Full text] [doi: 10.1007/s11552-014-9642-x] [Medline: 25414603]
2. Ahmad Z, Siddiqui N, Malik SS, Abdus-Samee M, Tytherleigh-Strong G, Rushton N. Lateral epicondylitis: a review of pathology and management. Bone Joint J 2013;95-B(9):1158-1164. [doi: 10.1302/0301-620X.95B9.29285] [Medline: 23997125]
3. Fan ZJ, Silverstein BA, Bao S, Bonauto DK, Howard NL, Spielholz PO, et al. Quantitative exposure-response relations between physical workload and prevalence of lateral epicondylitis in a working population. Am J Ind Med 2009;52(6):479-490. [doi: 10.1002/ajim.20700] [Medline: 19347903]
4. Cho Y, Hsu W, Lin L, Lin Y. Kinesio taping reduces elbow pain during resisted wrist extension in patients with chronic lateral epicondylitis: a randomized, double-blinded, cross-over study. BMC Musculoskelet Disord 2018;19(1):193 [FREE Full text] [doi: 10.1186/s12891-018-2118-3] [Medline: 29921250]
5. Boden AL, Scott MT, Dalwadi PP, Mautner K, Mason RA, Gottschalk MB. Platelet-rich plasma versus tenex in the treatment of medial and lateral epicondylitis. J Shoulder Elbow Surg 2019;28(1):112-119. [doi: 10.1016/j.jse.2018.08.032] [Medline: 30551782]

XSL•FO
RenderX

6. Silverstein B, Viikari-Juntura E, Kalat J. Use of a prevention index to identify industries at high risk for work-related musculoskeletal disorders of the neck, back, and upper extremity in Washington state, 1990-1998. Am J Ind Med 2002;41(3):149-169. [doi: 10.1002/ajim.10054] [Medline: 11920960]
7. Seiça EC, Buruian A, Gameiro D, Peixoto D, Pereira C. Lateral epicondylitis: current concepts in pathology, investigation, and management. SN Compr Clin Med 2023;5. [doi: 10.1007/s42399-023-01410-6]
8. Kroslak M, Pirapakaran K, Murrell GAC. Counterforce bracing of lateral epicondylitis: a prospective, randomized, double-blinded, placebo-controlled clinical trial. J Shoulder Elbow Surg 2019;28(2):288-295. [doi: 10.1016/j.jse.2018.10.002] [Medline: 30658774]
9. Kim JH, Hoy JF, Smith SR, Sabet A, Fernandez JJ, Cohen MS, et al. Recalcitrant lateral epicondylitis: a systematic review on current nonoperative and operative treatment modalities. JBJS Rev 2024;12(8). [doi: 10.2106/JBJS.RVW.24.00059] [Medline: 39106325]
10. Nirschl RP, Ashman ES. Elbow tendinopathy: tennis elbow. Clin Sports Med 2003;22(4):813-836. [doi: 10.1016/s0278-5919(03)00051-6] [Medline: 14560549]
11. Kim CH, Park YB, Lee JS, Jung HS. Platelet-rich plasma injection vs. operative treatment for lateral elbow tendinosis: a systematic review and meta-analysis. J Shoulder Elbow Surg 2022;31(2):428-436. [doi: 10.1016/j.jse.2021.09.008] [Medline: 34656779]
12. Li S, Yang G, Zhang H, Li X, Lu Y. A systematic review on the efficacy of different types of platelet-rich plasma in the management of lateral epicondylitis. J Shoulder Elbow Surg 2022;31(7):1533-1544. [doi: 10.1016/j.jse.2022.02.017] [Medline: 35337955]
13. Riff AJ, Saltzman BM, Cvetanovich G, Frank JM, Hemu MR, Wysocki RW. Open vs percutaneous vs arthroscopic surgical treatment of lateral epicondylitis: an updated systematic review. Am J Orthop (Belle Mead NJ) 2018;47(6). [doi: 10.12788/ajo.2018.0043] [Medline: 29979804]
14. Day JM, Lucado AM, Uhl TL. A Comprehensive rehabilitation program for treating lateral elbow tendinopathy. Intl J Sports Phys Ther 2019;14(5):818-829. [doi: 10.26603/ijspt20190818]
15. Taylor L, Wiebusch M, Bisset LM, Coombes BK. Adherence to exercise in lateral elbow tendinopathy, a scoping review. Musculoskelet Sci Pract 2024;72:102978 [FREE Full text] [doi: 10.1016/j.msksp.2024.102978] [Medline: 38820868]
16. MacDermid JC, Wojkowski S, Kargus C, Marley M, Stevenson E. Hand therapist management of the lateral epicondylosis: a survey of expert opinion and practice patterns. J Hand Ther 2010;23(1):18-30. [doi: 10.1016/j.jht.2009.09.009] [Medline: 19959328]
17. Nagrale AV, Herd CR, Ganvir S, Ramteke G. Cyriax physiotherapy versus phonophoresis with supervised exercise in subjects with lateral epicondylalgia: a randomized clinical trial. J Man Manip Ther 2009;17(3):171-178 [FREE Full text] [doi: 10.1179/jmt.2009.17.3.171] [Medline: 20046624]
18. Shirato R, Aoki M, Iba K, Wada T, Hidaka E, Fujimiya M, et al. Effect of wrist and finger flexion in relation to strain on the tendon origin of the extensor carpi radialis brevis: a cadaveric study simulating stretching exercises. Clin Biomech (Bristol) 2017;49:1-7. [doi: 10.1016/j.clinbiomech.2017.08.008] [Medline: 28826010]
19. Wang RH, Kenyon LK, McGilton KS, Miller WC, Hovanec N, Boger J, et al. The time is now: a FASTER approach to generate research evidence for technology-based interventions in the field of disability and rehabilitation. Arch Phys Med Rehabil 2021;102(9):1848-1859. [doi: 10.1016/j.apmr.2021.04.009] [Medline: 33992634]
20. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. Hoboken, NJ: Pearson/Prentice Hall; 2008.
21. Nowotny J, El-Zayat B, Goronzy J, Biewener A, Bausenhart F, Greiner S, et al. Prospective randomized controlled trial in the treatment of lateral epicondylitis with a new dynamic wrist orthosis. Eur J Med Res 2018;23(1):43 [FREE Full text] [doi: 10.1186/s40001-018-0342-9] [Medline: 30219102]
22. Karbach LE, Elfar J. Elbow instability: anatomy, biomechanics, diagnostic maneuvers, and testing. J Hand Surg Am 2017;42(2):118-126 [FREE Full text] [doi: 10.1016/j.jhsa.2016.11.025] [Medline: 28160902]
23. Jones SJ, Miller JMM. Spurling test. Treasure Island (FL): StatPearls Publishing; 2025.
24. Henni S, Hersant J, Ammi M, Mortaki F, Picquet J, Feuilloy M, et al. Microvascular response to the roos test has excellent feasibility and good reliability in patients with suspected thoracic outlet syndrome. Front Physiol 2019;10:136 [FREE Full text] [doi: 10.3389/fphys.2019.00136] [Medline: 30846945]
25. Miller TA, Pardo R, Yaworski R. Clinical utility of reflex studies in assessing cervical radiculopathy. Muscle Nerve 1999;22(8):1075-1079. [doi: 10.1002/(sici)1097-4598(199908)22:8<1075::aid-mus11>3.0.co;2-u]
26. Valdes K, LaStayo P. The value of provocative tests for the wrist and elbow: a literature review. J Hand Ther 2013;26(1):32-42; quiz 43. [doi: 10.1016/j.jht.2012.08.005] [Medline: 23062797]
27. Akkurt HE, Kocabaş H, Yılmaz H, Eser C, Şen Z, Erol K, et al. Comparison of an epicondylitis bandage with a wrist orthosis in patients with lateral epicondylitis. Prosthet Orthot Int 2018;42(6):599-605. [doi: 10.1177/0309364618774193] [Medline: 29806569]
28. Rompe JD, Overend TJ, MacDermid JC. Validation of the patient-rated tennis elbow evaluation questionnaire. J Hand Ther 2007;20(1):3-10; quiz 11. [doi: 10.1197/j.jht.2006.10.003] [Medline: 17254903]

XSL•FO
RenderX

29. Farzad M, MacDermid JC, Shafiee E, Beygi AS, Vafaei A, Varahra A, et al. Clinimetric testing of the persian version of the patient-rated tennis elbow evaluation (PRTEE) and the Disabilities of the Arm, Shoulder, and Hand (DASH) questionnaires in patients with lateral elbow tendinopathy. Disabil Rehabil 2022;44(12):2902-2907. [doi: 10.1080/09638288.2020.1844318] [Medline: 33180555]
30. IQWiG (Institute for Quality and Efficiency in Health Care). Tennis Elbow: Learn More – Tennis Elbow: Strengthening and Stretching Exercises. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2025.
31. Chourasia AO, Buhr KA, Rabago DP, Kijowski R, Lee KS, Ryan MP, et al. Relationships between biomechanics, tendon pathology, and function in individuals with lateral epicondylosis. J Orthop Sports Phys Ther 2013;43(6):368-378 [FREE Full text] [doi: 10.2519/jospt.2013.4411] [Medline: 23508267]
32. Stratford PW, Levy DR. Assessing valid change over time in patients with lateral epicondylitis at the elbow. Clin J Sport Med 1994;4(2):88-91. [doi: 10.1097/00042752-199404000-00004]
33. Hill CE, Heales LJ, Stanton R, Kean CO. Pain-free grip strength in individuals with lateral elbow tendinopathy: between- and within-session reliability of one versus three trials. Physiother Theory Pract 2023;39(5):1007-1015. [doi: 10.1080/09593985.2022.2030445] [Medline: 35114892]
34. Zhang C, Jia Z, Li J, Wang X, Yang S. Impact of lifestyle and clinical factors on the prognosis of tennis elbow. Sci Rep 2024;14(1):3063 [FREE Full text] [doi: 10.1038/s41598-024-53669-x] [Medline: 38321129]

Abbreviations

ECRB: extensor carpi radialis brevis
FASTER: Framework for Accelerated and Systematic Technology-based Intervention Development and Evaluation Research
LE: lateral epicondylitis
NRS: Numerical Rating Scale
PFG: pain-free grip strength
PPT: pressure pain threshold
PRP: platelet-rich plasma
PRTEE: Patient-Rated Tennis Elbow Evaluation
RCT: randomized controlled trial
ROM: range of motion

Edited by S Munce; submitted 09.Jul.2025; peer-reviewed by K Songur, J Blasco-Abadía; comments to author 03.Oct.2025; revised version received 29.Dec.2025; accepted 13.Jan.2026; published 04.Mar.2026.

Please cite as:
Ríos Rincón AM, Guptill C, Tran A, Kamran R, Alshammari S, Miguel Cruz A
Evaluation of a Stretching Forearm Sleeve for Lateral Epicondylitis: Repeated Measures Study
JMIR Rehabil Assist Technol 2026;13:e80400
URL: https://rehab.jmir.org/2026/1/e80400
doi:10.2196/80400
PMID:41780923

©Adriana M Ríos Rincón, Christine Guptill, Ann Tran, Rija Kamran, Salamah Alshammari, Antonio Miguel Cruz. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 04.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability, Acceptability, and Feasibility of a Personalized Adaptive Mirror Therapy for Upper-Limb Poststroke Rehabilitation Using Immersive Virtual Reality and Myoelectric Control: Single-Arm Pre-Post Study

Daniela De Bartolo[1,2*], PhD; Paolo De Pasquale[3*], PhD; Marta Russo[4,5], PhD; Denise Jennifer Berger[4,6], PhD; Antonella Maselli[4,7], PhD; Daniele Borzelli[4,7], PhD; Christian Nissler[8], PhD; Markus Nowak[8], PhD; Claudio Castellini[8,9], PhD; Giovanni Morone[1,10], MD, PhD; Andrea d'Avella[4,11], PhD

[1]Clinical Laboratory of Experimental Neurorehabilitation, IRCCS Fondazione Santa Lucia, Roma, Italy

[2]Faculty of Medicine and Psychology, Department of Social and Developmental Psychology, Sapienza University of Rome, Roma, Lazio, Italy

[3]IRCCS Centro Neurolesi Bonino-Pulejo, Messina, Italy

[4]Laboratory of Neuromotor Physiology, IRCCS Fondazione Santa Lucia, Rome, Italy

[5]Institute of Cognitive Sciences and Technologies, Consiglio Nazionale delle Ricerche (CNR), Rome, Italy

[6]Center of Space Biomedicine, Department of Systems Medicine, University of Rome Tor Vergata, Rome, Italy

[7]Department of Biomedical Sciences, Dentistry, and Morpho-Functional Imaging, University of Messina, Messina, Italy

[8]Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Weßling, Germany

[9]Assistive Intelligent Robotics Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

[10]Department of Life, Health and Environmental Sciences,, University of L'Aquila, L'Aquila, Italy

[11]Department of Biology, University of Rome Tor Vergata, Rome, Italy

[*]these authors contributed equally

Corresponding Author:
Daniela De Bartolo, PhD
Clinical Laboratory of Experimental Neurorehabilitation
IRCCS Fondazione Santa Lucia
Via Ardeatina 306/354
Roma
Italy
Phone: 39 0651501077
Email: daniela.debartolo@uniroma1.it

Abstract

Background: Stroke remains a primary cause of long-term disability worldwide, with upper-limb deficits affecting up to 80% of survivors acutely and 40% chronically. These deficits lead to considerable effects on their independence and overall quality of life. Conventional rehabilitation therapies are most effective when initiated shortly after a stroke, yet many patients face barriers to ongoing therapy post discharge. Recent advancements in low-cost rehabilitation systems, particularly those using virtual reality (VR) technologies, offer promising alternatives for enhancing upper-limb recovery.

Objective: Given the burden on health care systems and the limitations in access to high-intensity postdischarge rehabilitation, this study aimed to evaluate the feasibility, acceptability, and usability of an upper-limb adaptive mirror therapy using VR and myoelectric control for the rehabilitation of patients with chronic stroke developed through a user-centered design approach.

Methods: In this study, a total of 12 community-dwelling survivors of chronic stroke (mean age 52.9, SD 16.0 years; 4 female) with moderate to severe upper-limb impairments were enrolled. Participants were stratified by age (young: 18-55 years; older: 56-80 years) and impairment level (Fugl-Meyer Assessment-Upper Extremity score: 18-36=severe; 37-54=moderate). Acceptability was assessed for each session by patient self-evaluation of satisfaction and motivation through a visual analog scale, while the therapist assessed the patient's participation in therapy using the Pittsburgh Participation Rehabilitation Scale. Usability was measured with the User Satisfaction Evaluation Questionnaire scale and feasibility through the NASA (National Aeronautics and Space Administration) Task Load Index cognitive workload indices.

XSL•FO
RenderX

Results: Patients reported a significant increase in satisfaction from the intermediate to the final assessment (T1: 72% vs T2: 85%; *P*=.01) and stable high motivation levels. Differences in participation and motivation were observed based on impairment levels, with no effect of age. Usability ratings remained high (>80%) throughout the intervention, with no significant differences between baseline and end line (*P*=.56). Cognitive workload assessments showed a significant reduction over time, in perceived cognitive (*P*=.04), performance (*P*=.007), and effort demands (*P*<.001). Impairment level significantly influenced perceived workload, with participants with more severe impairment reporting higher cognitive, physical, temporal, and effort demands (all *P*<.001), while age did not contribute to variability in acceptability, usability, or workload measures.

Conclusions: VR therapy was found to be feasible, under adaptive task conditions, ensuring stable performance across patients. The protocol was usable and acceptable among patients with chronic stroke, especially those with moderate impairment, supporting its potential as a user-centered digital rehabilitation tool, warranting further investigation in controlled and home-based settings.

Trial Registration: ClinicalTrials.gov NCT07103122; https://clinicaltrials.gov/study/NCT07103122

(JMIR Rehabil Assist Technol 2026;13:e81894) doi:10.2196/81894

KEYWORDS

assistive technology; digital health; mirror therapy; rehabilitation; stroke; users experience; virtual reality

Introduction

Stroke is one of the leading causes of disability worldwide. In Europe, 1.5 million people are diagnosed with stroke every year, costing European societies more than €60 billion [1]. As survival rates improve, life expectancy in general increases, therefore the burden of caring for stroke survivors is likely to increase [2,3]. About 80% of stroke survivors experience upper-limb deficits acutely, while 40% chronically [4,5], and of those with minimal movement on admission, only 11.6%-14.0% regain full function [6,7]. Motor impairments of the upper limb are particularly common and enduring for years after their onset. For this reason, the lack of upper-limb recovery results in significant loss of independence and a reduced quality of life, thus leading to reduced psychological well-being following stroke [8]. To be effective, conventional therapy must be delivered in the first months after the stroke onset with a high rate of intensity [9]. The literature shows that recovery of the upper limb and hand requires long periods of rehabilitation programs that need to be goal-oriented and involve cognitive stimulation [10].

However, due to several factors [11,12], the lack of therapy upon discharge home has resulted in challenges delivering the amount of rehabilitation necessary to optimize recovery [13]. For this reason, in recent years, there has been a focus on the development of low-cost rehabilitation systems that the patient can learn to use during hospitalization and can then use independently to continue the rehabilitation process at home [14]. In this setting, digital applications embedded in virtual reality (VR) systems have been suggested as rehabilitation tools to improve upper-limb recovery [15]. The advantage of these technologies is to provide a motivating treatment without requiring a long therapist contact time [16,17]. The enjoyable and challenging nature of such activities may help address issues of boredom, and, in addition, the ability to provide feedback may enhance motor learning [18], motivation, and exercise adherence [19], and therefore help provide the high-intensity, repetitive practice necessary to drive neuroplasticity-dependent functional recovery [20,21].

While VR rehabilitation systems may enhance the effectiveness of rehabilitation therapy, it remains to be established whether these new technologies can be suitable for every patient, namely, if they can be accepted and used as part of their recovery [22]. Addressing acceptability is not an easy task, as the user's journey toward technological acceptance is complex and often nonlinear. A patient who decides to try a device may not use it in the long run; similarly, someone who has stopped using a system could go back to it later. Developers and researchers must address these issues in the design and evaluation process of new health care technologies, and across different stages of the users' experience of the system.

A second aspect linked to acceptability is the feasibility of using a technological device for patients after stroke, according to motor severity. Several systematic reviews have already reported that both commercially available [23,24] and bespoke versions of VR systems [25] are feasible to use with positive effects on upper-limb recovery following stroke, for those with moderate and mild upper-limb deficits. Evidence in support of the use in patients with severe impairment is less convincing, with studies aiming to include patients with moderate to severe upper-limb impairments showing nonsignificant levels of improvement [26]. Moreover, many studies [27-29], including participants with moderate to severe upper-limb impairments, have used robotics or physical assistance from therapists in addition to VR, suggesting issues with the feasibility of the VR systems when used alone [27] or limited feasibility in the community of patients with stroke [26,30]. While critical to exercise adherence, few studies [21-23] have considered patient evaluation of VR devices, and when such evaluation has been performed, there has often been a lack of analytical rigor.

This study builds upon the technological framework and rehabilitation protocol that integrates VR with myoelectric control to establish a novel approach to mirror therapy (MT) for the rehabilitation of upper-limb neuromotor functions in patients after stroke [31]. The system architecture, control algorithms, and preliminary motor outcomes were previously reported in the framework of the user-centered design approach [32], primarily focusing on the technical validation of the system and its potential impact on motor performance. In contrast, this manuscript addresses a complementary and conceptually distinct

XSL•FO
RenderX

objective, namely, the systematic evaluation of patients' experience conceptualized as the feasibility, usability, and acceptability of the intervention.

In De Pasquale et al [31], a VR-based MT protocol was developed by adapting a virtual therapy arm system originally designed for the treatment of phantom limb pain in upper-limb amputees [33]. The system, based on a head-mounted VR display, motion trackers, and surface electromyography (sEMG) sensors, was repurposed for use in patients with stroke, with particular attention to the feasibility across a wide range of motor impairment levels. The protocol was designed so that the exercises are based on the patient's level of motor proficiency and are thought to restore movements of daily activity. In fact, the virtual scene is the reproduction of different locations of a house in which the patient performs simple exercises that mimic the actions of daily life, such as reaching and grasping a virtual object through myoelectric control.

The integration of these characteristics into a rehabilitation protocol enables the effective use of the beneficial effects of MT, such as the motivational support provided by viewing the more affected limb moving [34], thus overcoming its weak aspects, such as its repetitive nature, limited therapeutic dosage, the necessity for specialized equipment, and professional supervision [35]. While conventional MT requires minimal physical equipment, it typically depends on continuous professional supervision to ensure correct execution and sustained engagement. The proposed VR-based system involves more specialized technological equipment but automates task delivery, feedback, and difficulty modulation, thereby reducing the need for constant supervision during therapy sessions once the system is set up. The addition of myoelectric interfaces to VR for MT rehabilitative protocols aids in interpreting the patient's intentions by analyzing the residual electromyographic signals from the affected limb, so that patients can initiate voluntary movements by using their intact corticospinal pathways. This interaction enables the provision of feedback, such as realistic visual representations of a virtual limb, thereby creating a closed-loop system that fosters motor relearning and stimulates active engagement. As a result, this approach enhances motor coordination, boosts muscle strength, and alleviates spasticity [36,37].

Despite the established benefits of VR technology in enhancing motor functionality through immersive learning environments, evidence supporting the effectiveness of VR combined with MT protocols remains limited, primarily due to small sample sizes, suboptimal research designs, and low-intensity training [38]. In response to these challenges, we have developed a cutting-edge system designed to enhance the efficacy of traditional MT by harnessing the capabilities of VR with myoelectric control.

Accordingly, this study does not aim to assess clinical efficacy or motor recovery outcomes, but rather to investigate whether the proposed VR-based MT can be realistically adopted, sustained, and tolerated by patients with chronic stroke, including those with moderate to severe upper-limb impairment.

In this context, stratification by impairment severity was introduced to explore potential differences in feasibility, acceptability, and perceived workload across functional profiles, rather than to support definitive between-group comparisons. This exploratory perspective is particularly relevant in early-stage evaluations of complex rehabilitation or monitoring technologies [39], where understanding usability across levels of impairment represents a prerequisite for subsequent efficacy trials.

Methods

Ethical Considerations

This experimental study is part of the registered randomized clinical trial (ClinicalTrials.gov NCT07103122) and therefore uses an exploratory, single-arm pilot design focused on user experience and feasibility outcomes, without a control group. It was conducted in full compliance with the Declaration of Helsinki for research involving humans. The local ethical committee of the Fondazione Santa Lucia approved this research on December 2, 2019, (protocol CE/Prog. 790). Patients and caregivers were informed about the study procedures prior to enrollment into the study, and patients provided written informed consent.

Participant privacy and confidentiality were ensured in accordance with standard data protection procedures. All data were anonymized prior to analysis, stored on secure institutional servers, and access was restricted to authorized research personnel only. No financial compensation or reimbursement was provided to participants for their participation in the study.

Participants

Based on our pilot study [31], we decided to enroll only patients with chronic stroke, who typically exhibit greater clinical stability, allowing for clearer assessment of treatment effects without the fluctuations seen in acute or subacute patients. Therefore, patients who had experienced a stroke more than a year before the first evaluation were enrolled for this study.

This study was designed as an exploratory pilot focusing on feasibility and user-experience outcomes rather than hypothesis testing. Therefore, the aim was to estimate usability/workload and feasibility parameters rather than to test efficacy; thus, a priori power analysis was not performed. Sample size was determined to provide adequate precision of key continuous user experience measures and to inform the design of a future adequately powered trial.

Criteria for enrollment were (1) upper-limb deficits with sufficient level of muscle power such that movement is possible with gravity eliminated (Medical Research Council [MRC] grade 2) or against gravity (MRC grade 3), as assessed using the MRC scale (grades ≥1) [40]; (2) scores ranging from 18/66 (27%) to 54/66 (80%) of upper-limb function, as assessed using the upper-limb extremity section of the Fugl-Meyer Assessment (FMA) scale [40]; (3) absence of severe linguistic impairments that may limit understanding of instructions; and (4) absence of cognitive impairments, as assessed with the Addenbrooke's Cognitive Examination-Revised version for the Italian population [41], visual deficit, or other neurologic disease in comorbidity that may affect the patient's ability to interact with the VR environment.

XSL•FO
RenderX

Clinical motor scales were collected at baseline to characterize the sample and stratify participants by impairment level; however, they were not analyzed as outcome measures in this manuscript. This choice reflects the exploratory nature of the study and the limited sample size, which suggested promising trends, although further investigation is needed to confirm rehabilitation efficacy.

Since age-related [42] and severity-related [7,26] differences were found in motor outcomes following VR upper-limb treatments, we created a blocked randomization list with 2 levels of stratification (age and severity), which guaranteed a balanced within-group split throughout the data collection process [43]. Therefore, participants aged between 18 and 55 years were considered young participants according to previous studies [44-46], while those aged between 56 and 80 years were considered older. The level of impairment was computed using 2 FMA-Upper Extremity–based cutoffs selected for functional relevance rather than normative severity thresholds. Patients scoring between 18/66 (27%) and 36/66 (54%) were classified as having severe impairment of the motor performance. This range reflects a markedly reduced voluntary motor control, typically associated with limited functional use of the upper limb.

Conversely, scores between 37 (56%) and 54 (80%) were classified as moderate impairment, as an indicative range of partial motor recovery and greater residual motor capacity, allowing for more consistent execution of goal-directed movements. The distinction between these 2 categories was chosen to ensure a clinically meaningful separation between patients with profoundly compromised motor function and those with a moderate, yet functionally relevant, level of recovery.

Given the limited sample size within each impairment subgroup, stratification was not intended to enable powered statistical comparisons, but rather to ensure representation of multiple functional levels of motor impairment, thus supporting an exploratory evaluation of feasibility-related outcomes.

We used the sealed envelope tool, whose implementation is based on Pocock [44], to create the randomization list of participants as a function of the stratification number of randomization blocks (age and level of impairment), which revealed the need to enroll at least 12 patients. This sample size was also used in previous work [16] about the effectiveness of the VR system on the rehabilitation of the upper limb in patients with stroke.

Although the rehabilitation protocol and technological platform are shared with the previous study [31], this analysis focuses on a different set of outcome measures, enrolled patients, and research questions. Specifically, the data reported here concern user experience, participation, usability, and perceived workload, which were not the primary outcomes of the earlier publication.

Apparatus and Equipment

A detailed description of the equipment is described in previous publications [31,33]. Briefly, the system consists of an HTC Vive VR System, including a head-mounted display, 2 motion trackers, 2 controllers, 2 base stations, which emit infrared light and synchronization signals to enable accurate spatial tracking of the head-mounted display, controllers, and trackers, and 2 Myo armbands (Thalmic Labs). Each tracker of HTC and each Myo armband were positioned respectively on the hand dorsum and on the forearm of each participant. The tracker system provides the position and orientation of the patient's hand. The Myo armband, consisting of 8 sEMG sensors, measures the forearm electric muscle activity of the user on the skin surface. The movement of the virtual hands is directly inferred from the kinematic data captured by the trackers, whereas the motion of the fingers is estimated based on forearm muscle activity. This estimation relies on a model trained using a machine learning approach consisting of an iterative version of random Fourier features ridge regression [47]. The acquired signals (tracker kinematic and sEMG) are transmitted wirelessly via Bluetooth. All processing, which includes signal acquisition, filtering, machine learning model training, and prediction, is carried out on an off-the-shelf laptop (Dell Alienware 15) with a dedicated GPU (NVIDIA GeForce GTX 1070).

Software and the Virtual Environment

The rehabilitation software runs in Unity, a game engine platform commonly used for the creation of interactive 3D applications and VR environments. It consists of 3 different virtual environments in which a table in front of the patients matches the position and dimensions of a rehabilitation table. For each patient, an anonymous form is created and filled with the participant's ID, date of birth, more affected side, and dominant hand. Before starting the training session, the program requires the acquisition of the positions of the table corners with 1 of the 2 controllers in use for the VR session for calibration, so that the virtual model of the table is spatially aligned with the corresponding physical table where the patient is performing the exercise. This was necessary to provide correct visuo-haptic feedback during the VR exposure.

Procedure

After the clinical evaluation, each patient was enrolled for experimental training with the VR-MT. Before starting, a researcher explained the task to facilitate the patient's interaction with the VR system. For all the patients, this was the first experience with a VR system in which myoelectric control must be exploited to achieve a task. The patient performed VR treatment while sitting on a chair with the possibility of resting their hands on a table. At the start of each session, the researcher carefully positioned the 2 Myo armbands over the brachioradialis muscles on each arm and attached the trackers to the back of the patient's hands, aligning the z-axis between the index and middle fingers. As the system supports bimanual rehabilitation tasks by leveraging VR capabilities to provide mirrored assistance, assessment of proximal (limb movement) and distal control (hand gestures).

Prior to each session, to personalize the training exercises by calibrating the reachable workspace the patient was asked to (1) assume a resting pose, with a comfortable position of the hand on the table, (2) extend the less impaired arm to the maximum range in the anteroposterior and mediolateral directions, as shown in Figure 1, and (3) extend the less affected limb in the same directions with the more affected limb, as much

XSL•FO
RenderX

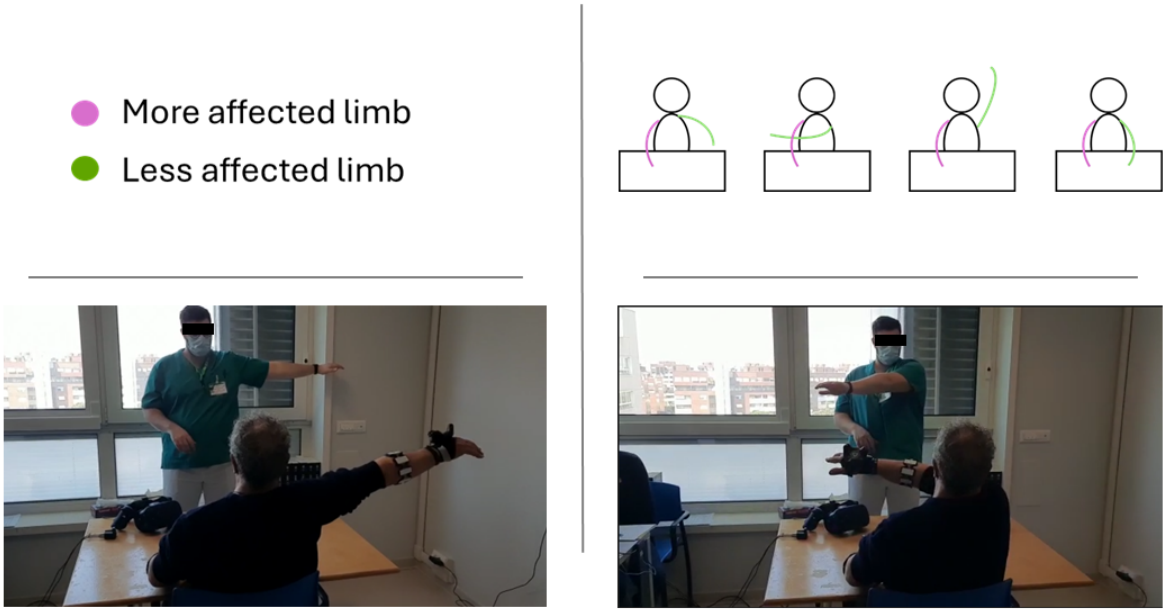
as they could. For each arm, the software evaluates the reachable area.

Although the calibration procedure visually emphasizes the movements of the less affected limb, the reachable space of the more affected limb is estimated based on the patient's attempted movements with the more affected arm during the calibration phase. Specifically, patients were instructed to extend, first the less affected, and then the more affected limb in the anteroposterior, vertical, and mediolateral directions, with the latter moving as far as their residual motor capabilities allowed. The system records the maximum spatial extent achieved by the more affected limb during these attempts, which is then used as an estimate of its reachable workspace.

A secondary assessment was dedicated to calibrating the visualization of the opening and closing of a virtual hand (gesture) according to the electromyography signals recorded by the Myo armband. The patient is asked to (1) assume a resting pose with a comfortable position of the hand on the table and minimal muscle activity, (2) reproduce the gesture of holding a ball with the fingers extended, and (3) reproduce the gesture of holding the handles of a concertina with finger flexed.

After acquiring these 3 independent recordings, the researcher verifies the quality of gesture recognition and decides if the neural network needs more training or not. Otherwise, to improve the gesture reconstruction, it is possible to add new gesture repetitions, discard all the recordings of a gesture, or discard them all and redo the training gesture from scratch.

Figure 1. Example of a calibration procedure. In the right-upper part, a mannequin depicts the different steps involved in the calibration; in the lower part, 2 pictures show a patient performing the procedure before the beginning of the rehabilitation session.

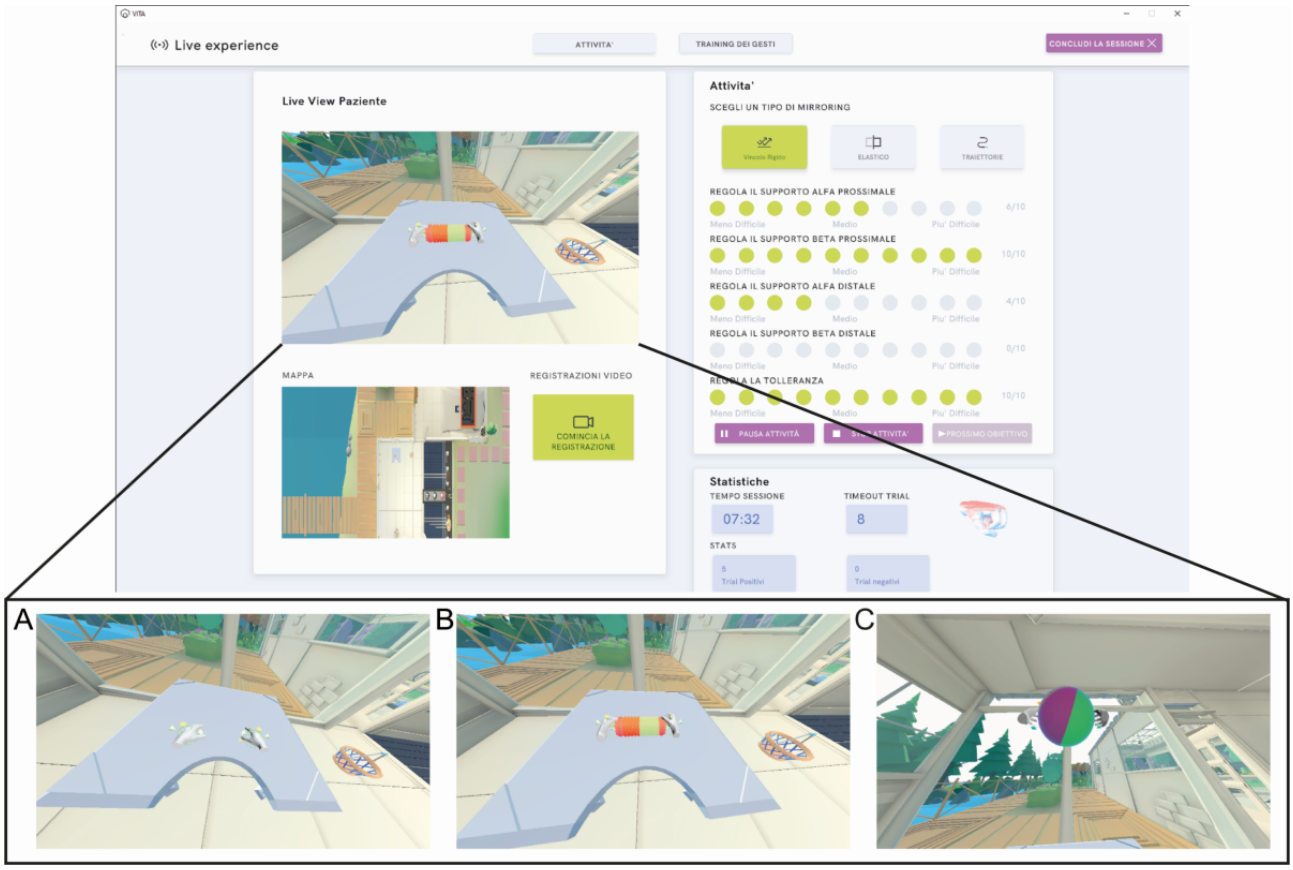


Training Protocol

The training comprised 24 sessions, each lasting about 20 minutes in a quiet room using a height-adjustable table and ergonomic seating to allow participants to perform reaching tasks in a physiologically appropriate posture, independently of body size. In the first session, the researcher encouraged the patient to visually explore the environment to facilitate familiarization with the VR scene, also avoiding possible distractions during the training. Then, the patient receives verbal instructions to perform the first tasks. The training involves 2 reaching virtual objects (a ball or a concertina) presented alternately to elicit different hand gestures. Specifically, interacting with the concertina requires performing a fist gesture (closing the fingers), while the ball requires a hand extension gesture (opening the fingers). Both the concertina and the ball feature 2 handles, positioned on opposite sides, to guide correct hand placement during the task, as depicted in Figure 2.

XSL•FO
RenderX

Figure 2. Overview of the Virtual Therapy Arm (VITA) interface and example of a session. The top panel shows the main user interface. On the left, the "Live View Patient" displays the real-time 3D visualization of the virtual environment from the participant's perspective. On the right, the "Activity" panel summarizes session information, including task progression, performance indicators (green dots), and overall statistics (eg, total time and score). The bottom panels are representative of exercises available: (A) resting pose, (B) the fist gesture needed to grasp the concertina, and (C) the open hand gesture as to take a balloon.



A reaching movement is considered successful when both virtual hands are positioned within 5 cm of the respective target handles. Real-time feedback on hand placement and gesture execution is provided within the VR environment. When both virtual hands are within the required tolerance, the handles change color from orange to green. Additionally, an on-screen gesture icon indicates the required hand posture and disappears once the correct gesture is performed. After both the position and gesture requirements are met, the target object gradually turns green, visually signaling the required holding time before task completion. Auditory feedback is also provided to enhance user awareness. A positive sound indicates successful task completion, while a negative sound is played if the trial exceeds the maximum allowed duration of 20 seconds. The position of each target varies in height and lateral deviation with respect to the initial resting position of the hands, which is recorded at the beginning of the session with the hands placed on the table. The maximum distance at which the target can appear is tailored to the maximum arm extension, while the minimum matches the resting position, both recorded by the patient prior to each session, for a total of 12 possible target positions. Targets were presented in a randomized order to prevent the repeated presentation of highly demanding trials, such as targets positioned near the limits of the participant's reachable workspace or requiring more effortful hand gestures, which could otherwise induce premature fatigue. All training sessions were supervised by a physiotherapist, with specific attention to movement quality and prevention of compensatory strategies, such as excessive shoulder elevation or trunk displacement during reaching. Verbal feedback was provided to encourage controlled arm extension while maintaining proper posture. The patient is guided by the therapist to follow the instructions in the VR environment on how to place their hands inside the 2 spheres that appear next to the target object, with open hands in the case of the ball, with closed fists in the case of the accordion.

The first training session is dedicated to finding the right match among parameters according to the level of ability of the patient; therefore, a baseline level, which also serves as a familiarization phase to the experimental training. It consists of alternatively setting α and β parameters to 1 (independence by the less affected limb) and 0 (totally dependent). The baseline is intended to identify the optimal level of performance at the VR task for each patient; it must be difficult enough to create a challenging exercise but not excessively engaging, so that it can be sustainable for the duration of the training session without frustration. At the end of each block, consisting of 12 reaching movements, the therapist adjusts the proximal and distal agency and capability parameters based on the patient's performance score. This score, calculated from the success rate within the block, guides the modulation of task difficulty for the subsequent block. When the patient achieves a success rate of >90%, corresponding to at least 11 successful trials, task difficulty is increased to promote greater motor engagement. If the success

XSL•FO
RenderX

rate falls within 70%-90% (9-10 successful trials), task difficulty remains unchanged, ensuring the patient continues to perform within an optimal challenge level. Conversely, if the success rate is below 70% (≤8 successful trials), task difficulty is reduced to keep the task within the patient's current capabilities and to prevent frustration or disengagement. This adaptive strategy was adopted to ensure that task performance remained within a sustainable range (ie, not falling below approximately 80%), thereby operationally defining feasibility as the ability to maintain engagement and performance over repeated sessions rather than complete task autonomy. Indeed, this framework is intended to sustain motivation and maximize motor improvement by continuously tailoring task difficulty to the patient's performance, with the goal of maintaining a success rate within the target range of 70%-90%. The specific parameter to be adjusted, as well as the direction of adjustment, is determined by the therapist, who makes these decisions based on clinical expertise and real-time observation of the patient's motor performance, fatigue, and overall condition. A complete description of how to set these 2 parameters to obtain virtual assistance is described in Multimedia Appendix 1.

Acceptability, Usability, and Feasibility Evaluation

A total of 3 assessments with clinical scales and questionnaires were provided. A motor baseline at T0 (enrollment) was performed by a physiotherapist, which evaluated the patient's muscle power through the MRC classification, and the residual motor proficiency with the FMA scale. Assessment was further performed at the intermediate evaluation (T1, after the completion of 12 sessions), and at the end of the therapy (T2, after the completion of 24 sessions). If the patients were discharged or expressed a willingness to interrupt the trial, the T2 evaluation was anticipated. Since acceptability reflects participants' subjective evaluation of the intervention after experiencing it, a meaningful baseline assessment at T0 was not considered conceptually appropriate. At T0, participants had not yet interacted with the VR system and therefore could not provide an informed evaluation of satisfaction or motivation related to the intervention. For this reason, changes in acceptability were evaluated between the intermediate (T1) and final (T2) assessments, when participants had already gained direct experience with the therapy.

Acceptability is a multifaceted construct that reflects the extent to which people delivering or receiving a health care intervention consider it to be appropriate [48]. Here, we assessed this construct according to the emotional responses the patients exhibited during the intervention using 2 visual analog scales for motivation and satisfaction. As in a previous study [16], patients rated their level of motivation to attend the therapy and the satisfaction of doing the therapy for each session on a visual analog scale ranging from 0 to 10 (with a higher score indicating a higher level of motivation/satisfaction). In addition, the therapist/researcher assessed patient adherence to therapy, reporting on a Likert scale, ranging from 1 to 6 (where 6 stands for the highest level), the patient's participation in the exercise through the Pittsburgh Participation to Rehabilitation Scale [49].

Usability, intended as the users' potential to achieve their goals with effectiveness, efficiency, and satisfaction in a specified context of use of technology [50], was assessed using specific tools. Precisely, the User Satisfaction Evaluation Questionnaire (USEQ) [51] is a questionnaire properly designed to evaluate the satisfaction of the user in virtual rehabilitation systems. It is composed of 6 items assessing (1) the user's experience with a VR system with respect to the enjoyment associated with the experience, (2) the success rate in using it, (3) the perceived control of the system, (4) the clarity of information provided by the system, (5) the level of comfort experienced, (6) and the evaluation of its impact for the patient's rehabilitation. We performed this assessment at the first (T0) and last (T2) sessions of each patient journey, to evaluate their first impression of the VR system versus the final considerations acquired after multiple uses of the system.

Finally, in the framework of the person-based approach, we assessed the feasibility of our VR intervention as a workload [52] associated with the experimental VR therapy. For this assessment, we used the NASA (National Aeronautics and Space Administration) Task Load Index (NASA-TLX) [53] for the multidimensional subjective assessment rates of perceived workload associated with the task. Patients compiled this scale at the end of each training session, making an evaluation of the task according to mental, physical, temporal, effort, and frustration demands. Each domain is scored on a straight line that goes from 0 to 10 points (with 10 as the highest perceived workload), and the patients mark the point that corresponds to their evaluation.

Statistical Analysis

The statistical analysis was performed using IBM SPSS statistical software (version 23), while customized algorithms were implemented in MATLAB (version 2020b; MathWorks) environments for data visualization. Data are reported in terms of means and SDs for continuous measures, as medians and IQRs for ordinal measures, and percentages for reporting variables expressed as frequency. The Kolmogorov-Smirnov test revealed a deviation from a normal distribution of all variables, thus nonparametric tests were adopted for comparing performance and ratings across the different times of evaluation (T0, T1, and T2). The Wilcoxon signed-rank test for related samples was used to assess the between-group analysis of differences in the self-assessed evaluation of patients after T1 and T2 sessions (after 12 and 24 treatment sessions, respectively). As a post hoc analysis, to assess differences within group (age and impairment), the Mann-Whitney *U* test was performed. The Spearman coefficient was further used to evaluate effect size in post hoc analysis to further test the contribution of impairment to the self-assessed cognitive workload domains. Analyses involving impairment level were conducted with an exploratory intent, aimed at identifying descriptive trends and potential patterns relevant to feasibility and user experience, rather than at testing formal hypotheses between subgroups.

Longitudinal usability outcomes were analyzed using linear mixed-effects models to account for repeated measurements and incomplete observations across sessions. Separate models were fitted for motivation, satisfaction, and participation after normalization to their respective theoretical maxima (motivation

XSL•FO
RenderX

and satisfaction: 10; participation: 6), resulting in bounded outcomes in the range 0-1.

For each outcome y_{ij}, measured for subject i at session j, the following model was specified:



where $\text{Session}_{c,ij}$ denotes the session index centered around the sample mean, Group_i represents the impairment group, $b_{0i} \sim N(0, \sigma^2_b)$ is a subject-specific random intercept, and $\varepsilon_{ij} \sim N(0, \sigma^2)$ is the residual error term.

Models were estimated using maximum likelihood. Fixed effects were evaluated to characterize overall temporal trends, baseline group differences, and exploratory group-by-session interactions. Subject-specific random intercepts were included to account for within-subject correlation induced by repeated measures. Missing data were handled implicitly by the model under a missing-at-random assumption, without imputation. Given the limited sample size within impairment subgroups, all linear mixed-effects model analyses were considered exploratory and hypothesis-generating.

Only data from participants who completed a sufficient number of training sessions (≥10 sessions) were included in the usability analyses. This threshold was selected as a pragmatic criterion corresponding to approximately 3 weeks of repeated use, considered sufficient to move beyond initial familiarization effects and to obtain stable user experience measures, in line with usability research indicating that reliable usability estimates emerge after repeated interactions rather than during early exposure [54].

For all the tests, the significance was set at 0.05, except for post hoc analyses for which the α level of significance, according to Bonferroni correction, was divided by the number of possible comparisons (equal to 8), thus set at 0.006.

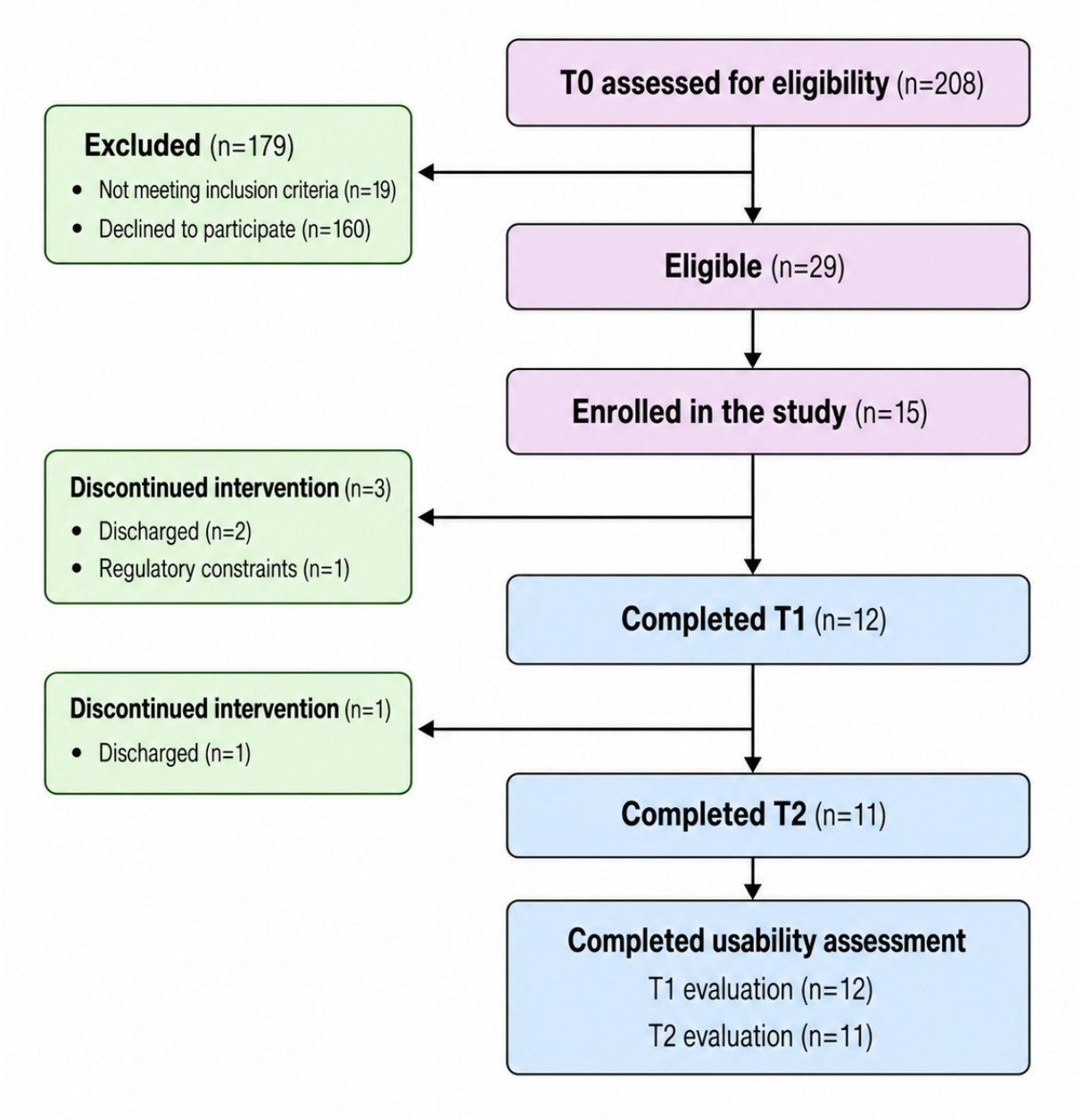
Results

Overview

A total of 15 community-dwelling stroke survivors were initially recruited for the study. Of these, 12 community-dwelling stroke survivors (4/12, 33% females) aged between 27 and 72 (mean 52.9, SD 16.0) years completed a sufficient number of training sessions and were therefore included in the usability analyses. Three participants were excluded from the usability analyses due to insufficient exposure to the intervention, having completed fewer than 10 sessions. A detailed overview of participant flow is provided in the participant flow diagram (Figure 3).

XSL•FO
RenderX

Figure 3. Participants flowchart. Research workflow at the 3 time points of assessment corresponding to the first evaluation (T0), after half of the treatment (T1), and the last (T2).



Stroke chronicity was between 1 and 28 (median 2.9, IQR 1.8-8.4) years, and participants were classified as having moderate to severe motor impairment according to the FMA-Upper Extremity score. A total of 4/12 (33%) patients reported weakness following the stroke, on the right side of the body, and only 3/12 (25%) had a mild linguistic disorder (eg, anomia and dysarthria) that did not compromise either the patient's ability to communicate or comprehend the verbal instructions given during the intervention sessions. Demographic and clinical information of all participants are presented in detail in Table 1.

XSL•FO
RenderX

Table 1. Demographic characteristics of patients with stroke. The data are presented according to age and motor classification. Motor impairment is quantified by the score obtained through the evaluation of the more affected limb with the Fugl-Meyer Assessment (FMA) scale, normalized to the total value of the upper extremity section and expressed as a percentage, as in a previous study [55]. Muscle power is reported as Medical Research Council (MRC) quality grade both for flexion (-f) and extension (-e) assessment of the more affected limb. The time since stroke (TSS) is reported in years; muscle power is the quality grade obtained with the MRC scale, assessed both for flexion and extension manipulative tasks. The side of weakness refers to the most affected side by stroke. In the "Training" column, the amount of training completion for each patient is reported.

Patient ID	Age		Motor impairment		Sex	TSS (years)	Muscle power		More affected limb	Linguistic deficit	Training (%)
Sub01	old	72	severe	45	Male	12	3	3	Left	No	100
Sub02	old	72	severe	29	Female	2	2	2	Left	No	100
Sub03	old	64	severe	41	Male	1	4	4	Left	No	96
Sub04	old	70	moderate	62	Male	7	4	4	Right	No	83
Sub05	old	64	moderate	82	Male	2	4	3	Right	Yes	100
Sub06	old	58	moderate	73	Male	1	3	3	Left	No	92
Sub07	young	48	severe	35	Male	28	2	2	Left	No	83
Sub08	young	50	severe	52	Male	3	3	3	Left	No	100
Sub09	young	27	severe	36	Female	4	3	3	Right	No	42
Sub10	young	38	moderate	68	Female	25	4	4	Left	Yes	92
Sub11	young	39	moderate	80	Male	2	4	4	Right	Yes	96
Sub12	young	33	moderate	61	Female	3	4	4	Left	No	88

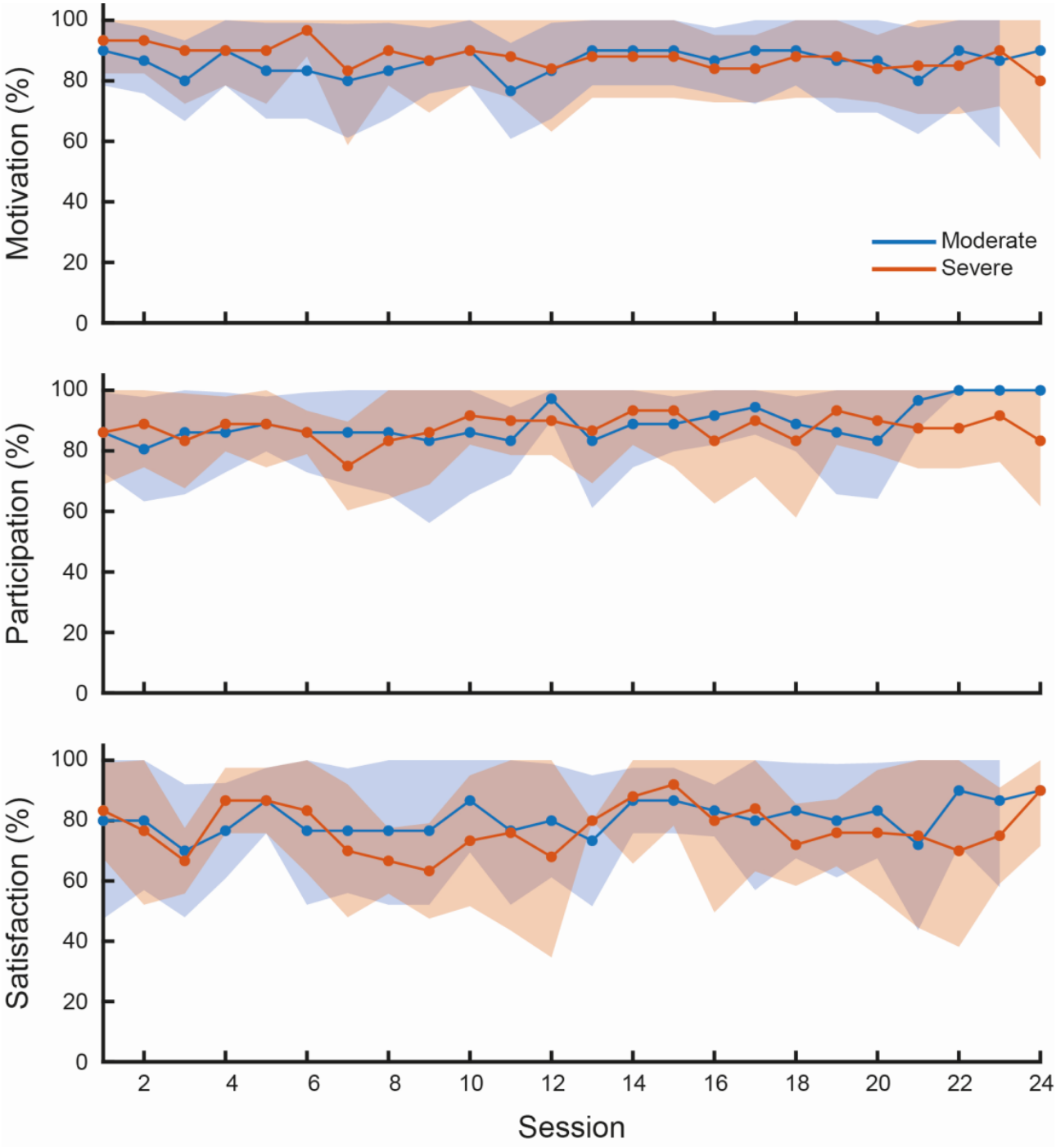
All participants completed the whole experiment except for Sub09, who dropped out prior to the intermediate evaluation (T1); therefore, this participant was not included in the statistical analysis for T2 but was included for analysis at the other assessment points. Linear mixed-effects models showed consistently high levels of motivation, satisfaction, and participation across participants throughout the training period. Motivation exhibited a high baseline level and remained stable over sessions, with no significant effects of impairment group or session, and a nonsignificant trend toward a group-by-session interaction (group: P=.55; session: P=.31; group × session: P=.07). Satisfaction similarly showed high baseline values and remained stable over time, with no significant main effects or interactions (group: P=.72; session: P=.41; group × session: P=.86). In contrast, participation showed a significant positive effect of session, indicating a progressive increase over time, while no significant effects of impairment group or group-by-session interaction were observed (session: P=.01; group: P=.96; group × session: P=.37).

Acceptability

In line with the overall longitudinal trends, acceptability outcomes were further explored by comparing intermediate (T1) and final (T2) assessments to characterize clinically meaningful changes over time and across impairment levels. Comparisons of reported scores at T1 and T2 revealed a significant difference (P=.01, Wilcoxon signed-rank test) in the mean level of satisfaction declared by patients between the T1 (72%) and T2 (85%), as well as a significant difference (P=.01, Wilcoxon signed-rank test) in the mean level of patient participation assessed by the therapist at T1 (93%) and T2 (89%), but not in the level of motivation (P=.38, Wilcoxon signed-rank test) which for moderate patients always remained very high as shown in Figure 4. Within-group analyses did not highlight differences explained by age for any of the parameters investigated. Indeed, adherence to therapy, assessed as the level of motivation expressed by patients and participation assessed by the researcher/therapist, was different for patients according to the level of impairment (motivation: P=.02; r=0.142; participation: P=.045, r=–0.124; Wilcoxon signed-rank test and Spearman coefficient of correlation, respectively).

XSL•FO
RenderX

Figure 4. Session-wise trajectories of acceptability assessment. The longitudinal evolution of normalized scores of acceptability assessment, across the 24 training sessions, separately for patients with moderate and severe impairment. Scores were normalized to the theoretical maximum of each scale (motivation and satisfaction: 10; participation: 6) and are expressed as percentages. Solid lines represent the mean score at each session, while shaded areas indicate the 95% CIs. Missing sessions due to early discontinuation were handled by computing session-wise averages based on available observations only.

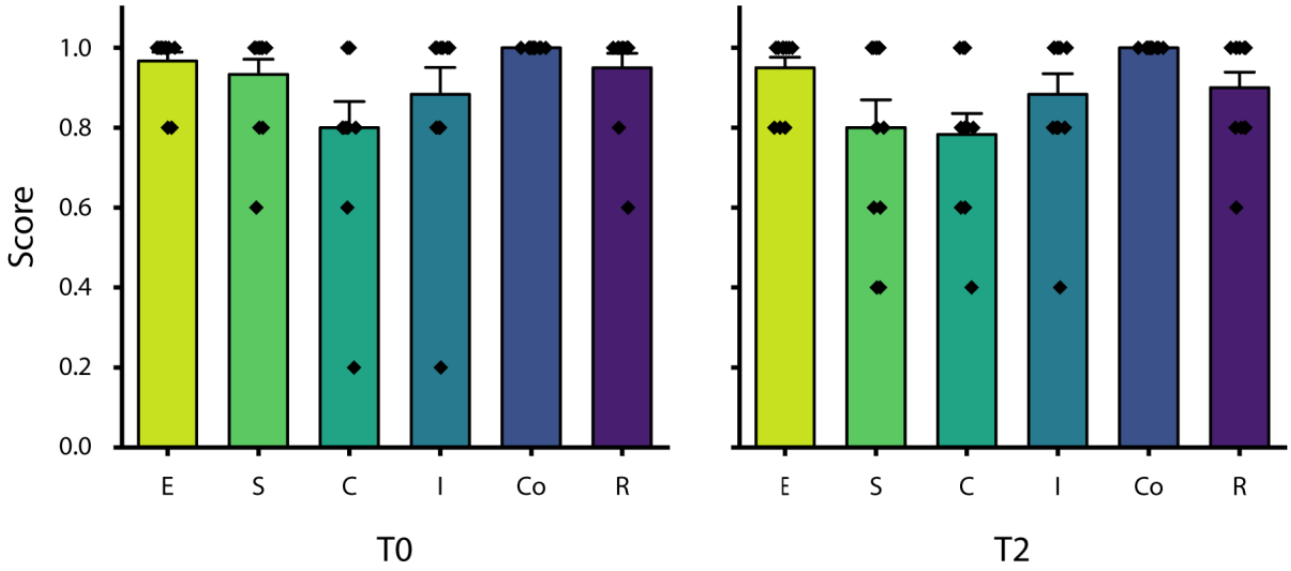


Usability

We asked patients to evaluate the device used for VR motor training according to the item of the USEQ scale reported in Figure 5. The patient performed this evaluation after the completion of the first training session (baseline, T0) and after the last one (end line, T2). All domains assessed with this scale scored at least more than 80% up to 95% in the first session, with small variation in the last one. Therefore, we did not find any difference between the two assessments (P=.56, Wilcoxon signed-rank test).

XSL•FO
RenderX

Figure 5. Usability assessment. The score of the User Satisfaction Evaluation Questionnaire obtained at baseline (T0, left panel) and end line (T2, right panel). Domains of usability were assessed as enjoyment (E), success (S), control (C), information (I), comfort (Co), and helpful for rehabilitation (R). Each bar is the mean score obtained for each domain of the User Satisfaction Evaluation Questionnaire, with error bars computed based on SD. Dots are representative of users' evaluations for each domain.

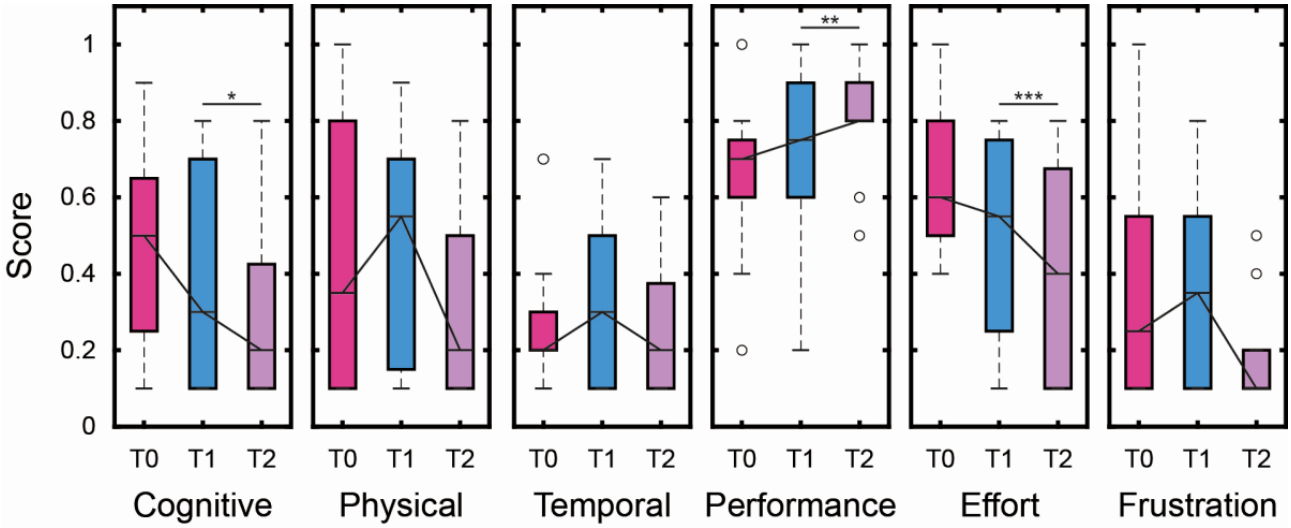


Importantly, the increase in satisfaction over time reflects patients' growing engagement and acceptance of the therapy, whereas the stable usability scores indicate that the system did not require a prolonged learning phase and was perceived as usable from the outset.

Feasibility

Patients' evaluation of cognitive load associated with the VR-MT training was assessed through the NASA-TLX cognitive index as shown in Figure 6, in which differences between the intermediate (T1) and the final evaluation (T2) were found for cognitive (P=.04, Wilcoxon signed-rank test), performance (P=.007, Wilcoxon signed-rank test), and effort domains (P<.001, Wilcoxon signed-rank test). As for acceptability measures, within-group analyses showed that age did not contribute to the variance of NASA domains in the longitudinal evaluation of the treatment.

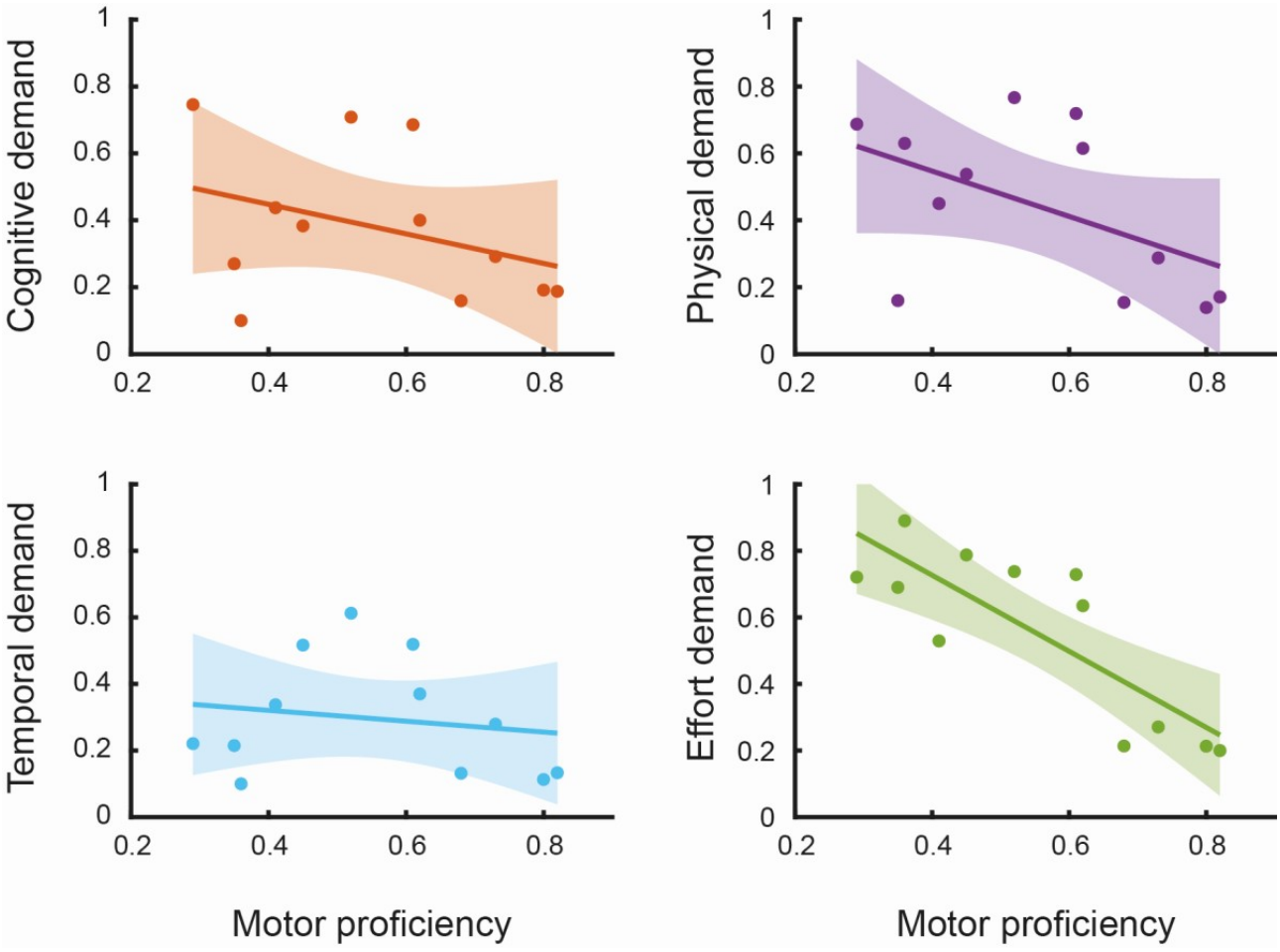
Figure 6. NASA Cognitive Workload Index Assessment. Each panel shows a domain of workload assessed through the rehabilitation protocol, with the three boxplots reporting the distributions at each time of the assessment (T0, T1, and T2). Median values have been linked with a solid line to highlight the trend of each domain across the three evaluations. The whisker edges represent the first and third quartiles of each distribution, while empty dots represent the outlier values. Statistically significant differences between the intermediate (T1) and the final assessment (T2) are indicated by brackets above the corresponding boxplots. Significant improvements were observed in the cognitive, performance, and effort domains (Wilcoxon signed-rank test, *P<.05, **P<.01, ***P<.001).



On the contrary, post hoc analyses with Mann-Whitney U test for independent samples and Spearman correlation coefficient as depicted in Figure 7, clearly show that cognitive workload is perceived differently according to the level of impairment for cognitive (P<.001; r=–0.462), physical (P<.001; r=–0.422), temporal (P<.001; r=–0.408) and effort domains (P<.001; r=–0.317).

XSL•FO
RenderX

Figure 7. Post hoc comparison of NASA Task Load Index scores. The figure shows, in clockwise order from the first panel on the top left, the scatter plot between motor proficiency and the NASA cognitive, physical, temporal, and effort demands. Each point represents the patient's overall assessment of the training according to the considered demand; the solid line is the fitted linear model with the variable motor proficiency as a predictor of the scores for the individual NASA demands. The shaded areas represent the CIs of the model set at 95%.



Notably, differences between impairment subgroups were more evident in measures related to perceived workload and participation, particularly in patients with severe impairment. While these observations cannot support inferential conclusions, they are consistent with clinical expectations and reinforce the relevance of impairment severity as a key factor in feasibility evaluations.

Discussion

The integration of immersive VR-MT in neurorehabilitation has shown promising results for enhancing motor recovery in patients with upper-limb impairments following neurological injuries. However, the effectiveness of training in immersive VR versus conventional therapy is still confined to expert opinion due to heterogeneity of study design, small sample sizes, and differences in the level of impairment of patients enrolled [56,57]. Recent studies suggest that rehabilitation delivered through VR systems may facilitate upper-limb motor and functional recovery and improve health-related quality of life after stroke, while also showing promising results in terms of feasibility and acceptability in populations with chronic stroke [16,18,21,25,38,58]. More precisely, VR protocols designed to embed MT principles have also reported good results in eliciting psychophysical effects such as tingling or paresthesia in the affected limb during the VR experience [58].

The aim of this study was to provide meaningful insights into the feasibility, usability, and acceptability of the adaptive MT using VR and myoelectric control for the rehabilitation of stroke survivors with moderate to severe upper-limb impairment. From a rehabilitation perspective, patients with severe upper-limb impairment represent the most critical test case for the feasibility of VR-based interventions. The exploratory patterns observed in this subgroup, particularly in terms of workload and participation, support the feasibility of the proposed system even under more challenging functional conditions, which is a key requirement for real-world adoption.

Our findings indicated that patients were satisfied with the training they underwent, maintaining a high level of motivation and participation in therapy, although this varied greatly across sessions and with specifics depending on the level of motor impairment. The system's usability assessment, similarly, maintained a very high rating since the first session. Conversely, the workload assessment associated with the task revealed a different perception in terms of cognitive, physical, time demand, and effort for severely impaired patients, who scored higher than moderately impaired patients.

XSL•FO
RenderX

In the context of increasing stroke incidence and improving survival rates, the need for sustainable and accessible rehabilitation options is fundamental [59]. Considering the demographic transition toward an aging population and an increasing prevalence of chronic conditions, including stroke, the inclusion of innovative technological solutions in clinical and home-based rehabilitation pathways is becoming more urgent and frequent. Furthermore, the long-lasting sequelae following stroke demand a sustainable and motivating therapeutic approach that addresses the patient's therapeutic needs and, at the same time, keeps a stable motivation to continue the rehabilitation program.

Although this study was conducted in a supervised clinical setting, the findings provide relevant insights into the potential feasibility of future home-based use. The absence of physical assistance during task execution, together with stable usability and sustained participation, suggests that selected patients, particularly those with moderate impairment, may be considered for future semisupervised or remotely monitored applications. The system was designed to automate task delivery, feedback, and difficulty adaptation, which are key prerequisites for reducing reliance on continuous therapist supervision. However, these results should not be interpreted as evidence of feasibility in a fully autonomous home setting. Indeed, future studies will be needed to support fully autonomous home use and indicate the need for additional safety mechanisms and structured therapist oversight before implementation in unsupervised settings.

Our study aligns with and extends current literature on the need for technological tools integration into poststroke rehabilitation settings [60,61], particularly by quantifying user experience in terms of usability of the VR motor program for upper-limb recovery, as well as of motivation and satisfaction.

Our data showed encouraging outcomes regarding user satisfaction and therapy participation, with significant improvement in both parameters between the intermediate and final evaluations (P=.01 for both). This result underscores the potential of VR systems not only as therapeutic tools but also as motivational enhancers throughout the recovery process. While motivation levels remained high across the study period regardless of time (P=.38), the fact that satisfaction increased significantly indicates that continued exposure and growing familiarity with the system positively influenced the patients' overall rehabilitation experience. These findings support earlier observations [62] that the engaging, interactive nature of VR environments can sustain user interest and potentially enhance neuroplastic changes through intensive, repetitive activity.

Crucially, our data also emphasizes that VR-based rehabilitation therapies must be adaptable to patients' functional capabilities. In fact, in our study, we included patients with different levels of motor proficiency to assess how this aspect could significantly influence adherence to therapy and the associated cognitive workload. The observed effects of impairment level on therapy participation (P=.045; r=–0.124) and motivation (P=.02; r=–0.142) point to a nuanced relationship between physical function and engagement, echoing concerns raised by Laver et al [26] regarding the limited benefits of VR interventions for individuals with more severe motor limitations unless additional assistive elements (eg, robotics or therapist involvement) are introduced. Our results are consistent with those of Chen et al [25] and Kim et al [30], who found that feasibility and therapeutic outcomes are significantly modulated by the severity of disability, and that user-centered design is imperative to address this diversity.

In this study, task difficulty was deliberately personalized through adaptive modulation of assistance and capability parameters, resulting in different absolute task demands across participants.

Importantly, as in similar feasibility research [52,63,64], NASA-TLX scores were interpreted as subjective indicators of perceived workload and feasibility rather than as proxies of motor improvement. While reductions in workload over time may partially reflect a system-related learning curve, they occurred within an adaptively regulated training protocol in which task difficulty was progressively adjusted. Therefore, longitudinal changes in NASA-TLX likely reflect the participants' evolving balance between increasing task demands and improved familiarity and efficiency in interacting with the system, rather than a reduction in task challenge or direct motor recovery.

Consequently, NASA-TLX scores should not be interpreted as a direct reflection of task difficulty per se, but rather as an index of the individual's perceived effort relative to their own functional capacity, cognitive resources, and motor impairment.

From this perspective, variability in NASA-TLX scores across participants is expected and informative, as it captures differences in perceived cognitive and physical demand associated with performing a personalized task. Our results show that impairment severity has a significant impact on perceived workload as assessed through the NASA-TLX scale. Differences in cognitive, physical, temporal, and effort demands were all significantly influenced by impairment level, particularly in post hoc comparisons (eg, cognitive: P<.001; r=–0.462), where higher motor proficiency is associated with a lower level of perceived workload. The observed association between impairment level and perceived workload suggests that, even under adaptive conditions, participants with more severe impairment experienced higher cognitive and physical demands. This finding provides meaningful insight into feasibility, highlighting the need to carefully balance task challenges and user burden when designing adaptive VR-based rehabilitation interventions. While the therapy was generally well-received, its design and implementation must remain sensitive to the cognitive and physical demands placed on users with greater motor impairment. Systems that fail to tailor difficulty and feedback mechanisms to the user's ability level risk inducing cognitive overload, decreased motivation, and ultimately discontinuation of use. The inverse relationship we found between impairment severity and perceived cognitive load, frustration, and effort suggests that the simplified interfaces and personalized training pathways delivered through our approach played a critical role in fostering sustained engagement in this group of patients.

XSL•FO
RenderX

Despite the additional challenge faced by patients with severe impairment, usability scores across all domains remained above 80%, both at the beginning and end of treatment. Patients gave high scores to the system in all domains assessed by the USEQ scale, with most of them maintaining 80%-95% approval from baseline to the end of the intervention. The stability of these results indicates that the system is not only initially user-friendly but remains so throughout the rehabilitation course. These findings are in line with a review of literature [65] addressing the problem of whether patients with varying degrees of motor impairment benefited from technology-assisted interventions, particularly when those interventions were intuitive and adjustable according to patient capabilities.

Although many interventions are designed with a primary focus on functional gains, fewer consider how the patient perceives the treatment, whether they enjoy it, feel comfortable using it, and believe it contributes to their recovery. This focus on user satisfaction and emotional response is increasingly recognized in the literature as essential for optimizing outcomes in digital health interventions [22,66]. As demonstrated in our findings, user satisfaction and motivation increased over the course of treatment, even if no significant changes in the objective interface or content of the application were made. This suggests that familiarity with the tool, confidence gained through usage, and a sense of mastery may contribute to increased acceptability over time. This finding aligns with more recent evidence suggesting that immersive VR applications can effectively facilitate adherence in neurological rehabilitation by providing a sense of agency (ie, perceived control over the virtual limb) and progression in the rehabilitation path [67,68].

We did not find any significant age-related differences in acceptability, usability, or cognitive load perception, suggesting that older adults can also effectively engage with and benefit from VR-based rehabilitation when the system is designed appropriately. These findings contribute to a growing body of work [69,70] that supports the notion that age itself is not a barrier to the acceptance of digital health technologies; rather, usability and personalization are the defining factors for acceptance.

Importantly, our study showed that the level of participation as assessed by the therapist remained high across both intermediate and final evaluations, with only a modest decrease. This suggests that while the initial novelty of the technology may taper over time, the structured and tailored nature of the VR training maintains its overall therapeutic value. Other studies have similarly shown that tailored rehabilitation technologies designed according to the principles of user-centered design yield higher satisfaction and usability scores than standardized or generic programs [16,25,31].

Finally, the broader implications of our study support recent discussions advocating for the inclusion of technology-based rehabilitation in standard care protocols, particularly given the limitations of conventional rehabilitation services in terms of accessibility, cost, and therapist availability post discharge [71,72].

This study has several limitations that should be acknowledged. First, postintervention clinical motor outcomes were not included, which limits the ability to draw conclusions regarding rehabilitation efficacy. This choice was intentional and reflects the role of this study within a broader user-centered design process guiding the development of the intervention. In this framework, the primary objective was to assess feasibility, usability, and acceptability from the user's perspective, as foundational steps in the iterative refinement of a technologically complex therapeutic system, rather than to quantify motor recovery outcomes, which fall within subsequent phases of clinical evaluation.

Second, given the relatively small sample size and the single-arm exploratory design limitation, the findings should be interpreted with appropriate caution. Nevertheless, this study was not intended to provide definitive evidence of effectiveness, but to generate preliminary, hypothesis-generating insights into user experience across different levels of motor impairment. In this context, the inclusion of participants with heterogeneous motor severity represents a strength, as it allowed the identification of differential patterns of participation, motivation, and perceived workload that may inform future personalization strategies. Accordingly, subgroup-related findings should be interpreted as exploratory and hypothesis-generating, and not as evidence of differential effectiveness between impairment levels.

Indeed, therapist assistance time, learning curves, and levels of independence during training were not explicitly quantified. Feasibility was intentionally operationalized within a patient-centered framework focused on sustained engagement, stable task performance under adaptive conditions, and acceptable perceived workload. Accordingly, the absence of age-related differences in feasibility outcomes should be interpreted within these boundaries. Future studies will be needed to integrate clinical motor outcomes and to quantify therapist involvement, autonomy progression, and learning trajectories in order to determine the effectiveness and real-world clinical impact of the intervention.

Finally, task difficulty and virtual assistance were dynamically adjusted to accommodate fatigue or reduced motor control, thereby limiting excessive physical strain. Although no adverse events were observed, safety outcomes were not systematically quantified. Future developments, particularly aimed at unsupervised or home-based use, should integrate automated detection of compensatory or unsafe movements, for example, by monitoring head or trunk position, to enhance safety and support independent training. Although we are still far from thinking that a VR-based therapy can completely replace the therapy performed by a physiotherapist, some important goals have been achieved. Unlike previous studies [26-29], the proposed system was designed to accommodate patients with severe motor impairment through adaptive assistance and task modulation, without requiring additional robotic devices or physical assistance from the therapist during task execution, as reported in another recent study [73]. Our study contributes to offering a more comprehensive view of the user experience by including patient-reported outcomes, therapist evaluations, and cognitive load assessments. This information, together with the evaluation of motor performance carried out with a novel therapeutic approach, will contribute to defining the best

XSL•FO
RenderX

practices for the use of this system in rehabilitation protocols that can be tailored to the specific needs of the patient.

Acknowledgments

We sincerely thank the physiotherapists Francesco La Commare and Nicoletta Polo for their dedication and essential contribution to the completion of this study, and all the health care professionals who supported patient recruitment and management during the challenging COVID-19 period. Their commitment was instrumental to the successful execution of this work.

Funding

This work was partially supported by the German Federal Ministry of Education and Research under the Robotics Institute Germany; by the VVITA project, awarded to Deutsches Zentrum für Luft- und Raumfahrt e.V. (Germany) by Helmoltz Association e.V.; by the Next Generation EU Project National Recovery and Resilience Plan, project MNESYS (PE0000006)—a multiscale integrated approach to the study of the nervous system in health and disease (DN 1553 874 11.10.2022), Spoke 1; and by Current Research Funds 2026, Ministry of Health, Italy.

Data Availability

The data that support the findings of this study are available on reasonable request from the corresponding author, DDB.

Authors' Contributions

Conceptualization: all authors
Methodology: all authors
Software: CN, MN, CC
Validation: DDB, PDP, MR, DJB, CC, GM, AdA
Formal analysis: DDB
Investigation: DDB
Resources: CC, GM, AdA
Data curation: DDB
Writing – original draft: DDB
Writing – review & editing: all authors
Visualization: DDB, PDP
Supervision: GM, AdA
Project administration: PDP, DDB
Funding acquisition: CC, GM, AdA

Conflicts of Interest

None declared.

Multimedia Appendix 1
Virtual assistance.
[PNG File , 696 KB - rehab_v13i1e81894_app1.png]

References

1. Luengo-Fernandez R, Violato M, Candio P, Leal J. Economic burden of stroke across Europe: a population-based cost analysis. Eur Stroke J 2020;5(1):17-25 [FREE Full text] [doi: 10.1177/2396987319883160] [Medline: 32232166]
2. Pucciarelli G, Rebora P, Arisido MW, Ausili D, Simeone S, Vellone E, et al. Direct cost related to stroke: a longitudinal analysis of survivors after discharge from a rehabilitation hospital. J Cardiovasc Nurs 2020;35(1):86-94. [doi: 10.1097/JCN.0000000000000620] [Medline: 31804249]
3. Wafa HA, Wolfe CDA, Emmett E, Roth GA, Johnson CO, Wang Y. Burden of stroke in Europe: thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years. Stroke 2020;51(8):2418-2427 [FREE Full text] [doi: 10.1161/STROKEAHA.120.029606] [Medline: 32646325]
4. Cramer SC, Nelles G, Benson RR, Kaplan JD, Parker RA, Kwong KK, et al. A functional MRI study of subjects recovered from hemiparetic stroke. Stroke 1997;28(12):2518-2527. [doi: 10.1161/01.str.28.12.2518] [Medline: 9412643]
5. Ali AS, Kumaran DS, Unni A, Sardesai S, Prabhu V, Nirmal P, et al. Effectiveness of an intensive, functional, and gamified rehabilitation program on upper limb function in people with stroke (EnteRtain): a multicenter randomized clinical trial. Neurorehabil Neural Repair 2024;38(4):243-256. [doi: 10.1177/15459683231222921] [Medline: 38284559]

XSL•FO
RenderX

6. Nakayama H, Jørgensen HS, Raaschou HO, Olsen TS. Recovery of upper extremity function in stroke patients: the Copenhagen stroke study. Arch Phys Med Rehabil 1994;75(4):394-398. [doi: 10.1016/0003-9993(94)90161-9] [Medline: 8172497]
7. Kwakkel G, Kollen BJ, van der Grond J, Prevo AJ. Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. Stroke 2003;34(9):2181-2186. [doi: 10.1161/01.STR.0000087172.16305.CD] [Medline: 12907818]
8. Chleboun SM, Brady K, Zelenak J. Quality of life following stroke: a qualitative study across 30 years. Perspect ASHA SIGs 2021;6(1):151-162. [doi: 10.1044/2020_persp-20-00206]
9. Hayward KS, Kramer SF, Dalton EJ, Hughes GR, Brodtmann A, Churilov L, et al. Timing and dose of upper limb motor intervention after stroke: a systematic review. Stroke 2021;52(11):3706-3717. [doi: 10.1161/STROKEAHA.121.034348] [Medline: 34601901]
10. Winstein CJ. Goal-oriented action: new perspectives with special emphasis on neurorehabilitation. In: Progress in Motor Control. Cambridge, Massachusetts: Academic Press; 2024:349-387.
11. Vluggen TPMM, van Haastregt JCM, Tan FES, Kempen GIJM, Schols JMGA, Verbunt JA. Factors associated with successful home discharge after inpatient rehabilitation in frail older stroke patients. BMC Geriatr 2020;20(1):25 [FREE Full text] [doi: 10.1186/s12877-020-1422-6] [Medline: 31973729]
12. Pui Kei C, Mohd Nordin NA, Abdul Aziz AF. The effectiveness of home-based therapy on functional outcome, self-efficacy and anxiety among discharged stroke survivors. Medicine (Baltimore) 2020;99(47):e23296 [FREE Full text] [doi: 10.1097/MD.0000000000023296] [Medline: 33217861]
13. Mistiaen P, Francke AL, Poot E. Interventions aimed at reducing problems in adult patients discharged from hospital to home: a systematic meta-review. BMC Health Serv Res 2007;7:47 [FREE Full text] [doi: 10.1186/1472-6963-7-47] [Medline: 17408472]
14. Arntz A, Weber F, Handgraaf M, Lällä K, Korniloff K, Murtonen K, et al. Technologies in home-based digital rehabilitation: scoping review. JMIR Rehabil Assist Technol 2023;10:e43615 [FREE Full text] [doi: 10.2196/43615] [Medline: 37253381]
15. Kenea CD, Abessa TG, Lamba D, Bonnechère B. Immersive virtual reality in stroke rehabilitation: a systematic review and meta-analysis of its efficacy in upper limb recovery. J Clin Med 2025;14(6):1783 [FREE Full text] [doi: 10.3390/jcm14061783] [Medline: 40142592]
16. Warland A, Paraskevopoulos I, Tsekleves E, Ryan J, Nowicky A, Griscti J, et al. The feasibility, acceptability and preliminary efficacy of a low-cost, virtual-reality based, upper-limb stroke rehabilitation device: a mixed methods study. Disabil Rehabil 2019;41(18):2119-2134. [doi: 10.1080/09638288.2018.1459881] [Medline: 29644897]
17. Clough BA, Casey LM. Technological adjuncts to increase adherence to therapy: a review. Clin Psychol Rev 2011;31(5):697-710. [doi: 10.1016/j.cpr.2011.03.006] [Medline: 21497153]
18. Levin MF, Weiss PL, Keshner EA. Emergence of virtual reality as a tool for upper limb rehabilitation: incorporation of motor control and motor learning principles. Phys Ther 2015;95(3):415-425 [FREE Full text] [doi: 10.2522/ptj.20130579] [Medline: 25212522]
19. da Cruz MMA, Ricci-Vitor AL, Borges GLB, da Silva PF, Turri-Silva N, Takahashi C, et al. A randomized, controlled, crossover trial of virtual reality in maintenance cardiovascular rehabilitation in a low-resource setting: impact on adherence, motivation, and engagement. Phys Ther 2021;101(5):pzab071. [doi: 10.1093/ptj/pzab071] [Medline: 33625515]
20. Cramer SC, Sur M, Dobkin BH, O'Brien C, Sanger TD, Trojanowski JQ, et al. Harnessing neuroplasticity for clinical applications. Brain 2011;134(Pt 6):1591-1609 [FREE Full text] [doi: 10.1093/brain/awr039] [Medline: 21482550]
21. Karamians R, Proffitt R, Kline D, Gauthier LV. Effectiveness of virtual reality- and gaming-based interventions for upper extremity rehabilitation poststroke: a meta-analysis. Arch Phys Med Rehabil 2020;101(5):885-896. [doi: 10.1016/j.apmr.2019.10.195] [Medline: 31821799]
22. Leonardsen ACL, Hardeland C, Helgesen AK, Grøndahl VA. Patient experiences with technology enabled care across healthcare settings-a systematic review. BMC Health Serv Res 2020;20(1):779 [FREE Full text] [doi: 10.1186/s12913-020-05633-4] [Medline: 32838784]
23. Parker J, Powell L, Mawson S. Effectiveness of upper limb wearable technology for improving activity and participation in adult stroke survivors: systematic review. J Med Internet Res 2020;22(1):e15981 [FREE Full text] [doi: 10.2196/15981] [Medline: 31913131]
24. Alarcón-Aldana AC, Callejas-Cuervo M, Bo APL. Upper limb physical rehabilitation using serious videogames and motion capture systems: a systematic review. Sensors (Basel) 2020;20(21):5989 [FREE Full text] [doi: 10.3390/s20215989] [Medline: 33105845]
25. Chen J, Or CK, Chen T. Effectiveness of using virtual reality-supported exercise therapy for upper extremity motor rehabilitation in patients with stroke: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022;24(6):e24111 [FREE Full text] [doi: 10.2196/24111] [Medline: 35723907]
26. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. Cochrane Database Syst Rev 2017;11(11):CD008349 [FREE Full text] [doi: 10.1002/14651858.CD008349.pub4] [Medline: 29156493]
27. Byl NN, Abrams GM, Pitsch E, Fedulow I, Kim H, Simkins M, et al. Chronic stroke survivors achieve comparable outcomes following virtual task specific repetitive training guided by a wearable robotic orthosis (UL-EXO7) and actual task specific

XSL•FO
RenderX

repetitive training guided by a physical therapist. J Hand Ther 2013;26(4):343-52; quiz 352. [doi: 10.1016/j.jht.2013.06.001] [Medline: 23911077]
28. Klamroth-Marganska V, Blanco JB, Campen K, Curt A, Dietz V, Ettlin T, et al. Three-dimensional, task-specific robot therapy of the arm after stroke: a multicentre, parallel-group randomised trial. Lancet Neurol 2014;13(2):159-166 [FREE Full text] [doi: 10.1016/S1474-4422(13)70305-3] [Medline: 24382580]
29. Abd El-Kafy EMA, Alshehri MA, El-Fiky AARE, Guermazi MA, Mahmoud HM. The effect of robot-mediated virtual reality gaming on upper limb spasticity poststroke: a randomized-controlled trial. Games Health J 2022;11(2):93-103. [doi: 10.1089/g4h.2021.0197] [Medline: 35100025]
30. Kim WS, Cho S, Ku J, Kim Y, Lee K, Hwang HJ, et al. Clinical application of virtual reality for upper limb motor rehabilitation in stroke: review of technologies and clinical evidence. J Clin Med 2020;9(10):3369 [FREE Full text] [doi: 10.3390/jcm9103369] [Medline: 33096678]
31. De Pasquale P, De Bartolo D, Russo M, Berger DJ, Maselli A, Borzelli D, et al. User-centered development of a personalized adaptive mirror therapy for upper-limb post-stroke rehabilitation using virtual reality and myoelectric control. Front Bioeng Biotechnol 2025;13. [doi: 10.3389/fbioe.2025.1655416]
32. Greenbaum J, Kyng M. Design at Work: Cooperative Design of Computer Systems. Boca Raton, FL: CRC Press; 2020.
33. Nissler C, Nowak M, Connan M, Büttner S, Vogel J, Kossyk I, et al. VITA-an everyday virtual reality setup for prosthetics and upper-limb rehabilitation. J Neural Eng 2019;16(2):026039. [doi: 10.1088/1741-2552/aaf35f] [Medline: 30864550]
34. Gandhi DB, Sterba A, Khatter H, Pandian JD. Mirror therapy in stroke rehabilitation: current perspectives. Ther Clin Risk Manag 2020;16:75-85 [FREE Full text] [doi: 10.2147/TCRM.S206883] [Medline: 32103968]
35. Hali K, Manzo MA, Koucheki R, Wunder JS, Jenkinson RJ, Mayo AL, et al. Use of virtual reality for the management of phantom limb pain: a systematic review. Disabil Rehabil 2024;46(4):629-636. [doi: 10.1080/09638288.2023.2172222] [Medline: 36724203]
36. Song R, Tong KY, Hu X, Zhou W. Myoelectrically controlled wrist robot for stroke rehabilitation. J Neuroeng Rehabil 2013;10:52 [FREE Full text] [doi: 10.1186/1743-0003-10-52] [Medline: 23758925]
37. Sarasola-Sanz A, Irastorza-Landa N, López-Larraz E, Shiman F, Spüler M, Birbaumer N, et al. Design and effectiveness evaluation of mirror myoelectric interfaces: a novel method to restore movement in hemiplegic patients. Sci Rep 2018;8(1):16688 [FREE Full text] [doi: 10.1038/s41598-018-34785-x] [Medline: 30420779]
38. Hsu HY, Kuo LC, Lin YC, Su F, Yang T, Lin C. Effects of a virtual reality-based mirror therapy program on improving sensorimotor function of hands in chronic stroke patients: a randomized controlled trial. Neurorehabil Neural Repair 2022;36(6):335-345. [doi: 10.1177/15459683221081430] [Medline: 35341360]
39. De Bartolo D, Morone G, Lupo A, Aloise F, Baricich A, Di Francesco D, et al. From paper to informatics: the Post Soft Care-App, an easy-to-use and fast tool to help therapists identify unmet needs in stroke patients. Funct Neurol 2018;33(4):200-205. [Medline: 30663966]
40. Paternostro-Sluga T, Grim-Stieger M, Posch M, Schuhfried O, Vacariu G, Mittermaier C, et al. Reliability and validity of the Medical Research Council (MRC) scale and a modified scale for testing muscle strength in patients with radial palsy. J Rehabil Med 2008;40(8):665-671 [FREE Full text] [doi: 10.2340/16501977-0235] [Medline: 19020701]
41. Cecchi F, Carrabba C, Bertolucci F, Castagnoli C, Falsini C, Gnetti B, et al. Transcultural translation and validation of Fugl-Meyer Assessment to Italian. Disabil Rehabil 2021;43(25):3717-3722. [doi: 10.1080/09638288.2020.1746844] [Medline: 32356509]
42. Pigliautile M, Ricci M, Mioshi E, Ercolani S, Mangialasche F, Monastero R, et al. Validation study of the Italian Addenbrooke's Cognitive Examination Revised in a young-old and old-old population. Dement Geriatr Cogn Disord 2011;32(5):301-307. [doi: 10.1159/000334657] [Medline: 22262124]
43. Dong Y, Liu X, Tang M, Huo H, Chen D, Du X, et al. Age-related differences in upper limb motor performance and intrinsic motivation during a virtual reality task. BMC Geriatr 2023;23(1):251 [FREE Full text] [doi: 10.1186/s12877-023-03970-7] [Medline: 37106330]
44. Pocock SJ. Clinical Trials: A Practical Approach. New Jersey: Wiley; 2013.
45. Kaps M, Grittner U, Jungehülsing G, Tatlisumak T, Kessler C, Schmidt R, SIFAP1 Investigators. Clinical signs in young patients with stroke related to FAST: results of the SIFAP1 study. BMJ Open 2014;4(11):e005276 [FREE Full text] [doi: 10.1136/bmjopen-2014-005276] [Medline: 25380809]
46. Pinter D, Enzinger C, Gattringer T, Eppinger S, Niederkorn K, Horner S, et al. Prevalence and short-term changes of cognitive dysfunction in young ischaemic stroke patients. Eur J Neurol 2019;26(5):727-732 [FREE Full text] [doi: 10.1111/ene.13879] [Medline: 30489673]
47. Patel GK, Nowak M, Castellini C. Exploiting knowledge composition to improve real-life hand prosthetic control. IEEE Trans Neural Syst Rehabil Eng 2017;25(7):967-975 [FREE Full text] [doi: 10.1109/TNSRE.2017.2676467] [Medline: 28278474]
48. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]

XSL•FO
RenderX

49. Iosa M, Galeoto G, De Bartolo D, Russo V, Ruotolo I, Spitoni G, et al. Italian version of the Pittsburgh rehabilitation participation scale: psychometric analysis of validity and reliability. Brain Sci 2021;11(5):626 [FREE Full text] [doi: 10.3390/brainsci11050626] [Medline: 34068212]
50. Bevan N. Usability. In: Liu L, Özsu MT, editors. Encyclopedia of Database System. Cham: Springer; 2009:1-5.
51. Gil-Gómez JA, Manzano-Hernández P, Albiol-Pérez S, Aula-Valero C, Gil-Gómez H, Lozano-Quilis J. USEQ: a short questionnaire for satisfaction evaluation of virtual rehabilitation systems. Sensors (Basel) 2017;17(7):1589 [FREE Full text] [doi: 10.3390/s17071589] [Medline: 28686174]
52. De Bartolo D, Baleca L, De Angelis D, Nocentini U, Iosa M. The effects of secondary motor and cognitive tasks on gait depend on functional walking ability in non-traumatic neurological patients: a feasibility pilot study. Appl Sci 2026;16(3):1484 [FREE Full text] [doi: 10.3390/app16031484]
53. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Advances in Psychology. Amsterdam, Netherlands: North-Holland Press; 1988:139-183.
54. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. San Francisco: Morgan Kaufmann; 2016:24-30.
55. Squeri V, Masia L, Giannoni P, Sandini G, Morasso P. Wrist rehabilitation in chronic stroke patients by means of adaptive, progressive robot-aided therapy. IEEE Trans Neural Syst Rehabil Eng 2014;22(2):312-325. [doi: 10.1109/TNSRE.2013.2250521] [Medline: 23508271]
56. Henderson A, Korner-Bitensky N, Levin M. Virtual reality in stroke rehabilitation: a systematic review of its effectiveness for upper limb motor recovery. Top Stroke Rehabil 2007;14(2):52-61. [doi: 10.1310/tsr1402-52] [Medline: 17517575]
57. Donati D, Pinotti E, Mantovani M, Casarotti S, Fini A, Tedeschi R, et al. The role of immersive virtual reality in upper limb rehabilitation for subacute stroke: a review. J Clin Med 2025;14(6):1903 [FREE Full text] [doi: 10.3390/jcm14061903] [Medline: 40142711]
58. Weber LM, Nilsen DM, Gillen G, Yoon J, Stein J. Immersive virtual reality mirror therapy for upper limb recovery after stroke: a pilot study. Am J Phys Med Rehabil 2019;98(9):783-788 [FREE Full text] [doi: 10.1097/PHM.0000000000001190] [Medline: 30964752]
59. Heinrich C, Morkisch N, Langlotz T, Regenbrecht H, Dohle C. Feasibility and psychophysical effects of immersive virtual reality-based mirror therapy. J Neuroeng Rehabil 2022;19(1):107 [FREE Full text] [doi: 10.1186/s12984-022-01086-4] [Medline: 36207720]
60. Feigin VL, Owolabi MO, World Stroke Organization–Lancet Neurology Commission Stroke Collaboration Group. Pragmatic solutions to reduce the global burden of stroke: a World Stroke Organization-lancet neurology commission. Lancet Neurol 2023;22(12):1160-1206. [doi: 10.1016/S1474-4422(23)00277-6] [Medline: 37827183]
61. Mitchell J, Shirota C, Clanchy K. Factors that influence the adoption of rehabilitation technologies: a multi-disciplinary qualitative exploration. J Neuroeng Rehabil 2023;20(1):80 [FREE Full text] [doi: 10.1186/s12984-023-01194-9] [Medline: 37340496]
62. Hadjipanayi C, Banakou D, Michael-Grigoriou D. Virtual reality exergames for enhancing engagement in stroke rehabilitation: a narrative review. Heliyon 2024;10(18):e37581 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e37581] [Medline: 39318803]
63. Ramirez-Zamora JD, Dominguez-Ramirez OA, Sepulveda-Cervantes G, Ramos-Velasco LE, Fernandez-Ramirez JM. NASA-TLX assessment for a haptic adaptive platform for upper extremity motor rehabilitation. 2022 Presented at: 2022 XXIV Robotics Mexican Congress (COMRob); 2022 November 9-11; Mineral de la Reforma, Hidalgo, Mexico p. 36-41. [doi: 10.1109/comrob57154.2022.9962313]
64. Jäggi S, Wachter A, Adcock-Omlin M. Feasibility and effects of cognitive-motor exergames on physical and cognitive functioning in people with Parkinson disease: a pilot randomized controlled study. J Neuroeng Rehabil 2023;20(1):52. [doi: 10.1186/s12984-023-01234-z]
65. Colamarino E, Morone G, Toppi J, Riccio A, Cincotti F, Mattia D, et al. A scoping review of technology-based approaches for upper limb motor rehabilitation after stroke: are we really targeting severe impairment? J Clin Med 2024;13(18):5414 [FREE Full text] [doi: 10.3390/jcm13185414] [Medline: 39336901]
66. Alshammari HM, Almutairi RI, Ghazwani AY, Aditya RS. Evaluating user satisfaction and engagement in mHealth: insights from the Integrated Digital Health Engagement Model (IDHEM). Digit Health 2025;11:20552076251346698 [FREE Full text] [doi: 10.1177/20552076251346698] [Medline: 40718399]
67. Perez-Marcos D, Bieler-Aeschlimann M, Serino A. Virtual reality as a vehicle to empower motor-cognitive neurorehabilitation. Front Psychol 2018;9:2120 [FREE Full text] [doi: 10.3389/fpsyg.2018.02120] [Medline: 30450069]
68. Bermúdez IBS, Fluet GG, Llorens R, Deutsch JE. Virtual reality for sensorimotor rehabilitation post stroke: design principles and evidence. In: Neurorehabilitation Technology. Cham: Springer International Publishing; 2016:573-603.
69. Huvila I, Cajander Å, Moll J, Enwald H, Eriksson-Backa K, Rexhepi H. Technological and informational frames: explaining age-related variation in the use of patient accessible electronic health records as technology and information. ITP 2021;38(5):1-22. [doi: 10.1108/itp-08-2020-0566]
70. Mair JL, Hashim J, Thai L, Tai ES, Ryan JC, Kowatsch T, et al. Understanding and overcoming barriers to digital health adoption: a patient and public involvement study. Transl Behav Med 2025;15(1):ibaf010. [doi: 10.1093/tbm/ibaf010] [Medline: 40167046]

XSL•FO
RenderX

71. Wang RH, Kenyon LK, McGilton KS, Miller WC, Hovanec N, Boger J, et al. The time is now: a FASTER approach to generate research evidence for technology-based interventions in the field of disability and rehabilitation. Arch Phys Med Rehabil 2021;102(9):1848-1859. [doi: 10.1016/j.apmr.2021.04.009] [Medline: 33992634]
72. Rampioni M, Leonzi S, Antognoli L, Mura A, Stara V. Poststroke eHealth technologies-based rehabilitation for upper limb recovery: systematic review. J Med Internet Res 2025;27:e57957 [FREE Full text] [doi: 10.2196/57957] [Medline: 40053744]
73. Munoz-Novoa M, Kristoffersen MB, Sunnerhagen KS, Naber A, Ortiz-Catalan M, Alt Murphy M. Myoelectric pattern recognition with virtual reality and serious gaming improves upper limb function in chronic stroke: a single case experimental design study. J Neuroeng Rehabil 2025;22(1):6 [FREE Full text] [doi: 10.1186/s12984-025-01541-y] [Medline: 39825410]

Abbreviations

FMA: Fugl-Meyer Assessment
MRC: Medical Research Council
MT: mirror therapy
NASA-TLX: NASA (National Aeronautics and Space Administration) Task Load Index
sEMG: surface electromyography
USEQ: User Satisfaction Evaluation Questionnaire
VR: virtual reality

Edited by S Munce; submitted 05.Aug.2025; peer-reviewed by MB Kristoffersen, H Wang, C Perciballi; comments to author 19.Dec.2025; revised version received 13.Mar.2026; accepted 23.Mar.2026; published 04.May.2026.

Please cite as:
De Bartolo D, De Pasquale P, Russo M, Berger DJ, Maselli A, Borzelli D, Nissler C, Nowak M, Castellini C, Morone G, d'Avella A Usability, Acceptability, and Feasibility of a Personalized Adaptive Mirror Therapy for Upper-Limb Poststroke Rehabilitation Using Immersive Virtual Reality and Myoelectric Control: Single-Arm Pre-Post Study
JMIR Rehabil Assist Technol 2026;13:e81894
URL: https://rehab.jmir.org/2026/1/e81894
doi:10.2196/81894
PMID:42082173

©Daniela De Bartolo, Paolo De Pasquale, Marta Russo, Denise Jennifer Berger, Antonella Maselli, Daniele Borzelli, Christian Nissler, Markus Nowak, Claudio Castellini, Giovanni Morone, Andrea d'Avella. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 04.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding Education and Training Needs in Rehabilitation Technologies Among Those With Lived Experience, Rehabilitation Health and Care Professionals, and Technology Professionals: Qualitative Study

Angela Thornton[1,2], PhD; Matthew Horrocks[1], PhD; Victoria Abbott-Fleming[3], LLB; Holly Blake[1,2,4], PhD

[1]School of Health Sciences, University of Nottingham, Nottingham, England, United Kingdom

[2]Arthritis UK Pain Centre, University of Nottingham, Nottingham, England, United Kingdom

[3]British Pain Society, London, England, United Kingdom

[4]NIHR Biomedical Research Centre: Nottingham, Nottingham University Hospitals NHS Trust, Nottingham, England, United Kingdom

Corresponding Author:
Holly Blake, PhD
School of Health Sciences
University of Nottingham
Queen's Medical Centre
Derby Road
Nottingham, England, NG7 2HA
United Kingdom
Phone: 44 115 9515151
Email: holly.blake@nottingham.ac.uk

Abstract

Background: Demand for rehabilitation services and associated technologies is increasing globally due to population aging and rising chronic disease burden. Despite their potential to enhance access, efficiency, and outcomes, adoption of rehabilitation technologies in clinical practice is variable. Barriers include limited workforce capability and confidence, resource constraints, fragmented governance, and challenges integrating technologies into existing systems. Collaboration among health and care professionals, educators, technology professionals, and individuals with lived experience is essential to identify training needs and strengthen capability for technology-enabled rehabilitation.

Objective: This study aimed to develop an in-depth understanding of education and training needs for rehabilitation technologies across 3 stakeholder groups—rehabilitation professionals, technology professionals, and individuals with lived experience to inform future training initiatives.

Methods: A qualitative study was conducted, primarily across England, using purposive sampling to recruit stakeholders from each group. Data were collected through online focus groups and individual interviews and analyzed using reflexive thematic analysis. Although thematic analysis is inherently interpretive, the initial analysis focused primarily on the descriptive content of participants' accounts. Subsequently, deeper theoretical interpretation was undertaken through application of the capability, opportunity, motivation–behavior (COM-B) model, selected domains from the Theoretical Domains Framework (TDF), and the Consolidated Framework for Implementation Research (CFIR). COM-B was used to identify behavioral determinants, with selected domains from the TDF used to clarify behavioral influences. CFIR was then applied to situate behavioral determinants within organizational and system-level contexts.

Results: A total of 50 participants contributed to 7 focus groups and 11 individual interviews. Following this, 3 overarching themes were identified: (1) perceived capability and confidence in technology engagement, (2) system-level resource constraints, and (3) fragmented governance structures and organizational decision-making. COM-B mapping indicated that capability-related influences were most prominent among individuals with lived experience, whereas opportunity-related and structural constraints were more influential for rehabilitation and technology professionals. CFIR analysis highlighted that these behavioral determinants were embedded within, and often constrained by, infrastructure, procurement processes, and wider National Health Service (NHS) policy.

XSL•FO
RenderX

Conclusions: Individuals with lived experience require accessible, ongoing support to build confidence and skills in using rehabilitation technologies. Rehabilitation professionals and technology professionals emphasized the need for structured education on available technologies and practical training on integrating them into clinical workflows, particularly for telerehabilitation and emerging digital and AI-enabled tools. Findings indicate that training initiatives must extend beyond individual skill development to address organizational resources, infrastructure, and governance processes. Coordinated, cross-stakeholder education and training strategies are essential to support equitable and sustainable adoption of rehabilitation technologies within the NHS. Training programs should be co-designed across stakeholder groups, embed behavioral and implementation considerations, and align with NHS infrastructure and procurement pathways. Addressing both capability and system-level constraints will be critical for achieving widespread, confident, and sustained use of rehabilitation technologies.

(JMIR Rehabil Assist Technol 2026;13:e97602) doi:10.2196/97602

KEYWORDS

rehabilitation technology; assistive technology; telerehabilitation; education and training; implementation; qualitative research; thematic analysis; stakeholders

Introduction

Background

Rehabilitation technologies encompass a wide range of tools from mobility aids (eg, wheelchairs) to robotic exoskeletons to virtual reality applications and are used to support rehabilitation across hospital, community, and home settings. For individuals with disabilities, such technologies have the potential to enhance functional recovery, promote independence, and extend access to rehabilitation services [1,2]. Current applications such as telerehabilitation can support underresourced health and care professionals within the National Health Service (NHS) by enabling more flexible models of rehabilitation delivery [3,4]. In the United Kingdom, NHS policy has increasingly promoted digitally enabled models of care and workforce development to support the integration of technologies across health services, including rehabilitation [5-7]. However, translation of these ambitions into routine practice has been inconsistent across organizations and regions, and adoption into clinical practice remains generally low [8]. Barriers include limited education and training for health and care staff, and low professional confidence, situated within a complex, challenging and fragmented environment [8-11]. Education and training represent potentially modifiable factors that may influence how different stakeholders engage with rehabilitation technologies in practice.

Prior work used the Implementation Stakeholder Engagement Model (I-STEM) to co-design an inclusive stakeholder engagement strategy for rehabilitation technology training [12]. Through engagement with rehabilitation professionals, educators, technology innovators, and individuals with lived experience, the study identified key implementation priorities, including accessibility, instructional design, technological integration, stakeholder engagement, and staff competence and training. Stakeholders consistently highlighted gaps in awareness, confidence, and training, identifying education and training as a critical and potentially modifiable determinant of rehabilitation technology adoption and use. While the I-STEM study established priorities for stakeholder engagement and training development, it did not explore in depth how education and training needs differed across stakeholder groups. The present qualitative study was therefore undertaken to examine these perspectives in greater depth and identify priorities for future rehabilitation technology education, training, and implementation initiatives.

To our knowledge, this study is the first to conduct a detailed comparative analysis of the training needs for rehabilitation technologies across individuals with lived experience, rehabilitation professionals, and technology professionals. While grounded in inductive thematic analysis, by integrating behavioral and implementation frameworks, the study also clarifies how individual training needs are shaped by organizational context. The findings illuminate challenges associated with achieving widespread adoption of digital technologies in the NHS following the acceleration of interest and uptake during the COVID-19 pandemic [13], developments which underpin the NHS 10 Year Health Plan (2025) Fit for the Future [14].

Research Objectives

This study aimed to explore in-depth the experiences, perceptions, and needs regarding the use of technology in rehabilitation, with a focus on identifying gaps in education and training, preferred learning formats, and perceived barriers to digital adoption.

Methods

Study Design

This study was guided by a cocreated stakeholder engagement strategy, structured using the I-STEM [12]. The authors recommended a subsequent stage of qualitative research. Hence, this qualitative study was designed to generate richer, deeper insights into the education and training needs in rehabilitation technology among a diverse group of stakeholders. It comprised homogeneous focus groups of between 3 and 9 individuals. Individual interviews were also conducted to accommodate participants' schedules, maximize participation, and increase accessibility. The groups and interviews were moderated using a semistructured topic guide covering key concepts while also allowing the researcher flexibility and autonomy to explore other relevant topics [15]. The guides for each stakeholder were developed by the project team (HB, AT, MH, and VAF) and evolved iteratively during fieldwork (Multimedia Appendices 1-3). After each group or interview, participants were given the opportunity to complete a short semistructured questionnaire

XSL•FO
RenderX

that sought their views on the research experience. Reporting aligns with the consolidated criteria for reporting qualitative research (COREQ) [16] (Multimedia Appendix 4).

Participants

This study was conducted primarily across England. Three key stakeholder groups were represented:

1. Rehabilitation professionals, defined as health or care professionals working in rehabilitation settings (n=19).
2. Technology professionals, individuals who innovate, develop, or supply rehabilitation technology (n=9).
3. Lived experience, individuals with different levels and types of injury, illness, or disability with lived experience of rehabilitation (n=22). We use the term lived experience throughout as it reflects the majority (n=17), while also representing 2 participants with both lived and care experience and 3 caregivers.

Participants were purposively sampled through a variety of routes to ensure inclusion of stakeholders with direct experience of rehabilitation technologies and representing a wide range of lived experiences, clinical expertise, and technology types. Recruitment occurred through the National Institute for Health and Care Research (NIHR) HealthTech Research Centre for Rehabilitation's website and targeted emails to the researchers' academic and clinical networks as well as posts on professional platforms such as LinkedIn. Participants were also invited to share study information within their networks (snowball sampling).

Study information was disseminated via the project website, which hosted downloadable participant information sheets, consent forms, and demographic questionnaires.

Data Collection

Fieldwork was conducted between December 11, 2025, and February 2, 2026. Interviews and group discussions were conducted online by project researchers (AT and MH) using Microsoft Teams. The researchers who conducted the interviews and focus groups (AT and MH) had no prior professional or personal relationships with participants before study commencement. Their only contact with participants occurred during recruitment and data collection. With participants' consent, sessions were audio and video-recorded and automatically transcribed using Microsoft Teams; transcripts were subsequently checked for accuracy by one of the researchers (AT). Moreover, 2 individual interviews (1 lived experience and 1 technology professional) were not audio-recorded due to technical issues. Detailed notes were taken during these interviews and expanded immediately afterwards. These accounts were included in the analysis alongside the verified transcripts and field notes recorded during data collection. Mean duration of groups and interviews varied by stakeholder type (lived experience: mean 86, SD 13.53 minutes; rehabilitation professionals: mean 52, SD 21.86 minutes; technology professionals: mean 61, SD 26.38 minutes), with group discussions generally lasting longer than individual interviews, reflecting the additional time for debate and discussion among group participants.

Data Analysis

A total of 18 datasets were collected, comprising 7 group discussions (3 lived experience, and 2 each for rehabilitation and technology professionals) and 11 individual interviews (3 with lived experience, 6 with rehabilitation professionals, and 2 with technology professionals).

Inductive data analysis was conducted following the principles of reflexive thematic analysis as described by Braun and Clarke [17]. Transcripts from focus groups and individual interviews were read by the researcher (AT), who conducted most of the focus groups and interviews to maintain close familiarity with the data. Initial codes were generated through AT's iterative engagement with the dataset, capturing features of the data relevant to the research questions. Transcripts from focus groups and individual interviews were analyzed together within a single analytic process, enabling comparison of perspectives across stakeholder groups while retaining attention to differences between participant types. This analysis informed the development of preliminary themes reflecting patterns of shared meaning both within and across participant groups. Theme development was strengthened through comparison across stakeholder groups (individuals with lived experience, rehabilitation professionals, and technology professionals). This process was led by AT, but the themes were reviewed and refined by MH and further informed through ongoing discussion with other members of the research team (HB and VAF), who contributed to the refinement and interpretation of themes through critical reflection and who contributed additional disciplinary perspectives. Themes were then further refined by the research team and descriptively named and checked to ensure that each theme was internally coherent, conceptually distinct, and organized around a unifying concept [18]. A 15-point checklist by Braun and Clarke [19] was used to support the quality and rigor of the thematic analysis (Multimedia Appendix 5). Analysis was conducted manually in Microsoft Word. This approach supported close engagement with the data during coding and theme development and is consistent with reflexive thematic analysis, in which software acts primarily as a data management tool [20,21].

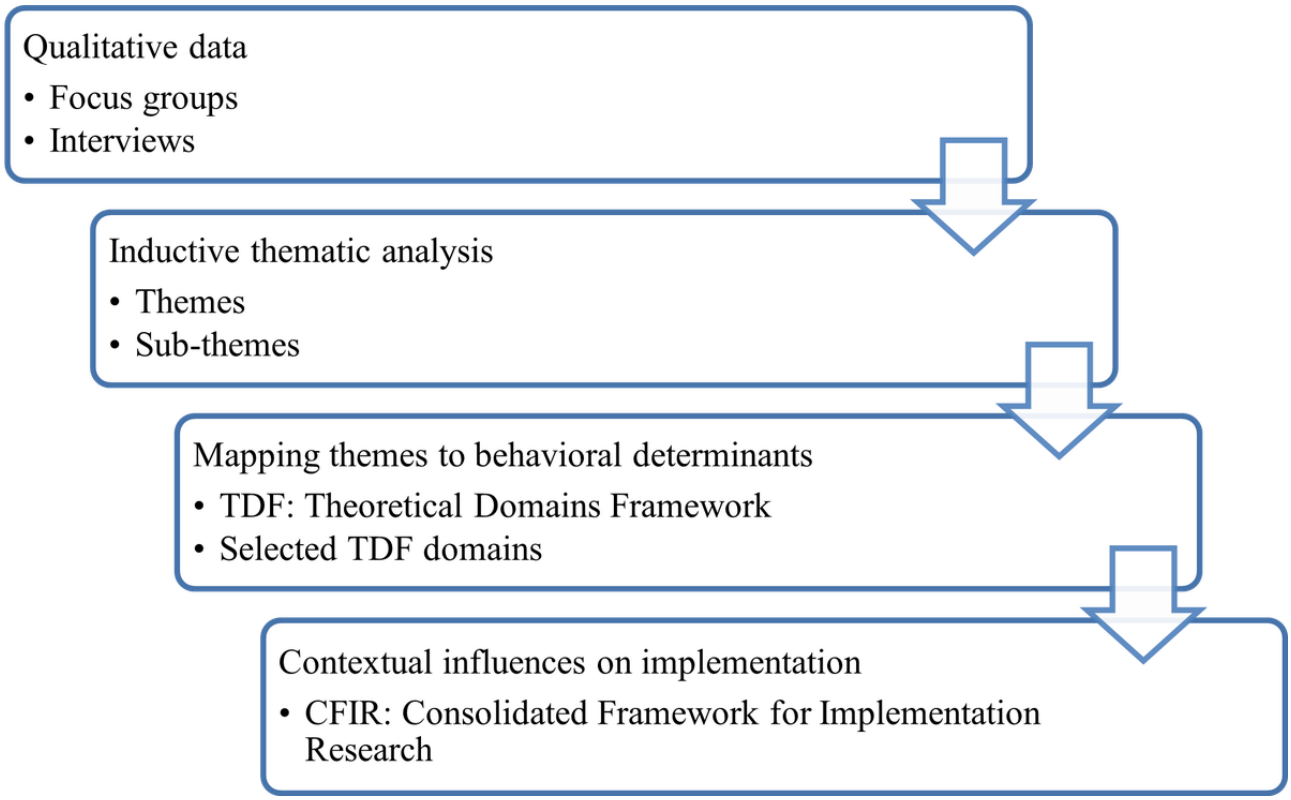
Although reflexive thematic analysis is inherently interpretive, the initial analysis focused primarily on participants' explicit accounts and patterns of shared meaning. Following inductive thematic analysis, themes were subsequently interpreted using theoretical frameworks. First, themes were mapped onto the capability, opportunity, motivation–behavior (COM-B) model [22] to identify behavioral determinants influencing the use of rehabilitation technology, as COM-B provides a structured approach to understanding behavior. Where further clarification was required, relevant domains from the Theoretical Domains Framework (TDF) [23] were used to identify underlying behavioral constructs, with only domains relevant to the emerging themes applied. Finally, findings were interpreted using the Consolidated Framework for Implementation Research (CFIR) [24], enabling behavioral determinants to be considered alongside organizational, technological, and system-level influences (refer to Figure 1 for analysis process). The 3 frameworks were selected because they provide complementary levels of analysis. COM-B identifies the broad behavioral

XSL•FO
RenderX

determinants underpinning stakeholders' education and training needs, TDF provides greater detail on these behavioral determinants, and CFIR contextualizes these. Theoretical frameworks were applied during interpretation rather than initial analysis to preserve the inductive nature of the thematic analysis.

Due to project timelines and the interpretive nature of reflexive thematic analysis, participants were not asked to review the thematic findings. Instead, analytic rigor was supported through comparison of perspectives across stakeholder groups and ongoing reflexive discussions within the research team.

Figure 1. Qualitative analysis process illustrating the progression from inductive thematic analysis to identification of behavioral determinants (capability, opportunity, motivation–behavior with selected TDF domains) and interpretation within implementation context using CFIR. CFIR: Consolidated Framework for Implementation Research; TDF: Theoretical Domains Framework.



Reflexivity

Reflexive thematic analysis recognizes that themes are actively developed through the researchers' interpretive engagement with the data rather than emerging independently of the analytic process [17]. The research team brought complementary perspectives including behavioral science (HB), psychology and psychotherapy practice (MH), industry-based health technology research (AT), and lived experience (VAF). These perspectives informed interpretation during coding and theme development through reflexive discussions in which assumptions, interpretations, and candidate themes were critically examined and refined.

Ethical Considerations

The University of Nottingham Faculty of Medicine and Health Sciences Research Ethics Committee granted ethics approval on October 30, 2025 (FMHS 16-1025). Minor amendments were noted to file on November 25, 2025. Interested individuals emailed completed documentation to the research team. Written informed consent and demographic data were obtained prior to participation, with verbal consent reconfirmed and audio-recorded at the start of each group or interview.

Results

Participant Characteristics

A total of 50 participants took part in the study—19 rehabilitation professionals, 9 technology professionals, and 22 lived experience participants. We use the term lived experience to represent this stakeholder group, as it reflects the majority (n=17), while also encompassing 2 participants with both lived and care experience and 3 caregivers. Participants were recruited from across England, with individual representation from Northern Ireland in the lived experience group. Demographic characteristics are summarized in Table 1. Lived experience participants demonstrated diversity in age, gender identity, ethnicity, and disability status. The majority identified as having a disability or long-term health condition. Rehabilitation and technology professionals were predominantly female and White and largely did not identify as having a disability. Participant demographics are presented by stakeholder group in Tables 1-3.

XSL•FO
RenderX

Table 1. Demographic characteristics of lived experience participants (n=22).

Characteristic[a]	Value[b]
Age (y), mean (SD)[c]	37.3 (11.6)
Gender identity, n (%)	
Woman	11 (50)
Man	7 (32)
Nonbinary	4 (18)
Ethnicity, n (%)	
White	8 (42)
Black	5 (26)
Mixed	5 (26)
Asian	1 (5)
Missing or nonresponse	3 (14)
Identifies as having disability or long-term condition, n (%)	17 (85)

[a]Geographic spread: England and Northern Ireland.

[b]Percentages are calculated using the number of respondents to each individual question. Denominators therefore vary across demographic characteristics due to nonresponse. Percentages for missing or nonresponse categories are calculated using the total sample size.

[c]Mean ages and SDs were estimated using the midpoints of the reported age bands, as exact ages were not collected.

Table 2. Demographic characteristics of rehabilitation professionals (n=19).

Characteristic[a]	Value[b]
Age (y), mean (SD)[c]	42.1 (11.9)
Gender identity, n (%)	
Woman	14 (74)
Man	5 (26)
Nonbinary	0 (0)
Ethnicity, n (%)	
White	16 (84)
Black	1 (5)
Mixed	2 (11)
Asian	0 (0)
Identifies as having disability or long-term condition, n (%)	2 (11)

[a]Geographic spread: England (South, Midlands, North).

[b]Percentages are calculated using the number of respondents to each individual question. Denominators therefore vary across demographic characteristics due to non-response. Percentages for missing or nonresponse categories are calculated using the total sample size.

[c]Mean ages and SDs were estimated using the midpoints of the reported age bands, as exact ages were not collected.

XSL•FO
RenderX

Table 3. Demographic characteristics of technology professionals (n=9).

Characteristic[a]	Value[b]
Age (y), mean (SD)[c]	46.3 (16.8)
Gender identity, n (%)	
Woman	6 (67)
Man	3 (33)
Nonbinary	0 (0)
Ethnicity, n (%)	
White	6 (75)
Black	0 (0)
Mixed	1 (12.5)
Asian	1 (12.5)
Missing or nonresponse	1 (12.5)
Identifies as having disability or long-term condition, n (%)	2 (22)

[a]Geographic spread: England (South, Midlands, North).

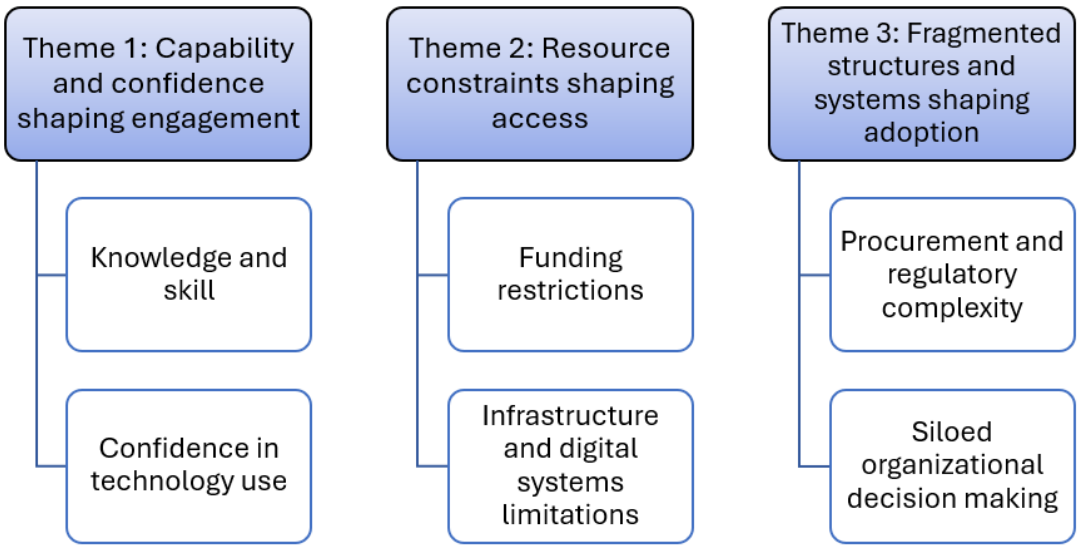
[b]Percentages are calculated using the number of respondents to each individual question. Denominators therefore vary across demographic characteristics due to non-response. Percentages for missing or nonresponse categories are calculated using the total sample size.

[c]Mean ages and SDs were estimated using the midpoints of the reported age bands, as exact ages were not collected.

Thematic Analysis

As presented in Figure 2, 3 main themes (each with 2 subthemes) were identified, which to varying degrees are relevant to all 3 stakeholder types. Each theme, along with its subthemes, is presented in turn. Although analytically distinct, subthemes were interconnected in participants' accounts and are therefore presented together. Participant quotations are labeled using anonymized identifiers showing participant number and stakeholder group (eg, ID12, lived experience) to preserve confidentiality while providing context.

Figure 2. Themes and subthemes identified through reflexive thematic analysis.



Theme 1: Capability and Confidence Shaping Engagement

Overview

Theme 1 consisted of following two subthemes: (1) knowledge and skill, and (2) confidence in technology use.

The extent to which perceived capability and confidence emerged as a theme varied by stakeholder type.

Lived Experience

Participants commonly described limited familiarity with newer rehabilitation technologies, contributing to reduced confidence to engage. Gaps in knowledge and skills led to uncertainty about how technologies functioned and how they might support recovery. This was compounded by what they perceived as limited education and follow-up from health professionals.

XSL•FO
RenderX

Participants were particularly concerned about using technologies incorrectly, especially where opportunities for feedback or clarification were lacking. Engagement was therefore shaped not only by access, but by the availability of professional support to build knowledge, skills, and confidence.

> *You have to give guidance so that the patient knows when things might not be working. Instead, patients just have to go out into the Wild West and figure it out for themselves.* [ID19, lived experience]

> *Because the rehab center was understaffed, we couldn't see the physiotherapist every time so some things you just have to figure out on your own or from others in your shoes.* [ID12, lived experience]

These accounts highlight how confidence was closely linked to ongoing professional support. Both technology professionals and rehabilitation professionals acknowledged that some patients lacked confidence or motivation to engage, yet rehabilitation staff did not overtly reflect on their own role in shaping this.

Rehabilitation Professionals

Rehabilitation professionals expressed interest in rehabilitation technologies but positioned engagement as constrained primarily by wider NHS pressures (themes 2 and 3).

> *Technology will make rehab easier and more effective.* [ID30, rehabilitation professional]

> *Every week a new piece of tech comes out that we don't get access to, or we can't have access to.* [ID56, rehabilitation professional]

Training opportunities were described as limited and ad hoc, with time pressure restricting engagement. However, unlike lived experience participants, professionals did not emphasize their own knowledge or skill gaps as primary barriers.

> *And then there's the training issues and the time that that's going to take.* [ID61, rehabilitation professional]

Technology Professionals

Technology professionals frequently attribute limited uptake to gaps in clinician capability and confidence. Clinicians were perceived as hesitant to adopt unfamiliar technologies, particularly where these added to existing workload pressures.

> *People…are overstretched in their work, which is what we're trying to alleviate. But when you go and talk to someone… they're like I can't do what I've got on my plate already.* [ID38, technology professional]

> *Clinicians still have a lack of confidence around digital tech.* [ID44, technology professional]

Capability and confidence were experienced differently across stakeholder groups. For lived experience participants, they directly influenced engagement. For rehabilitation professionals, NHS constraints predominated while technology professionals framed clinician engagement partly as a capability issue but also acknowledged wider NHS constraints.

Theme 2: Resource-Constrained Environments Shaping Access

Overview

Theme 2 consisted of following two subthemes: (1) funding restrictions, and (2) infrastructure and digital system limitations.

Across stakeholder groups, engagement was constrained by NHS system-level resource limitations, including restricted funding, which resulted in limited access to rehabilitation technology as well as infrastructure and system limitations.

Lived Experience

Participants primarily experienced these resource constraints as limited access to rehabilitation technologies through the NHS, with many unable to self-fund.

> *Most of the technology is not under the NHS…it's really expensive and I would not say that it is accessible.* [ID18, lived experience]

Engagement was therefore dependent not only on motivation and confidence, but on NHS provision.

Rehabilitation Professionals

Rehabilitation professionals identified funding constraints as a major barrier to adoption, limiting both investment in new technologies and maintenance of existing technology. Even where rehabilitation technologies were available, they were often insufficient to meet patient demand, limiting wider uptake.

> *But you know yourself, the NHS is broke.* [ID62, rehabilitation professional]

Participants described the need to demonstrate cost-effectiveness, for example, through measurable improvements in patient outcomes to justify investment, with rehabilitation competing against more immediate service pressures such as patient care.

> *I think you have to have a really robust case.* [ID61, rehabilitation professional]

Infrastructure limitations further constrained use. Outdated and fragmented IT systems, unreliable Wi-Fi, and limited technical support reduced usability and confidence in technologies that required connectivity.

> *I think there's something like 300 different electronic systems within the NHS.* [ID61, rehabilitation professional]

Limited physical space—which had intensified following the repurposing of facilities during the COVID-19 pandemic—also restricted use, particularly for larger technologies

> *Our physical space then stops us for some of the bigger technology.* [ID55, rehabilitation professional]

Technology Professionals

Technology professionals echoed these constraints, highlighting limited budgets, infrastructure constraints, competing priorities, and difficulties demonstrating the value needed to encourage investment.

XSL•FO
RenderX

> *The NHS cannot cope with the current demands and so anything new basically goes to the bottom of the list.* [ID44, technology professional]

Infrastructure challenges, particularly fragmented IT systems and aligning new technologies with these, were described as major barriers, and unreliable connectivity, including hospital Wi-Fi, further complicated use.

> *Aligning more with workflows…that is much more challenging.* [ID46, technology professional]

Overall, resource constraints limited access and adoption. Lived experience participants experienced restricted access directly, while rehabilitation and technology professionals emphasized funding, equipment, and infrastructure limitations. These constraints operated at a system level, distinct from capability (theme 1) and fragmented structures and systems (theme 3), although closely related.

Theme 3: Fragmented Structures and Systems Shaping Adoption

Overview

Theme 3 consisted of following two subthemes: (1) procurement and regulatory complexity, and (2) siloed organizational decision-making.

Participants described governance structures within the NHS as shaping whether and how rehabilitation technologies were adopted. Governance encompassed formal approval pathways, procurement processes, regulatory requirements, and organizational decision-making mechanisms.

These issues were only raised by rehabilitation and technology professionals; lived experience participants did not contribute to this theme, reflecting their distance from the approval and purchasing processes that determine technology availability within the NHS.

Rehabilitation Professionals

Participants described procurement processes as slow, inconsistent, and difficult to navigate, with substantial variation across NHS Trusts (NHS Trusts are publicly funded health care organizations responsible for delivering hospital, community, mental health, and ambulance services to local populations). This meant that there was no standardized route for rolling out technologies across NHS Trusts.

> *We're under the same umbrella of NHS, but we work very differently and have our own systems in place.* [ID56, rehabilitation professional]

A small minority also stated data governance requirements added an additional layer of complexity.

> *It's hard to overcome some of the natural governance issues… about how data is going to be collected, stored, used.* [ID62, rehabilitation professional]

Technology Professionals

Technology professionals similarly described regulatory approval as time-consuming and costly. Developers were required to navigate multiple NHS Trust-specific procurement processes, each with distinct requirements and timelines. This fragmentation constrained widespread adoption, even where technologies had demonstrated value. From this perspective, barriers to adoption reflected siloed organizational decision-making.

> *They (NHS Trusts) each have a different set up and they each have a different kind of protocol."* [ID43, technology professional]

Governance structures shaped adoption independently of capability (theme 1) and resource constraints (theme 2). While resources determined what could be supported, governance processes determined what could be approved and implemented. These operated at a system level beyond the control of individual rehabilitation and technology stakeholders and did not directly impact those with lived experience.

Summary of Stakeholder Training Needs

Although the nature of training needs differed across stakeholder groups, several clear priorities emerged. Individuals with lived experience emphasized the need for practical instruction, ongoing professional support, and confidence-building in the use of rehabilitation technologies. Rehabilitation professionals highlighted the need for additional education on available technologies and practical training to support integration into clinical workflows. Technology professionals emphasized the need for training on how to communicate effectively with the NHS and to navigate complex procurement pathways.

Mapping Themes to Behavioral Determinants (COM-B Informed by TDF)

Given the study aim of identifying education and training needs, themes were further interpreted using the COM-B model [22], which conceptualizes behavior as shaped by capability (eg, knowledge and skills), opportunity (eg, environmental and organizational factors), and motivation, including both reflective (eg, beliefs and intentions) and automatic (eg, habits and emotions) components. Selected domains from the TDF [23] were used to clarify the nature of influences evident in stakeholder accounts. Only domains clearly reflected in the data were applied, and mapping was conducted separately for each stakeholder group. Table 4 shows the mapping applied.

XSL•FO
RenderX

Table 4. Mapping themes to behavioral determinants using the capability, opportunity, motivation–behavior model informed by selected Theoretical Domains Framework domains.

Stakeholder group	Behavior	COM-B[a] components	Determinants (TDF[b] domains)
Lived experience	Engage with and use rehabilitation technologies	Capability and opportunity, with motivation influences	• Limited knowledge and practical skills (knowledge; skills → capability) • Doubts about competence and confidence (beliefs about capabilities → reflective motivation) • Emotional responses (eg, anxiety and uncertainty; emotion → automatic motivation) • Limited access to technologies, guidance, and support (environmental context and resources → opportunity)
Rehabilitation professionals	Adopt and integrate technologies into practice	Opportunity (predominant), with capability and motivation influences	• Limited organizational resources (eg, funding and equipment; environmental context and resources → opportunity) • Infrastructure constraints (eg, IT systems and connectivity; environmental context and resources → opportunity) • Complex procurement pathways (environmental context and resources → opportunity) • Fragmented organizational decision-making (environmental context and resources → opportunity)
Technology professionals	Secure approval and enable adoption of rehabilitation technologies within the NHS[c]	Opportunity (structural), with reflective motivation influences	• Regulatory and approval requirements; variable procurement pathways (environmental context and resources → opportunity) • Infrastructure and system integration challenges (environmental context and resources → opportunity) • Perceived clinician training and confidence gaps (beliefs about capabilities → reflective motivation)

[a]COM-B: capability, opportunity, motivation–behavior.

[b]TDF: Theoretical Domains Framework.

[c]NHS: National Health Service.

For lived experience participants, influences relating to capability were prominent. Participants described gaps in knowledge and practical skills, which created uncertainty about correct use of rehabilitation technologies. These concerns reflected doubts about competence, reducing their confidence and willingness to engage. Opportunity-related influences took the form of limited NHS availability and the inability to purchase technology privately. Limited guidance and support from clinical staff further shaped engagement. In this context, motivation appeared closely tied to confidence and professional support.

For rehabilitation professionals, the key behavior was adopting and integrating rehabilitation technologies into clinical practice. Clinicians noted that education and training opportunities were limited, but engagement was largely constrained by funding, infrastructure, and decision-making processes within the NHS. In practice, this meant restricted availability of equipment, outdated and poorly integrated IT systems, and complex procurement pathways across NHS Trusts. Although clinicians expressed interest in the potential of rehabilitation technologies, motivation appeared closely linked to these practical constraints.

For technology professionals, the principal behavior concerned securing approval for and enabling the use of rehabilitation technologies within the NHS. Engagement was shaped by regulatory requirements, procurement pathways, and variation across NHS Trusts, which determined whether technologies could be adopted. Infrastructure challenges, including integration with existing NHS systems, further limited uptake. Developers perceived clinicians to lack confidence and readiness to use rehabilitation technologies, which they attributed to limited education and training.

Across stakeholder groups, the emphasis of influences differed. Capability-related factors were most visible among lived experience participants, whereas organizational and system-level constraints were more prominent in accounts from rehabilitation and technology professionals.

Contextualizing Behavioral Determinants Within Organizational Systems (CFIR)

While the COM-B interpretation highlighted the influence of capability, opportunity, and motivation on stakeholder engagement, these influences were not independent of the organizational context for rehabilitation technologies. Accounts from rehabilitation and technology professionals pointed to wider NHS systems that shaped what was practically possible. To situate these influences, findings were interpreted in relation to relevant constructs from the CFIR [24], with a primary focus on organizational (inner setting) and external policy (outer setting) factors. Subthemes were aligned with CFIR constructs through a deductive mapping process, with constructs that were clearly or meaningfully reflected in the data prioritized.

Constructs relating to intervention characteristics (eg, evidence strength and quality) were only indirectly represented—primarily through participants' references to cost-effectiveness and the need to demonstrate value—and are therefore indicated as secondary and indirect in Table 5. Similarly, elements of the process domain relating to implementation—primarily stakeholder engagement—were

XSL•FO
RenderX

evident in accounts of clinician hesitancy and competing demands. However, other process constructs (eg, planning, evaluation, and execution) were not substantively represented; where relevant, engagement is therefore indicated as a secondary and partial construct in Table 5. Constructs relating to individual characteristics were not mapped, as these were more appropriately examined through COM-B and TDF (Theme 1). Table 5 presents the mapping of themes 2 and 3 to relevant CFIR domains and constructs. Theme 1 is not included, as it reflects individual-level influences examined through COM-B and TDF.

Table 5. Mapping of themes 2 and 3 to relevant Consolidated Framework for Implementation Research domains and constructs.

Theme	Stakeholder group	CFIR[a] domain	CFIR constructs	Illustrative determinants
Theme 2: resource-constrained environments shaping access	Lived experience	Inner setting	• Available resources	Limited NHS[b] provision; inability to self-fund technology
Theme 2: resource-constrained environments shaping access	Rehabilitation professionals	Inner setting	• Available resources; • Primary: available resources; compatibility • Secondary: evidence strength and quality (indirect)	Restricted funding; need to demonstrate cost-effectiveness; insufficient equipment; IT integration challenges; unreliable Wi-Fi; limited technical support
Theme 2: resource-constrained environments shaping access	Technology professionals	Inner setting	• Primary: available resources; compatibility • Secondary: evidence strength and quality (indirect)	Infrastructure constraints; fragmented IT systems; limited investment capacity; challenges demonstrating value
Theme 3: fragmented structure and systems shaping adoption	Rehabilitation professionals	Outer setting or inner setting	• Primary (outer setting): external policies and incentives • Primary (inner setting): structural characteristics	Procurement variability; inconsistent NHS Trust approval processes; information governance requirements
Theme 3: fragmented structure and systems shaping adoption	Technology professionals	Outer setting or inner setting; process (partial)	• Primary (outer setting): external policies and incentives • Primary (inner setting): structural characteristics • Secondary (process): engagement (partial)	Regulatory approval requirements; fragmented NHS Trust–level procurement; dispersed decision-making authority; clinician hesitancy; limited capacity to engage with new technologies

[a]CFIR: Consolidated Framework for Implementation Research.

[b]NHS: National Health Service.

Theme 2—resource-constrained environments shaping access—aligned primarily with CFIR constructs relating to available resources and compatibility within the inner setting. Limited funding, insufficient equipment, and infrastructure constraints shaped the feasibility of technology use, constraining opportunity regardless of individual capability or motivation.

Theme 3—fragmented structures and systems shaping adoption—aligned with CFIR constructs across the outer and inner setting domains, particularly external policies and incentives (outer setting) and structural characteristics (inner setting). Regulatory requirements, procurement pathways, and variation across NHS Trusts shaped approval and adoption processes, often constraining implementation before individual behavior became relevant. In this way, governance structures influenced the opportunity for engagement.

Overall, these findings indicate that behavioral determinants identified through COM-B were embedded within organizational and policy contexts. Education and training may strengthen knowledge and confidence, but their impact is likely to depend on resource capacity and governance processes within NHS systems.

Across stakeholder groups, several shared issues relevant to training and education were identified. Individuals with lived experience identified gaps in both the education and training they received from rehabilitation professionals, but their experience was largely limited to low- and mid-technology devices, making it difficult to identify or prioritize training needs related to more advanced technologies. Rehabilitation professionals reported exposure to a broader range of technologies, although access varied but training was described as limited and ad hoc rather than embedded within continuing professional development. Technology professionals also emphasized the importance of training for rehabilitation staff. These findings have direct implications for the design of training and implementation strategies.

Discussion

Key Results

Previous research has typically focused on specific rehabilitation technologies and/or single stakeholder groups. A key novel contribution of this study is the integrated comparative analysis of education and training needs across 3 distinct stakeholder groups—rehabilitation professionals, technology professionals,

XSL•FO
RenderX

and individuals with lived experience. By examining these groups together, this study identifies specific training needs for each stakeholder group and how these interact to influence the adoption of rehabilitation technologies within the NHS.

This study identified 3 interconnected themes—capability and confidence, resource-constrained environments shaping access, and fragmented structure and systems shaping adoption. By applying behavioral (COM-B or TDF) and implementation (CFIR) frameworks to inductive thematic analysis, this study shows that training needs are shaped by both individual capabilities and wider organizational and system factors within the NHS. Although some constructs span different models, the frameworks were applied at different levels of interpretation. COM-B and TDF were used to understand determinants of individual behavior, whereas CFIR was used to interpret organizational and system-level influences. While capability-related influences were most visible among individuals with lived experience, opportunity-related and structural constraints predominated in professional and developer accounts. Nevertheless, all stakeholder groups recognize the prospective benefits of rehabilitation technologies, particularly their potential to address resource constraints within the NHS and improve patient outcomes and quality of care.

For individuals with lived experience, engagement with rehabilitation technologies was linked to knowledge, confidence, and perceived competence, and participants identified gaps in the education and training they received from rehabilitation professionals. They found it challenging to identify and prioritize training by technology type since their experience of rehabilitation technologies typically focused on low- or mid-tech devices, including mobility aids, activity and health and well-being trackers, and apps, with little experience of smart assessment and training devices. The study findings highlight an opportunity to support individuals with lived experience to better understand the role of newer technologies in rehabilitation, particularly approaches such as telerehabilitation for monitoring and support in the community, where access to specialist services is often limited.

In contrast, rehabilitation professionals had experience of a wide range of rehabilitation technology, although access varied notably both across and within NHS Trusts, reflecting differences in resources, funding, and organizational priorities. Training in emerging technologies was perceived to be limited and ad hoc rather than embedded in continuing professional development (CPD). Interest in technologies such as telerehabilitation, digital patient education and engagement tools, and AI-enabled documentation systems was prioritized, driven by their perceived potential to address workforce resource and capacity pressures.

Technology professionals also highlighted the need for clinician training but confirmed that there are significant barriers to adoption, including resource constraints such as limited funding together with fragmented systems and structures within the NHS.

Implications for Training and Practice

Mapping these findings to the COM-B model enables us to provide actionable, stakeholder-specific recommendations for the design of rehabilitation technology training that extends beyond generic skills development.

For individuals with lived experience, training should prioritize practical hands-on learning to build confidence and capability and include ongoing clinical support and opportunities for a bidirectional feedback loop between users and health and care professionals.

For rehabilitation professionals, technologies such as telerehabilitation, digital patient education and engagement tools, and AI-enabled documentation systems are a priority, and training should address capability and focus on practical integration within clinical workflows and NHS systems and structures. Motivation to adopt technologies appeared to be linked to their perceived value in improving efficiency and patient care, suggesting that training should emphasize practical benefits alongside skill development. However, the effectiveness of training is likely to depend on organizational contexts within the NHS, including resource availability and system infrastructure [25,26].

For technology professionals, including those involved in the development and supply of rehabilitation technologies, training needs are related to the tools to equip them to communicate effectively with the NHS about potential new technologies and to navigate regulatory and procurement processes and decision-making [24,25,27,28]. The main challenge is designing training that can be applied to the NHS as a whole while still accounting for the variability across individual NHS Trusts [29,30]. These system-level barriers may be most effectively addressed by facilitating collaboration between groups of rehabilitation professionals, technology professionals, and NHS representatives from different NHS Trusts and centralized operations where they exist. Rehabilitation engineers may occupy an overlapping position between rehabilitation and technology professionals, with some working as part of multidisciplinary rehabilitation or assistive technology teams [31]. Across all stakeholder groups, training should adopt a multimodal format combining online and experiential learning.

Implications for Adoption of Rehabilitation Technologies

Our findings suggest that while education and training in rehabilitation technologies are necessary, their impact on adoption is likely to be constrained by organizational and system-level factors within the NHS. This highlights that training alone is unlikely to drive widespread adoption without concurrent changes to the contexts in which technologies are implemented. In this context, the NHS 10 Year Health Plan (2025) “Fit for the Future,” which outlines a shift toward a “digital-first” health care system, further emphasizes the need to align training with system readiness to support meaningful integration of rehabilitation technologies into clinical pathways [14].

XSL•FO
RenderX

Strengths and Limitations

The sample of 50 is relatively large for a qualitative study [20,32,33], and this study is novel in that it provides an in-depth comparative analysis of the training needs of diverse stakeholders. Furthermore, by applying behavioral and implementation frameworks to inductive thematic analysis, the study situates individual training needs within the context of the NHS, thereby informing real-world decisions on how best to design and deliver training in this environment.

Nevertheless, several limitations should be acknowledged. Participants with lived experience and rehabilitation professionals were overrepresented compared with technology professionals. However, despite smaller numbers in the technology group, themes relating to constraints in introducing and using rehabilitation technologies within the NHS were consistent across both professional groups, suggesting sufficient depth and richness of data to support the analytic interpretations.

Group sizes varied (3-9 participants), and both focus groups and individual interviews were conducted. These methodological variations may have influenced discussion dynamics and the extent of individual contribution; however, combining methods is a pragmatic strategy to maximize participation and provide complementary insights [34,35]. The recurrence of themes across both group and interview data reduces concerns that the manner of data collection materially influenced the core findings.

Although lived experience participants reflected diversity in age, ethnicity, and disability status, comparable diversity was less apparent among rehabilitation and technology professionals, and these differences may have influenced the range of views represented. As the study was not designed to examine how demographic characteristics shaped participants' perspectives, no conclusions can be drawn regarding the influence of ethnicity, gender identity, or disability on the findings. Furthermore, the demographic questionnaire was designed for use across all stakeholder groups. Consequently, detailed professional characteristics (eg, profession, clinical role, or years of experience) were not collected for rehabilitation or technology professionals, limiting further characterization of these participants. Further research should collect more detailed information on professional background, clinical role, and experience, alongside seeking greater representation of younger professionals, male stakeholders, those identifying as nonbinary, and individuals from a wider range of ethnic backgrounds.

Although recruitment was UK-wide, the sample was concentrated in England; only 1 participant was based in Northern Ireland, and none were recruited from Scotland or Wales. Given differences across the devolved health care systems, the absence of representation from Scotland and Wales may limit the transferability of these findings beyond the English NHS. Future research should seek greater representation from Scotland, Wales, and Northern Ireland to determine the extent to which these findings apply in these settings.

Finally, 2 individual interviews were not audio-recorded. Although detailed notes were taken, the absence of verbatim transcripts may have reduced the depth and nuance captured from those participants. However, these participants represented a small proportion of the total sample and themes discussed in these interviews were consistent with those identified across the wider dataset [36].

Comparison With Prior Work

Research that specifically addresses the education and training needs for rehabilitation technology across multiple stakeholders such as rehabilitation professionals, technology professionals, and individuals with lived experience and how to collaborate effectively to design effective training initiatives is limited. Training needs for rehabilitation professionals specifically tend to focus on a specific area, such as assistive technology [37] and virtual reality [38,39], or feedback-based technology [40], wireless technologies [41], and telerehabilitation [42]. Research focusing on technology professionals tends to investigate their role in the process rather than explicit training needs [43,44]. The needs of individuals with disabilities are typically viewed through the lens of accessibility and inclusivity and emphasize the importance of user-centered design [9,45-47]. Hence, a key strength of this study is identifying priorities for training in rehabilitation technology across diverse stakeholders and by identifying behavioral factors that determine engagement and situating them within the context of the NHS.

Conclusions

This study highlights that training is necessary but insufficient on its own; workforce development strategies must be accompanied by organizational, infrastructural, and governance changes if rehabilitation technologies are to be adopted sustainably across NHS services. This novel study speaks to the importance of collaboration in the fast-developing multidisciplinary area of rehabilitation science [48].

Acknowledgments

We thank all stakeholders who participated in this activity, and the National Institute for Health and Care Research (NIHR) HealthTech Research Centre for Rehabilitation for administrative support with focus groups, in particular, Jacqueline Swift.

Data Availability

Due to the sensitive nature of qualitative data and to protect participant confidentiality, transcripts cannot be publicly shared. Deidentified excerpts are available within the manuscript.

XSL•FO
RenderX

Funding

This public engagement and consultation activity was supported by the NIHR HealthTech Research Centre for Rehabilitation (NIHR205288), which funds a proportion of the corresponding author's time.

Generative AI Use

The authors used ChatGPT (OpenAI, GPT-5.5; accessed February-March 2026) during manuscript preparation to support language refinement, improve readability, and explore behavioral and implementation science frameworks that might be relevant for further interpretation following completion of the inductive thematic analysis. ChatGPT did not analyze qualitative data, generate codes or themes, determine framework mapping, or the study design or conclusions. Suggestions generated by ChatGPT were critically evaluated by the authors against the implementation science literature and the study findings before the capability, opportunity, motivation–behavior model, selected Theoretical Domains Framework domains, and Consolidated Framework for Implementation Research were chosen and applied. All data analysis, interpretation, framework mapping, and manuscript content were undertaken by the authors, who take full responsibility for the final manuscript. The relevant ChatGPT prompts are provided in Multimedia Appendix 6.

Authors' Contributions

Conceptualization: HB and VAF
Data curation: AT
Formal analysis: AT
Funding acquisition: HB
Investigation: AT and MH
Methodology: HB and AT
Project administration: HB, AT, and MH
Resources: HB and AT
Writing—original draft preparation: AT and HB
Writing—review and editing: MH, and VAF
All authors have reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guide Groups and Interviews – Lived Experience.
[DOCX File , 197 KB - rehab_v13i1e97602_app1.docx]

Multimedia Appendix 2
Guide Groups and Interviews – Rehabilitation Professionals.
[DOCX File , 197 KB - rehab_v13i1e97602_app2.docx]

Multimedia Appendix 3
Guide Groups and Interviews – Technology Professionals.
[DOCX File , 199 KB - rehab_v13i1e97602_app3.docx]

Multimedia Appendix 4
Consolidated Criteria for Reporting Qualitative Research (COREQ).
[DOCX File , 33 KB - rehab_v13i1e97602_app4.docx]

Multimedia Appendix 5 [DOCX File , 24 KB - rehab_v13i1e97602_app5.docx]

Multimedia Appendix 6
ChatGPT Prompt Log February and March 2026.
[DOCX File , 43 KB - rehab_v13i1e97602_app6.docx]

References

1. Martinez-Martin E, Cazorla M. Rehabilitation technology: assistance from hospital to home. Comput Intell Neurosci 2019;2019:1431509 [FREE Full text] [doi: 10.1155/2019/1431509] [Medline: 31281333]

XSL•FO
RenderX

2. Willingham TB, Stowell J, Collier G, Backus D. Leveraging emerging technologies to expand accessibility and improve precision in rehabilitation and exercise for people with disabilities. Int J Environ Res Public Health 2024;21(1):79 [FREE Full text] [doi: 10.3390/ijerph21010079] [Medline: 38248542]
3. Howroyd F, Earle N, Weblin J, McWilliams D, Raven M, Duggal NA, et al. Transitioning to a virtual post-intensive care rehabilitation service in response to the COVID-19 pandemic: results of multidisciplinary focus-groups. Front Med (Lausanne) 2024;11:1513121 [FREE Full text] [doi: 10.3389/fmed.2024.1513121] [Medline: 39830386]
4. Adams R, Jones A, Lefmann S, Sheppard L. Rationing is a reality in rural physiotherapy: a qualitative exploration of service level decision-making. BMC Health Serv Res 2015;15(1):121 [FREE Full text] [doi: 10.1186/s12913-015-0786-3] [Medline: 25880469]
5. NHS England. The NHS Long Term Plan. London, UK: NHS England; 2019.
6. Department of Health and Social Care, NHS England. A Plan for Digital Health and Social Care. London, UK: Department of Health and Social Care; 2022.
7. NHS England. NHS Long Term Workforce Plan. London, UK: NHS England; 2023.
8. Asthana S, Jones R, Sheaff R. Why does the NHS struggle to adopt eHealth innovations? A review of macro, meso and micro factors. BMC Health Serv Res 2019;19(1):984 [FREE Full text] [doi: 10.1186/s12913-019-4790-x] [Medline: 31864370]
9. Mitchell J, Shirota C, Clanchy K. Factors that influence the adoption of rehabilitation technologies: a multi-disciplinary qualitative exploration. J Neuroeng Rehabil 2023;20(1):80 [FREE Full text] [doi: 10.1186/s12984-023-01194-9] [Medline: 37340496]
10. Morris J, Jones M, Thompson N, Wallace T, DeRuyter F. Clinician perspectives on mRehab interventions and technologies for people with disabilities in the United States: a national survey. Int J Environ Res Public Health 2019;16(21):4220 [FREE Full text] [doi: 10.3390/ijerph16214220] [Medline: 31683536]
11. Reddy KJ. Integrating technology into clinical practice In: Innovations in neurocognitive rehabilitation: harnessing technology for effective therapy. Cham: Springer Nature Switzerland; 2025:329-350.
12. Blake H, Abbott-Fleming V, Abdalrahim A, Horrocks M. Co-Designing an inclusive stakeholder engagement strategy for rehabilitation technology training using the I-STEM model. Int J Environ Res Public Health 2025;23(1):13 [FREE Full text] [doi: 10.3390/ijerph23010013] [Medline: 41595807]
13. Peek N, Sujan M, Scott P. Digital health and care in pandemic times: impact of COVID-19. BMJ Health Care Inform 2020;27(1):e100166 [FREE Full text] [doi: 10.1136/bmjhci-2020-100166] [Medline: 32565418]
14. Department of Health and Social Care. Fit for the future: the 10 year health plan for England London. England London: Department of Health and Social Care; 2025.
15. Adeoye‐Olatunde OA, Olenik NL. Research and scholarly methods: semi‐structured interviews. J Am Coll Clin Pharm 2021;4(10):1358-1367. [doi: 10.1002/jac5.1441]
16. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
17. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
18. Byrne D. A worked example of braun and clarke's approach to reflexive thematic analysis. Qual Quant 2021;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
19. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
20. Mattimoe R, Hayden M, Murphy B. J Approaches to analysis of qualitative research data: a reflection on the manual and technological approaches. Finance: Account Finance Gov Rev; 2021:1-16.
21. Thompson SM. Barrett PA Summary oral reflective analysis: a method for interview data analysis in feminist qualitative research. ANS Adv Nurs Sci 1997;20(2):55-65. [Medline: 9394035]
22. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
23. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
24. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4(1):50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
25. Wakeling J, New H, Hall A. Christie L Addressing the learning and development needs of the NHS support workforce. Br J Healthc Manag 2025;31(6):1-12. [doi: 10.12968/bjhc.2024.0081]
26. Imison C, Buchan J. Xavier S NHS workforce planning: limitations and possibilities. London: The King's Fund; 2009.
27. Evans TD, Ahmad O, Alderman JE, Bailey G, Bannister P, Barlow N, et al. The role of procurement frameworks in responsible AI innovation in the National Health Service: a multi-stakeholder perspective. Front Health Serv 2025;5:1608087 [FREE Full text] [doi: 10.3389/frhs.2025.1608087] [Medline: 40538862]

XSL•FO
RenderX

28. Kyratsis Y, Ahmad R, Holmes AH. Technology adoption and implementation in organisations. BMJ Open 2012;2(2):e000872 [FREE Full text] [doi: 10.1136/bmjopen-2012-000872] [Medline: 22492183]
29. Collins B Adoption and spread of innovation in the NHS. London: The King's Fund; 2018.
30. Marjanovic S, Altenhofer M, Hocking L, Chataway J, Ling T. Innovating for improved healthcare. Sci Public Policy 2020;47(2):283-297. [doi: 10.1093/scipol/scz051]
31. Institute of Physics and Engineering in Medicine. Report on the rehabilitation engineering workforce. Institute of Physics and Engineering in Medicine 2016.
32. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research. Soc Sci Med 2022;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
33. Boddy CR. Sample size for qualitative research. Qual Mark Res 2016;19(4):426-432. [doi: 10.1108/qmr-06-2016-0053]
34. Lambert SD, Loiselle CG. Combining individual interviews and focus groups. J Adv Nurs 2008;62(2):228-237. [doi: 10.1111/j.1365-2648.2007.04559.x] [Medline: 18394035]
35. Caillaud S, Flick U, Morgan D. In: Barbour RS, Morgan DL, editors. Focus groups in triangulation contexts. London: Palgrave Macmillan UK; 2017:155-177.
36. Rutakumwa R, Mugisha JO, Bernays S, Kabunga E, Tumwekwase G, Mbonye M, et al. Conducting in-depth interviews with and without voice recorders: a comparative analysis. Qual Res 2020;20(5):565-581 [FREE Full text] [doi: 10.1177/1468794119884806] [Medline: 32903872]
37. Manship S, Hatzidimitriadou E, Moore J, Stein M, Towse D, Smith R. The experiences and perceptions of health-care professionals regarding assistive technology training: a systematic review. Assist Technol 2024;36(2):123-146 [FREE Full text] [doi: 10.1080/10400435.2023.2219294] [Medline: 37399518]
38. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251 [FREE Full text] [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
39. Dumitrescu C. Use of emerging technologies in rehabilitation education and practice. Front Rehabil Sci 2025;6:1655454 [FREE Full text] [doi: 10.3389/fresc.2025.1655454] [Medline: 40861010]
40. Hamilton C, Lovarini M, McCluskey A, Folly de Campos T, Hassett L. Experiences of therapists using feedback-based technology to improve physical function in rehabilitation settings: a qualitative systematic review. Disabil Rehabil 2019;41(15):1739-1750. [doi: 10.1080/09638288.2018.1446187] [Medline: 29513052]
41. Dicianno BE, Salh A, Morris L, Xiang Y, Ding D. Rehabilitation clinicians' use of mainstream wireless technologies in practice: a scoping review. Disabil Rehabil Assist Technol 2024;19(8):2742-2760. [doi: 10.1080/17483107.2024.2316891] [Medline: 38349177]
42. Stark AL, Krayter S, Dockweiler C. Competencies required by patients and health professionals regarding telerehabilitation: a scoping review. Digit Health 2023;9:20552076231218841 [FREE Full text] [doi: 10.1177/20552076231218841] [Medline: 38107985]
43. DiGiovine CP, Donahue M, Bahr P, Bresler M, Klaesner J, Pagadala R, et al. Rehabilitation engineers, technologists, and technicians: vital members of the assistive technology team. Assist Technol 2023;35(1):23-34. [doi: 10.1080/10400435.2018.1454713] [Medline: 29953331]
44. Rowley BA, Mitchell DF, Weber C. Educating the rehabilitation engineer as a service provider. Assist Technol 1997;9(1):62-69. [doi: 10.1080/10400435.1997.10132297] [Medline: 10168026]
45. Zhou C, Ai Y, Wang S, Yuan Y, Zhang A, Hu H, et al. Barriers and facilitators to participation in electronic health interventions in older adults with cognitive impairment: an umbrella review. BMC Geriatr 2024;24(1):1037 [FREE Full text] [doi: 10.1186/s12877-024-05645-3] [Medline: 39725926]
46. Blake H, Abbott-Fleming V, Greaves S, Somerset S, Chaplin WJ, Wainwright E, et al. Five years of patient and public involvement and engagement (PPIE) in the development and evaluation of the Pain-at-Work toolkit to support employees' self-management of chronic pain at work. Res Involv Engagem 2025;11(1):81 [FREE Full text] [doi: 10.1186/s40900-025-00757-5] [Medline: 40665420]
47. Alt Murphy M, Pradhan S, Levin M, Hancock NJ. Uptake of technology for neurorehabilitation in clinical practice: a scoping review. Phys Ther 2024;104(2):pzad140 [FREE Full text] [doi: 10.1093/ptj/pzad140] [Medline: 37856528]
48. Winstein CJ, Requejo PS, Zelinski EM, Mulroy SJ, Crimmins EM. A transformative subfield in rehabilitation science at the nexus of new technologies, aging, and disability. Front Psychol 2012;3:340 [FREE Full text] [doi: 10.3389/fpsyg.2012.00340] [Medline: 23049517]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
COM-B: capability, opportunity, motivation–behavior
CPD: Continuing Professional Development
COREQ: Consolidated Criteria for Reporting Qualitative Research
I-STEM: Implementation Stakeholder Engagement Model

XSL•FO
RenderX

NHS: National Health Service
NIHR: National Institute for Health and Care Research
TDF: Theoretical Domains Framework

Edited by J Sarvestan; submitted 08.Apr.2026; peer-reviewed by C Kiekens, L Herdiman; comments to author 15.Jul.2026; revised version received 16.Sep.2026; accepted 05.Oct.2026; published 13.Nov.2026.

Please cite as:
Thornton A, Horrocks M, Abbott-Fleming V, Blake H
Understanding Education and Training Needs in Rehabilitation Technologies Among Those With Lived Experience, Rehabilitation Health and Care Professionals, and Technology Professionals: Qualitative Study
JMIR Rehabil Assist Technol 2026;13:e97602
URL: https://rehab.jmir.org/2026/1/e97602
doi:10.2196/97602
PMID:

©Angela Thornton, Matthew Horrocks, Victoria Abbott-Fleming, Holly Blake. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 13.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development, Usability, and Feasibility of a Pain-Stratified Smartphone-Based Home Intervention System for Patients With Adhesive Capsulitis: Exploratory Feasibility and Usability Trial

Hsun-Han Hsu[1,2], PT, MS; Kai-Chia Cheng[3], PT, PhD; Ya-Lan Chiu[4,5], PT, PhD

[1]Department of Physical Therapy and Assistive Technology, National Yang Ming Chiao Tung University, Taipei, Taiwan

[2]Glow Physical Therapy Center, Kaohsiung, Taiwan

[3]National Center for Geriatrics and Welfare Research, National Health Research Institutes, Yunlin, Taiwan

[4]Department of Physical Therapy, College of Medicine, National Cheng Kung University, No.1, University Road, Tainan, Taiwan

[5]Physical Therapy Center, National Cheng Kung University Hospital, Tainan, Taiwan

Corresponding Author:
Ya-Lan Chiu, PT, PhD
Department of Physical Therapy, College of Medicine, National Cheng Kung University, No.1, University Road, Tainan, Taiwan

Abstract

Background: Frozen shoulder causes prolonged pain and restricted range of motion that impair daily function. Home exercise is an indispensable part of self-management, but adherence is often suboptimal under traditional printed or verbal instruction because of the lack of objective monitoring and individualized progression. Existing rehabilitation apps offer sensor-based assessment and feedback but generally lack an adaptive mechanism that tailors the exercise prescription to a patient's dynamically changing clinical status.

Objective: This study aimed to develop a personalized, pain-stratified telerehabilitation app for patients with frozen shoulder; to evaluate its usability, feasibility, and safety as an add-on to usual care; and to explore the system's clinical outcomes on a descriptive basis.

Methods: An Android app integrating smartphone sensor-based range of motion self-assessment, pain-based exercise stratification, usage records, and exercise reminders was developed. In an exploratory, assessor-blinded, randomized controlled feasibility trial, 10 patients with frozen shoulder were randomly allocated to an app group (n=5) or a handout group (n=5), and both underwent a 4-week home exercise intervention with identical content. Primary outcomes included system usability (System Usability Scale) and feasibility (retention and exercise adherence). Secondary outcomes focused on the exploratory clinical signs included pain and function (Shoulder Pain and Disability Index) and range of motion (measured by goniometer). Given the small sample size, continuous variables were analyzed nonparametrically and interpreted strictly as descriptive clinical trends without inferential hypothesis testing.

Results: All 10 participants completed the trial (attrition rate 0%). The mean System Usability Scale score in the app group was 87.5 (SD 5.86), corresponding to an "excellent" usability rating. Over the 28-day intervention, the app's usage logs objectively documented a mean of 18.0 (SD 5.0) of 28 exercise days (18/28, 64.3%) in the app group, demonstrating the feasibility of objective home monitoring. Both groups improved in Shoulder Pain and Disability Index scores and range of motion. Notably, all participants in the app group achieved a descriptive within-group improvement in the clinically challenging shoulder external rotation, whereas the handout group showed smaller marginal gains; however, between-group comparisons across all parameters were interpreted descriptively without significance testing due to the sample size. On the other clinical outcomes, the handout group showed numerically comparable or larger improvements.

Conclusions: This study shows that a pain-stratified, personalized telerehabilitation app is highly usable and feasible as an add-on mobile health intervention for patients with frozen shoulder and that it provides objective home monitoring of exercise adherence that traditional handouts cannot. These exploratory findings highlight the need for a fully powered confirmatory trial. To robustly evaluate clinical efficacy, future study designs must include a common objective adherence measure, stratification variables, and an extended follow-up period.

(JMIR Rehabil Assist Technol 2026;13:e107314) doi:10.2196/107314

KEYWORDS

frozen shoulder; adhesive capsulitis; telerehabilitation; mobile health; mHealth; smartphone app; home exercise; exercise adherence; range of motion

XSL•FO
RenderX

Introduction

Frozen shoulder, also clinically termed adhesive capsulitis, is pathologically characterized by excessive scar tissue formation and fibrotic adhesions within the shoulder joint capsule, leading to severe pain, joint stiffness, and progressive functional impairment of the shoulder [1]. The literature reports that the prevalence of this condition is approximately 3% to 5% in the general population, with a particular predilection for middle-aged and older women aged 40 to 60 years [2]. Although the precise pathogenesis remains incompletely understood, established clinical risk factors include female sex, diabetes mellitus, and thyroid disease [3-5].

Clinically, frozen shoulder presents primarily with spontaneous pain and global restriction of shoulder range of motion, most notably in the planes of flexion, abduction, internal rotation, and external rotation [6-8]. The condition not only impairs shoulder function but also leads to reduced quality of life, work disability, and sleep disturbance [9], imposing a substantial socioeconomic burden due to direct medical costs and work absence [10]. While clinical treatment strategies are diverse [1-3], early professional medical intervention is strongly recommended to substantially improve prognosis and shorten the disease course [2,3,11]. According to the clinical practice guidelines issued by the American Physical Therapy Association (APTA) [3], patient education and home exercise training play an indispensable role in overall self-management, with specific program designs, such as incorporating scapulothoracic strengthening, showing marked benefits [12]. However, previous studies strongly caution that overly aggressive physical interventions can overstimulate the inflamed joint capsule, leading to secondary injury [13,14]. Conversely, performing self-directed exercises within a pain-free range achieves significantly better long-term functional outcomes [13]. Therefore, in line with APTA guidelines, exercise prescriptions must be strictly stratified according to the patient's tissue irritability level [3]. For patients with high irritability, treatment should focus on active-assisted mobility exercises within a pain-free range rather than high-intensity tissue stretching, underscoring the necessity of personalized training designs that allow patients to dynamically adjust their routine according to their symptoms.

Nevertheless, although a well-structured and personalized home exercise design is critical to the prognosis of frozen shoulder, the practical realization of these clinical benefits remains highly dependent on patient exercise adherence in an unsupervised environment [15,16]. When patients perceive treatment as ineffective or experience a sense of helplessness, exercise discontinuation frequently ensues [17]. These challenges highlight 3 major gaps in traditional home programs that rely heavily on printed or verbal instructions. First, health care providers are unable to monitor patient adherence in real time [16,17]. Second, patients often forget movement details in the home environment; without immediate professional feedback, they risk secondary injury caused by improper compensation [16]. Third, the lack of a dynamic monitoring mechanism prevents therapists from tracking daily changes in pain and joint angles, making it difficult to fine-tune prescriptions in a timely manner [18].

With the rapid advancement of mobile health (mHealth) and digital health technologies, smartphone apps have become an innovative solution to overcome the spatial and temporal constraints of traditional health care [19,20]. These systems can automatically record usage, using built-in sensors or image-based methods to measure shoulder joint angles and provide real-time visual feedback [21]. Numerous scholars have developed frozen shoulder apps equipped with range of motion measurement and audio-visual guidance [22,23]; however, these systems mostly provide one-size-fits-all static prescriptions, lacking the clinical decision-making logic for personalized, stratified adjustments based on the patient's dynamic pain status [22-25]. This disconnect between self-assessment and exercise prescription renders them unable to adapt to the dynamically changing tissue irritability of patients across different stages, limiting the substantive clinical benefits of digital telerehabilitation. Crucially, neither traditional printed handouts nor reminder-only apps can dynamically adjust exercise intensity in response to a patient's fluctuating tissue irritability. Furthermore, while existing sensor-equipped apps provide basic range of motion tracking, they typically deliver static exercise prescriptions decoupled from the day-to-day pain variations that clinical guidelines use to grade rehabilitation intensity. Therefore, the present system was specifically designed to bridge this clinical gap by closing the "self-assessment-to-prescription" loop in an unsupervised home setting.

Accordingly, the objectives of this study were 3-fold. The primary aim was to develop a personalized, pain-stratified smartphone intervention system capable of dynamically tailoring exercise prescriptions to the patient's real-time pain status. The second aim was to evaluate the system's usability and feasibility as an add-on to usual care. Finally, the third aim was to explore its clinical usefulness and generate essential variance parameters to inform the design of future, adequately powered confirmatory trials.

Methods

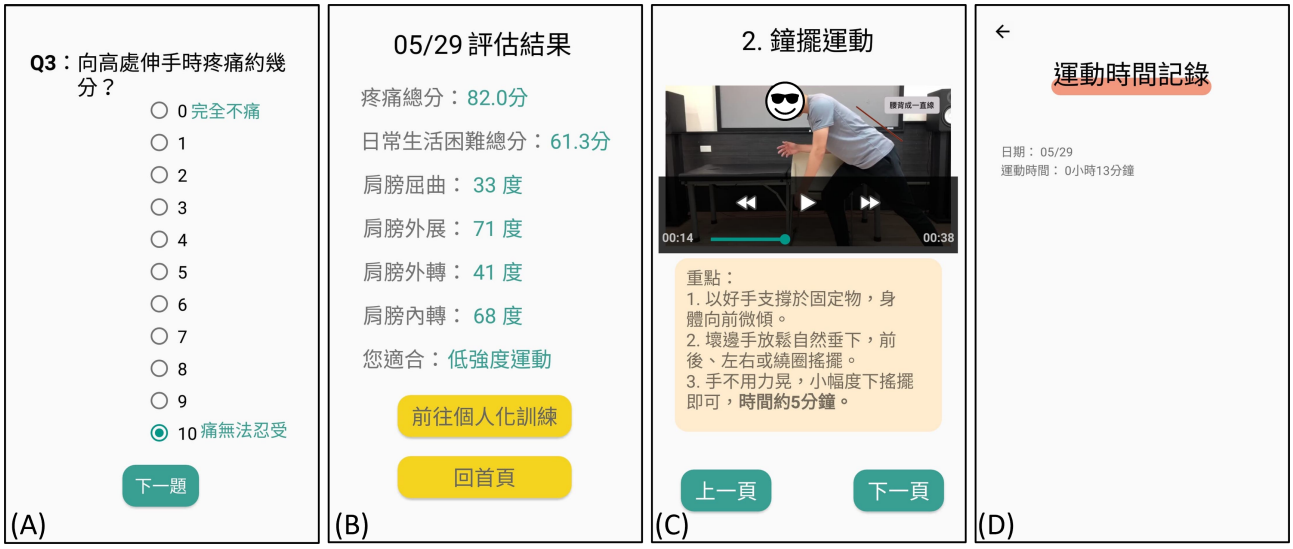
App Development

Development Environment and App Functions

This study selected Google's Android operating system as the platform for smartphone apps because Android is the most widely used smartphone operating system in Taiwan. The software development tool used in this study was Android Studio, with Java and XML as the programming languages. The app comprised 5 main functions: frozen shoulder educational content, self-assessment, personalized exercise, usage records, and an exercise reminder. Screenshots of the app are shown in Figure 1.

XSL•FO
RenderX

Figure 1. Screenshots of the Android frozen shoulder telerehabilitation app: (A) Shoulder Pain and Disability Index (SPADI)–based pain self-assessment, (B) daily self-assessment result with the system-recommended exercise intensity, (C) video-based exercise instruction, and (D) exercise-time record.



Frozen Shoulder Educational Content

According to recommendations from previous studies, enhancing the educational knowledge of patients with frozen shoulder can help them manage their symptoms and reduce fear [3,14]. By informing patients of the causes of frozen shoulder and symptoms that change over time, patients can gain advance understanding of the symptom trajectory, and further recommendations can be made regarding exercises appropriate for different stages, along with relevant precautions. The educational content on frozen shoulder incorporated in this app includes an introduction to frozen shoulder, disease staging, and treatment options, enabling patients to better understand their own condition through this educational content and to ease their fear of the symptoms.

Self-Assessment

The assessment items included pain, difficulty with activities of daily living, and range of motion. The personalized exercise program designed in this study used the pain level to determine the intensity of the personalized exercise subsequently recommended by the system.

Assessments of pain and difficulty in activities of daily living were adapted from the Shoulder Pain and Disability Index (SPADI), chosen for its established validity and responsiveness [26]. Administered as a digital questionnaire, it evaluated the affected shoulder's status over the past week through 2 sections: a 5-item pain subscale (rated from 0, "no pain at all," to 10, "unbearable pain") and an 8-item functional difficulty subscale (rated from 0, "no difficulty at all," to 10, "complete inability"). To ensure data completeness, the app required users to provide an answer for each item sequentially before proceeding to the next.

The range of motion assessment comprised 4 shoulder movements: shoulder flexion, shoulder abduction, shoulder external rotation, and shoulder internal rotation. Movement descriptions and demonstration videos were provided for users to view and review. The videos included precautions for each movement—such as the starting position during measurement and avoiding shoulder shrugging—prompting users to complete the assessment in the correct posture. The smartphone was secured to the user's wrist with an arm strap, allowing the sensor to detect the 3-axis rotation angles of the phone in accordance with the movement as a measure of shoulder range of motion.

After the user tapped Start, the first movement began. There was a 3-second preparation period, and upon hearing the start tone, the user began the movement and was required to reach the maximum extent of that movement (full range) within 5 seconds; the movement stopped upon hearing the stop tone, at which point the app displayed the change in range of motion on the screen in real time. Shoulder range of motion was calculated using the smartphone's built-in rotation vector sensor: a built-in function was used to convert the sensor's quaternion data into a rotation matrix, and a further function converted the rotation matrix into the smartphone's rotation angles about the *x*-, *y*-, and *z*-axes. Azimuth is the angle of rotation of the phone about the *z*-axis, pitch is the angle of rotation about the *x*-axis, and roll is the angle of rotation about the *y*-axis. In this study, the smartphone's rotation angle about the *z*-axis was used as the measurement value for shoulder flexion and abduction, and its rotation angle about the *x*-axis was used as the measurement value for shoulder external and internal rotation.

As for the dedicated app developed in this study, its reliability and validity relative to a conventional universal goniometer were evaluated by our research team in a separate technical validation study. Specifically, the smartphone app demonstrated excellent within-day intrarater reliability across all movements (intraclass correlation coefficients >0.95; SEM 2.5°-5.0°) and adequate between-day intrarater reliability (intraclass correlation coefficients ranging from 0.83 to 0.95). Furthermore, criterion-related validity showed a highly significant correlation with the universal goniometer (Pearson correlation coefficients >0.91) across all 4 shoulder motions. The concordance correlation coefficients showed high concordance for flexion (0.98), abduction (0.95), and internal rotation (0.92), with slightly lower concordance for external rotation (0.84), which is appropriately attributed to possible trunk compensatory movements. Given this established technical foundation, the

XSL•FO
RenderX

present trial focuses primarily on clinical feasibility and user experience following system intervention. The reliability and validity of these sensor-based measurements were evaluated in a separate study; the corresponding indices are provided in Multimedia Appendix 1.

After all 4 ranges of motion were measured, the final page presented the user's assessment results for that day, including the recommended exercise intensity. The assessment results were automatically saved within the usage records function.

Personalized Exercise Training

According to the 2013 APTA guidelines [3], physical therapists should determine appropriate exercise content on the basis of the patient's overall irritability level. However, considering that this system is a telerehabilitation home-based intervention without real-time supervision by health care personnel, it is difficult for the smartphone sensor to fully exclude trunk compensation caused by patient pain when measuring range of motion. In addition, verification of this system's sensor accuracy and consistency is currently under peer review (Y-L Chiu, PhD, PT, unpublished data, October 2026). To ensure absolute clinical safety of home-based rehabilitation and to prevent the system from prescribing an excessively high exercise intensity based on erroneous range of motion values measured under compensation, the system's personalization algorithm currently adopts a conservative strategy, using the patient's subjective pain level as the sole decision parameter for progression of prescription intensity. Pain was captured with the 5-item SPADI pain subscale, each item rated from 0 (no pain at all) to 10 (unbearable pain), and the mean of the 5 items—yielding a single 0 to 10 pain score—was used as the stratification indicator. On the basis of this mean pain score, the app classified the recommended exercise intensity into 3 levels: low intensity for severe pain (mean pain score ≥7), medium intensity for moderate pain (4≤ mean pain score ≤6), and high intensity for mild pain (mean pain score ≤3). For the clinical outcome analysis and on the app's results screen (Figure 1B), the same subscale is expressed as a 0 to 100 percentage; the 0 to 10 mean and the 0 to 100 percentage are equivalent, differing only by a factor of 10.

The exercise content was designed with reference to previous literature and included active-assisted exercises, stretching exercises, muscle strengthening exercises, and correct-posture reminders, with each intensity level containing different exercise content [3,13,27]. For the user interface design presenting the exercise content, reference was made to the presentation approach of previous similar apps; to prevent users from forgetting the exercise movements and to allow them to view the content at any time, the exercise content in this app was presented in video form with accompanying key reminders [21-23]. The detailed exercise list is provided in Multimedia Appendix 2.

Usage Records and Exercise Reminder

The record-keeping function was implemented using SQLite, the database built into Android Studio, and contained both assessment results and exercise time data. The app was designed with an alarm reminder to encourage users to exercise regularly each day. Using sound and vibration as the notification method, notifications popped up at 9:00 AM and 9:00 PM each day.

Intervention Trial

Study Design

This study used an exploratory, assessor-blinded, randomized controlled feasibility trial. To comprehensively evaluate the benefits of the digital health intervention, the study design was organized around 2 main axes: first, examining the system usability and feasibility of the smartphone app as a home-based rehabilitation tool, and second, using random allocation to divide participants into an app group and a handout group to undergo 4 weeks of home exercise intervention with identical content, in order to evaluate exercise adherence as a core feasibility metric, while assessing pain levels and shoulder range of motion strictly as exploratory secondary outcomes rather than to establish comparative clinical efficacy. Participant recruitment was conducted from March to June 2019, and the subsequent intervention and follow-up period extended to July 2019. Participants were recruited from rehabilitation clinics and communities in Beitou District, Taipei City, and were screened to meet the inclusion criteria.

Participants

The inclusion criteria were shoulder pain symptoms for at least 4 months, restricted shoulder range of motion (shoulder external rotation <45° or restricted movement in 2 or more planes), and the ability to operate a smartphone. Exclusion criteria included normal passive range of motion, glenohumeral osteoarthritis, neuropathic pain, and corticosteroid injection within the preceding 3 months [23]. A total of 10 participants (n=5 per group) completed the evaluation. While unpowered for clinical efficacy, this sample size aligns with methodological guidelines for designing and evaluating early-stage feasibility and usability trials [28,29]. A pragmatic cohort of 10 to 12 participants is generally considered sufficient to identify primary usability issues, evaluate user acceptability, and assess core feasibility parameters prior to a larger trial [30]. Consequently, the clinical component of this study is strictly intended for exploring clinical usefulness (exploratory) and generating variance estimates to inform an adequately powered future definitive trial, rather than to conduct inferential hypothesis testing or establish comparative efficacy.

Study Instruments

In the intervention trial, the app group installed the frozen shoulder app developed in this study; if a participant used a smartphone running a non-Android operating system, the researchers lent them a Zenfone 4 (ASUS) smartphone. In addition, the researchers provided participants in the app group with a smartphone arm strap to assist with self-assessment. Participants allocated to the handout group were given a copy of a frozen shoulder educational handout produced by the research team (Multimedia Appendices 3 and 4).

Study Procedure

After participants met the eligibility criteria and provided consent, the researchers first recorded their baseline demographic data and then allocated them in a 1:1 ratio to the

XSL•FO
RenderX

app group or the handout group. Allocation used sealed, opaque, and outwardly identical lots (5 per group) that had been prepared and shuffled in advance by a researcher not involved in outcome assessment; each participant drew 1 lot at enrollment, and the identical appearance of the lots concealed the allocation until assignment. After group allocation, a pretest was conducted, comprising pain level, level of difficulty in activities of daily living, and shoulder range of motion. To ensure blinding, the pretests and posttests of both groups were administered by a physical therapist who was blinded to group allocation, covering the questionnaires and shoulder range of motion, with range of motion measured using a conventional goniometer. For the app group, in addition to the conventional measurement method, another researcher assisted in installing the app on the participant's smartphone on the same day and, based on the pretest assessment results, assisted in confirming the training program recommended by the system for the first week's exercise intensity, which was automatically generated according to the assessment results. The researchers instructed participants in the app group on how to use the app and taught those in the handout group the educational exercises.

The intervention lasted 4 weeks in total, and each participant was required to perform exercise training at least once daily; in addition, the app group was required to perform self-assessment at least once per week in order to continuously update their personalized training program. Each week, the researchers contacted participants via a messaging app to determine whether they were experiencing any difficulty operating the smartphone app or performing the exercises. In the handout group, the frequency and duration of that week's exercise were recorded using 2 questions: "How many times did you perform the exercises this week?" and "About how long did each exercise session last?" Participants in the app group were instead asked to provide screenshots of the app's built-in training logs. If a participant had questions, the researchers scheduled a time to discuss the issues in person. After 4 weeks, a posttest was conducted, with the same items and procedure as the pretest.

Intervention Content and Usual Care Co-Interventions

Both groups performed a home exercise program with identical content. The exercise intensity was divided into 3 levels (low, medium, and high); in the app group, the system automatically assigned the training difficulty according to the assessment results, whereas in the handout group, the difficulty was set by the researchers. All 3 levels included correct-posture reminders. Low-intensity training was to be performed at least once daily for at least 15 minutes per session, whereas the other 2 intensity levels were to be performed at least once daily for at least 20 minutes per session.

Importantly, during the 4-week intervention, participants in both groups continued to receive their routine medical care for frozen shoulder under the National Health Insurance system. This usual care included physical therapy modalities and pharmacological treatments. However, the specific types, dosages, and frequencies of these co-interventions were not quantitatively recorded or controlled for as covariates in this study.

Ethical Considerations

The experimental protocol of this study was approved by the Institutional Review Board of National Yang-Ming University (approval number YM107116F). Written informed consent was obtained from all participants prior to enrollment. To maintain privacy and data confidentiality, all participant information was deidentified. Participants did not receive any compensation. This trial was not registered in a public registry. This is acknowledged as a limitation (see the *Limitations* section). No changes were made to the trial design, eligibility criteria, or study procedures after the trial commenced.

Outcome Measures

Primary Outcome

The primary outcomes focused on trial feasibility. Indicators included participant retention rate and intervention safety (adverse event incidence). To collect exercise adherence data in a realistic clinical setting, researchers contacted both groups weekly. The handout group provided interview-based self-reports of their exercise frequency, whereas the app group was asked to provide screenshots of the app's built-in training logs, rather than verbally report their exercise frequency. Because the 2 methods differ fundamentally, the resulting adherence values are not comparable between groups. Adherence was therefore summarized separately for each group as a descriptive feasibility metric, with the app group's logs illustrating the app's capacity for objective home monitoring. Each participant's exercise adherence was tabulated after 4 weeks, calculated as follows:

Exercise adherence (%) = (total number of exercise days) / 28 days × 100

To evaluate the usability of the frozen shoulder app in this study, the System Usability Scale (SUS) was used, with a maximum total of 100 points, where higher scores indicate greater satisfaction with system usability [31]. In addition to the SUS scores, each participant in the app group provided brief open-ended feedback on their experience of using the app at the end of the intervention. These responses were collected and summarized to identify common views regarding the user experience.

Secondary Outcomes

Pain and difficulty in activities of daily living were assessed using the SPADI, which comprises 13 self-rated items—5 for the pain subscale and 8 for the activities-of-daily-living difficulty subscale. Each item was rated from 0 to 10, with 0 indicating no pain at all or no difficulty at all and 10 indicating unbearable pain or extreme difficulty. The pain and activities-of-daily-living difficulty subscales were scored separately, and each subscale score was converted to a percentage ranging from 0 to 100, calculated as follows:

Pain subscale score (%) = (sum of the 5 pain item scores) / 50 × 100

Activities-of-daily-living difficulty subscale score (%) = (sum of the 8 difficulty item scores) / 80 × 100

XSL•FO
RenderX

Active shoulder range of motion was assessed with a goniometer for 4 shoulder movements: shoulder flexion, shoulder abduction, shoulder external rotation with the arm at the side, and shoulder internal rotation. The goniometric measurement methods followed those published by Norkin and White [32] in 2016. Shoulder flexion and shoulder abduction were measured in the standing position, with the fulcrum at the acromion, measuring the angle between the vertical midline of the body and the midline of the upper arm; shoulder external rotation was measured in the standing position, with the starting position being the upper arm against the body and the elbow flexed at 90°, the fulcrum at the midpoint of the olecranon process of the ulna, measuring the change in angle of the midline of the ulna; shoulder internal rotation was measured in the supine position, with the fulcrum at the olecranon process, measuring the angle between the midline of the ulna and the vertical to the ground. Each movement was measured 3 times at full range, and the mean of the 3 measurements was calculated after recording. All outcome measures and their assessment procedures were defined before enrollment began, and no changes were made to the outcomes or their measurement methods after the trial commenced.

Statistical Analysis

Consistent with the aims of a usability and feasibility study, feasibility indicators—recruitment, retention, adherence, system usability, and safety—were defined as the primary outcome domain and are summarized with descriptive statistics. Given the sample size (n=5 per group), clinical outcomes (pain, function, and range of motion) were treated as exploratory and are presented at the individual-participant level, with group means and SDs reported for context; no inferential hypothesis testing was performed on the clinical outcomes, and no *P* values are reported for these comparisons. Given the small sample size and in line with CONSORT (Consolidated Standards of Reporting Trials) recommendations, baseline characteristics are presented descriptively without significance testing.

Results

Participant Demographic Characteristics

Of the 20 individuals assessed for eligibility (recruited from rehabilitation outpatient clinics and a community activity center in the Beitou District of Taipei), 10 did not proceed to enrollment—because they did not meet the inclusion criteria (n=4) or declined to participate owing to time constraints (n=6)—and the remaining 10 patients with frozen shoulder were enrolled and randomly allocated to the app group (n=5) or the handout group (n=5); the participant flow is summarized in Figure 2 (CONSORT flow diagram). Group allocation was concealed until assignment: each participant drew 1 sealed, opaque, and identical lot at enrollment, yielding a 1:1 allocation (5 per group). All 10 participants completed the 4-week intervention, with no withdrawals (retention 100%). Baseline demographic characteristics, presented descriptively in Table 1, were broadly comparable between groups in age, sex, affected side, and symptom duration. One baseline difference is, however, relevant to the interpretation of the outcomes reported below: shoulder external rotation was considerably lower in the app group than in the handout group at entry (mean 20.6° vs 35.8°), leaving the app group with more room for improvement in that plane (Table 2).

XSL•FO
RenderX

Figure 2. Participant flow (CONSORT [Consolidated Standards of Reporting Trials] diagram) for the assessor-blinded comparison of a smartphone-based home exercise system with an educational handout in adhesive capsulitis: screening, exclusions, 1:1 allocation, and completion (n=10).

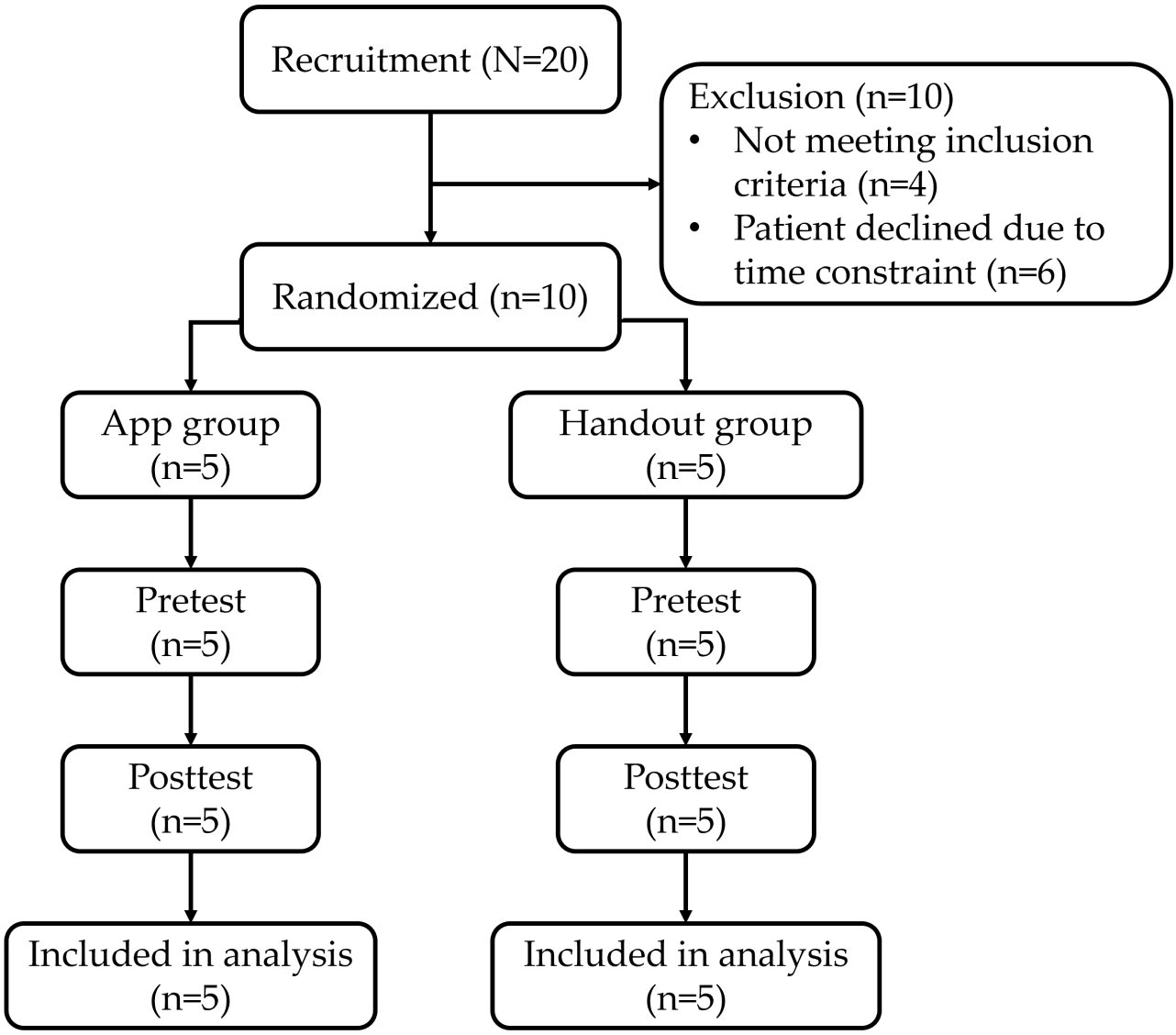


Table . Baseline demographic and clinical characteristics of the 10 participants with adhesive capsulitis.

Characteristic	App group (n=5)	Handout group (n=5)
Age (y), mean (SD)	49.4 (5.0)	51.0 (9.4)
Female, n (%)	4 (80)	5 (100)
Affected side—right, n (%)	1 (20)	1 (20)
Symptom duration (mo), mean (SD)	7.67 (3.8)	7.40 (2.6)

XSL•FO
RenderX

Table . Clinical outcomes by group: pre-post values and mean change.

Outcome and group	Pre, mean (SD)	Post, mean (SD)	Change, mean
SPADI[a] pain (%)			
App	44.4 (19.9)	30.0 (26.5)	−14.4
Handout	50.4 (13.6)	32.0 (28.0)	−18.4
SPADI function (%)			
App	40.5 (16.6)	26.2 (17.1)	−14.3
Handout	43.5 (19.8)	23.3 (21.9)	−20.2
Flexion (°)			
App	136.2 (28.7)	144.1 (20.8)	+7.9
Handout	142.3 (28.5)	153.3 (19.9)	+11.0
Abduction (°)			
App	121.8 (38.6)	128.1 (29.1)	+6.3
Handout	122.4 (35.6)	139.5 (28.6)	17.1
External rotation (°)			
App	20.6 (5.6)	30.5 (6.2)	+9.9
Handout	35.8 (21.8)	38.1 (19.6)	+2.3
Internal rotation (°)			
App	38.5 (10.7)	43.4 (7.2)	+4.9
Handout	47.5 (13.5)	51.5 (14.8)	+4.0

[a]SPADI: Shoulder Pain and Disability Index.

Feasibility and Exercise Adherence

Feasibility, the primary outcome domain, was supported across all indicators. Retention was 100% (10/10) with no withdrawals, and no adverse events or exercise-induced pain exacerbations related to the home exercise program were reported in either group during the 4-week intervention. Over the 28-day intervention, exercise adherence was captured objectively through the app's usage logs in the app group and through weekly interview-based self-reports in the handout group. Because these instruments differ fundamentally in how they capture behavior, the adherence values are reported descriptively as feasibility metrics and are not compared between groups. In the app group, the usage logs documented a mean of 18.0 (SD 5.0) exercise days (18/28, 64.3% adherence), illustrating the app's capacity for continuous, objective home monitoring. The handout group's self-reported values are shown in Table 3 for completeness only. The individual participants' adherence is shown in Figure 3.

Table . Exercise adherence over the 4-week home intervention, reported separately for each group by data source[a].

Group: data source	Adherence (d), mean (SD)	Completion rate (%)	Median (IQR)	Range	% meeting ≥3×/week
App group (n=5): app usage log[b]	18.0 (5.0)	64.3	21 (16-21)	10-22	80
Handout group (n=5): weekly self-report[c]	11.8 (3.9)	42.1	11 (9-13)	8-18	40

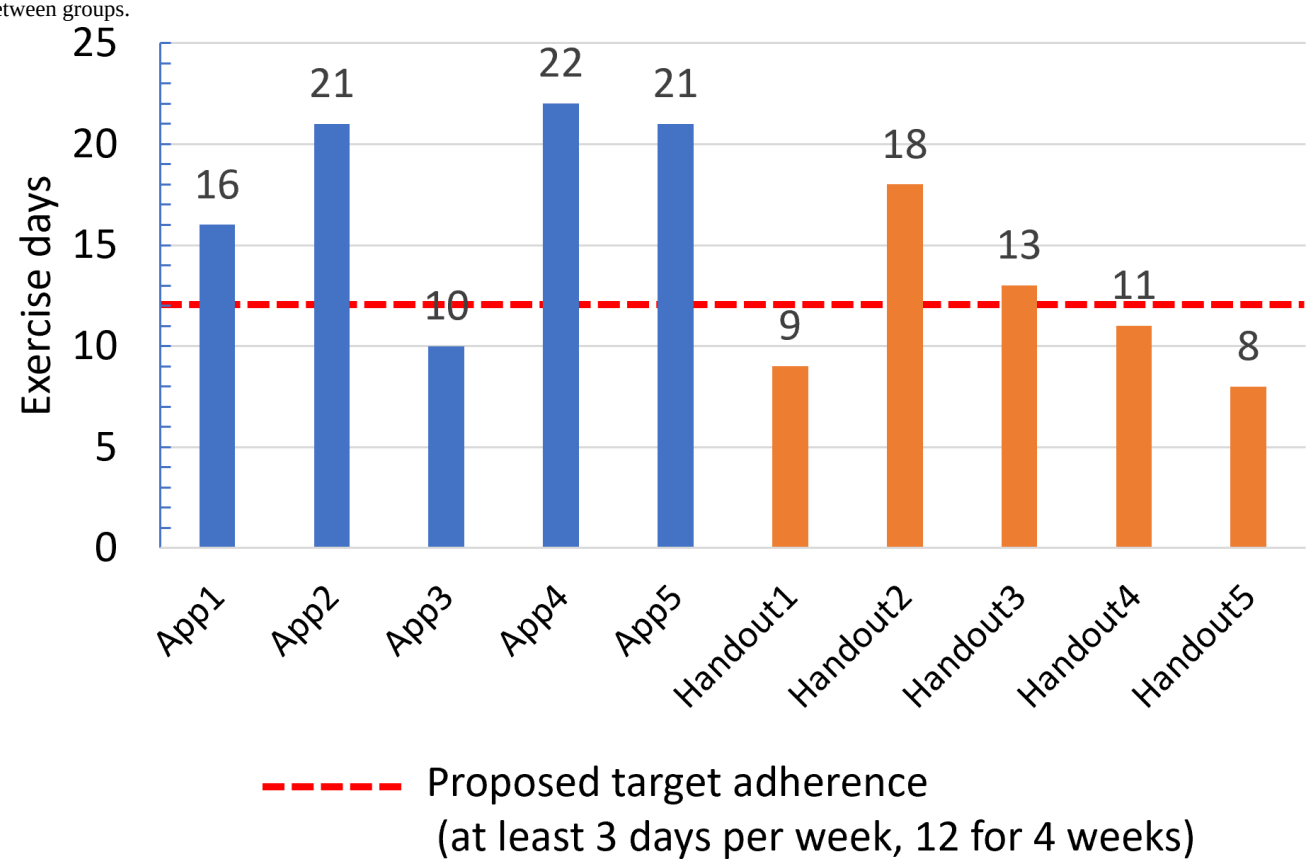
[a]Adherence was measured with different instruments in the 2 groups; values are presented descriptively and are not directly comparable between groups.

[b]Recorded in the app's built-in usage log and submitted weekly as screenshots.

[c]Reported by participants in weekly interviews (recall-based self-report).

XSL•FO
RenderX

Figure 3. Individual participant exercise adherence over a 4-week home-based intervention for frozen shoulder, both app group and handout group. The red horizontal dashed line at 12 days indicates the proposed clinically reasonable target of completing home exercises at least 3 days per week. Adherence was derived from app usage logs in the app group and from weekly self-report in the handout group; values are not directly comparable between groups.



System Usability

Overview

System usability was evaluated in the app group (n=5). The mean SUS score was 87.5 (SD 5.86), well above the established benchmark of 68 and corresponding to an "excellent" rating on the adjective rating scale of Bangor et al [31]. Individual scores ranged from 77.5 to 92.5; per-participant scores and open-ended feedback are presented in Table 4. Furthermore, qualitative review of the weekly follow-up communications revealed a complete absence of technical or operational troubleshooting requests. Specifically, participants reported no difficulties across the core practical domains of notifications, exercise performance, self-assessment execution, or daily usage recording. This lack of operational friction indicates that the initial face-to-face setup and tutorial effectively preempted potential technical barriers, further substantiating the system's high practical feasibility and intuitive interface in a home setting. Recurring themes in the qualitative feedback were that the video demonstrations helped participants remember the exercises, the reminder function supported a regular daily routine, and the usage records helped participants track their progress; 1 participant requested a more customizable reminder time.

Table . System Usability Scale (SUS) scores and open-ended feedback from the 5 app–arm participants who completed the evaluation.

Participants	SUS score[a]	User feedback
1	87.5	The video format is excellent; I do not forget the exercise content.
2	77.5	The alarm reminder function could be made more customizable.
3	90.0	I found this approach very helpful for rehabilitation.
4	92.5	It helps me perform the rehabilitation exercises consistently every day.
5	90.0	It is helpful for tracking records, and I do not forget the movements.

[a]Mean 87.5 (SD 5.86).

XSL•FO
RenderX

Clinical Outcomes (Exploratory)

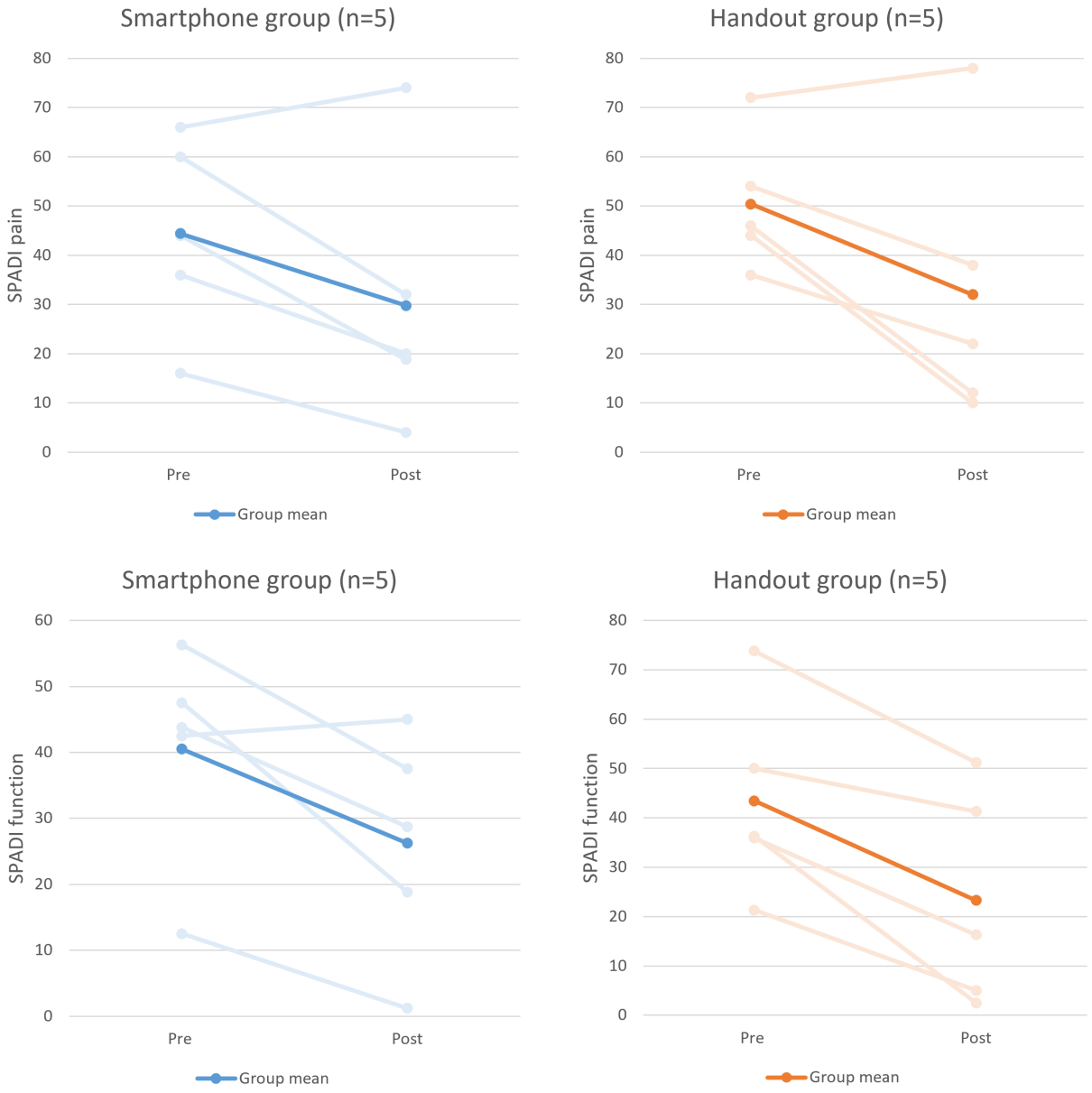
Both groups improved on all clinical outcomes over the 4-week intervention. Because this was a small sample trial of 5 participants per group, clinical outcomes are presented descriptively at the individual level (Figures 4 and 5), with group means summarized in Table 2. Consistent with the exploratory framing, no inferential testing was performed, and no *P* values are reported. The magnitude and direction of the between-group differences—the contrast most relevant to whether the app outperformed the handout—are summarized in Table 2; across outcomes, the handout group showed numerically larger improvements in most measures, and no consistent advantage of the app was evident.

XSL•FO
RenderX

Figure 4. Individual and group mean shoulder range of motion trajectories (flexion, abduction, and external and internal rotation) from preintervention to postintervention over 4 weeks, by arm (app, n=5; handout, n=5); exploratory, descriptive.

XSL•FO
RenderX

Figure 5. Individual and group mean Shoulder Pain and Disability Index (SPADI) pain and function trajectories from preintervention to postintervention over 4 weeks, by arm (app, n=5; handout, n=5); exploratory, descriptive.



Pain and Function (SPADI)

Both SPADI pain and function decreased in the 2 groups, with numerically larger improvements in the handout group (Table 2); no descriptive advantage for the app was evident.

Shoulder Range of Motion

Shoulder flexion, abduction, and internal rotation increased in both groups, with the handout group again showing numerically larger or comparable gains and no consistent between-group advantage (Table 2; individual trajectories in Figure 4). For shoulder external rotation, all 5 app-group participants improved (Figure 4), and the mean gain was larger in the app group (+9.9° vs +2.3°). This apparent difference must, however, be interpreted with caution: the app group started from a substantially lower baseline (20.6° vs 35.8°), and despite the larger gain, it remained lower in absolute range of motion after the intervention (30.5° vs 38.1°). The individual trajectories in Figure 4 illustrate this pattern, and no clear advantage of the app over the handout is evident once baseline is taken into account.

Discussion

Principal Findings

This study aimed to develop and evaluate the usability, feasibility, and exploratory clinical outcomes of a personalized smartphone app for patients with frozen shoulder. The core contribution of this work lies in system development: we successfully engineered an integrated tool that unifies sensor-based self-assessment, pain-driven exercise personalization, and objective usage logging, overcoming the limitations of prior static, one-size-fits-all interventions. The trial confirmed that this system is safe and highly usable in an unsupervised home setting. Furthermore, the app's usage logs provided continuous, objective documentation of home exercise, a monitoring capability that printed handouts cannot offer. We

XSL•FO
RenderX

also observed a descriptive trend of improvement in shoulder external rotation. Consistent with the feasibility design, this outcome is best interpreted as a hypothesis-generating signal that warrants further investigation. Ultimately, this study establishes a robust technological foundation and practical feasibility to justify a future adequately powered definitive trial.

System Usability and Feasibility

A key strength of this trial was the confirmation that the smartphone app possesses high clinical feasibility and system usability in the population of patients with frozen shoulder. Notably, the trial achieved a 100% retention rate with no dropouts during the 4-week intervention. Furthermore, a qualitative review of the weekly follow-up communications revealed a complete absence of technical or operational troubleshooting requests; no participants reported difficulties with system manipulation, notifications, or daily usage, indicating a highly intuitive interface.

Quantitative usability assessments further supported these observations. The SUS score of the app group reached 87.5 (SD 5.86), well above the established average SUS benchmark of 68 and corresponding to an “excellent” rating on the adjective rating scale reported by Bangor et al [31]. It should be neutrally noted, however, that these SUS evaluations were provided exclusively by the app-arm participants who completed the trial. While this reflects high satisfaction among active users, it inherently excludes the perspectives of those who initially declined participation, and thus represents the usability among completers rather than universal acceptance.

This high usability rating is highly consistent with the pilot study by Stütz et al [22], which likewise developed a dedicated home-based rehabilitation app for frozen shoulder and reported an SUS score of 88. By comparison, through more comprehensive cross-validation of quantitative and qualitative feedback, the present study further confirmed the practical advantages of the system. Qualitative participant feedback indicated that the intuitive video-based exercise demonstrations effectively overcame a key limitation of traditional printed educational handouts, with which movement details are easily forgotten. Users reported that the videos helped them review movements at any time and that the built-in usage records successfully tracked their daily rehabilitation status. Regarding the exercise reminder function, participants affirmed that it helped establish a regular rehabilitation habit; additionally, specific feedback suggested that allowing users to personally customize the alarm times in future iterations could further optimize the system’s flexibility for busy schedules. Together, these findings fully confirm that this integrated system can effectively fulfill the supervisory function of home-based telerehabilitation.

Exercise Adherence

Exercise adherence was a primary feasibility outcome. Adherence was captured with different instruments in the 2 groups: through the app’s built-in objective usage logs in the app group and through weekly interview-based self-reports in the handout group. Because these instruments are fundamentally not equivalent, adherence values cannot be compared between groups and are interpreted only as descriptive feasibility metrics. Over the 28-day program, the app group’s logs documented a mean of 18.0 (SD 5.0) exercise days (18/28, 64.3% adherence). Interpreted on its own terms, the app group’s data represent a substantial level of unsupervised home engagement and demonstrate an objective self-monitoring capability that printed handouts structurally cannot provide.

When attempting to contextualize the app group’s logged adherence rate (18/28, 64.3%) in the literature, it is evident that a directly comparable reference point is currently lacking. For instance, while Chen et al [25] reported high adherence rates (96.6% and 85.2%), their trial used daily text messages and a substantially different overall treatment approach. Furthermore, their adherence data were collected via self-report over a 2-week interval, whereas our app group relied on continuous, automated daily system logs. Given the fundamental differences in both intervention modalities and measurement timeframes, a direct numerical comparison of their findings with ours is methodologically unsuitable. Other related trials have relied on subjective numerical rating scales without calculating specific adherence percentages, precluding direct numerical comparisons [18,23]. Conversely, although Stütz et al [22] used an objective, app-based tracking mechanism similar to ours, their study was a small, single-arm evaluation (n=5). Consequently, while the app group’s logged adherence appears numerically lower than rates reported in recall-based studies, there is currently no literature comparable to our methodology that can serve as a direct, objective benchmark.

Nevertheless, this level of adherence highlights a critical clinical reality: there is a gap between theoretical effectiveness and real-world adherence. To achieve better outcomes, patients need to engage in the intervention actively, yet many factors influence adherence. For instance, patients with a 9-to-5 job may lack the time and energy to perform rehabilitation after work, and requiring daily exercise may feel pressured and overwhelming. Therefore, providing patients with a specific, practical adherence target—such as completing exercises at least 3 days per week (approximately 42% compliance)—may be more useful and realistic. Setting these targets in a progressive (graded) manner in future digital interventions may better sustain long-term patient motivation without imposing undue psychological pressure.

Furthermore, this asymmetry in measurement highlights an important methodological lesson for future research. To accurately evaluate whether a digital app improves adherence and to examine adherence as a mediator linking system use to clinical change, future definitive trials must apply a single, objective adherence measure across all arms—such as electronically time-stamped exercise logging for all participants.

Exploratory Clinical Outcomes and System Refinement Insights

After 4 weeks of home training, both groups demonstrated descriptive improvements in pain reduction (SPADI) and shoulder range of motion (flexion, abduction, and internal rotation). For these parameters, the handout group showed numerically larger or comparable mean gains, and the descriptive data did not indicate a consistent between-group

XSL•FO
RenderX

advantage for the app. As a small-sample feasibility evaluation, these descriptive trajectories align with previous short-term interventions and confirm the clinical safety of deploying the app as an add-on mHealth tool. Crucially, the absence of a clear descriptive between-group difference in overall pain and general range of motion may have been confounded by concurrent medical co-interventions. Because participants continued to receive routine care under the national health insurance system (including physical therapy modalities and pharmacological treatments), the substantial efficacy of these routine treatments may have produced a ceiling effect in the short term, obscuring the marginal benefit of the home exercise. This observation reinforces the intended positioning of the system: it serves as a reinforcing digital aid for objective home monitoring, rather than a replacement for routine medical care.

A specific descriptive pattern of practical interest emerged in shoulder external rotation, a clinically challenging movement often severely restricted and pain-provoking in frozen shoulder. All 5 app-group participants showed a consistent within-group improvement, with a numerically larger mean gain than the handout group (+9.9° vs +2.3°). While the app operationalizes a pain-stratified adjustment mechanism designed to safely guide patients through these challenging movements, this observed difference must be interpreted with strict caution. As noted previously, the app group began with a substantially lower baseline in external rotation (20.6° vs 35.8°), providing a larger window for improvement and rendering the data highly susceptible to regression to the mean. Therefore, rather than being interpreted as evidence of clinical superiority, this external-rotation pattern serves strictly as a hypothesis-generating signal to be tested in future adequately powered trials using stratified randomization to control for baseline severity.

Finally, the exploratory clinical data provide direct functional insights for future system refinement. For instance, the app group showed a more modest descriptive gain in shoulder abduction compared to the handout group (+6.3° vs +17.1°). Because the app's current personalized algorithm adopts a highly conservative intensity to prioritize home safety, it did not emphasize reaching the extreme end range of motion during abduction exercises. This outcome highlights a clear technical direction for the next iteration of the system: refining the dynamic exercise algorithm to safely and progressively guide patients toward greater end-range loading once their tissue irritability decreases.

Limitations

This study has several limitations. First, this study was a usability and feasibility trial, and its small overall enrollment (n=10) inherently limits the precision and generalizability of the clinical estimates. The clinical outcomes are strictly exploratory and hypothesis-generating; adequately powered randomized controlled trials are required to establish clinical efficacy. Second, the fixed-time push notifications (9 AM and 9 PM) lacked the flexibility to integrate more deeply with patients' daily routines, potentially limiting their behavior-change potential. Third, concurrent usual care under the National Health Insurance system, including physical therapy modalities and pharmacological treatments, was not quantitatively tracked. Neither the frequency and type of physical therapy sessions nor the types and doses of medications were recorded in either group, confounding the evaluation of the smartphone intervention's isolated clinical contribution. Future trial protocols should use a standardized weekly cointervention log for all arms (physical therapy sessions and modalities, medications and doses, and injections), ideally verified against National Health Insurance records. Fourth, the 4-week study period precludes the assessment of sustained adherence and long-term benefits, which are critical given the 1- to 2-year natural course of frozen shoulder. Fifth, prioritizing patient safety in an unsupervised setting, the current personalization algorithm stratifies intensity based solely on subjective pain. Future iterations should comprehensively integrate objective range of motion data once compensatory-movement filtering algorithms mature. Finally, this trial was not registered in a public trial registry, either prospectively or retrospectively, which reduces transparency and limits compliance with international trial reporting standards. Any future efficacy trial based on this prototype will be prospectively registered in a public trial registry.

Conclusions

In summary, this study successfully developed and demonstrated the feasibility and usability of a telerehabilitation system for frozen shoulder that integrates several key functions, including digital education, personalized irritability-based self-assessment and stratification, objective monitoring via an SQLite database, and scheduled exercise reminders. Its core value lies in moving beyond the traditional one-size-fits-all home rehabilitation model and putting home-based precision rehabilitation into practice. The results of this study confirmed that when used as an add-on mHealth intervention to usual care, the system possesses excellent usability (SUS score of 87.5), can effectively overcome the drawback of traditional printed education—with which movements are easily forgotten—and enables objective tracking of day-to-day home exercise through its usage log. Over the 4-week home intervention, the usage logs documented home exercise on a mean of 18.0 (SD 5.0) of 28 (64.3%) days, and all app-group participants showed a within-group improvement in the most clinically challenging plane of shoulder external rotation. Because this was a small-sample trial, no inferential between-group testing was undertaken. The descriptive data, shaped by the small sample, the baseline imbalance in external rotation, and unmonitored usual care (physical therapy modalities and pharmacological treatments), should not be considered evidence of clinical superiority; nevertheless, the system demonstrated clear value in usability, feasibility, and objective home monitoring as an assistive tool for home-based telerehabilitation. Future work should expand the sample size, incorporate an intelligent personalized dynamic reminder mechanism, and combine long-term follow-up to establish its long-term evidence-based standing in the field of digital health.

XSL•FO
RenderX

Generative AI Use

Generative AI (OpenEvidence, Claude, and Google Gemini) was used to assist with the literature search and to improve the language and readability of portions of the manuscript. The authors reviewed and edited all AI-assisted text and take full responsibility for the content.

Funding

The authors declared no financial support was received for this work.

Data Availability

The study protocol and deidentified data are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Reliability and validity of range of motion assessment tool.
[PDF File, 65 KB - rehab_v13i1e107314_app1.pdf]

Multimedia Appendix 2
Training list.
[PDF File, 87 KB - rehab_v13i1e107314_app2.pdf]

Multimedia Appendix 3
Educational handout English version.
[PDF File, 87 KB - rehab_v13i1e107314_app3.pdf]

Multimedia Appendix 4
Educational handout Chinese version.
[PDF File, 569 KB - rehab_v13i1e107314_app4.pdf]

Checklist 1
CONSORT extension pilot and feasibility trials checklist.
[PDF File, 265 KB - rehab_v13i1e107314_app5.pdf]

Checklist 2
CONSORT-EHEALTH V1.6 (2011) checklist.
[PDF File, 1107 KB - rehab_v13i1e107314_app6.pdf]

References

1. Le HV, Lee SJ, Nazarian A, Rodriguez EK. Adhesive capsulitis of the shoulder: review of pathophysiology and current clinical treatments. Shoulder Elbow 2017 Apr;9(2):75-84. [doi: 10.1177/1758573216676786] [Medline: 28405218]
2. Uppal HS, Evans JP, Smith C. Frozen shoulder: a systematic review of therapeutic options. World J Orthop 2015 Mar 18;6(2):263-268. [doi: 10.5312/wjo.v6.i2.263] [Medline: 25793166]
3. Kelley MJ, Shaffer MA, Kuhn JE, et al. Shoulder pain and mobility deficits: adhesive capsulitis. J Orthop Sports Phys Ther 2013 May;43(5):A1-A31. [doi: 10.2519/jospt.2013.0302] [Medline: 23636125]
4. Milgrom C, Novack V, Weil Y, Jaber S, Radeva-Petrova DR, Finestone A. Risk factors for idiopathic frozen shoulder. Isr Med Assoc J 2008 May;10(5):361-364. [Medline: 18605360]
5. Wong CK, Levine WN, Deo K, et al. Natural history of frozen shoulder: fact or fiction? A systematic review. Physiotherapy 2017 Mar;103(1):40-47. [doi: 10.1016/j.physio.2016.05.009] [Medline: 27641499]
6. Cyriax JH. Textbook of Orthopaedic Medicine: Diagnosis of Soft Tissue Lesions, 5th edition: Baillière, Tindall & Cassell; 1969.
7. Hannafin JA, Chiaia TA. Adhesive capsulitis. A treatment approach. Clin Orthop Relat Res 2000 Mar(372):95-109. [doi: 10.1097/00003086-200003000-00012] [Medline: 10738419]
8. Neviaser AS, Hannafin JA. Adhesive capsulitis: a review of current treatment. Am J Sports Med 2010 Nov;38(11):2346-2356. [doi: 10.1177/0363546509348048] [Medline: 20110457]

XSL•FO
RenderX

9. Ding H, Tang Y, Xue Y, et al. A report on the prevalence of depression and anxiety in patients with frozen shoulder and their relations to disease status. Psychol Health Med 2014;19(6):730-737. [doi: 10.1080/13548506.2013.873814] [Medline: 24382210]
10. Cohen C, Ejnisman B. Epidemiology of frozen shoulder. In: Itoi E, Arce G, Bain GI, Diercks RL, Guttmann D, Imhoff AB, et al, editors. Shoulder Stiffness: Current Concepts and Concerns: Springer; 2015:21-30. [doi: 10.1007/978-3-662-46370-3_2]
11. Wang W, Shi M, Zhou C, et al. Effectiveness of corticosteroid injections in adhesive capsulitis of shoulder: a meta-analysis. Medicine (Baltimore) 2017 Jul;96(28):e7529. [doi: 10.1097/MD.0000000000007529] [Medline: 28700506]
12. Celik D. Comparison of the outcomes of two different exercise programs on frozen shoulder. Acta Orthop Traumatol Turc 2010;44(4):285-292. [doi: 10.3944/AOTT.2010.2367] [Medline: 21252605]
13. Diercks RL, Stevens M. Gentle thawing of the frozen shoulder: a prospective study of supervised neglect versus intensive physical therapy in seventy-seven patients with frozen shoulder syndrome followed up for two years. J Shoulder Elbow Surg 2004;13(5):499-502. [doi: 10.1016/j.jse.2004.03.002] [Medline: 15383804]
14. Mertens MG, Meert L, Struyf F, Schwank A, Meeus M. Exercise therapy is effective for improvement in range of motion, function, and pain in patients with frozen shoulder: a systematic review and meta-analysis. Arch Phys Med Rehabil 2022 May;103(5):998-1012. [doi: 10.1016/j.apmr.2021.07.806] [Medline: 34425089]
15. Jack K, McLean SM, Moffett JK, Gardiner E. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228. [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
16. Sluijs EM, Kok GJ, van der Zee J. Correlates of exercise compliance in physical therapy. Phys Ther 1993 Nov;73(11):771-782. [doi: 10.1093/ptj/73.11.771] [Medline: 8234458]
17. Campbell R, Evans M, Tucker M, Quilty B, Dieppe P, Donovan JL. Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee. J Epidemiol Community Health 2001 Feb;55(2):132-138. [doi: 10.1136/jech.55.2.132] [Medline: 11154253]
18. Lambert TE, Harvey LA, Avdalis C, et al. An app with remote support achieves better adherence to home exercise programs than paper handouts in people with musculoskeletal conditions: a randomised trial. J Physiother 2017 Jul;63(3):161-167. [doi: 10.1016/j.jphys.2017.05.015] [Medline: 28662834]
19. Thompson D, Rattu S, Tower J, Egerton T, Francis J, Merolli M. Mobile app use to support therapeutic exercise for musculoskeletal pain conditions may help improve pain intensity and self-reported physical function: a systematic review. J Physiother 2023 Jan;69(1):23-34. [doi: 10.1016/j.jphys.2022.11.012] [Medline: 36528508]
20. Davergne T, Meidinger P, Dechartres A, Gossec L. The effectiveness of digital apps providing personalized exercise videos: systematic review with meta-analysis. J Med Internet Res 2023 Jul 13;25:e45207. [doi: 10.2196/45207] [Medline: 37440300]
21. Chiensriwimol N, Chan JH, Mongkolnam P, Mekhora K. Monitoring frozen shoulder exercises to support clinical decision on treatment process using smartphone. Procedia Comput Sci 2017;111:129-136. [doi: 10.1016/j.procs.2017.06.019]
22. Stütz T, Emsenhuber G, Huber D, et al. Mobile phone-supported physiotherapy for frozen shoulder: feasibility assessment based on a usability study. JMIR Rehabil Assist Technol 2017 Jul 20;4(2):e6. [doi: 10.2196/rehab.7085] [Medline: 28729234]
23. Choi Y, Nam J, Yang D, Jung W, Lee HR, Kim SH. Effect of smartphone application-supported self-rehabilitation for frozen shoulder: a prospective randomized control study. Clin Rehabil 2019 Apr;33(4):653-660. [doi: 10.1177/0269215518818866] [Medline: 30526016]
24. Lin YJ, Chung CY, Chen CPC, Hsieh YW, Wang CF, Chen CC. The feasibility and efficacy of remote app-guided home exercises for frozen shoulder: a pilot study. Healthcare (Basel) 2024 May 27;12(11):1095. [doi: 10.3390/healthcare12111095] [Medline: 38891171]
25. Chen HC, Chuang TY, Lin PC, Lin YK, Chuang YH. Effects of messages delivered by mobile phone on increasing compliance with shoulder exercises among patients with a frozen shoulder. J Nurs Scholarsh 2017 Jul;49(4):429-437. [doi: 10.1111/jnu.12308] [Medline: 28632975]
26. Staples MP, Forbes A, Green S, Buchbinder R. Shoulder-specific disability measures showed acceptable construct validity and responsiveness. J Clin Epidemiol 2010 Feb;63(2):163-170. [doi: 10.1016/j.jclinepi.2009.03.023] [Medline: 19683414]
27. Russell S, Jariwala A, Conlon R, Selfe J, Richards J, Walton M. A blinded, randomized, controlled trial assessing conservative management strategies for frozen shoulder. J Shoulder Elbow Surg 2014 Apr;23(4):500-507. [doi: 10.1016/j.jse.2013.12.026] [Medline: 24630545]
28. Lyndon H, Latour JM, Marsden J, Campbell S, Stevens K, Kent B. The holistic assessment and care planning in partnership intervention study (HAPPI): a protocol for a feasibility, cluster randomized controlled trial. J Adv Nurs 2019 Nov;75(11):3078-3087. [doi: 10.1111/jan.14106] [Medline: 31222778]
29. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 1;60(1):95-103. [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
30. Hwang W, Salvendy G. Number of people required for usability evaluation: the 10±2 rule. Commun ACM 2010 May;53(5):130-133. [doi: 10.1145/1735223.1735255]
31. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
32. Norkin CC, White DJ. Measurement of Joint Motion: A Guide to Goniometry, 5th edition: F.A. Davis Company; 2016.

XSL•FO
RenderX

Abbreviations

APTA: American Physical Therapy Association
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
mHealth: mobile health
SPADI: Shoulder Pain and Disability Index
SUS: System Usability Scale

Edited by J Sarvestan; submitted 17.Jul.2026; peer-reviewed by CC Chen, JY Lim; revised version received 06.Oct.2026; accepted 12.Oct.2026; published 05.Nov.2026.

Please cite as:
Hsu HH, Cheng KC, Chiu YL
Development, Usability, and Feasibility of a Pain-Stratified Smartphone-Based Home Intervention System for Patients With Adhesive Capsulitis: Exploratory Feasibility and Usability Trial
JMIR Rehabil Assist Technol 2026;13:e107314
URL: https://rehab.jmir.org/2026/1/e107314
doi:10.2196/107314

© Hsun-Han Hsu, Kai-Chia Cheng, Ya-Lan Chiu. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 5.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparing the Feasibility of S-Patch vs 24-Hour Holter in Frail Older People With Falls: Randomized Open-Label Crossover Pilot Study

Barbara Rosario[1], MBChB, MPhil; Vern Hsen Tan[2], MBBS; Vivian Cantiller Barrera[1], MD; Hnin Tun Mon[3], PhD; Clara Yi-En Seah[4], MBBS; Colin Yeo[2], MBBS

[1]Department of Geriatric Medicine, Changi General Hospital, 2 Simei Street 3, Singapore, Singapore

[2]Department of Cardiology, Changi General Hospital, Singapore, Singapore

[3]Health Services Research, Changi General Hospital, Singapore, Singapore

[4]Department of Emergency Medicine, Changi General Hospital, Singapore, Singapore

Corresponding Author:
Barbara Rosario, MBChB, MPhil
Department of Geriatric Medicine, Changi General Hospital, 2 Simei Street 3, Singapore, Singapore

Abstract

Background: Falls commonly cause injuries in older adults and can present with syncope or presyncope. Typically, frail older adults with cognitive impairment are underrepresented in research and experience difficulty tolerating conventional ECG monitoring devices. S-Patch is a novel, lightweight dual-lead device designed to provide continuous ECG monitoring for up to 72 hours with minimal interference to daily activities.

Objective: This study aimed primarily to evaluate the feasibility, tolerability, and adherence of the S-Patch compared with standard 24-hour Holter monitoring in older adults (aged ≥60 years) presenting with unwitnessed falls, syncope, or near-syncope, as assessed by wear duration, completion rates, and patient-reported experience. The secondary objective of this study was to assess device usability and compare arrhythmia detection rates between the devices in a real-world setting.

Methods: This prospective, single-center, randomized, open-label, crossover feasibility pilot study was conducted from March 8, 2020, to February 28, 2023. Participants were randomized 1:1 into group A (Holter followed by S-Patch) or group B (S-Patch followed by Holter), enabling up to 4 days of continuous monitoring.

Results: Seventy patients were enrolled, of whom 32 (45.7%) had cognitive impairment, and the majority were frail. Group A participants were older (86.9 vs 82.9 years) with higher use of dementia medications (6/35, 17.1% vs 3/35, 8.6%). Frailty was highly prevalent across the cohort, and fear of falling was significantly greater among frail patients (P=.04). Both devices were well tolerated, with no difference in overall device preference, but S-Patch enabled substantially longer monitoring durations compared to Holter. Mean S-Patch wear time exceeded 60 hours and was significantly longer when the device was applied first (68.8 vs 62.9 h; P=.04), suggesting acceptance and adherence. Importantly, tolerability was maintained in frail patients with cognitive impairment. Arrhythmia detection rates were observed as comparable, with both devices identifying one clinically significant arrhythmia, resulting in pacemaker implantation.

Conclusions: S-Patch is a feasible, well-tolerated cardiac monitoring device in frail older adults and those with cognitive impairment. Compared with Holter, S-Patch supported longer continuous monitoring without compromising patient acceptability, an important consideration in real-world clinical practice.

Trial Registration: ClinicalTrials.gov NCT06977789; https://clinicaltrials.gov/ct2/show/NCT06977789

(JMIR Rehabil Assist Technol 2026;13:e81761) doi:10.2196/81761

KEYWORDS

falls; frailty; geriatrics; arrhythmia; syncope; monitoring

Introduction

Falls are the most common cause of injury in older people, with 10% of falls resulting in serious injuries and injury-related disability [1-3]. One-third of community dwellers over 65 years of age will fall each year, and half of those will fall again [4]. Falls-related injuries are the fifth leading cause of death in older people [5] and one of the most frequent reasons for admission to long-term care institutions in previously independent older people [6].

Long-term monitoring studies suggest that arrhythmias are present in a substantial proportion of patients with unexplained

XSL•FO
RenderX

falls, with a subset attributable to treatable rhythm disturbances [2]. However, establishing this relationship is often challenging in clinical practice, particularly in unwitnessed events or when cognitive impairment limits history taking.

Syncope may present as a fall in the absence of a witness account, particularly when coupled with amnesia [2]. First-time syncope is associated with an 80% increased 1-year risk of fall-related injuries, such as fractures and traumatic head injuries, compared with an age- and sex-matched group [7]. Age and multiple comorbidities, including cardiac arrhythmias, are independent risk factors for falls and collectively increase the likelihood of unexplained falls [8,9].

In a large cohort of emergency department (ED) trauma cases (N=5420), 3.3% of falls were attributed to syncope [10], and independent predictors of cardiac syncope were advanced age (>65 y), coronary artery disease, and the presence of pathologic Q waves [10]. Syncope and falls clinic reports identified cardiac causes (rhythm or conduction disorders) in 44% of patients [11], and after identification of an underlying cause, permanent pacemaker (PPM) implantation was associated with a reduced falls risk [12]. Although multifactorial interventions for falls have been less effective in preventing falls-related injuries [13], falls related to arrhythmias remain potentially modifiable factors.

Current guidelines recommend cardiovascular assessment, including ambulatory ECG monitoring, in the evaluation of falls and syncope [14-17]. However, a major limitation in older populations is the tolerability and practicality of monitoring devices. Conventional Holter monitors are relatively bulky, require multiple leads, and impose restrictions on daily activities such as bathing, which may reduce adherence, especially in frail individuals with cognitive impairment. While implantable loop recorders provide prolonged monitoring, their invasive nature limits acceptability in many patients.

As a result, there remains an unmet need for noninvasive, well-tolerated monitoring solutions that can be used effectively in real-world older populations. The S Patch3-Cardio/S-Patch Ex (S-Patch) is a lightweight, novel continuous heart rhythm recording device with a single wire. The S-Patch has been compared with Holter and telemetry [18], and this study aims to compare the tolerability, feasibility, and adherence of the S-Patch with the standard Holter device in an older population. We postulate that it will be as tolerable and efficacious as the Holter device. To our knowledge, this is the first study comparing the S-Patch with the Holter device in an older population in Singapore.

Methods

Hypothesis

The hypothesis was that S-Patch is at least as preferred as Holter.

Primary Objective

The primary objective of this study was to evaluate tolerability and adherence of the S-Patch compared with Holter, assessed by wear duration, completion rates, and patient-reported experience.

Secondary Objective

The secondary objective of this study was to assess device usability and to compare arrhythmia detection rates between devices in a real-world situation.

Recruitment, Consent, and Randomization

This prospective, single-center, randomized, open-label, crossover pilot study was performed between March 8, 2020, and February 28, 2023, at a tertiary hospital in Singapore. Eligible patients included older adults aged 60 years and older, attending the ED for an unwitnessed fall, syncope, or presyncope, and referred for Holter testing. Exclusion criteria included patients with known significant bradyarrhythmia, PPM, or inability to tolerate Holter monitoring.

Eligible patients were randomized into 2 groups using a block randomization method with a 1:1 ratio and were randomly allocated to either group A (Holter for 24 h followed sequentially by S-Patch for 72 h) or group B (S-Patch for 72 h followed sequentially by Holter for 24 h), with both monitors worn for up to 4 days. An events diary and a questionnaire were used to evaluate their preference between Holter and the S-Patch recorder. The visit schedule is shown in Multimedia Appendix 1, the user preference questionnaire in Multimedia Appendix 2, and the event diary in Multimedia Appendix 3.

Device Information, Data Recording, and Analysis

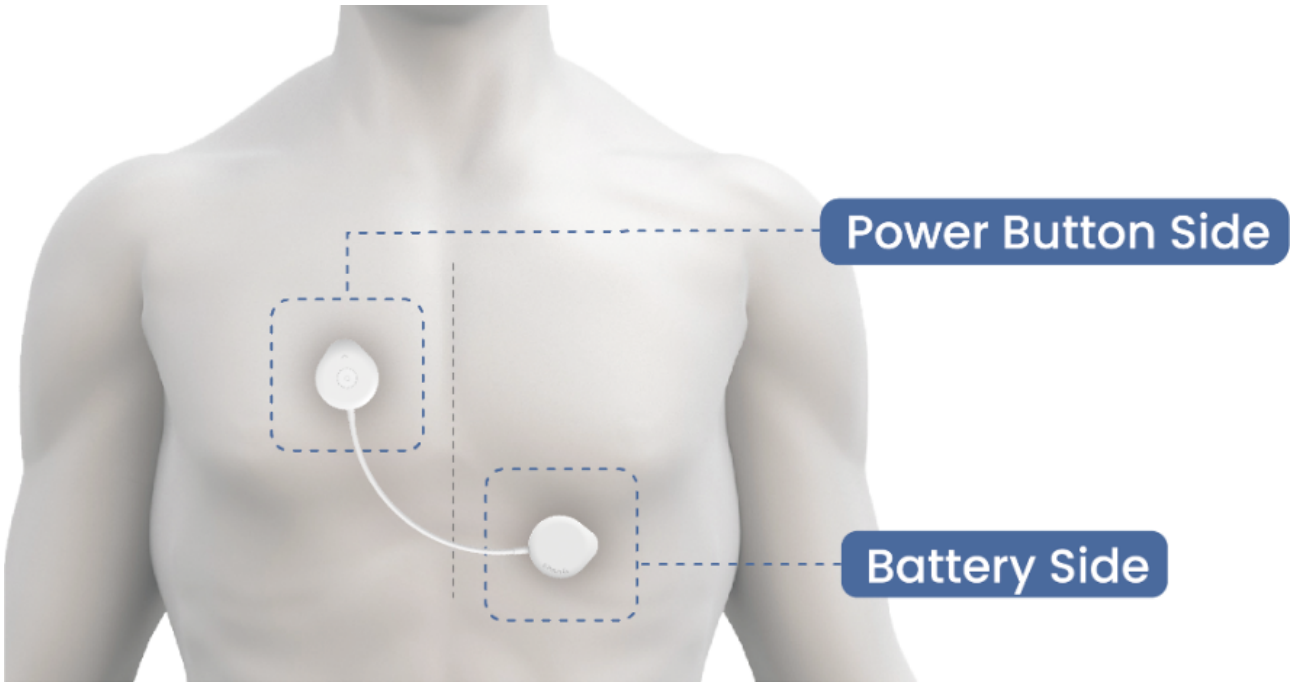
The S-Patch is a lightweight device containing a bioprocessor that continuously collects ECG data (for up to 72 h) and transmits them in real time via a Bluetooth connection to an Android smartphone app. The initial coupling requires an internet connection that pairs the device with a precreated patient profile. The device is easily applied with only 2 standard ECG electrodes placed typically in the V2-V3 position (Figure 1). The S-Patch sends data to the phone (to be kept within 3 m) via Bluetooth continuously and does not require an internet connection. The app allows the patient to be able to add symptoms or activities either by using a precreated list or by entering their own information. At the end of the test period, the recorded information is uploaded to a cloud-based database via an internet connection and is analyzed by an algorithm, after which a report is generated and reviewed by a consultant cardiologist according to the following clinical criteria for the definition of a major abnormality:

1. Ventricular tachycardia of 3 or more consecutive premature ventricular beats
2. Sinus or ventricular pauses of 3 or more seconds
3. Bradycardia with a heart rate of less than 30 beats per minute
4. Mobitz type II atrioventricular block
5. Complete heart block
6. The definition or clinical criteria for minor abnormalities include the following:
    a. Multiple ventricular ectopic beats at a rate of 100 or more premature beats per hour
    b. Paroxysmal supraventricular tachycardia for 100 or more beats per hour
    c. Bradycardia with a heart rate of 30 to 39 beats per minute

XSL•FO
RenderX

d. Mobitz type I atrioventricular block

Figure 1. S-Patch electrode placement. Permission was obtained for reproduction of the S-Patch3-Cardio/S-Patch Ex placement image in this manuscript.



Results from the Holter were reviewed and analyzed as per current clinical practice. The data comparison recorded is shown in Multimedia Appendix 4.

Data Collection, Power Calculation, and Statistical Analysis

Data collection included demographic data on age, sex, race, BMI, and comorbidities, including an existing or new diagnosis of dementia or mild cognitive impairment. Cognitive screening was undertaken using the Abbreviated Mental Test (AMT), with a score of ≥7 considered normal, and functional status was assessed using the Barthel score [19] and categorized into independent (20), mild (16-19), moderate (11-15), and severe disability (≤10). The Short Physical Performance Battery (SPPB) assessment [20] was conducted, with scores ranging from 0 (worst performance) to 12 points (best performance). Falls and fear of falling were ascertained by screening questions, “Do you have fear of falling?” and “How many falls have you sustained in the last year?” Frailty was assessed using the Clinical Frailty Scale (CFS) [21], a widely used judgment-based frailty screening tool and nationally agreed community frailty screening tool [22]. Outcomes derived from hospitalization data included length of stay (LOS); ED reattendance at 7 days, 30 days, or 6 months; emergency hospital readmission at 30 days or 6 months; and inpatient mortality.

Power calculation determined that 116 patients were required to achieve 80% power to detect a difference in abnormality rates of 20%, assuming a significance level of 5%. Assuming a 20% dropout rate, the study aimed to recruit 140 patients.

Participant characteristics at baseline and secondary outcomes were examined according to their statistical distributions. Parametric continuous data were compared using 2-tailed *t* tests and are presented as mean (SD), while nonparametric data were analyzed with Mann-Whitney *U* tests and are presented as median (IQR). Categorical data comparing participant characteristics and assigned groups were evaluated using chi-square tests or Fisher exact tests and are presented as proportions. The significance level was set at α=.05, and all tests were 2-tailed. Statistical analysis was performed using Stata statistical software (version 16; StataCorp).

Ethical Considerations

The study protocol was approved by the SingHealth Centralized Institutional Review Board (CIRB; ethics approval number 2019/2969) and Clinical Trial Registration 2019/2960. Informed consent was obtained from all participants, and if patients lacked capacity to make decisions, informed consent was obtained from their legally appointed representatives.

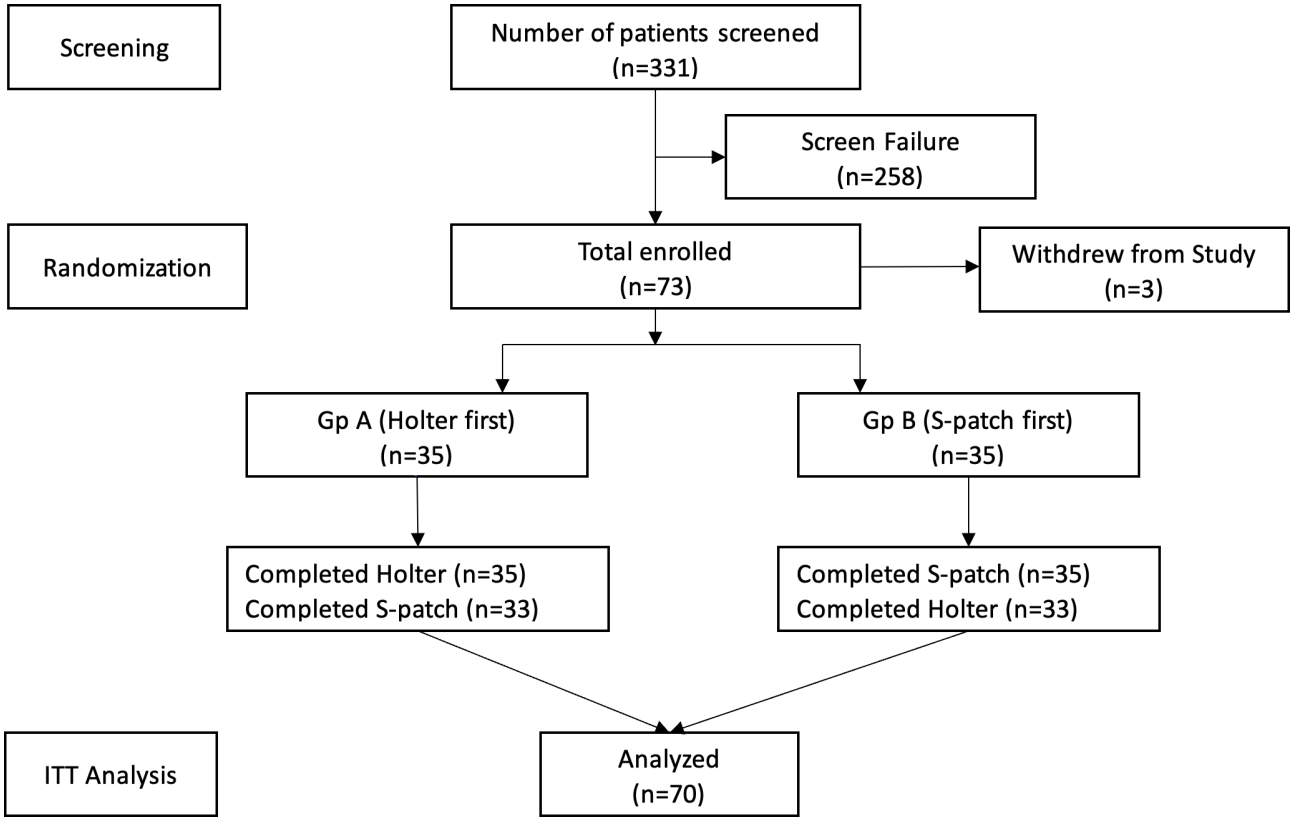
Results

Screening and Recruitment

The study recruited 73 patients, 3 of whom withdrew from the study, and 70 patients completed the study interventions. Group A received Holter first, followed immediately by S-Patch, and Group B received S-Patch first, followed immediately by Holter, providing up to 4 days of continuous monitoring. Two patients received only Holter (one declined S-Patch and the other had no recording of S-Patch as data were not transferred from the Android smartphone app). Two patients refused Holter monitoring and only had S-Patch monitoring. There were 35 patients in both groups, of whom 33 in both groups completed sequential Holter and S-Patch monitoring or vice versa. The majority of patients completed cardiac monitoring during their inpatient stay. Screening and recruitment are shown in Figure 2.

XSL•FO
RenderX

Figure 2. figure Screening and recruitment. Gp A: group A; Gp B: group B; ITT: intention-to-treat.



Demographics

Group A participants were older than Group B participants (mean age 86.9, SD 5.8 y vs 82.9, SD 7.1 y), with a similar proportion of females in both groups. There were more Chinese patients (30/35, 85.7% vs 22/35, 62.9%) and more Indian patients (4/35, 11.4% vs 2/35, 5.7%) in Group A and more Malay patients in Group B (0% vs 8/35, 22.9%; Table 1). BMI was similar in both groups, as was the prevalence of hypertension (29/35, 82.9% vs 30/35, 85.7%) and diabetes mellitus (17/35, 48.6% vs 19/35, 54.3%), but Group B had more patients with atrial fibrillation (AF; 5/35, 14.3% vs 10/35, 28.6%) and stroke (8/35, 22.9% vs 14/35, 40%), although there were no statistical differences in comorbidities between the 2 groups. Few patients were active smokers across the groups (2/35, 5.7% vs 2/35, 5.7%). Only a small proportion of patients (1/35, 2.9% vs 2/35, 5.7%) had left ventricular ejection fraction below 40%. There were no differences in the use of antiarrhythmic or dementia medications (6/35, 17.1% vs 3/35, 8.6%; Table 1).

XSL•FO
RenderX

Table . Demographics, comorbidities, cognitive impairment, frailty and functional assesssments, device tolerance, duration and arrhythmia detection, and hospitalization usage.

Characteristic	Group A (Holter first), n=35	Group B (S-Patch first), n=35	*P* value
Gender, n (%)			>.99
Female	22 (62.9)	21 (60)	
Male	13 (37.1)	14 (40)	
Race, n (%)			.005
Chinese	30 (85.7)	22 (62.9)	
Malay	0 (0)	8 (22.9)	
Indian	4 (11.4)	2 (5.7)	
Others	1 (2.9)	3 (8.6)	
Age (y), mean (SD)	86.9 (5.8)	82.9 (7.1)	.01
BMI (34 vs 34), mean (SD)	22.1 (4.0)	22.9 (4.5)	.41
Comorbidities, n (%)			
Diabetes mellitus	17 (48.6)	19 (54.3)	.81
Hypertension	29 (82.9)	30 (85.7)	>.99
CCF[a]	2 (5.7)	0 (0)	.49
Stroke/TIA[b]	8 (22.9)	14 (40)	.20
AF[c]	5 (14.3)	10 (28.6)	.24
IHD[d]	9 (25.7)	11 (31.4)	.79
COPD[e]	0 (0)	1 (2.9)	>.99
Parkinson disease	4 (11.4)	4 (11.4)	>.99
Smoker, n (%)	2 (5.7)	2 (5.7)	>.99
LVEF[f] (%; 30 vs 28), mean (SD)	57.8 (7.3)	56.8 (7.3)	.58
LVEF <40, n (%)			>.99
No	34 (97.1)	33 (94.3)	
Yes	1 (2.9)	2 (5.7)	
Medications, n (%)			
None	8 (22.9)	15 (42.9)	.13
Beta blocker	13 (37.1)	15 (42.9)	.81
Calcium channel blocker	16 (45.7)	10 (28.6)	.22
Digoxin	1 (2.9)	1 (2.9)	>.99
Dementia (donepezil and memantine)	6 (17.1)	3 (8.6)	.48
Cognition, n (%)			
Dementia			.47
No	17 (48.6)	21 (60)	
Yes	18 (51.4)	14 (40)	
Abbreviated Mental Test (AMT; 28 vs 29), mean (SD)	6.5 (3.0)	6.5 (3.3)	.98
Rockwood Clinical Frailty Score (CFS), mean (SD)	4.8 (1.3)	4.5 (1.5)	.31
CFS groups, n (%)			.81
Not frail/robust (1-3)	6 (17.1)	7 (20)	

XSL•FO
RenderX

Characteristic	Group A (Holter first), n=35	Group B (S-Patch first), n=35	*P* value
Mildly frail (4-5)	17 (48.6)	19 (54.3)	
Moderate/severe frail (6-9)	12 (34.3)	9 (25.7)	
Barthel Index (35 vs 35), mean (SD)	15.8 (4.6)	16.5 (4.4)	.54
Barthel Index subgroups, n (%)			.37
Independent (20)	9 (25.7)	14 (40)	
Mild disability (16-19)	13 (37.1)	10 (28.6)	
Moderate disability (11-15)	6 (17.1)	8 (22.9)	
Severe disability (≤10)	7 (20)	3 (8.6)	
Short Physical Performance Battery Test (SPPB; 24 vs 25), mean (SD)	3.7 (2.8)	3.9 (3.0)	.84
Fear of falling, n (%)			.21
No	18 (60)	14 (42.4)	
Yes	12 (40)	19 (57.6)	
Number of falls in the past 12 mo, mean (SD)	2.2 (1.9)	2.0 (1.9)	.67
Duration of device (h), mean (SD)			
Holter (35 vs 33)	22.6 (1.3)	22.5 (1.1)	.52
S-Patch (33 vs 35)	62.9 (13.9)	68.8 (9.5)	.04
S-Patch analysis time (33 vs 35; h)	53.1 (13.4)	58.9 (13.6)	.08
Symptoms, n (%)			.72
Unwitnessed fall	21 (60)	17 (48.6)	
Syncope	5 (14.3)	6 (17.1)	
Near syncope	0 (0)	0 (0)	
NA	9 (25.7)	12 (34.3)	
Arrhythmia events by device[g], n (%)			
Holter (Yes)	10 (28.6)	15 (42.9)	.32
S-Patch (Yes)	9 (25.7)	9 (25.7)	>.99
Holter: sinus node disease, n (%)			>.99
No	31 (88.6)	29 (87.9)	
Yes	4 (11.4)	4 (12.1)	
Holter: AV[h] node disease			>.99
No	32 (91.4)	31 (93.9)	
Yes	3 (8.6)	2 (6.1)	
Holter: Tachyarrhythmia, n (%)			.13
None	31 (88.6)	23 (69.7)	
SVT[i]	0 (0)	1 (3)	
AF/Flutter	4 (11.4)	8 (24.3)	
VT[j]/VF[k]	0 (0)	1 (3)	
S-Patch: sinus node disease, n (%)			—[l]
No	33 (100)	35 (100)	
Yes	0 (0)	0 (0)	

XSL•FO
RenderX

Characteristic	Group A (Holter first), n=35	Group B (S-Patch first), n=35	*P* value
S-Patch: AV node disease, n (%)			.67
No	30 (90.9)	33 (94.3)	
Yes	3 (9.1)	2 (5.7)	
S-Patch: Tachyarrhythmia, n (%)			.41
None	27 (81.8)	28 (80)	
SVT	3 (9.1)	1 (2.9)	
AF/Flutter	3 (9.1)	6 (17.1)	
VT/VF	0 (0)	0 (0)	
Preference (Holter vs S-Patch), n (%)			.36
Holter	3 (8.8)	7 (21.2)	
S-Patch	22 (64.7)	17 (51.5)	
No preference	9 (26.5)	9 (27.3)	
None of the device	0 (0)	0 (0)	
Prefer Holter, n (%)			
User friendly and ease of operation	2 (5.7)	3 (8.6)	>.99
Comfortable	1 (2.9)	0 (0)	>.99
Does not affect daily activities	1 (2.9)	3 (8.6)	.61
Others, please specify	2 (5.7)	2 (5.7)	>.99
Prefer S-Patch, n (%)			
User friendly and ease of operation	13 (37.1)	9 (25.7)	.44
Comfortable	16 (45.7)	12 (34.3)	.46
Does not affect daily activities	11 (31.4)	7 (20)	.41
Others, please specify	1 (2.9)	4 (11.4)	.36
No preference, n (%)			
User friendly and ease of operation	3 (8.6)	2 (5.7)	>.99
Comfortable	4 (11.4)	8 (22.9)	.34
Does not affect daily activities	4 (11.4)	4 (11.4)	>.99
Others, please specify	1 (2.9)	0 (0)	>.99
Neither device, n (%)			—
User friendly and ease of operation	0 (0)	0 (0)	
Comfortable	0 (0)	0 (0)	
Does not affect daily activities	0 (0)	0 (0)	
Others, please specify	0 (0)	0 (0)	
Holter satisfaction, n (%)			>.99
Extremely satisfied	0 (0)	0 (0)	
Very satisfied	2 (66.7)	4 (57.1)	
Moderately satisfied	1 (33.3)	3 (42.9)	
Not so satisfied	0 (0)	0 (0)	
Dissatisfied	0 (0)	0 (0)	

XSL•FO
RenderX

Characteristic	Group A (Holter first), n=35	Group B (S-Patch first), n=35	*P* value
S-Patch satisfaction, n (%)			.07
Extremely satisfied	6 (27.3)	1 (5.9)	
Very satisfied	14 (63.6)	10 (58.9)	
Moderately satisfied	2 (9.1)	6 (35.3)	
Not so satisfied	0 (0)	0 (0)	
Dissatisfied	0 (0)	0 (0)	
Hospitalization usage			
Hospital admission at index visit, n (%)			>.99
No	2 (5.7)	3 (8.6)	
Yes	33 (94.3)	32 (91.4)	
Total number of hospitalization days (34 vs 35), mean (SD)	0.68 (0.91)	0.77 (1.48)	.75
Length of stay (days; 33 vs 32), mean (SD)	17.9 (16.2)	16.8 (20.8)	.80
ED[m] reattendance (6 mo), n (%)			
Yes	13 (37.1)	7 (20)	.19
Hospital readmission (6 mo), n (%)			
Yes	12 (34.3)	4 (11.4)	.04
Mortality, n (%)	0 (0)	1 (2.9)	>.99

[a]CCF: congestive cardiac failure.

[b]TIA: transient ischemic attack.

[c]AF: atrial fibrillation.

[d]IHD: ischemic heart disease.

[e]COPD: chronic obstructive pulmonary disease.

[f]LVEF: left ventricular ejection fraction.

[g]Arrhythmia events relate to events detected in individual patients but patients may have more than one reported arrhythmia event.

[h]AV: atrioventricular.

[i]SVT: supraventricular tachycardia.

[j]VT: ventricular tachycardia.

[k]VF: ventricular fibrillation.

[l]Not applicable.

[m]ED: emergency department.

Tolerability of S-Patch vs Holter

Holter device was tolerated well in both groups and was worn for similar periods in group A vs group B (mean 22.6, SD 1.3 h vs 22.5, SD 1.1 h; *P*=.52; Table 1). S-Patch was also tolerated well in both groups; however, the device was worn for longer in group B, where S-Patch was administered first (mean 62.9, SD 13.4 h vs 68.8, SD 9.5 h; *P*=.04; Table 1). The overall S-Patch analysis time was longer in group B (mean 53.1, SD 13.4 h vs 58.9, SD 13.6 h; *P*=.08) but did not reach statistical significance. Although there was no difference in the patient's device preference between groups, a higher proportion of participants in group A (3 vs 22) vs group B (7 vs 17) preferred S-Patch, with 9 participants in both groups having no preference (Table 1). Compliance with completion of preference questionnaires was high in both groups (group A, 34/35, 97% and group B, 33/35, 94%; Table 1).

Cognition and Dementia

Overall, 45.7% (32/70) of patients had dementia or cognitive impairment, either known or newly diagnosed during the index admission, with similar numbers of patients with cognitive impairment across both groups (group A: 18/35, 51.4% vs group B: 14/35, 40%; Table 1). Only a small proportion of patients (9/70, 12.9%) were taking anticholinesterase inhibitors (ACHi; group A: 6/35, 17.1% vs group B: 3/35, 8.6%). None of the patients on ACHi had symptomatic bradyarrhythmia. Average AMT score was 6.5 for both groups (Table 1). In terms of the duration of monitoring undertaken in those with or without cognitive impairment, both devices were tolerated well, and Holter was worn for 22.5 (SD 1.4) hours vs 22.6 (SD 1.1 hours; *P*=.66) and S-Patch for 67.5 (SD 9.5) hours vs 64.6 (SD 11.3 hours; *P*=.33; Table 2).

XSL•FO
RenderX

Table . Cognitive impairment and dementia status (N=70).

Characteristic	No known dementia, n=38 (54.3%)	Yes dementia, n=32 (45.7%)	*P* value
Barthel Index group, n (%)			.001
Independent (20)	17 (44.7)	6 (18.8)	
Mild disability (16-19)	16 (42.2)	7 (21.9)	
Moderate disability (11-15)	4 (10.5)	10 (31.3)	
Severe disability (≤10)	1 (2.6)	9 (28.0)	
CFS[a] groups, n (%)			.001
Not frail/robust (1-3)	13 (34.2)	0 (0)	
Mildly frail (4-5)	23 (60.5)	13 (40.6)	
Moderate/severe frail (6-9)	2 (5.3)	19 (59.4)	
AMT[b], n (%)			<.001
Score (≥7)	26 (68.4)	7 (21.9)	
Score (0 - 6)	5 (13.2)	19 (59.4)	
Not done	7 (18.4)	6 (18.8)	
Duration of device (h), mean (SD)			
Holter	22.6 (1.1)	22.5 (1.4)	.66
S-Patch	64.6 (13.8)	67.5 (9.5)	.33
S-Patch analysis time (h)	57.8 (14.6)	54.0 (12.4)	.26
Fear of falling, n (%)			.13
No	21 (60)	11 (39.3)	
Yes	14 (40)	17 (60.7)	
Unable to answer	0 (0)	0 (0)	
Number of falls in the past 12 mo, mean (SD)	1.9 (1.8)	2.4 (2.0)	.28
ED attendance, n (%)			—
No	0 (0)	0 (0)	
Yes	37 (100)	32 (100)	
Total number of hospitalization days, mean (SD)	0.49 (0.9)	1.0 (1.5)	.08
Length of stay (d), mean (SD)	13.2 (9.4)	21.7 (24.0)	.07
Mortality, n (%)	0 (0)	1 (3.1)	.46
Hospital readmit (6 mo), n (%)			
Yes	6 (15.8)	10 (31.3)	.16
ED readmit (6 mo), n (%)			
Yes	9 (23.7)	11 (34.4)	.43

[a]CFS: Clinical Frailty Scale.

[b]AMT: Abbreviated Mental Test.

Frailty, Falls, and Functional Status

Frailty assessed using CFS was similar between group A and group B, with 34.3% (12/35) vs 25.7% (9/35) of participants being moderately or severely frail, 48.6% (17/35) vs 54.3% (19/35) being very mildly or mildly frail, and 17.1% (6/35) vs 20% (7/35) not being frail. There were more patients with cognitive impairment in the moderate or severely frail groups (19/35, 59.4% vs 0/35, 0%) compared to nonfrail (*P*=.001; Table 2). There was a trend toward more falls with increasing frailty (5/13, 38.5% in CFS 1 - 3 vs 19/36, 52.8% in CFS 4 and 5 vs 14/21, 66.7% in CFS 6 - 9), which did not reach statistical significance (Table 3). However, fear of falling was significantly higher in frail patients (4/13, 30.8% in CFS 1 - 3 vs 14/32, 43.8% in CFS 4 and 5 vs 13/18, 72.2% in CFS 6 - 9; *P*=.04; Table 3) but did not translate to increased hospital usage.

XSL•FO
RenderX

Table . Barthel status and the impact of function on tolerance of the device, fear of falls, and hospital usage.

Characteristics	Independent (20), n=23 (32.9%)	Mild disability (16-19), n=23 (32.9%)	Moderate disability (11-15), n=14 (24.3%)	Severe disability (≤10), n=10 (14.3%)	*P* value
CFS[a] groups, n (%)					.001
Not frail/robust (1-3)	10 (43.5)	3 (13)	0 (0)	0 (0)	
Mildly frail (4-5)	13 (56.5)	19 (82.6)	4 (28.6)	0 (0)	
Moderate/severe frail (6-9)	0 (0)	1 (4.4)	10 (71.4)	10 (100)	
AMT[b] groups, n (%)					.05
Normal (≥7)	12 (52.2)	14 (60.9)	6 (42.9)	1 (10)	
Cognitive impairment (0 - 6)	5 (21.7)	8 (34.8)	5 (35.7)	6 (60)	
Not done	6 (26.1)	1 (4.4)	3 (21.4)	3 (30)	
Dementia, n (%)					.001
No	17 (73.9)	16 (69.6)	4 (28.6)	1 (10)	
Yes	6 (26.1)	7 (30.4)	10 (71.4)	9 (90)	
Duration of device (h), mean (SD)					
Holter	22.9 (1.3)	22.4 (1.2)	22.2 (1.1)	22.8 (1.2)	.30
S-Patch	64.6 (14.2)	65.8 (12.0)	70.8 (4.3)	62.9 (13.9)	.39
S-Patch analysis time (h)	56.8 (14.1)	55.9 (13.8)	58.2 (12.2)	52.3 (15.5)	.78
Fear of falling, n (%)					.40
No	13 (65)	11 (50)	5 (41.7)	3 (33.3)	
Yes	7 (35)	11 (50)	7 (58.3)	6 (66.7)	
Unable to answer	0 (0)	0 (0)	0 (0)	0 (0)	
Number of falls in the past 12 mo, mean (SD)	1.8 (1.4)	1.7 (1.5)	3.4 (2.9)	1.7 (1.3)	.03
ED[c] attendance, n (%)					.96
No	0 (0)	0 (0)	0 (0)	0 (0)	
Yes	22 (100)	21 (100)	17 (100)	9 (100)	
Total number of hospitalization days, mean (SD)	0.6 (0.9)	0.4 (0.5)	1.1 (2.0)	1.1 (1.4)	.22
Length of stay (d), mean (SD)	17.1 (22.8)	13.5 (10.4)	26.6 (23.5)	13.9 (9.2)	.21
Mortality, n (%)	0 (0)	0 (0)	1 (7.1)	0 (0)	.35
Hospital readmit (6 mo), n (%)					
Yes	7 (30.4)	5 (21.7)	0 (0)	4 (40)	.06
ED readmit (6 mo), n (%)					
Yes	10 (43.5)	5 (21.7)	0 (0)	5 (50)	.007

[a]CFS: Clinical Frailty Scale.

[b]AMT: Abbreviated Mental Test.

[c]ED: emergency department.

Both groups presented predominantly with an unwitnessed fall (21/35, 60% vs 17/35, 48.6%) and fewer with syncope (5/35, 14.3% vs 6/35, 17.1%) or near syncope (9/35, 25.7% vs 12/35, 34.3%). Falls in the preceding 12 months were similar in both group A and group B (2.2 vs 2.0, SD 1.9), with fear of falling also being similar (12/30, 40% vs 19/33, 57.6%). The SPPB

XSL•FO
RenderX

was performed in 68% (24/35) of group A and 71% (25/35) of group B participants, with a mean score of 3.7 (SD 2.8) vs 3.9 (SD 3.0; *P*=.84; Table 1), and functional impairment was mostly cited as the reason why patients were unable to perform the SPPB. Fall frequency in the preceding 12 months was higher in those with moderate disability (3.4 falls, SD 2.9) compared with those with severe disability (1.7 falls, SD 1.3), mild disability (1.7 falls, SD 1.5), and independent patients (1.8 falls, SD 1.4; *P*=.03; Table 3).

Observed Arrhythmia Detection

In terms of arrhythmia events in group A vs group B, Holter identified 28.6% (10/35) vs 42.9% (15/35; *P*=.32) and S-Patch identified 9 vs 9 (25/70, 35.7% vs 18/70, 25.7%; *P*>.99; Table 1). There were no statistically significant differences in sinus node disease, AV nodal disease, or tachyarrhythmia between the group A and group B participants for Holter and S-Patch (Table 1). Outcome comparison between the 2 treatment arms using McNemar test showed no differences in outcome events between Holter and S-Patch monitoring, 25 out of 70 (35.7%) vs 18 out of 70 (25.7%; *P*=.21; Table 4). For agreement analysis using paired data from the crossover design (n=70), a 2 × 2 contingency table (Table 4) was used to compare the detection of clinically significant arrhythmias by both devices. McNemar test showed no significant difference in arrhythmia detection between the 2 devices (χ^2_1=2.13, *P*=.21), and Cohen κ was 0.24 (*P*=.02), indicating fair agreement between devices. Using Holter as the reference standard, S-Patch sensitivity was 40% (10/25) and specificity was 82.2% (37/45), giving a positive predictive value of 55.6% (10/18) and a negative predictive value of 71.2% (37/52).

Table . Outcome comparison between 2 treatment arms and contingency table comparing detection of clinically significant arrhythmias for both devices.[a]

Contingency table	Holter: yes	Holter: no	Total
S-Patch: Yes	10	8	18
S-Patch: No	15	37	52

[a]McNemar test was used to compare detection of clinically significant arrhythmias between devices. Clinically significant arrhythmias included atrial fibrillation, atrial tachycardia, tachyarrhythmias, and bradyarrhythmias. S-Patch identified clinically significant arrhythmias in 25.7% (18/70) of participants and Holter monitoring identified clinically significant arrhythmias in 35.7% (25/70) of participants (*P*=.21).

S-Patch identified one symptomatic arrhythmia and a pause of >3 seconds, and Holter confirmed sinus arrest, with a long pause suggestive of sinus node disease for which a PPM was recommended. Another patient had an asymptomatic 2.8-second pause on S-Patch, which was not classified as an event.

There was a high level of agreement (κ=0.8584) between the 2 blinded raters (cardiologists) for AF, which was significantly greater than expected by chance (*P*<.001). The observed agreement for artifact was 37.54% and fell slightly below the expected agreement of 41.18%, resulting in a low κ coefficient of 0.0582, but this was not statistically significant (*P*=.21).

Hospital Usage and Mortality

ED attendance resulted in most patients being admitted to hospital (94.1% vs 91.4%; *P*>.99). Hospital LOS was similar in both groups (mean 17.9, SD 16.2 vs 16.8, SD 20.8; *P*=.80; Table 1). There was only 1 (2.9%, *P*>.99) mortality in group B during the study period (Table 1). Odds ratios for ED reattendance at 7 days, 30 days, or 6 months and for hospital readmission at 30 days or 6 months were similar between the 2 groups (Table 5).

XSL•FO
RenderX

Table . Comparison of hospital readmission and emergency department reattendance across study groups.

Outcome	Odds ratio or IRR[a] (95% CI)	*P* value
Hospital readmission[b] (OR[c])		
Tx B (S-Patch first)	0.33 (0.11-1.03)	.06
Hospital readmission[b] (30 d; IRR)		
Tx B (S-Patch first)	0.36 (0.12-1.14)	.08
Hospital readmission[b] (6 mo; IRR)		
Tx B (S-Patch first)	0.38 (0.13-1.12)	.08
ED[d] reattendance[b] (OR)		
Tx B (S-Patch first)	0.42 (0.14-1.24)	.12
ED reattendance[b] (7 d; IRR)		
Tx B (S-Patch first)	0.25 (0.02-2.76)	.26
ED reattendance[b] (30 d; IRR)		
Tx B (S-Patch first)	0.50 (0.18-1.41)	.19
ED reattendance[b] (6 mo) (IRR)		
Tx B (S-Patch first)	0.53 (0.22-1.27)	.15

[a]IRR: incidence rate ratio.

[b]Reference: Tx A (Holter first).

[c]OR: odds ratio.

[d]ED: emergency department.

ED reattendance within 6 months was high (5/10, 50%) in those whose Barthel score indicated severe disability, 0% (0/14) in those with moderate disability, 21.7% (5/23) in those with mild disability, and 43.5% (10/23) for independent patients (*P*=.007; Table 4). Those with moderate functional disability had a longer average LOS (26.6, SD 23.5 days) compared with those with severe disability (13.9, SD 9.2 days), mild disability (13.5, SD 10.4 days), and independent patients (17.1, SD 22.8 days; *P*=.21; Table 6).

XSL•FO
RenderX

Table . Clinical Frailty Scale.

CFS[a]	Not frail/robust (1-3), n=13 (18.6%)	Mildly frail (4-5), n=36 (51.4%)	Moderate/severe frail (6-9), n=21 (30.0%)	*P* value
Bathel Index group, n (%)				.001
Independent (20)	10 (76.9)	13 (36.1)	0 (0)	
Mild disability (16-19)	3 (23.1)	19 (52.8)	1 (4.8)	
Moderate disability (11-15)	0 (0)	4 (11.1)	10 (47.6)	
Severe disability (≤10)	0 (0)	0 (0)	10 (47.6)	
AMT[b] groups, n (%)				.001
Normal (≥7)	10 (76.9)	19 (52.8)	4 (19.1)	
Cognitive impairment (0 - 6)	0 (0)	13 (36.1)	11 (52.4)	
Not done	3 (23.1)	4 (11.1)	6 (28.6)	
Dementia, n (%)				.001
No	13 (100)	23 (63.9)	2 (9.5)	
Yes	0 (0)	13 (36.1)	19 (59.4)	
Duration of device (h), mean (SD)				
Holter	23.1 (0.8)	22.3 (1.3)	22.7 (1.2)	.10
S-Patch	61.3 (17.3)	66.9 (10.7)	66.9 (10.6)	.35
S-Patch analysis time (h)	56.6 (16.6)	56.7 (12.8)	54.8 (13.9)	.87
Fear of falling, n (%)				.04
No	9 (69.2)	18 (56.3)	5 (27.8)	
Yes	4 (30.8)	14 (43.8)	13 (72.2)	
Number of falls in the past 12 mo, mean (SD)	1.5 (1.0)	2.0 (1.9)	2.6 (2.2)	.26
ED[c] attendance, n (%)				>.99
No	0 (0)	0 (0)	0 (0)	
Yes	12 (100)	36 (100)	21 (100)	
Total number of hospitalization days, mean (SD)	0.4 (0.9)	0.6 (0.9)	1.0 (1.7)	.31
Length of stay (d), mean (SD)	9.6 (7.0)	17.9 (19.9)	21.2 (20.0)	.22
Mortality, n (%)	0 (0)	0 (0)	1 (4.8)	.48
Hospital readmit (6 mo), n (%)				>.99
Yes	3 (23.1)	8 (22.2)	5 (23.8)	
ED readmit (6 mo), n (%)				>.99
Yes	4 (30.8)	10 (27.8)	6 (28.6)	

[a]CFS: Clinical Frailty Scale.

[b]AMT: Abbreviated Mental Test.

[c]ED: emergency department.

Discussion

To our knowledge, this prospective single-center randomized open-label crossover pilot study is the largest to evaluate the tolerability and real-world use of the S-Patch in an older population, with a high prevalence of cognitive impairment and frailty. In this cohort, the S-Patch was well tolerated and enabled longer continuous monitoring durations compared with standard Holter monitoring. These findings address a key gap in clinical practice, where the success of ambulatory ECG monitoring is

XSL•FO
RenderX

often limited not by diagnostic capability but by patient adherence and device acceptability.

Compliance with monitoring was good for both devices, but S-Patch was tolerated for longer periods when S-Patch was administered first (group A [Holter first], 62.9, SD 13.4 h vs group B [S-Patch first] 68.8, SD 9.5 h; *P*=.04). In terms of durability, S-Patch also reports the available analysis time, which allows some interpretation of efficacy. Analysis time was 53.1 (SD 13.4) hours vs 58.9 (SD 13.6) hours. The difference between duration worn and available analysis time may be explained by loss of signal or due to interference during patient activities. The requirement for a paired mobile device within close proximity may have introduced practical challenges and may have contributed to reduced data capture in some patients.

This study is unique in its high proportion of vulnerable patients, including those with dementia (32/70, 45.7%) and frailty (57/70, 81.4%). Conventional monitoring approaches are frequently challenging in these groups due to device complexity, discomfort, and interference with daily activities. S-Patch offers an alternative to Holter monitoring, and its lightweight, minimally intrusive design likely contributed to sustained wear time. Importantly, tolerability was maintained despite the need for caregiver support in some cases.

Holter is considered the gold standard for short-term monitoring, and the diagnostic yield of clinically significant arrhythmias varies between 5.2% and 11% [23-25], with older age increasing the likelihood of significant arrhythmia identification [23,25]. Oversensing was more frequent with the S-Patch (75.76%) and, less so, undersensing (6%), which highlights the major challenges of analyzing complex and variable ECG signals. Other studies using S-Patch 2 have shown a cardiologist-level accuracy of 99.8% [26]. In terms of diagnostic performance, arrhythmia detection rates were comparable between devices, with one clinically significant arrhythmia identified by both modalities leading to pacemaker implantation. Although the S-Patch demonstrated lower sensitivity but reasonable specificity relative to Holter, the clinical implications of this finding should be interpreted cautiously. In older adults with falls, short-term monitoring has inherent limitations. Arrhythmias are often intermittent and may not coincide with the monitoring window, and detection in the absence of symptom correlation may not alter management. Conversely, a negative result does not exclude an arrhythmic cause.

In the Multi-Ethnic Study of Atherosclerosis, AF was first detected on days 3 to 14 of monitoring [27], highlighting that extending monitoring to 72 hours may still be insufficient for arrhythmia detection. The infrequent nature of syncopal events makes their relationship to falls challenging, but longer periods of ILR monitoring have shown that 71% of patients with unexplained falls had an underlying arrhythmia that was not initially present but was detected within 9 months using ILR [2]. However, their invasive nature limits widespread acceptability, reinforcing the importance of maximizing tolerability in noninvasive monitoring approaches.

Conversely, ambulatory recording in those with a history of recurrent falls [28] showed no causative arrhythmias and concluded that most arrhythmias detected did not correlate with symptoms [28]. The impact of falls is potentially devastating, and first-time syncope is associated with an 80% increased 1-year risk of fall-related injuries compared with an age- and sex-matched group [7]. AF is also associated with falls in older adults [9], and the prevalence of AF in this study was high at 16.5%, of whom 41% had paroxysmal AF, making detection more challenging [9]. In addition, even those with a history of AF have a 2.5 times greater risk of nonaccidental falls [29], and independent predictors of falls included AF, neurological disorders, and age ≤81 years [29].

Falls remain a multifactorial geriatric syndrome with significant negative consequences, including fear of falling, functional deterioration, anxiety, and depression [7], and older patients feel stigmatized by falls as a marker of aging [30]. This study found that those with moderate functional impairment had increased falls in the preceding 12 months (mean 3.4, SD 2.9; *P*=.03), and frail patients had a greater fear of falling compared with nonfrail patients (13/18, 72.2% vs 4/13, 30.8%; *P*=.04). Cardiac arrhythmias represent one of the few potentially reversible contributors to falls, and improving access to practical and acceptable monitoring strategies is clinically important.

Targeting higher-risk patients, such as those with unwitnessed falls, syncope, or established cardiovascular disease, may further improve diagnostic efficiency [8]. Independent predictors of cardiac syncope include age older than 65 years, the presence of coronary artery disease, and Q waves on the ECG [10]. ECG abnormalities were reported in 64% of syncope clinic attendees, where 26% of ECGs suggested an underlying cardiac cause [11], and high-risk ECG features or elevated cardiac enzymes can guide clinicians in the selection for onward referral or longer-term monitoring [31]. High-risk patients include those post-TAVI, with a high rate of arrhythmia detection, with new-onset AF in 19.6%, and recurrent episodes of AF were detected in those with a history of AF [32].

There were challenges and limitations within this study, which included that this was a single-site, open-label study with a crossover design to provide a comparison for tolerability, and blinding was not possible due to physical differences in the devices. Devices were placed by technicians, and analysis was undertaken by cardiologists not involved in the direct care of the patients to reduce the risk of bias. The randomized crossover design mitigated the risk of bias, and the 2 groups were well balanced. The study was underpowered as only 73 patients from an intended 116 were recruited, which limits the statistical power and increases the likelihood of type II errors. Recruitment of older patients was challenging, consistent with other literature [33]. Reasons for declining the study have been categorized into 4 groups. First, patient-related factors, such as older patients' reticence to participate in clinical trials, concern about tolerating a device (particularly in those with dementia), and fear of prolonging hospitalization, despite both devices being available for home recordings. Second, consent-related issues, which included the inability or lack of relatives to attend the hospital to provide legally appointed representative consent, particularly during the pandemic restrictions. Third, physician-related factors, in that Holter monitoring was already initiated or the physician preferred monitoring for a longer

XSL•FO
RenderX

period. Finally, the COVID-19 pandemic meant that research activities were suspended for long periods, and there was an upgrade and replacement of the S-Patch during which recruitment could not take place. Finally, the majority of Holter and S-Patch recordings were undertaken during hospitalization, as most patients declined home monitoring, and this was the commonest reason cited for withdrawing from the crossover part of the study. Monitoring for arrhythmia during hospitalization may not be reflective or representative of arrhythmias during normal day-to-day activity, which may explain some of the variability in the identification of arrhythmias between the 2 devices.

The sequential design (Holter followed by S-Patch, or vice versa) may also introduce a major confounding factor related to the temporal variability of arrhythmias and makes it challenging to decide if this is due to device failure or the absence of an arrhythmia during the specific monitoring window. This limitation makes a direct comparison of diagnostic performance ("efficacy") inherently problematic.

In conclusion, S-Patch compared with Holter supported longer continuous monitoring without compromising patient acceptability and was well tolerated in both frail older adults with cognitive impairment. S-Patch may be particularly useful in patients who are unable to tolerate conventional Holter monitoring or where extended monitoring is desirable but invasive approaches are not appropriate.

Acknowledgments

The authors thank the Clinical Measurement Unit staff, the Clinical Trials & Research Unit staff, Carolyn Yap, Geraldine Lim, and the Emergency Department staff at Changi General Hospital for their support in undertaking this study. The S-Patch3-Cardio/S-Patch Ex and the electrodes used in this study were sponsored by Samsung (manufacturer of S-Patch3-Cardio/S-Patch Ex). AI was not used to write the manuscript. ChatGPT 5.5 was used only to reword the abstract.

Funding

This study was supported by the Changi General Hospital Research Grant (CHF2019.05-P).

Data Availability

Anonymized data are available from the corresponding author upon reasonable request.

Authors' Contributions

CY, VHT, and BR contributed to the study conceptualization, data collection, data interpretation, results analysis, and manuscript writing. HTM contributed to data interpretation, results analysis, statistical analysis, and manuscript writing. VCB and CYES contributed to data collection, data interpretation, and manuscript writing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Assessment schedule.
[DOCX File, 13 KB - rehab_v13i1e81761_app1.docx]

Multimedia Appendix 2
User preference survey.
[DOCX File, 21 KB - rehab_v13i1e81761_app2.docx]

Multimedia Appendix 3
S-Patch event diary.
[DOCX File, 13 KB - rehab_v13i1e81761_app3.docx]

Multimedia Appendix 4
Comparison of Holter vs S-Patch data.
[DOCX File, 13 KB - rehab_v13i1e81761_app4.docx]

References

1. Gill TM, Murphy TE, Gahbauer EA, Allore HG. The course of disability before and after a serious fall injury. JAMA Intern Med 2013 Oct 28;173(19):1780-1786. [doi: 10.1001/jamainternmed.2013.9063] [Medline: 23958741]

XSL•FO
RenderX

2. Bhangu J, McMahon CG, Hall P, et al. Long-term cardiac monitoring in older adults with unexplained falls and syncope. Heart 2016 May;102(9):681-686. [doi: 10.1136/heartjnl-2015-308706] [Medline: 26822427]
3. Fuller GF. Falls in the elderly. Am Fam Physician 2000 Apr 1;61(7):2159-2168. [Medline: 10779256]
4. Soriano TA, DeCherrie LV, Thomas DC. Falls in the community-dwelling older adult: a review for primary-care providers. Clin Interv Aging 2007;2(4):545-554. [doi: 10.2147/cia.s1080] [Medline: 18225454]
5. Kannus P, Parkkari J, Niemi S, Palvanen M. Fall-induced deaths among elderly people. Am J Public Health 2005 Mar;95(3):422-424. [doi: 10.2105/AJPH.2004.047779] [Medline: 15727970]
6. Donald IP, Bulpitt CJ. The prognosis of falls in elderly people living at home. Age Ageing 1999 Mar;28(2):121-125. [doi: 10.1093/ageing/28.2.121] [Medline: 10350407]
7. Numé AK, Carlson N, Gerds TA, et al. Risk of post-discharge fall-related injuries among adult patients with syncope: a nationwide cohort study. PLoS One 2018;13(11):e0206936. [doi: 10.1371/journal.pone.0206936] [Medline: 30462687]
8. Wiseman T, Betihavas V. The association between unexplained falls and cardiac arrhythmias: a scoping literature review. Aust Crit Care 2019 Sep;32(5):434-441. [doi: 10.1016/j.aucc.2018.08.003] [Medline: 30420209]
9. Hung CY, Wu TJ, Wang KY, et al. Falls and atrial fibrillation in elderly patients. Acta Cardiol Sin 2013 Sep;29(5):436-443. [Medline: 27122741]
10. Bhat PK, Pantham G, Laskey S, Como JJ, Rosenbaum DS. Recognizing cardiac syncope in patients presenting to the emergency department with trauma. J Emerg Med 2014 Jan;46(1):1-8. [doi: 10.1016/j.jemermed.2013.04.064] [Medline: 24063875]
11. de Ruiter SC, Wold JFH, Germans T, Ruiter JH, Jansen RWMM. Multiple causes of syncope in the elderly: diagnostic outcomes of a Dutch multidisciplinary syncope pathway. Europace 2018 May 1;20(5):867-872. [doi: 10.1093/europace/eux099] [Medline: 28520944]
12. Brenner R, Ammann P, Yoon SI, et al. Reduction of falls and fractures after permanent pacemaker implantation in elderly patients with sinus node dysfunction. Europace 2017 Jul 1;19(7):1220-1226. [doi: 10.1093/europace/euw156] [Medline: 27702858]
13. Hopewell S, Copsey B, Nicolson P, Adedire B, Boniface G, Lamb S. Multifactorial interventions for preventing falls in older people living in the community: a systematic review and meta-analysis of 41 trials and almost 20 000 participants. Br J Sports Med 2020 Nov;54(22):1340-1350. [doi: 10.1136/bjsports-2019-100732] [Medline: 31434659]
14. Brignole M, Moya A, de Lange FJ, et al. 2018 ESC Guidelines for the diagnosis and management of syncope. Eur Heart J 2018 Jun 1;39(21):1883-1948. [doi: 10.1093/eurheartj/ehy037] [Medline: 29562304]
15. Shen WK, Sheldon RS, Benditt DG, et al. 2017 ACC/AHA/HRS guideline for the evaluation and management of patients with syncope: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. Circulation 2017 Aug 1;136(5):e60-e122. [doi: 10.1161/CIR.0000000000000499] [Medline: 28280231]
16. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):1-36. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
17. Falls: assessment and prevention in older people and in people 50 and over at higher risk. : National Institute for Health and Care Excellence (NICE); 2013 URL: https://www.nice.org.uk/guidance/ng249/resources/falls-assessment-and-prevention-in-older-people-and-in-people-50-and-over-at-higher-risk-pdf-66143964997573 [accessed 2026-06-23]
18. Upadhyayula S, Kasliwal RR. Wellysis S-PAtch Cardio versus conventional holter ambulatory electrocardiographic monitoring (the PACER trial): preliminary results. J Clin Prev Cardiol 2019;8(4):173-182. [doi: 10.4103/JCPC.JCPC_43_19]
19. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988;10(2):61-63. [doi: 10.3109/09638288809164103] [Medline: 3403500]
20. Lauretani F, Ticinesi A, Gionti L, et al. Short-Physical Performance Battery (SPPB) score is associated with falls in older outpatients. Aging Clin Exp Res 2019 Oct;31(10):1435-1442. [doi: 10.1007/s40520-018-1082-y] [Medline: 30515724]
21. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005 Aug 30;173(5):489-495. [doi: 10.1503/cmaj.050051] [Medline: 16129869]
22. Frailty Strategy Policy Report. : Ministry of Health; 2023 URL: https://www.moh.gov.sg/others/resources-and-statistics/reports-frailty-strategy-policy-report/ [accessed 2026-07-08]
23. Kühne M, Schaer B, Moulay N, Sticherling C, Osswald S. Holter monitoring for syncope: diagnostic yield in different patient groups and impact on device implantation. QJM 2007 Dec;100(12):771-777. [doi: 10.1093/qjmed/hcm104] [Medline: 18089543]
24. Freund O, Caspi I, Shacham Y, et al. Holter ECG for syncope evaluation in the internal medicine department-choosing the right patients. J Clin Med 2022 Aug 16;11(16):4781. [doi: 10.3390/jcm11164781] [Medline: 36013018]
25. Kühne M, Schaer B, Sticherling C, Osswald S. Holter monitoring in syncope: diagnostic yield in octogenarians. J Am Geriatr Soc 2011 Jul;59(7):1293-1298. [doi: 10.1111/j.1532-5415.2011.03486.x] [Medline: 21718271]
26. Jeon E, Oh K, Kwon S, et al. A lightweight deep learning model for fast electrocardiographic beats classification with a wearable cardiac monitor: development and validation study. JMIR Med Inform 2020 Mar 12;8(3):e17037. [doi: 10.2196/17037] [Medline: 32163037]

XSL•FO
RenderX

27. Heckbert SR, Austin TR, Jensen PN, et al. Yield and consistency of arrhythmia detection with patch electrocardiographic monitoring: the Multi-Ethnic Study of Atherosclerosis. J Electrocardiol 2018;51(6):997-1002. [doi: 10.1016/j.jelectrocard.2018.07.027] [Medline: 30497763]
28. Davison J, Brady S, Kenny RA. 24-hour ambulatory electrocardiographic monitoring is unhelpful in the investigation of older persons with recurrent falls. Age Ageing 2005 Jul;34(4):382-386. [doi: 10.1093/ageing/afi108] [Medline: 15901576]
29. Sanders NA, Ganguly JA, Jetter TL, et al. Atrial fibrillation: an independent risk factor for nonaccidental falls in older patients. Pacing Clin Electrophysiol 2012 Aug;35(8):973-979. [doi: 10.1111/j.1540-8159.2012.03443.x] [Medline: 22694347]
30. Koh VJW, Matchar DB, Chan AWM, et al. Reducing falls among community-dwelling older adults from clinicians' perspectives: a systems modeling approach. Innov Aging 2023;7(7):igad077. [doi: 10.1093/geroni/igad077] [Medline: 37694132]
31. Sandhu RK, Sheldon RS. Syncope in the emergency department. Front Cardiovasc Med 2019;6:180. [doi: 10.3389/fcvm.2019.00180] [Medline: 31850375]
32. Miki T, Senoo K, Ohkura T, et al. Importance of monitoring cardiac arrhythmias using 14-day patch electrocardiography in super older patients who underwent transcatheter aortic valve replacement. Am J Cardiol 2023 Dec 15;209:57-59. [doi: 10.1016/j.amjcard.2023.09.090] [Medline: 37858594]
33. Jansen S, van der Velde N. Syncope in older adults: challenges, approach and treatment. Age Ageing 2024 Feb 1;53(2):1-4. [doi: 10.1093/ageing/afad245] [Medline: 38331395]

Abbreviations

AF: atrial fibrillation
AMT: Abbreviated Mental Test
CFS: Clinical Frailty Scale
CIRB: Centralized Institutional Review Board
ED: emergency department
LOS: length of stay
PPM: permanent pacemaker
SPPB: Short Physical Performance Battery

Edited by S Munce; submitted 03.Aug.2025; peer-reviewed by B Cahill, HMS Ajmal, Y Cho; revised version received 05.May.2026; accepted 01.Jun.2026; published 14.Aug.2026.

Please cite as:
Rosario B, Tan VH, Barrera VC, Mon HT, Seah CYE, Yeo C
Comparing the Feasibility of S-Patch vs 24-Hour Holter in Frail Older People With Falls: Randomized Open-Label Crossover Pilot Study
JMIR Rehabil Assist Technol 2026;13:e81761
URL: https://rehab.jmir.org/2026/1/e81761
doi:10.2196/81761

© Barbara Rosario, Vern Hsen Tan, Vivian Cantiller Barrera, Hnin Tun Mon, Clara Yi-En Seah, Colin Yeo. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 14.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Reliability and Discriminant Ability of an Instrumented Timed Up and Go Test in People With Postsurgical Orthopedic Conditions: Quantitative Study

Marica Giardini[1], PT; Ilaria Arcolin[1], PT, MSc; Valerio Antonio Arcobelli[2], PhD; Michela Picardi[3], PT, PhD; Sabato Mellone[2], PhD; Marco Godi[1], PT, MSc

[1]Department of Physical Medicine and Rehabilitation, Institute of Veruno, Istituti Clinici Scientifici Maugeri IRCCS, Gattico-Veruno, Italy

[2]Department of Electrical, Electronic, and Information Engineering–Guglielmo Marconi (DEI), University of Bologna, Viale del Risorgimento 2, Bologna, Italy

[3]Department of Neurorehabilitation Sciences, Casa di Cura Igea, Milano, Italy

Corresponding Author:
Valerio Antonio Arcobelli, PhD
Department of Electrical, Electronic, and Information Engineering–Guglielmo Marconi (DEI), University of Bologna, Viale del Risorgimento 2, Bologna, Italy

Abstract

Background: The Timed Up and Go (TUG) test is widely used to assess mobility and fall risk in older adults and orthopedic patients. Its instrumented variant (iTUG), based on inertial measurement units, enables an objective quantification of motor performance and can even be implemented using smartphone technology. However, its broader clinical adoption remains limited by concerns about reliability, feasibility, and the interpretability of the extracted parameters.

Objective: This study aimed to evaluate the test-retest reliability of variables derived from a single-sensor iTUG in orthopedic inpatients undergoing rehabilitation and to determine whether a subset of reliable sensor-based metrics can support a multidimensional assessment of functional mobility and discriminate among common orthopedic conditions.

Methods: We recruited 104 inpatients at discharge from a rehabilitation ward after total hip arthroplasty, total knee arthroplasty, or femur fracture. Each participant performed the iTUG test on 2 consecutive days using a smartphone-based solution consisting of an inertial measurement unit placed on the lower back. From 100 extracted variables, those with excellent test-retest reliability (intraclass correlation coefficient ≥0.75) were retained. Exploratory factor analysis was used to identify underlying mobility domains, and linear discriminant analysis with 10-fold cross-validation tested their ability to classify diagnostic groups.

Results: Out of 100 iTUG-derived variables, 36 demonstrated excellent test-retest reliability, and 25 were retained for multivariate analysis. Exploratory factor analysis identified 5 factors—walking ability, pace or rhythm, sit-to-walk smoothness, turning ability, and mediolateral angular stability—explaining 80.8% of the total variance. These factors showed good classification accuracy (68%) and achieved an area under the curve of 0.86 and an overall mean accuracy of 0.68 (SD 0.14) for distinguishing among total hip arthroplasty, total knee arthroplasty, and femur fracture. In contrast, total iTUG duration alone yielded an area under the curve of 0.62. All patients used walking aids, and gait variables were more reliable than jerk-based or coordination metrics.

Conclusions: The single-sensor iTUG provides reliable and clinically informative metrics that go beyond traditional stopwatch timing, enabling a multidimensional view of functional mobility in orthopedic patients. The approach is feasible, interpretable, and compatible with real-world mobile health apps, supporting personalized rehabilitation monitoring and future integration into digital decision support systems.

(JMIR Rehabil Assist Technol 2026;13:e82632) doi:10.2196/82632

KEYWORDS

instrumented Timed Up and Go; iTUG test; inertial measurement unit; IMU; reliability; orthopedic patient; wearable sensor

Introduction

The Timed Up and Go (TUG) test is widely recognized for its simplicity and clinical applicability [1]. It is considered a fundamental tool for assessing functional mobility, including elements of balance, walking ability, and transitional movements in older adults [2-4]. The test involves rising from a chair, walking 3 m, performing a 180° turn, returning to the chair, and sitting down, with the total time recorded using a stopwatch [5].

Its widespread clinical use stems from its practicality and robust clinimetric properties. Within neurological practice, the TUG test is useful in identifying fall risk among people with

XSL•FO
RenderX

conditions such as stroke and Parkinson disease [6]. In orthopedics, the test has been recommended for evaluating patients with hip fractures [7] and for distinguishing between future recurrent fallers and nonfallers in frail older adults post hip fracture [8]. Additionally, the TUG has been demonstrated as a valuable test in predicting recovery trajectories: shorter recovery stays in patients with total hip arthroplasty (THA) [9] and early identification of postsurgical rehabilitation needs in patients with total knee arthroplasty (TKA) [10].

The advent of inertial measurement units (IMUs) has expanded the TUG's potential through detailed and objective kinematic analysis of its subcomponents. The instrumented TUG (iTUG) enables the detection of specific functional impairments [11], improves fall risk sensitivity [12], supports rehabilitation monitoring, and predicts clinical outcomes [13-15].

Indeed, a recent review has examined different methods for analyzing the iTUG, focusing on specific aspects such as key variables, clinical applications, and predictive value [1]. While this review underscores the clinical potential of the iTUG, its psychometric properties remain insufficiently explored. Existing research has primarily addressed the reliability of parameters in healthy older adults or individuals with neurological conditions [11,16,17].

To the best of our knowledge, no studies have analyzed iTUG reliability in patients with orthopedic conditions. This study aimed to evaluate the test-retest reliability of the iTUG in orthopedic inpatients in a rehabilitative setting. Specifically, we targeted 3 common orthopedic conditions: THA, TKA, and femur fractures. Additionally, we sought to identify the most reliable iTUG variables and assess their ability to discriminate between these patient groups using exploratory factor analysis (EFA) and linear discriminant analysis (LDA). We hypothesize that iTUG-derived variables can effectively distinguish between the 3 orthopedic groups and that domains related to dynamic mobility tasks, such as walking, will provide the highest discriminative value.

The objective is not to enable the diagnosis of those orthopedic conditions; rather, the hypothesis is that, based on measures derived from the iTUG, it may be possible to discriminate among these conditions in the phases immediately prior to discharge, given the distinct biomechanical implications underlying each. This initial evidence would support the use of the iTUG as a promising tool for patient follow-up, whether at home or during subsequent check-up visits.

Methods

Patients

Between April 2023 and December 2023, all patients admitted to the musculoskeletal ward of the Department of Physical Medicine and Rehabilitation at the Istituti Clinici Scientifici Maugeri Istituto di Ricovero e Cura a Carattere Scientifico (Gattico-Veruno, Piedmont, Italy) for rehabilitation following orthopedic surgery were screened for inclusion during the last 3 to 4 days of hospitalization. Patients were eligible if their primary admitting diagnosis was unilateral THA, TKA, or a femur fracture surgically repaired with internal fixation. Inclusion criteria included being aged 18 years or older, the ability to walk at least 10 meters without assistance (Functional Ambulation Category score ≥3), using the walking aid used in the final days of hospitalization, and the cognitive ability to understand instructions (Functional Independence Measure cognitive subscale ≥29).

Exclusion criteria included the presence of major associated injuries of the lower or upper limbs, contraindications to mobilization and weight-bearing on the lower extremities, clinical instability, disorders of bone metabolism other than osteoporosis, neurological or major cardiac diseases, pathological fractures, a prior history of overt dementia, terminal illness (life expectancy <6 mo), and prehospitalization dependence on nursing home care or inability to walk independently.

Ethical Considerations

The study was approved by the local ethics committee (2678 CE) and was conducted in accordance with the World Medical Association's Declaration of Helsinki. All participants provided informed consent for study participation. To ensure participants' privacy and confidentiality, all data were anonymized and coded using unique identification numbers. No personally identifiable information was stored in the research database or in the investigational application, and access to the data was restricted to authorized study personnel. Participants did not receive any financial compensation for their participation.

Setting and Procedures

iTUG Test

All evaluations were conducted in the Laboratory of Posture and Movement at the ICS Maugeri of Veruno. The laboratory provided a spacious, quiet area with adequate lighting, suitable for performing iTUG tests. The iTUG test was recorded following established protocols developed by Caronni et al [18,19]. During the test, participants wore a commercial IMU secured with a belt around their lower back at the level of the third lumbar vertebra (L3).

All participants were instructed to perform the iTUG test, which consisted of standing up from a chair, walking 3 m, pivoting, returning to the chair, and sitting down. An ordinary chair with a seat height of 44 cm was used. To mark the turning point, a yellow and black tape was placed on the floor at a distance of 3 m from the chair.

Participants initiated the task upon receiving a start signal from the iTUG commercial software and were instructed to maintain a comfortable and safe walking pace throughout the test.

Instrumentation

The iTUG system was supplied by mHealth Technologies srl (Monte San Pietro, Italy) in a portable briefcase containing the following components: a motion sensor with its charging station and power supply; 2 fabric belts of different sizes, equipped with Velcro closures, to secure the sensor to the patient; an Android smartphone with its charging cable; a user manual; a quick guide; and a declaration of conformity. The motion sensor is a lightweight IMU (22 g) with dimensions of 54 mm×33

XSL•FO
RenderX

mm×14 mm. It incorporates a triaxial accelerometer, a triaxial gyroscope, and a triaxial magnetometer, offering a maximum sampling frequency of 200 Hz. The sensor is positioned at the third lumbar vertebra (L3) using a Velcro-secured belt, with the buttons facing downward. Proper positioning is critical, as incorrect placement results in unreliable data.

The system connects the IMU to a smartphone via Bluetooth, with the smartphone acting as the central processing and control unit. During the iTUG test, the IMU captures movement signals, which are processed by the smartphone app "mTUG," developed by mHealth Technologies srl. The app manages the system and generates a report on the patient's kinematic performance during the test. The IMU signals provide the total duration of the TUG test and, through validated algorithms [16,17], automatically segment the test into four distinct phases: (1) sit-to-walk, (2) walk, (3) turn 180°, and (4) turn-to-sit. From each test, a total of 100 metrics is obtained (Multimedia Appendix 1).

Procedure

The iTUG procedures were conducted by 4 licensed physical therapists who underwent a dedicated half-day training session focused on standardized test administration, including sensor placement, task instructions, and use of the data acquisition and analysis software. Each participant was evaluated twice: first on the third-to-last day of hospitalization and then 24 hours later. A rater, randomly selected from the pool of 4 trained raters, conducted the evaluation. The rater provided instructions, demonstrated the procedure, secured the belt with the sensor, and initiated the measurement by activating the software to give the start signal. To ensure familiarity with the test, participants performed an initial trial that was not recorded and excluded from the analysis, following established guidelines [11]. After the familiarization trial, participants completed 2 recorded trials. During each trial, the rater manually recorded the time to complete the TUG test using a stopwatch in addition to the automated measurements provided by the iTUG system. The entire procedure, including the familiarization trial and the 2 recorded trials, was repeated 24 hours later by a different randomly selected trained rater.

Data Analysis

Reliability Measures

The best iTUG trial, that is, the trial with the shortest time recorded by the stopwatch, recorded on the first day was compared with the best trial recorded on the second day. These iTUG tests were used to calculate the test-retest reliability of the 100 variables collected for each participant, expressed as intraclass correlation coefficients (ICCs). In particular, we used an ICC 2,1 (model 2, form 1) [20] because the rater was not fixed and the data used to calculate the ICC were the measures of a single walking trial. An ICC less than 0.40 was considered to indicate poor reliability, between 0.40 and 0.59 indicated moderate reliability, between 0.60 and 0.74 indicated good reliability, and 0.75 or greater indicated excellent reliability [21]. All variables that obtained an ICC value less than 0.75 during reliability evaluation were excluded from the subsequent analysis.

Only after completing the variable reduction process, and thus after analyzing the correlation matrix, was it decided to calculate other measures of agreement for the resulting variables. Indeed, the SE of measurement (SEM), to assess the absolute error of the instrument, and the minimum detectable change (MDC), which represents the smallest change considered significant beyond the measurement error of an individual, were employed [20,22]. The SEM was calculated by multiplying the SD of the measurements by the square root of 1 minus the ICC (SD × √1-ICC). The MDC at 90% was derived using the formula MDC=1.65×2×SEM.

Correlation Matrix

For simplicity, we chose to perform all analyses on the variables collected during the first trial on the first day of testing. The variables included in the correlation matrix were tested for normal distribution using the Shapiro-Wilk test and Q-Q plot. For variables with nonnormal distributions, a logarithmic transformation was applied, followed by a reassessment of normality. Subsequent analyses were performed using the transformed variables where appropriate.

The data were screened for factorability using several well-described criteria, including the Kaiser-Meyer-Olkin (KMO) measure for sampling adequacy, the Bartlett test of sphericity, an anti-image correlation matrix [23-26], and the Pearson correlation and determinant score.

- The KMO statistic ranges from 0 to 1, with values above 0.50 considered acceptable for individual variables and a total value above 0.70 deemed adequate for the overall correlation matrix [20].
- The Bartlett test of sphericity evaluates the null hypothesis that the original correlation matrix is equal to an identity matrix (one with all 1's on the diagonal). A significant Bartlett test (P<.05) suggests that the variables are suitable for factor analysis. This condition prevents the application of factor analysis because it seeks items that are sufficiently correlated for the extraction of the factors or dimensions [20].
- The anti-image correlation matrix was inspected to test whether partial correlations among variables were small; we calculated the percentage of elements with values exceeding 0.30.
- From the correlation matrix, created to explore the relationships between variables, only variables with Pearson correlation coefficients ≥0.20 but <0.90 were considered, in order to avoid the presence of both isolated variables and multicollinearity [20].
- The determinant of a correlation matrix assesses issues related to extreme multicollinearity (ie, questionnaire items with high correlations) and singularity (questionnaire items with perfect correlations). These issues can hinder the application of exploratory factor analysis, indicating the need to exclude problematic items. Since the determinant of a correlation matrix can range from 0 to 1, it should be greater than 0.00001 to mitigate the potential problems associated with multicollinearity and singularity [27].

XSL•FO
RenderX

Exploratory Factor Analysis

We conducted an EFA using principal component extraction with varimax rotation. While principal component analysis is not inherently a latent variable model, it is commonly employed in EFA as a factor extraction method when the objective is to identify underlying constructs [23]. The resulting factors were interpreted as latent dimensions of iTUG performance. iTUG variables were first standardized and mean-centered. The number of factors to retain was determined using the following standard criteria [28]:

- Kaiser criterion [28,29], which retains only factors with eigenvalues ≥1
- Cattell scree test [30], based on the visual inspection of the plot of eigenvalues
- Horn parallel analysis [31], where random eigenvalues are compared to the observed data to identify meaningful factors

Only variables with factor loadings ≥|0.40| were considered relevant contributors to each factor. The factor scores for each individual were then computed and used in subsequent analyses, representing the extent to which each participant expressed the corresponding latent dimension [32].

Linear Discriminant Analysis

The factor scores derived from the EFA were entered into an LDA with 10-fold cross-validation to evaluate the ability of iTUG variables to classify patients into the 3 diagnostic groups (THA, TKA, and femur fracture). Model performance was assessed using receiver operating characteristic (ROC) analysis, and discrimination was quantified using the area under the curve (AUC). The AUC of the proposed model was compared with that obtained from manually timed TUG performance (stopwatch) and with AUC values derived from an LDA model, including all iTUG variables with good reliability (ICC>0.75). AUC values were computed and reported separately for the THA, TKA, and femur fracture groups.

AUC values were interpreted as follows: poor discrimination, AUC<0.70; moderate discrimination, AUC between 0.70 and 0.89; and high discrimination, AUC≥0.90. Classification accuracy, calculated from the confusion matrix, was also reported separately for each diagnostic group.

Statistical Analysis

Mean (SD) values were used for descriptive statistics, and mean (SE) values were used for the figures. One-way ANOVA was used for the direct comparison between 2 variables (“Total Duration till initial contact with the chair” vs TUG test duration by chronometer). The level of significance was set at $P<.05$. Statistical analysis was conducted using R software (R Foundation for Statistical Computing), with the packages *blandr*, *EFAtools*, *FactoMineR*, *graphics*, *irr*, *lattice*, *MASS*, *methods*, *multiUS*, *psych*, and *stats*.

Sample Size Calculation

The sample size required for the reliability study was calculated based on the ICC method [33]. With 2 assessors using the iTUG system and assuming a reliability of at least 0.7, a sample of 104 patients was determined to be adequate.

Results

Participants

Of the 266 inpatients screened, 112 (42%) met the inclusion criteria and were enrolled in the study. The remaining patients were excluded from the study for the following reasons: no walking ability, cognitive impairments, clinical instability, other types of fractures besides femur fractures, or severe comorbidities. Among the 112 recruited participants, 3 did not undergo the second-day trial due to clinical reasons, 2 did not return for reassessment, and for 3 others, iTUG data could not be extracted due to software malfunctions. Thus, the final sample comprised 104 participants. Of these, 47% were hospitalized for femur fractures, 36% for THA, and 17% for TKA. Table 1 presents the clinical characteristics of the assessed patients.

XSL•FO
RenderX

Table . Characteristics of the recruited sample (N=104).

Characteristics	Value
Demographic	
Age (y), mean (SD)	73.6 (10.9)
Sex, n	
Female	69
Male	35
Clinical	
BMI, mean (SD)	22.8 (3.5)
Interval between surgery and assessment (d), mean (SD)	26.4 (6.8)
Injury side, n	
Right	54
Left	50
Type of surgical intervention, n	
TKA[a]	18
THA[b]	37
Femur fracture by osteosynthesis	20
Femur fracture by THA	29

[a]TKA: total knee arthroplasty.

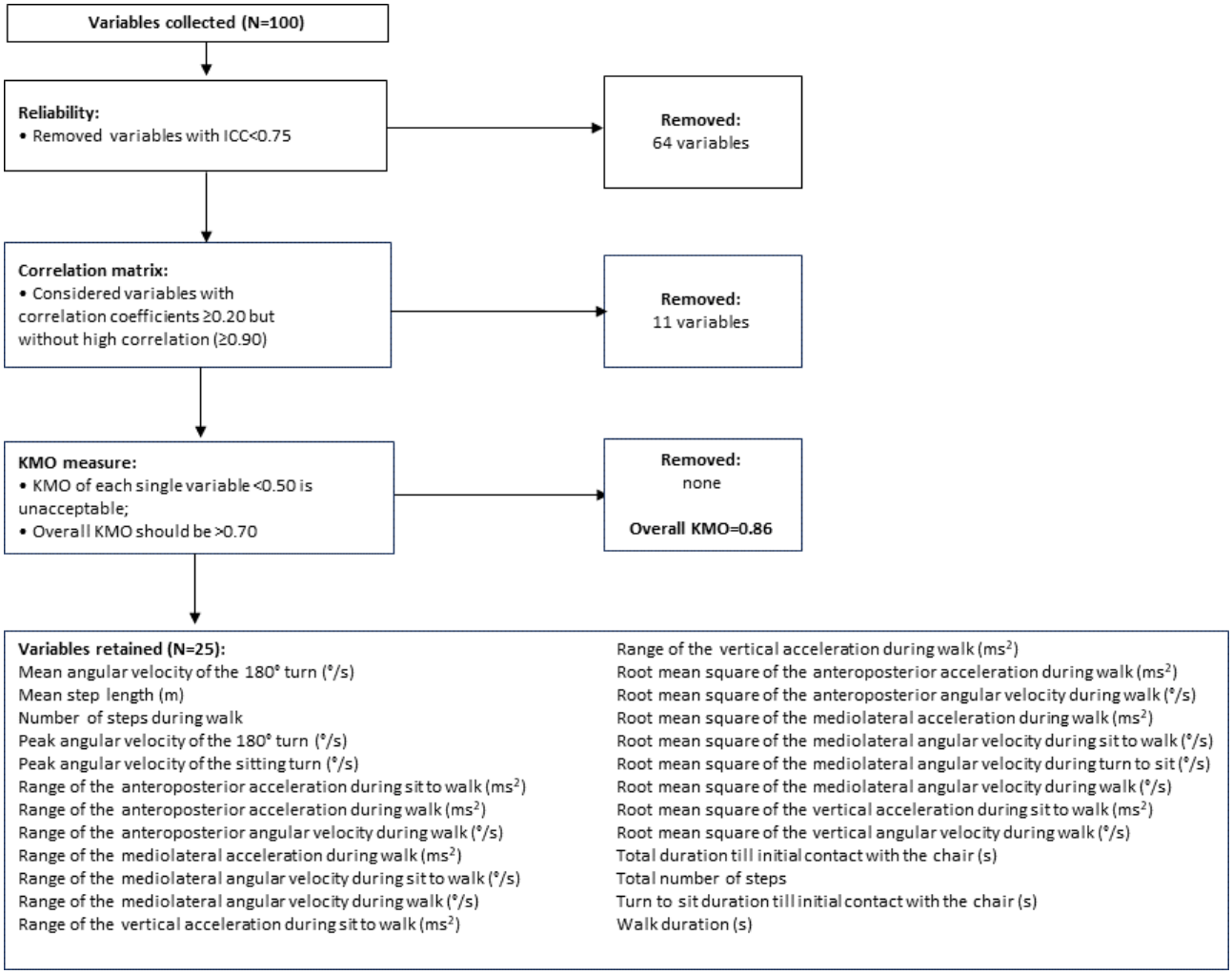
[b]THA: total hip arthroplasty.

Reliability Measures

The total iTUG duration recorded using the stopwatch was highly comparable to the iTUG-derived variable "Total Duration till initial contact with the chair" (correlation: y = 0.99x + 0.24; R^2=0.99). The mean time was 17.5 (SD 6.3) seconds for manual versus 17.6 (SD 6.3) seconds for iTUG, with no significant differences among orthopedic groups (ANOVA; *P*=.09). The ICC for this variable was 0.90 (95% CI 0.849 - 0.928), and it was therefore included in the subsequent EFA. Figure 1 illustrates the stepwise selection process for reliable iTUG variables.

XSL•FO
RenderX

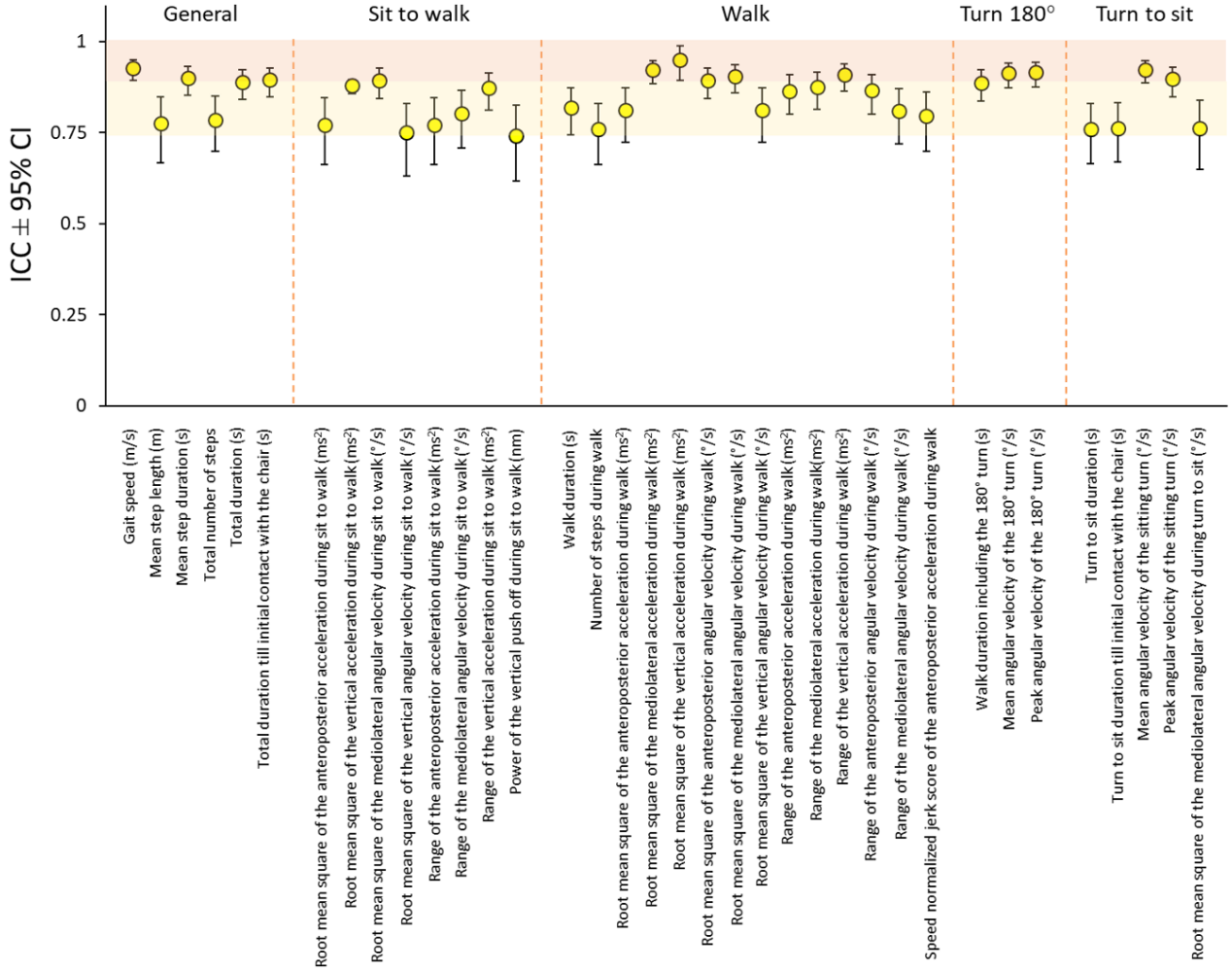
Figure 1. Flowchart describing the initial selection of variables from the instrumented Timed Up and Go (iTUG) test. ICC: intraclass correlation coefficient; KMO: Kaiser-Meyer-Olkin.



ICC values for all 100 iTUG variables initially considered were calculated. Subsequently, only variables with high ICC values were retained, resulting in a total of 36 variables. Figure 2 illustrates the ICC of these variables, providing an assessment of test-retest reliability: each point represents the ICC value for a variable, with error bars indicating the 95% CIs around the estimate. These intervals reflect the precision of the ICC measurement; narrower intervals denote higher confidence in the reliability estimate. With these 36 reliable variables, a correlation matrix was calculated.

XSL•FO
RenderX

Figure 2. Representation of the test-retest reliability of the 36 variables, divided into the different tasks of the instrumented Timed Up and Go (iTUG) test, with intraclass correlation coefficient (ICC) values >0.75. The top orange bar indicates excellent ICCs, meaning >0.90, while the yellow bar indicates ICCs >0.75, used as a cutoff to identify highly reliable variables.

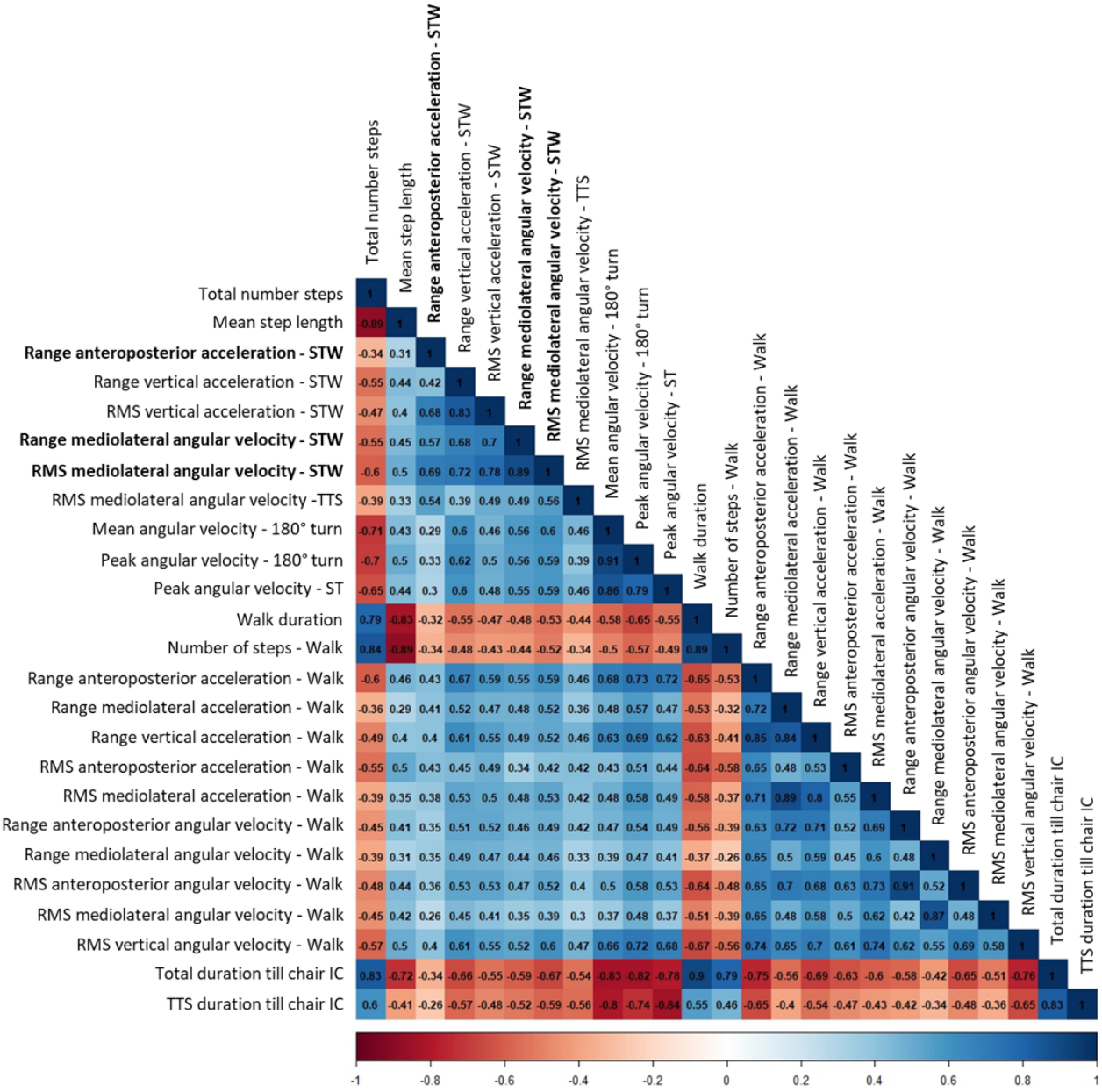


Correlation Matrix and Factor Analysis

Of the 36 retained variables, 6 were normally distributed; the remaining variables were log-transformed, and their distributions were then evaluated again for normality. Based on the predefined criteria for the correlation matrix (see "Methods" section), 11 variables were eliminated from the initial group of 36 variables (Figure 3).

XSL•FO
RenderX

Figure 3. Correlation matrix, shown as heatmap, between the final 25 variables of the instrumented Timed Up and Go (iTUG) test. Data from all participants were used to prepare the matrix (N=104). Variables shown in bold were already normally distributed and did not require logarithmic transformation. IC: initial contact; RMS: root mean square; ST: sitting turn; STW: sit-to-walk; TTS: turn-to-sit.



Sampling adequacy was confirmed (KMO=0.86), and the Bartlett test indicated significant correlations among variables (χ^2_{300}=3199.92; P<.001). The determinant of the correlation matrix was 7.36×10^{-16}, indicating no issues with multicollinearity or singularity. The percentage of partial correlations below the threshold of 0.3 in the anti-image correlation matrix was 16%, a value that can be considered adequate for proceeding with factor analysis.

Table 2 shows the correlations among the 25 selected variables, with all pairwise correlations exceeding the minimum acceptable threshold of 0.2. The table also reports the factor loadings for each variable on the 5 extracted factors, which were labeled according to their conceptual meaning and the naming convention proposed by Coni et al [34]: “walking ability,” “pace or rhythm,” “sit-to-walk smoothness,” “turning ability,” and “mediolateral angular stability.” The factor loadings were derived from a PCA with varimax rotation, which preserves orthogonality and retains the full variance of the original data. PCA identified 5 components with eigenvalues ≥1 (Kaiser criterion), and this was confirmed by the Horn parallel analysis. Together, these 5 components explained 80.8% of the total variance in the iTUG dataset, with individual contributions of 0.54, 0.09, 0.07, 0.06, and 0.05 of the variance, respectively.

XSL•FO
RenderX

Table . Factor loadings of the instrumented Timed Up and Go (iTUG) variables after principal component analysis[a].

	Walking ability	Pace or rhythm	Sit-to-walk smoothness	Turning ability	Mediolateral angular stability
Range of the anteroposterior acceleration during walk (ms^2)	*0.529*	—[b]	—	0.489	0.410
Range of the anteroposterior angular velocity during walk (°/s)	*0.811*	—	—	—	—
Range of the mediolateral acceleration during walk (ms^2)	*0.828*	—	—	—	—
Range of the vertical acceleration during walk (ms^2)	*0.704*	—	—	0.403	—
Root mean square of the anteroposterior acceleration during walk (ms^2)	*0.474*	0.492	—	—	—
Root mean square of the anteroposterior angular velocity during walk (°/s)	*0.791*	—	—	—	—
Root mean square of the mediolateral acceleration during walk (ms^2)	*0.782*	—	—	—	—
Root mean square of the vertical angular velocity during walk (°/s)	*0.530*	—	—	0.469	—
Mean step length (m)	—	*0.903*	—	—	—
Number of steps during walk	—	*–0.900*	—	—	—
Total duration till initial contact with the chair (s)	—	*–0.598*	—	–0.630	—
Total number of steps	—	*–0.771*	—	–0.451	—
Walk duration (s)	–0.405	*–0.786*	—	—	—
Range of the anteroposterior acceleration during sit-to-walk (ms^2)	—	—	*0.823*	—	—
Range of the mediolateral angular velocity during sit-to-walk (°/s)	—	—	*0.749*	—	—
Range of the vertical acceleration during sit-to-walk (ms^2)	—	—	*0.576*	0.422	—
Root mean square of the mediolateral angular velocity during sit-to-walk (°/s)	—	—	*0.801*	—	—
Root mean square of the mediolateral angular velocity during turn-to-sit (°/s)	—	—	*0.547*	—	—

XSL•FO
RenderX

	Walking ability	Pace or rhythm	Sit-to-walk smoothness	Turning ability	Mediolateral angular stability
Root mean square of the vertical acceleration during sit-to-walk (ms^2)	—	—	*0.795*	—	—
Mean angular velocity of the 180° turn (°/s)	—	—	—	*0.858*	—
Peak angular velocity of the 180° turn (°/s)	—	—	—	*0.743*	—
Peak angular velocity of the sitting turn (°/s)	—	—	—	*0.827*	—
Turn to sit duration till initial contact with the chair (s)	—	—	—	*–0.823*	—
Range of the mediolateral angular velocity during walk (°/s)	—	—	—	—	*0.845*
Root mean square of the mediolateral angular velocity during walk (°/s)	—	—	—	—	*0.857*

[a]Italicized values indicate the highest loading for each variable, representing the primary factor association. Only values greater than 0.4 are reported in the table. Variables shown in italics were already normally distributed and did not require logarithmic transformation.

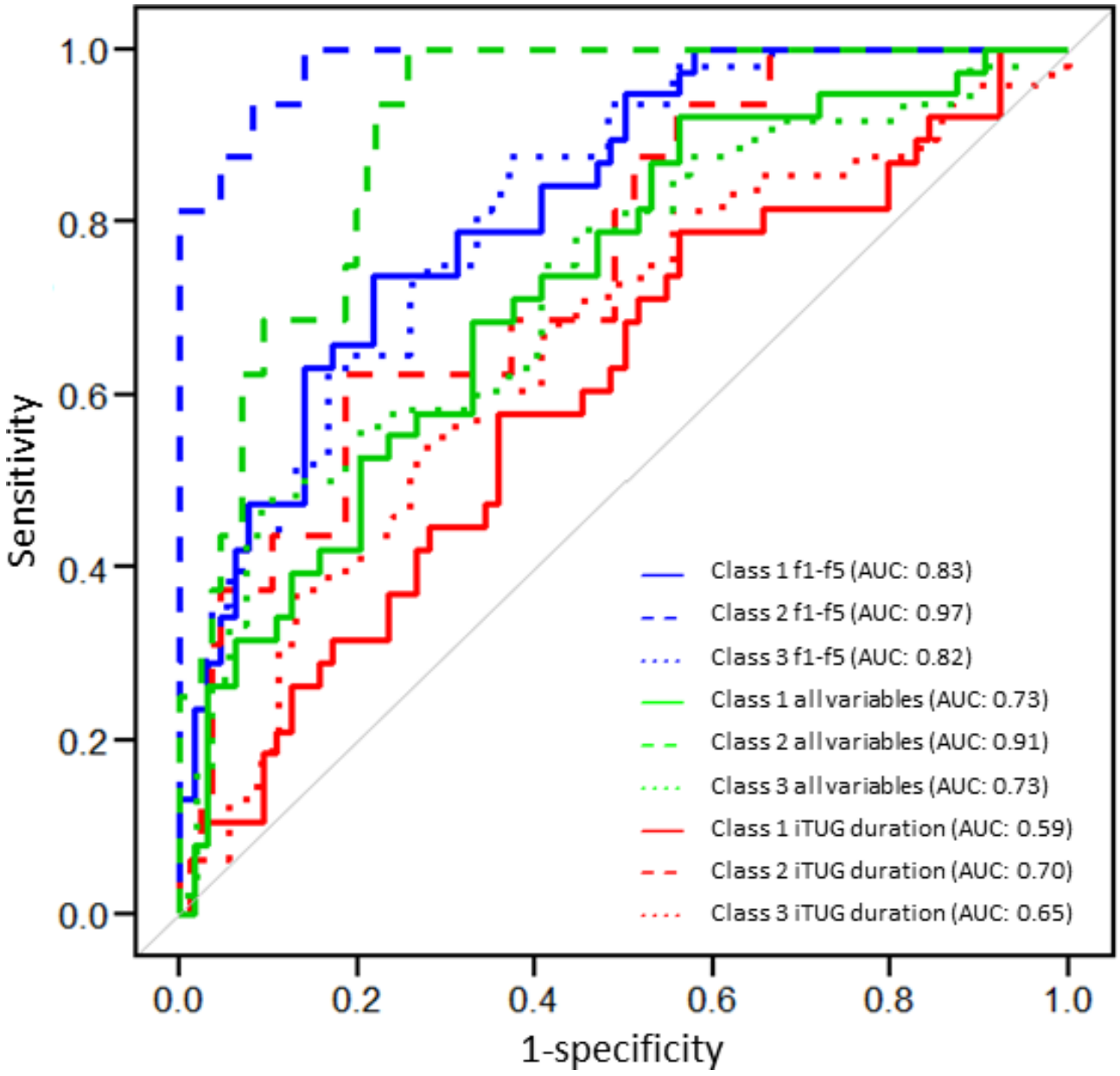
[b]Not applicable.

Linear Discriminant Analysis

LDA revealed that the first 5 factor scores, corresponding to the factors walking ability, pace or rhythm, sit-to-walk smoothness, turning ability, and mediolateral angular stability, successfully discriminated between the 3 diagnoses with an accuracy of 0.68. These 5 factor scores demonstrated a total AUC of 0.87, with an AUC of 0.83 for THA, 0.97 for TKA, and 0.82 for femur fractures (Figure 4).

XSL•FO
RenderX

Figure 4. Representation of receiver operating characteristic (ROC) curves. ROC compares the classification performance of instrumented Timed Up and Go (iTUG) duration (variable labeled as iTUG "total duration till initial contact with the chair"; red), factors (f1–f5; blue), and 5 randomly selected variables (green). Sensitivity is plotted against 1–specificity, illustrating the discriminative ability of each variable. The curves demonstrate the ability of different predictors to distinguish between patient diagnoses: total hip arthroplasty (THA; solid lines), total knee arthroplasty (TKA; long dashed lines), and femur fracture (short dashed lines). AUC: area under the curve.



The accuracy of the discriminant analysis model using the variable "total duration" was 0.50, with a total AUC of 0.65. In contrast, the model including all iTUG variables with good reliability achieved an accuracy of 0.54 and an AUC of 0.79.

Measures of Agreement

Table 3 shows the descriptive statistics, test-retest reliability indices, and measurement error parameters for the retained variables (n=25) across 2 sessions (day 1 and day 2). Each variable is accompanied by its mean and SD for both days, reflecting the overall consistency of the measures across time. The SEM provides an estimate of the measurement error attributable to random variations, offering insight into the precision of each variable. Lower SEM values reflect greater measurement precision. The MDC, derived from the SEM, represents the smallest change in a variable that can be interpreted as a true difference beyond measurement error. Larger MDC values indicate greater variability in detecting meaningful changes over time, whereas smaller values suggest higher sensitivity.

XSL•FO
RenderX

Table . Descriptive statistics and agreement measures of the retained variables (n=25).

Variable	Day 1, mean (SD)	Day 2		ICC[a]		
		Mean (SD)	ICC	95% CI	SEM[b]	MDC[c] (90%)
Total number of steps	15.298 (4.952)	14.644 (4.110)	0.779	0.689-0.846	2.138	4.973
Mean step length (m)	0.573 (0.192)	0.589 (0.167)	0.775	0.669-0.847	0.085	0.198
Range of the anteroposterior acceleration during sit-to-walk (ms^2)	9.842 (2.058)	9.702 (2.217)	0.784	0.681-0.853	0.993	2.310
Range of the vertical acceleration during sit-to-walk (ms^2)	4.724 (2.091)	4.733 (1.812)	0.805	0.712-0.868	0.862	2.005
Root mean square of the vertical acceleration during sit-to-walk (ms^2)	1.337 (0.512)	1.373 (0.505)	0.882	0.826-0.920	0.174	0.406
Range of the mediolateral angular velocity during sit-to-walk (°/s)	143.058 (53.462)	150.363 (52.583)	0.871	0.808-0.913	19.064	44.351
Root mean square of the mediolateral angular velocity during sit-to-walk (°/s)	52.204 (15.502)	54.974 (15.589)	0.890	0.830-0.927	5.173	12.035
Root mean square of the mediolateral angular velocity during turn-to-sit (°/s)	0.700 (0.305)	0.726 (0.340)	0.482	0.236-0.649	0.232	0.539
Mean angular velocity of the 180° turn (°/s)	66.099 (27.070)	72.208 (29.286)	0.901	0.834-0.938	8.911	20.730
Peak angular velocity of the 180° turn (°/s)	124.330 (42.056)	131.032 (45.224)	0.911	0.863-0.941	13.07	30.405
Peak angular velocity of the sitting turn (°/s)	130.800 (46.552)	140.803 (51.975)	0.889	0.823-0.928	16.489	38.359
Walk duration (s)	9.134 (3.804)	8.181 (3.073)	0.790	0.648-0.869	1.597	3.714

XSL•FO
RenderX

Variable	Day 1, mean (SD)	Day 2		ICC[a]		
		Mean (SD)	ICC	95% CI	SEM[b]	MDC[c] (90%)
Number of steps during walk	10.019 (4.098)	9.404 (3.287)	0.855	0.783-0.903	1.417	3.296
Range of the anteroposterior acceleration during walk (ms^2)	5.369 (2.001)	5.697 (2.093)	0.862	0.794-0.908	0.76	1.769
Range of the mediolateral acceleration during walk (ms^2)	5.868 (2.192)	6.338 (2.704)	0.868	0.797-0.913	0.896	2.085
Range of the vertical acceleration during walk (ms^2)	6.810 (2.722)	7.349 (2.670)	0.901	0.840-0.937	0.849	1.974
Root mean square of the anteroposterior acceleration during walk (ms^2)	1.438 (0.490)	1.553 (0.557)	0.810	0.711-0.874	0.23	0.534
Root mean square of the mediolateral acceleration during walk (ms^2)	0.915 (0.287)	0.982 (0.325)	0.912	0.846-0.946	0.091	0.212
Range of the anteroposterior angular velocity during walk (°/s)	54.028 (19.465)	56.768 (20.359)	0.863	0.797-0.907	7.372	17.150
Range of the mediolateral angular velocity during walk (°/s)	72.046 (27.695)	73.475 (27.123)	0.812	0.723-0.872	11.862	27.595
Root mean square of the anteroposterior angular velocity during walk (°/s)	9.390 (3.202)	9.997 (3.766)	0.888	0.828-0.926	1.172	2.726
Root mean square of the mediolateral angular velocity during walk (°/s)	13.355 (4.248)	14.005 (5.014)	0.901	0.852-0.934	1.46	3.396

XSL•FO
RenderX

Variable	Day 1, mean (SD)	Day 2		ICC[a]		
		Mean (SD)	ICC	95% CI	SEM[b]	MDC[c] (90%)
Root mean square of the vertical angular velocity during walk (°/s)	20.331 (6.604)	21.725 (6.561)	0.803	0.702-0.868	2.934	6.826
Total duration till initial contact with the chair (s)	17.671 (6.248)	15.838 (5.291)	0.853	0.619-0.929	2.241	5.214
Turn-to-sit duration till initial contact with the chair (s)	4.165 (1.931)	3.683 (1.655)	0.738	0.603-0.826	0.926	2.155

[a]ICC: intraclass correlation coefficient.

[b]SEM: SE of measurement.

[c]MDC: minimal detectable change.

Discussion

This study demonstrated that, in a sample of heterogeneous orthopedic patients discharged from a rehabilitative ward, only 25 out of the 100 original or exported iTUG variables were both reliable and capable of fitting into a robust assessment model with good indices. EFA on this model identified 5 domains that contribute to motor performance: walking ability, pace or rhythm, sit-to-walk smoothness, turning ability, and mediolateral angular stability. The ability of these factors to differentiate between the 3 orthopedic diagnoses is significantly higher than the total duration of the test. This finding underscores the clinical value of moving beyond single-duration metrics toward a multidimensional, sensor-based assessment framework.

A total of 36 out of the 100 variables derived from the iTUG exhibited excellent test-retest reliability. The fact that only one-third of the variables are reliable is consistent with findings previously reported in a population of elderly patients with femoral fractures [35]. In recent years, several studies have investigated the reliability of the iTUG in different populations [36-39], yet none have conducted a detailed analysis of individual variable reliability in patients with knee or hip replacements.

In our study, “Total Duration till initial contact with the chair” was found to be more reliable than the durations of the individual subphases. The most reliable variables were those related to the walking phase, whereas jerk-based metrics did not reach adequate levels of reliability. This finding is consistent with previous literature suggesting that the limited reliability of jerk-based features may be due to multiple concurrent factors, particularly during postural transitions such as sit-to-stand [39]. As reported by Weiss et al [40], outcome measures like jerk and range, which are highly sensitive to subtle variations and signal noise, may show greater variability and therefore reduced reliability, despite standardized acquisition protocols.

As expected, the measures that can be obtained using a stopwatch—that is, total duration and the gait variables derived from it (ie, gait speed, mean step duration, and cadence)—showed high reliability. It is worth noting that, although these variables can be derived from manual timing, the instrumented test additionally offers the advantage of being operator-independent [41].

Measures of stability, regularity, and coordination exhibited low reliability, likely due to the limited number of steps performed during the iTUG test. It is well established that reliable stability and coordination metrics typically require more than 20 to 25 steps [42,43]. In our sample, participants performed, on average, fewer than 15 steps, highlighting a limitation inherent to the test itself rather than to the measures.

Furthermore, all participants in our orthopedic patient sample used a walking aid, which appears to compromise jerk-related measures by disrupting the natural flow of gait [44]. As a result, only 1 jerk variable during the walking phase demonstrated good reliability.

A recent review has confirmed the reliability of single-point IMUs for gait metric analysis and their potential to support clinical applications, especially when placed over the L3-L4 vertebrae [45]. This approach has been shown to be valid for the iTUG as well [16]. Despite the simplicity of use and the demonstrated validity of using a single IMU placed over the lumbar region, the accuracy of spatiotemporal gait parameter estimation is significantly influenced by sensor placement [46]. For instance, although the average error across multiple strides remained consistent across different sensor positions, significant variations were observed in single-stride errors and variability parameters [47]. Practical considerations, such as ease of execution and time efficiency associated with different IMU

XSL•FO
RenderX

configurations, also play an important role. The single-sensor approach offers a clear advantage in terms of usability, as starting data collection is as straightforward as pressing "start" on a smartphone, comparable to using a stopwatch. On the other hand, while adding more sensors can enhance data richness and improve the accuracy of the estimation of some features, it might notably increase the time required to complete the evaluation, and its reliability would still be influenced by sensor placement [47,48]. Therefore, the iTUG test with only a single-sensor setup provides a good balance between sensor count and wearability, which is an optimal tradeoff between accuracy and practicality in real-world applications.

Furthermore, a single sensor for mobility assessment facilitates the transition of clinical evaluations from traditional supervised clinical settings to unsupervised home environments, thereby promoting and enabling continuity of care scenarios where the patient receives integrated, coordinated, and person-centered care, both during acute episodes and in the management of chronic conditions, reducing fragmentation in health care delivery [45].

Previous studies have demonstrated the use of iTUG measures in distinguishing preoperative functional status across various orthopedic conditions [49]. For instance, Bloomfield [50] categorized patients undergoing total knee replacement into moderate-functioning and low-functioning groups based on preoperative test duration. Similarly, Gasparutto et al [51] conducted a biomechanical analysis of the TUG test in total hip replacement patients, comparing functional deficits in each phase of the test in patients with those of a control group. However, these discriminative capabilities within relatively homogeneous orthopedic populations appear to be primarily driven by large differences in total test duration, rather than by the intrinsic discriminative power of the extracted sensor-based variables. In fact, both studies reported differences exceeding the 4-second minimal clinically important difference for orthopedic populations [52], potentially limiting the added value of sensor-based analysis.

In contrast, given the lack of significant differences in total TUG duration among the 3 patient groups, our findings suggest that this measure alone is insufficient to differentiate between groups in a sample of relatively high-functioning individuals. Instead, our iTUG-based model, which incorporates the 5 principal components, provided a significantly superior discriminative ability compared to total duration alone. Specifically, the iTUG model achieved an AUC greater than 0.8, whereas the model based solely on total duration yielded an AUC less than 0.7, indicating limited predictive power.

Overall, our results reinforce the added value and sensitivity of iTUG over the traditional stopwatch-based TUG. At the same time, they show that when sensor technology is not used and only standard variables are considered, the total duration recorded by the iTUG remains comparable to conventional measurements, as recently reported by Dos Santos et al [53]. This underscores the importance of integrating sensor-based analytics to refine mobility assessments and optimize rehabilitation strategies, ultimately improving clinical decision-making and patient care.

Although the LDA conducted in this study demonstrated good discriminant ability among groups (mean overall accuracy 0.68, SD 0.14), more advanced predictive models, such as machine learning algorithms, could perform even better. Machine learning approaches have already been applied to iTUG assessments in various populations and for different purposes, including evaluating levodopa responsiveness in individuals with Parkinson disease [54] and predicting falls across different pathologies [1]. In this perspective, the present study proposes a reliability-driven and interpretable framework for variable selection and dimensionality reduction, which may provide a preliminary layer of feature selection for the development of predictive models not only in orthopedic rehabilitation but also in other populations characterized by mobility impairments, such as neurological disorders. While the discriminative patterns identified here are specific to postsurgical orthopedic patients, the latent mobility domains derived from the iTUG (eg, walking ability, turning, sit-to-walk smoothness, and mediolateral stability) consist of fundamental components of functional mobility, which are also affected in other conditions, including neurological disorders.

This study has several clinical implications. First, the use of multidimensional, sensor-derived metrics, particularly those related to walking dynamics, offers a standardized and operator-independent approach to assessing mobility in orthopedic patients. This supports consistent tracking of rehabilitation progress across settings. Second, the factor structure identified through EFA provides clinicians with interpretable domains (eg, sit-to-walk smoothness and mediolateral angular stability) with strong discriminative performance (AUC>0.8) that can inform targeted therapy planning beyond total time. Third, the iTUG model's ability to discriminate between patients with similar test durations suggests that it can uncover latent performance differences that might otherwise remain undetected, especially in high-functioning individuals.

Selected variables showed stable and reproducible measurements in the present dataset. This suggests their potential suitability not only for cross-sectional group differentiation but also for longitudinal monitoring, where consistency of measurement is required to detect true change beyond measurement noise. From a clinical perspective, this capability is particularly relevant for stratifying and monitoring patients within the same orthopedic condition, such as those undergoing THA or TKA. The assessment and selection of sensor-based features represent an initial step toward the development of an evaluation model aimed at quantifying motor impairments that may distinguish the categories and reflect rehabilitation targets. The underlying objective is to capture functional dimensions that could inform treatment decisions; accordingly, these variables may serve as candidate outcome measures to track patient progress, evaluate treatment effectiveness, and support early adjustment of rehabilitation strategies. They may also contribute to finer personalization within each category by highlighting individual functional profiles.

Finally, once a reliable and clinically meaningful feature set is identified, the same variables could provide a basis for additional predictive models, such as fall-risk estimation or forecasting

XSL•FO
RenderX

functional decline. However, these findings should be considered exploratory. Validation in larger and more diverse cohorts will be necessary to assess the robustness, generalizability, and actual performance of both the assessment and predictive models and to determine their potential clinical use. Future work will therefore focus on increasing the sample size and conducting within-diagnosis analyses to identify distinct functional recovery profiles, enabling the distinction between patients who are recovering as expected and those showing subtle but clinically meaningful deviations from typical recovery trajectories. Such early identification may support timely, personalized rehabilitation adjustments before functional decline or plateau becomes evident using conventional time-based measures.

Moreover, the feasibility of a single-sensor setup reinforces the iTUG's potential for remote monitoring and continuity of care postdischarge. Reliable domains such as walking ability and pace or rhythm may represent valuable longitudinal indicators of recovery and could be integrated into digital tools for monitoring rehabilitation outcomes.

In conclusion, this study confirms the reliability of selected iTUG variables in orthopedic patients and reinforces their clinical value for a multidimensional assessment of motor performance. Factor analysis identified 5 meaningful components capable of distinguishing between common orthopedic conditions, addressing the limitations of relying solely on total test duration. These findings support the hypothesis that the development of predictive models and clinical decision support tools based on sensor-based functional tests, such as the iTUG, has the potential to improve clinical decision-making and optimize rehabilitation interventions.

Acknowledgments

The authors would like to express their sincere gratitude to all the people, patients, and physiotherapists who participated in this study.

Funding

This work was partially supported by the Ricerca Corrente funding scheme of the Ministry of Health, Italy.

Data Availability

The dataset used for this study is not publicly available due to patient privacy concerns but may be made available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: M Godi
Data curation: M Giardini, M Godi
Formal analysis: M Giardini, IA, VAA, M Godi
Investigation: M Giardini, IA, M Godi
Methodology: MP, SM
Writing–original draft: M Giardini, M Godi
Writing–review and editing: M Giardini, IA, VAA, MP, SM, M Godi
All authors participated in discussions, read and approved the final manuscript.

Conflicts of Interest

SM owns a share of mHealth Technologies srl, Monte San Pietro, Italy. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multimedia Appendix 1
The 100 parameters extracted from the instrumented Timed Up and Go (iTUG) test, categorized according to the specific functional phases: sit-to-walk, walking, turning, and turn-to-sit.
[PDF File, 143 KB - rehab_v13i1e82632_app1.pdf]

References

1. Ortega-Bastidas P, Gómez B, Aqueveque P, Luarte-Martínez S, Cano-de-la-Cuerda R. Instrumented Timed Up and Go Test (iTUG)-more than assessing time to predict falls: a systematic review. Sensors (Basel) 2023 Mar 24;23(7):3426. [doi: 10.3390/s23073426] [Medline: 37050485]
2. Podsiadlo D, Richardson S. The Timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
3. Herman T, Giladi N, Hausdorff JM. Properties of the "Timed Up and Go" test: more than meets the eye. Gerontology 2011;57(3):203-210. [doi: 10.1159/000314963] [Medline: 20484884]

XSL•FO
RenderX

4. Nightingale CJ, Mitchell SN, Butterfield SA. Validation of the Timed Up and Go test for assessing balance variables in adults aged 65 and older. J Aging Phys Act 2019 Apr 1;27(2):230-233. [doi: 10.1123/japa.2018-0049] [Medline: 30117359]
5. Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the “Get-Up and Go” test. Arch Phys Med Rehabil 1986 Jun;67(6):387-389. [Medline: 3487300]
6. Caronni A, Picardi M, Scarano S, et al. Pay attention: you can fall! The Mini-BesTest scale and the turning duration of the TUG test provide valid balance measures in neurological patients: a prospective study with falls as the balance criterion. Front Neurol 2023;14:1228302. [doi: 10.3389/fneur.2023.1228302] [Medline: 37745667]
7. McDonough CM, Harris-Hayes M, Kristensen MT, et al. Physical therapy management of older adults with hip fracture. J Orthop Sports Phys Ther 2021 Feb;51(2):CPG1-CPG81. [doi: 10.2519/jospt.2021.0301] [Medline: 33522384]
8. Wald P, Chocano-Bedoya PO, Meyer U, et al. Comparative effectiveness of functional tests in fall prediction after Hip Fracture. J Am Med Dir Assoc 2020 Sep;21(9):1327-1330. [doi: 10.1016/j.jamda.2020.02.008] [Medline: 32276783]
9. Oosting E, Kapitein PJC, de Vries SV, Breedveld E. Predicting short stay total hip arthroplasty by use of the Timed Up and Go-test. BMC Musculoskelet Disord 2021 Apr 16;22(1):361. [doi: 10.1186/s12891-021-04240-6] [Medline: 33863323]
10. Givens DL, Eskildsen S, Taylor KE, Faldowski RA, Del Gaizo DJ. Timed Up and Go test is predictive of patient-reported outcomes measurement information system physical function in patients awaiting total knee arthroplasty. Arthroplast Today 2018 Dec;4(4):505-509. [doi: 10.1016/j.artd.2018.07.010] [Medline: 30560183]
11. van Lummel RC, Walgaard S, Hobert MA, et al. Intra-rater, inter-rater and test-retest reliability of an instrumented Timed Up and Go (iTUG) test in patients with Parkinson’s disease. PLoS ONE 2016;11(3):e0151881. [doi: 10.1371/journal.pone.0151881] [Medline: 26999051]
12. Williams JM, Nyman SR. Association between the instrumented Timed Up and Go test and cognitive function, fear of falling and quality of life in community dwelling people with dementia. J Frailty Sarcopenia Falls 2018 Dec;3(4):185-193. [doi: 10.22540/JFSF-03-185] [Medline: 32300707]
13. Tan D, Pua YH, Balakrishnan S, et al. Automated analysis of gait and modified timed up and go using the Microsoft Kinect in people with Parkinson’s disease: associations with physical outcome measures. Med Biol Eng Comput 2019 Feb;57(2):369-377. [doi: 10.1007/s11517-018-1868-2] [Medline: 30123947]
14. Mangano GRA, Valle MS, Casabona A, Vagnini A, Cioni M. Age-related changes in mobility evaluated by the Timed Up and Go Test instrumented through a single sensor. Sensors (Basel) 2020 Jan 28;20(3):719. [doi: 10.3390/s20030719] [Medline: 32012930]
15. Bergquist R, Nerz C, Taraldsen K, et al. Predicting advanced balance ability and mobility with an instrumented Timed Up and Go Test. Sensors (Basel) 2020 Sep 3;20(17):4987. [doi: 10.3390/s20174987] [Medline: 32899143]
16. Mellone S, Tacconi C, Chiari L. Validity of a smartphone-based instrumented Timed Up and Go. Gait Posture 2012 May;36(1):163-165. [doi: 10.1016/j.gaitpost.2012.02.006] [Medline: 22421189]
17. Picardi M, Redaelli V, Antoniotti P, et al. Turning and sit-to-walk measures from the instrumented Timed Up and Go test return valid and responsive measures of dynamic balance in Parkinson’s disease. Clin Biomech (Bristol) 2020 Dec;80:105177. [doi: 10.1016/j.clinbiomech.2020.105177] [Medline: 32979787]
18. Caronni A, Sterpi I, Antoniotti P, et al. Criterion validity of the Instrumented Timed Up and Go test: a partial least square regression study. Gait Posture 2018 Mar;61:287-293. [doi: 10.1016/j.gaitpost.2018.01.015] [Medline: 29413799]
19. Caronni A, Picardi M, Aristidou E, et al. How do patients improve their Timed Up and Go test? Responsiveness to rehabilitation of the TUG test in elderly neurological patients. Gait Posture 2019 May;70:33-38. [doi: 10.1016/j.gaitpost.2019.02.010] [Medline: 30802642]
20. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice, 2nd edition: Prentice Hall; 2000.
21. McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires, 3rd edition: Oxford University Press; 2006.
22. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007 Jan;60(1):34-42. [doi: 10.1016/j.jclinepi.2006.03.012] [Medline: 17161752]
23. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Prac Assess Res Eval 2005;10(1):7. [doi: 10.7275/jyj1-4868]
24. Spicer J. Making Sense of Multivariate Data Analysis: An Intuitive Approach: SAGE Publications Inc; 2005. URL: https://books.google.co.in/books/about/Making_Sense_of_Multivariate_Data_Analys.html?id=ao_8EaoUuSwC&redir_esc=y [accessed 2026-03-28]
25. Matsunaga M. How to factor-analyze your data right: do’s, don’ts, and how-to’s. Int J psychol res 2010;3(1):97-110. [doi: 10.21500/20112084.854]
26. Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL. Practical considerations for using exploratory factor analysis in educational research. Prac Assess Res Eval 2013;18(1):6. [doi: 10.7275/qv2q-rk76]
27. Can S, van de Schoot R, Hox J. Collinear latent variables in multilevel confirmatory factor analysis: a comparison of maximum likelihood and Bayesian estimations. Educ Psychol Meas 2015 Jun;75(3):406-427. [doi: 10.1177/0013164414547959] [Medline: 29795827]
28. Williams B, Onsman A, Brown T. Exploratory factor analysis: a five-step guide for novices. Australas J Paramed 2010 Jan;8(3):1-13. [doi: 10.33151/ajp.8.3.93]

XSL•FO
RenderX

29. Gaskin CJ, Happell B. On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. Int J Nurs Stud 2014 Mar;51(3):511-521. [doi: 10.1016/j.ijnurstu.2013.10.005] [Medline: 24183474]
30. Cattell RB. The scree test for the number of factors. Multivariate Behav Res 1966 Apr 1;1(2):245-276. [doi: 10.1207/s15327906mbr0102_10] [Medline: 26828106]
31. Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika 1965 Jun;30:179-185. [doi: 10.1007/BF02289447] [Medline: 14306381]
32. Astephen JL, Deluzio KJ. Changes in frontal plane dynamics and the loading response phase of the gait cycle are characteristic of severe knee osteoarthritis application of a multidimensional analysis technique. Clin Biomech (Bristol) 2005 Feb;20(2):209-217. [doi: 10.1016/j.clinbiomech.2004.09.007] [Medline: 15621327]
33. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. Stat Med 2002 May 15;21(9):1331-1335. [doi: 10.1002/sim.1108] [Medline: 12111881]
34. Coni A, Mellone S, Colpo M, Bandinelli S, Chiari L. A factor analysis model of the instrumented Timed Up and Go test for physical capability assessment. Gait Posture 2018 Oct;66:S11-S12. [doi: 10.1016/j.gaitpost.2018.07.117]
35. Baracco S, Arcolin I, Corna S, Godi M, Giardini M. Instrumented Timed Up and Go test: a reliable tool for elderly with femur fracture. Gait Posture 2024 Oct;114:S4-S5. [doi: 10.1016/j.gaitpost.2024.08.018]
36. Wüest S, Massé F, Aminian K, Gonzenbach R, de Bruin ED. Reliability and validity of the inertial sensor-based Timed “Up and Go” test in individuals affected by Stroke. J Rehabil Res Dev 2016;53(5):599-610. [doi: 10.1682/JRRD.2015.04.0065] [Medline: 27898161]
37. Craig JJ, Bruetsch AP, Lynch SG, Horak FB, Huisinga JM. Instrumented balance and walking assessments in persons with multiple sclerosis show strong test-retest reliability. J Neuroeng Rehabil 2017 May 22;14(1):43. [doi: 10.1186/s12984-017-0251-0] [Medline: 28532417]
38. Moreno-Verdú M, Ferreira-Sánchez MDR, Martín-Casas P, Atín-Arratibel M. Imagined Timed Up and Go test (iTUG) in people with Parkinson’s disease: test-retest reliability and validity. Disabil Rehabil 2023 Mar 8;2023:1-11. [doi: 10.1080/09638288.2023.2185688] [Medline: 36890615]
39. Zhou J, Yao Q, Han R, et al. Reliability and validity of Instrumented Timed Up and Go test in typical adults and elderly: a systematic review. Arch Phys Med Rehabil 2025 Jul;106(7):1092-1107. [doi: 10.1016/j.apmr.2025.03.001] [Medline: 40054550]
40. Weiss A, Herman T, Plotnik M, Brozgol M, Giladi N, Hausdorff JM. An instrumented Timed Up and Go: the added value of an accelerometer for identifying fall risk in Idiopathic Fallers. Physiol Meas 2011 Dec;32(12):2003-2018. [doi: 10.1088/0967-3334/32/12/009] [Medline: 22094550]
41. Greene BR, Rutledge S, McGurgan I, et al. Assessment and classification of early-stage multiple sclerosis with inertial sensors: comparison against clinical measures of disease state. IEEE J Biomed Health Inform 2015 Jul;19(4):1356-1361. [doi: 10.1109/JBHI.2015.2435057] [Medline: 26087505]
42. Riva F, Bisi MC, Stagni R. Gait variability and stability measures: minimum number of strides and within-session reliability. Comput Biol Med 2014 Jul;50:9-13. [doi: 10.1016/j.compbiomed.2014.04.001] [Medline: 24792493]
43. Pasciuto I, Bergamini E, Iosa M, Vannozzi G, Cappozzo A. Overcoming the limitations of the harmonic ratio for the reliable assessment of gait symmetry. J Biomech 2017 Feb 28;53:84-89. [doi: 10.1016/j.jbiomech.2017.01.005] [Medline: 28104246]
44. Bateni H, Maki BE. Assistive devices for balance and mobility: benefits, demands, and adverse consequences. Arch Phys Med Rehabil 2005 Jan;86(1):134-145. [doi: 10.1016/j.apmr.2004.04.023] [Medline: 15641004]
45. Mobbs RJ, Perring J, Raj SM, et al. Gait metrics analysis utilizing single-point inertial measurement units: a systematic review. Mhealth 2022;8:9. [doi: 10.21037/mhealth-21-17] [Medline: 35178440]
46. Seo K, Jung J, Hwang J, Kim K, Kim SH. Assessing the impact of imu sensor location on spatio-temporal gait parameter estimation. Annu Int Conf IEEE Eng Med Biol Soc 2024 Jul;2024:1-4. [doi: 10.1109/EMBC53108.2024.10781929] [Medline: 40039154]
47. Küderle A, Roth N, Zlatanovic J, Zrenner M, Eskofier B, Kluge F. The placement of foot-mounted IMU sensors does affect the accuracy of spatial parameters during regular walking. PLoS ONE 2022;17(6):e0269567. [doi: 10.1371/journal.pone.0269567] [Medline: 35679231]
48. Höglund G, Grip H, Öhberg F. The importance of inertial measurement unit placement in assessing upper limb motion. Med Eng Phys 2021 Jun;92:1-9. [doi: 10.1016/j.medengphy.2021.03.010] [Medline: 34167702]
49. Böttinger MJ, Labudek S, Schoene D, et al. “TiC-TUG”: Technology in clinical practice using the instrumented Timed Up and Go test—a scoping review. Aging Clin Exp Res 2024 Apr 27;36(1):100. [doi: 10.1007/s40520-024-02733-7] [Medline: 38676844]
50. Bloomfield RA. In-Clinic Functional Measurement and Analysis of Knee Osteoarthritis Patients Undergoing Total Knee Replacement [Doctoral Dissertation]. : The University of Western Ontario (Canada); 2021 URL: https://ir.lib.uwo.ca/etd/7808/ [accessed 2026-03-18]
51. Gasparutto X, Gueugnon M, Laroche D, Martz P, Hannouche D, Armand S. Which functional tasks present the largest deficits for patients with Total Hip Arthroplasty before and six months after surgery? A study of the Timed Up-And-Go Test phases. PLoS ONE 2021;16(9):e0255037. [doi: 10.1371/journal.pone.0255037] [Medline: 34506498]

XSL•FO
RenderX

52. Overgaard JA, Kristensen MT. Minimal clinical important difference and responsiveness of performance-based and self-reported measures used in older adults following outpatient rehabilitation programme Post-Hip Fracture Surgery. Geriatr Orthop Surg Rehabil 2021:37. [doi: 10.1177/21514593211058764]
53. Santos ACAD, Matheus D, Battistella LR. Reference values for instrumented Timed Up and Go—outcomes of vertical accelerations and angular velocities in healthy subjects. J Int Soc Phys Rehabil Med 2025;8(1):1-5. [doi: 10.1097/ph9.0000000000000050]
54. He J, Wu L, Du W, et al. Instrumented Timed Up and Go test and machine learning–based levodopa response evaluation: a pilot study. J Neuroeng Rehabil 2024 Sep 18;21(1):163. [doi: 10.1186/s12984-024-01452-4] [Medline: 39294708]

Abbreviations

AUC: area under the curve
EFA: exploratory factor analysis
ICC: intraclass correlation coefficient
IMU: inertial measurement unit
iTUG: instrumented Timed Up and Go
KMO: Kaiser-Meyer-Olkin
LDA: linear discriminant analysis
MDC: minimum detectable change
PCA: principal component analysis
ROC: receiver operating characteristic
SEM: SE of measurement
THA: total hip arthroplasty
TKA: total knee arthroplasty
TUG: Timed Up and Go

Edited by S Munce; submitted 19.Aug.2025; peer-reviewed by M Tramontano, O Vickers, S Erel; accepted 26.Jan.2026; published 01.Apr.2026.

Please cite as:
Giardini M, Arcolin I, Arcobelli VA, Picardi M, Mellone S, Godi M
Reliability and Discriminant Ability of an Instrumented Timed Up and Go Test in People With Postsurgical Orthopedic Conditions: Quantitative Study
JMIR Rehabil Assist Technol 2026;13:e82632
URL: https://rehab.jmir.org/2026/1/e82632
doi:10.2196/82632

© Marica Giardini, Ilaria Arcolin, Valerio Antonio Arcobelli, Michela Picardi, Sabato Mellone, Marco Godi. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 1.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of an Exercise-Assisting Mobile App (Osteoarthritis-Rehabilitation Assistant [O-RA]) on Rehabilitation Outcomes in Older Adults: Randomized Controlled Parallel Clinical Trial

Pajaree Sornmayura[1], MD; Sintip Pattanakuhar[1,2], MD, PhD; Napaschol Intapan[3,4]; Krittipat Tragoolpua[3,4]; Kaewkla Sroykabkaew[3,4]; Rungkan Wangboon[4], MSc; Krittipong Wachirangkul[4], BSc; Sartita Wannarat[4], BFA; Jakkrit Klaphajone[1], MD

[1]Department of Rehabilitation Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[2]Department of Biomedical Engineering, Faulty of Engineering, University of Strathclyde, 106 Rottenrow East, Glasgow, United Kingdom

[3]Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand

[4]The Prince Royal's College, Chiang Mai, Thailand

Corresponding Author:
Sintip Pattanakuhar, MD, PhD
Department of Rehabilitation Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

Abstract

Background: Mobile apps and biofeedback using motion analysis have both been used separately to increase compliance with exercise programs. We developed a mobile app, Osteoarthritis-Rehabilitation Assistant (O-RA), that uses motion analysis technology in the mobile app to assist older adults with performing a knee exercise program.

Objective: This study aimed to evaluate the effects of the O-RA app on the compliance and correctness of the exercise program by older adults.

Methods: We conducted an assessor-blind, parallel-design, randomized controlled trial with 40 older adults (aged 60 years or older) who had no symptoms and no diagnosis of knee osteoarthritis. Participants were divided into 2 groups: O-RA app (intervention) group and standard treatment (control) group. Both groups were taught 4 types of exercise programs by a physical therapist for 15 minutes and were instructed to do exercises at home every day for 1 week. The number of exercises, the percentage between observed and prescribed exercises, the correctness of exercises, and overall pain during the program were assessed in both groups.

Results: The control group had significantly higher compliance with the exercise program than the intervention group (t_{38}=3.5044, P=.001). There was no statistically significant difference in the correctness of the exercise program between the intervention and control groups. The difficulty of use and satisfaction were 47 and 59, respectively, out of the full score of 100. The main problems were the instability and the difficulty using the app.

Conclusions: In older adults without knee osteoarthritis symptoms or diagnosis, the O-RA app was not a facilitator but a barrier to the lower extremity exercise program. An updated version, aiming to increase the stability and make it more user-friendly, should be developed; however, more comprehensive data, including qualitative user feedback and standardized usability metrics, will be needed to effectively guide its design.

Trial Registration: Thai Clinical Trial Record TCTR20240923002; https://www.thaiclinicaltrials.org/export/pdf/TCTR20240923002

(JMIR Rehabil Assist Technol 2026;13:e80971) doi:10.2196/80971

KEYWORDS

knee; osteoarthritis; rehabilitation; motion analysis; mobile app; compliance

Introduction

Osteoarthritis (OA) is a disease caused by the gradual degeneration of the joints [1]. There are many risk factors for the disease, such as increasing age, genetics, joint injuries, and abnormal knee structures or deformities [2]. The pathophysiology of OA involves an increase in force acting on the articular cartilage, causing a change in the synovial membrane and subchondral bone [3]. Knee joints are most commonly involved in OA. Currently, the treatment approach

XSL•FO
RenderX

for knee OA starts with nonpharmacological treatment followed by pharmacological treatment [4].

Exercise is the first-line conservative treatment for knee OA. A systematic review aiming to demonstrate the effects of exercise modalities found that several current guidelines for the treatment of OA of the knee, including the American College of Rheumatology; European Alliance of Associations for Rheumatology; Osteoarthritis Research Society International; and European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases [4,5], recommend aerobic exercise and leg muscle strengthening exercises; exercises to strengthen and endurance of the lower extremities, especially the knee extensors; exercises to increase knee flexibility or range; and low-impact aerobic exercises as a first-line nonpharmacological treatment with strong supporting evidence [4]. Focusing on the exercises for strengthening the knee extensors commonly used in clinical practice, Zacharias et al [6] conducted a systematic review and meta-analysis of the literature on strengthening exercises and found that they could increase knee extensor and flexor strength. This systematic review supports the beneficial effects of strengthening exercises on the muscle strength of knee extensors and flexors in patients with knee OA, although clinical outcomes such as pain reduction and functional improvement require further evaluation in larger clinical trials.

Although strengthening exercises are beneficial for patients with knee OA, emphasis should be placed on adherence to the exercises. A previous study found that only 60.5% of patients performed the exercises correctly after being taught by a physician, and this rate decreased to 48.8% at the 4-week follow-up [7]. Increasing adherence to self-administered home exercises by older adults is important.

Mobile health (mHealth) has emerged as a promising approach in health care, with the potential to support diverse activities across diverse populations including health education, increasing access to health care services, improving prevention strategies, enhancing treatment delivery, and supporting public health programs [8]. In the context of musculoskeletal rehabilitation, mHealth apps have been developed primarily to support treatment strategies and improve adherence to prescribed exercise programs [9]. The Osteoarthritis-Rehabilitation Assistant (O-RA) app described in this study was designed with this specific purpose: to assist older adults with performing knee exercise programs correctly and consistently through motion analysis technology integrated into a mobile app.

Several studies have used mobile apps to improve exercise efficiency in patients with knee OA. For example, Alasfour and Almarwani [10] studied the effects of an Arabic mobile app on exercise adherence and knee pain in patients with knee OA compared with standard care. They found that the app could significantly increase exercise rate and reduce pain. In addition, Biebl et al [11] found that the joint angle of exercise, measured using a mobile app, was not inferior to the joint angle of exercise measured by a physical therapist. In addition, a study by Thiengwittayaporn et al [12] that studied the effects of a Thai mobile app on the accuracy of home exercise in patients with knee OA found that the group using the mobile app had higher exercise accuracy.

One of the advantages of using a mobile app is that it can include a motion analysis tool in its features. Motion analyses during exercise can be used as a biofeedback method to improve the clinical results of the exercise program [13]. However, only the study by Biebl et al [11] integrated motion analysis into the mobile app. The Prince Royal College School and Chiang Mai University jointly developed the O-RA mobile app using the technique of motion analysis of the body with automatic learning of artificial neural networks (deep learning) technology to analyze exercise posture to increase the accuracy. We hypothesized that the O-RA app can enhance the exercise program by improving the compliance with and correctness of the exercise program. However, a phase 1 study aiming to prove the safety and efficacy of the intervention in older adults without disease should be conducted before using the app with patients with knee OA.

The primary objective of this study was to evaluate the effect of the O-RA app on improving compliance with the exercise program by comparing the number of times the exercise program was completed and percentage of the observed and prescribed numbers of exercises completed between a group using the O-RA app and a group using standard teaching. The secondary objectives were to evaluate the (1) effect of the O-RA app on the correctness of physical exercise and (2) safety of the exercise program by assessing whether participants in the intervention group experienced increased pain, measured using a visual analog scale (VAS), compared with the standard teaching group.

Methods

Study Design

This study was a parallel-design, assessor-blinded, phase 1 randomized controlled trial based on the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guideline. All parts of the study were conducted at the Outpatient Clinic of the Department of Rehabilitation Medicine, Maharaj Nakorn Chiang Mai Hospital, from January 2024 through November 2024.

Study Populations

Since this study was a phase 1 trial, the study population was older adult volunteers without knee pain. The inclusion criteria comprised of (1) age ≥60 years, (2) no previous formal instruction on physical exercises for knee OA, (3) the ability to participate in all study processes, (4) being able to read and write in the Thai language, (5) having a caregiver who can record the volunteers' daily physical exercises for the entire study, and (6) voluntary participation in the study for both volunteers and caregivers. The exclusions criteria were (1) having limitations in visibility or understanding how to use the O-RA app system; (2) the presence of contraindications to exercise, including heart disease, diabetes, or uncontrolled high blood pressure; (3) having knee pain or knee deformity or limited knee extension of more than 10 degrees or inability to bend the knee more than 90 degrees; and (4) having balance

XSL•FO
RenderX

problems, indicated by being unable to get up from a chair without using hands.

Sample Size Calculation

The sample size was calculated based on the primary outcome, which was compliance with the exercise program (ie, the percentage of the observed number of exercises compared with the number of prescribed exercises) according to the research by Alasfour and Almarwani [10] and calculated using the formula for comparing two independent means [14]. With an effect size of 28.05, an SD of the overall group of 28.54, an α of .05, and a β of .10, the sample size of each group should be 17. After accounting for 20% compensation for participants lost to follow-up, the final sample size was 40 (20 participants per group).

Study Protocol

We invited 40 older adult volunteers without knee pain by posting an announcement in Maharaj Nakorn Chiang Mai Hospital. Interested volunteers were contacted by the researchers, and an appointment was made to provide information and perform an assessment before participating in the research at the rehabilitation medicine outpatient examination room. In addition to a search within their medical history, they were asked about a history of chronic diseases and abnormal symptoms, including a physical examination to assess for abnormalities in the knee joints and balance, as well as to exclude those with contraindications for exercise.

After being invited to participate in the study and meeting the inclusion and exclusion criteria, participants were assessed for their ability to use the O-RA app correctly. If they were able to do so, they were asked to provide written informed consent. After that, the participants were randomized into either an intervention or control group using a computer-based simple randomization method. The randomization was conducted by a research assistant who was not involved in the exercise teaching and outcome assessment processes. The randomization sequence was stored in a sealed envelope in a locked drawer, and the assessor was blinded to it.

Volunteers were randomly divided into 2 groups: O-RA app group (intervention group) and standard teaching group (control group). Volunteers in the O-RA app group were taught how to use the O-RA app system on a dedicated smartphone on which the O-RA app had already been installed. They were then taught an exercise program for knee OA by a physical therapist for 15 minutes and were instructed to do the exercises using the O-RA app at home every day, 3 times a day for 1 week. The standard teaching group was taught an exercise program for knee OA by a physical therapist for 15 minutes and then instructed to do the exercises by themselves at home every day, 3 times a day for 1 week without using the O-RA app. Both groups were given an exercise logbook to record the number of exercises actually completed. The logbook was recorded by caregivers to ensure that the recording was not biased. After 1 week, the number of times and percentage of exercises in both groups were assessed using a specially developed assessment form and analyzed as completeness of the exercise program. Correctness of the exercise program was assessed by a rehabilitation physician in both groups. The score and percentage of correctness were then calculated. Pain during the exercise program was assessed using the pain VAS in both groups. Satisfaction with the use and ease of use were also assessed only in the O-RA app group. The researchers then conducted the statistical analyses to compare the completeness and correctness of the exercise program between the intervention and control groups.

Exercise Program

The exercise program in this study consisted of 6 types of lower extremity exercises specifically designed by one of the investigators who was a specialist in sports rehabilitation (JK; Figure 1).

Figure 1. Exercise programs in the Osteoarthritis-Rehabilitation Assistant (O-RA) app. Written informed consent was obtained from the individual for the publication of any potentially identifiable images or data included in this figure.



For the hip abduction exercise, the instructions were to stand up straight, hold the backrest of a chair or table edge, spread your legs about 30 - 45 degrees to the side, perform one side at a time, and alternate the legs. The benefits are strengthening the hip abductors and improving stability while standing.

The leg curl instructions were to stand up straight, hold the backrest of a chair or table edge, and alternate bending one knee

XSL•FO
RenderX

at a time. The benefits are strengthening the hamstrings and increasing stability while standing.

For the knee extension exercise, the instructions were to sit on the chair and alternate straightening one knee at a time. The benefits are strengthening the front thigh muscles and stretching the back thigh muscles.

The semi-squat instructions were to stand up straight, hold the backrest of a chair or table edge, slowly bend both knees until the knees are bent at about 45 degrees, and keep both feet on the ground. The benefits are strengthening the quadriceps and hip extensors and improving stability while standing; this exercise is more difficult than the knee extension.

O-RA App System

The O-RA app is a mobile app available only on the Android operating system. The main function of the O-RA app system is to create motion analysis from a camera that is built into the mobile phone (Figure 2). To obtain this function, a pose estimation group of artificial intelligence was used to detect the joint position of the human body, process it into mathematical values, and apply an algorithm to calculate the degree of the joint movement. In this study, we used Movenet, a deep learning model for detecting points on the human body that can capture 17 joint points on the body. For the script programming language, JavaScript and cascading style sheets (CSS) were used. For the user experience (UX) and user interface (UI) design, the Figma collaborative interface design tool was used. The minimum requirements of the mobile phone for installing and running the O-RA app are as follows: central processing unit (CPU), Qualcomm Snapdragon 730 Octa Core; graphics processing unit (GPU) Adreno 618; 6 GB RAM; and 64 GB ROM. In this study, we used 2 mobile phones, both Samsung Galaxy S23 Ultra models (CPU, Qualcomm Snapdragon 8 Gen 2 Octa Core; GPU, Adreno 740; 8 GB RAM; 256/512 GB ROM). Firebase Storage from Google Cloud Platform was used as memory storage.

Figure 2. Example of how the app works while doing exercises.





Technical Development of the Physiotherapy Module

For the physiotherapy component, an algorithm was developed to assess whether the user's exercise performance was correct. We used 3 adjacent keypoints from the MoveNet model output to calculate joint angles. First, the radian value was computed using the Math.atan2 function to determine the tangent of the angle (θ). This value was then multiplied by 180 and divided by π (3.1415), and the absolute value was obtained using Mathlabs, yielding the central angle between the 3 keypoints. After the model completed its prediction, it identified keypoints surrounding the center of the image, selecting the keypoint with the highest confidence score as the coordinate. Once MoveNet predicted the keypoint coordinates in the image, the developers fed the required coordinates into the mathematical formula to calculate the angle formed by 3 keypoints and verified whether

XSL•FO
RenderX

the resulting angle met the physiotherapy criteria for correct exercise performance.

Frontend Development (Mobile App)

The mobile app (Android) was developed using the React Native framework with JavaScript and CSS as the script programming languages. These languages enabled the integration of complex features into the mobile app, including smooth UX/UI design, user preference storage, and access to the mobile phone camera. React Native was selected because of its high performance in mobile environments and its structural similarity to the React framework, which facilitated the integration of the MoveNet model into the app. Additionally, the extensive and diverse package ecosystem of React Native provided the developers with multiple alternative solutions when encountering technical challenges during development.

For the app screen design, Figma was used as the collaborative interface design tool, enabling multiple users to work on the design simultaneously in real time and to export button graphics and other visual assets, which significantly streamlined the design process in React Native. For the in-app typography, the Fahkwang-Bold font was selected because its clear letterforms and distinct character heads facilitate readability for older adult users compared with standard fonts.

Backend and Database

Firestore from Google Cloud Platform was used as the database because of its fast data storage, retrieval, and transmission capabilities. Firestore is cost-effective for small-scale applications, with no charges incurred when data storage and transfer volumes remain low while being easily scalable as the user base grows. Firestore also provides high security levels, with access rights that can be restricted to individual users or administrators, and data encryption before storage in the cloud database. Firebase Storage from Google Cloud Platform was used to store user profile images, selected for its seamless integration with Firestore and comparable security features. Firebase Authentication was used for the login and registration system because of its multilayered data encryption and user management capabilities.

Speech Synthesis

For speech synthesis in the O-RA app, the *React-Native-tts* and *Expo-speech* packages were used to synthesize Thai-language voice instructions within the app, enhancing ease of use and convenience for older adult users.

System Testing

The O-RA app was developed iteratively with input from clinical experts, including physiotherapists and rehabilitation physicians. Prior to the trial, a pilot test was conducted with a small group of users to identify and resolve critical usability issues. However, a formal usability evaluation using validated instruments (eg, System Usability Scale or Technology Acceptance Model) was not conducted prior to the trial, which is acknowledged as a limitation of this study. The app includes short exercise demonstration videos and step-by-step diagrams to guide correct exercise performance. A notification system during exercising is also included to remind users to perform their exercises. We plan to include more customizable reminder options, such as opt-in text messages at user-chosen intervals (eg, daily, 3 times per week), in future versions.

Outcome Determination

The exercise program was considered complete when all 4 exercises were performed (on both sides for knee extension, leg curl, and hip abduction, and 2 times for semi-squat) at least 8 out of 10 times for each exercise. The correctness was measured according to the case record form and by calculating the score and percentage of the score achieved. The number of times and percentages of exercises were compared, and the scores and percentages of correctness of exercises were compared between the group using the O-RA app and the group using standard teaching methods. The VAS for pain during the exercise program was assessed as an adverse effect outcome. Satisfaction with and difficulty using the O-RA app were rated by the participants in the intervention group using a structured questionnaire with a VAS (0 - 100), and obstacles to using the system were also collected through open-ended narrative questions.

Statistical Analysis

Descriptive statistics, including frequencies and percentages for categorical variables or means (SDs) and medians (IQRs) for continuous variables, were used to describe demographic characteristics such as age, gender, and underlying diseases and other relevant variables of the study participants.

For continuous outcomes (compliance with the exercise program indicated by number of times and percentage of exercises, correctness of the exercise program indicated by exercise accuracy scores and percentage, pain VAS), inferential statistics were used. First, a Shapiro-Wilk test was applied to both outcomes to determine whether the distributions were normal. An independent *t* test was used if the parameter was normally distributed, whereas the Mann-Whitney *U* test was used if the parameter was not normally distributed. The Fisher exact test was used to compare categorical data, including the number of participants who exercised at least 80% of the time every day, the number of participants who correctly exercised, and the number of participants who experienced pain (as a safety monitoring measure), between the O-RA app group and the standard teaching group. For variables with non-normal distributions (as determined using the Shapiro-Wilk test), the median and IQR were reported; for normally distributed variables, the mean and SD were reported. The 2-tailed significance level was set at $P<.05$. The data were analyzed using an intention-to-treat method with a superiority design.

Ethical Considerations

This study was conducted in accordance with the ethical standards of the responsible institutional and national committees on human experimentation and with the World Medical Association Declaration of Helsinki (1975, as revised in 2013). The study protocol was approved by the Institutional Ethics Committee of the Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand (study code: REH-2566 - 0329).

XSL•FO
RenderX

Written informed consent was obtained from all participants prior to randomization and data collection. Before providing consent, prospective participants received verbal and written information from a member of the research team describing the purpose of the study, exercise intervention procedures, expected duration of involvement, potential benefits and risks, data handling arrangements, and right to withdraw at any time without penalty. Caregivers responsible for maintaining participants' exercise logbooks were informed of the study procedures and voluntarily agreed to their role.

All participant data were handled in accordance with the data protection policies of the Faculty of Medicine, Chiang Mai University, and with the Thai Personal Data Protection Act B.E. 2562 (2019) [15]. Study records were pseudonymized using participant identification codes at the point of data entry; the key linking identifiers to codes were stored separately from the study dataset in a locked, access-restricted location accessible only to named members of the research team. Paper case record forms and exercise logbooks were stored in locked cabinets within the Department of Rehabilitation Medicine, Maharaj Nakorn Chiang Mai Hospital. Electronic data collected through the O-RA app were transmitted over encrypted connections and stored on Firestore and Firebase Storage (Google Cloud Platform), with access restricted to named investigators and developers via Firebase Authentication. Where identifiable information is present in figures, we collected informed consent from the individual.

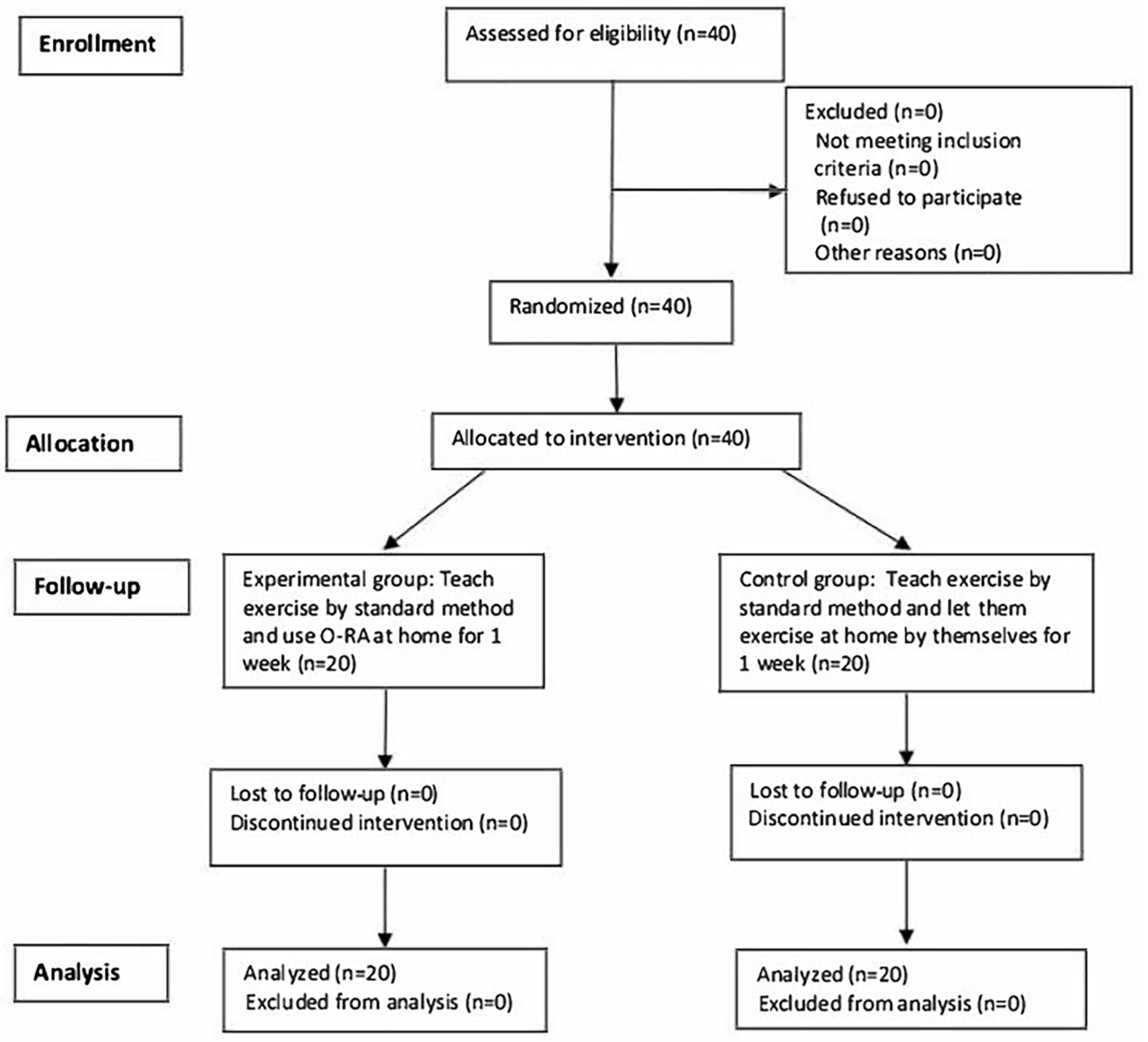
Each participant received 300 (US $9.21) as compensation for travel expenses associated with the in-person study visits.

Results

Recruitment of the Participants

During the 11-month data collection period, 40 participants were recruited. All participants completed the study protocol without loss to follow-up, and data from all participants were analyzed. Since there were no missing data, no imputation method was used. The CONSORT (Consolidated Standards of Reporting Trials) study flow diagram is presented in Figure 3.

Figure 3. Study flow according to the CONSORT (Consolidated Standards of Reporting Trials) diagram. O-RA: Osteoarthritis-Rehabilitation Assistant.



XSL•FO
RenderX

Baseline Characteristics of the Participants

Of the 40 recruited participants, the mean (SD) age was 67.2 (5.8) years. Most were female (33/40, 83%). The most common underlying disease was hypertension (20/40, 50%), followed by dyslipidemia (17/40, 43%). One-half of the participants had an education level higher than secondary school level (20/40, 50%), and 9 participants (9/40, 23%) were still working. When comparing the intervention and control groups, there were no statistically significant differences in age, sex, underlying disease, education level, and working status (Table 1).

Table . Baseline characteristics of the participants, with the significance level set at $P<.05$.

Characteristics	Overall (n=40)	Intervention group (n=20)	Control group (n=20)	Statistical test result (*df*)	*P* value
Age (years), mean (SD)	67.2 (5.8)	67.6 (6.0)	66.9 (5.6)	–0.0800 (38)[a]	.71
Sex (female), n (%)	33 (83)	17 (85)	16 (80)	—[b]	>.99
Underlying disease, n (%)					
Hypertension	20 (50)	9 (45)	11 (55)	—[b]	.75
Diabetes	8 (20)	3 (15)	5 (25)	—[b]	.70
Dyslipidemia	17 (43)	7 (35)	10 (50)	—[b]	.52
Others	13 (33)	6 (30)	7 (35)	—[b]	>.99
Highest education level, n (%)				—[b]	.48
No education	2 (5)	0 (0)	2 (10)		
Primary school	9 (22)	4 (20)	5 (25)		
Secondary school	7 (7)	3 (15)	4 (20)		
Higher degree	22 (55)	13 (65)	9 (45)		
Currently working, n (%)	9 (23)	5 (25)	4 (20)	—[b]	>.99

[a]Independent *t* test.

[b]2-tailed Fisher exact test.

Effects of the O-RA App on Compliance With the Exercise Program

Compliance with the exercise program was assessed using 3 outcome parameters, including the number of exercises per week, the percentage of observed number of exercises to the prescribed number of exercises, and the number of participants who completed at least 80% of the exercises every day. Notably, we used the percentage between the observed number of exercises and the prescribed number of exercises as a primary outcome since it was an outcome presented in the previous study we used as a reference for calculating the sample size [10].

Focusing on the primary outcome, participants in the intervention groups had significantly lower observed/prescribed number of exercise percentages than those in the control group for each type as well as all types of exercise (mean compliance: intervention, 75%, SD 22%; control, 94%, SD 8%; t_{38}=3.5044; P=.001). We also found that participants in the intervention group had a significantly lower number of exercises per week than those in the control group for each type as well as all types of exercise (intervention: mean 1264, SD 373; control, mean 1575, SD 135; t_{38}=3.5044; P=.001). In addition, the number of participants who completed all exercises every day, at least 80% of the time, was significantly lower in the intervention group than in the control group for hip abduction (intervention, 3/20, 15%; control, 10/20, 50%; P=.04) and leg curl (intervention, 3/20, 15%; control, 10/20, 50%; P=.04) exercise types but not significantly different for knee extension and semi-squat exercise types or for all types of exercise (Table 2).

XSL•FO
RenderX

Table . Comparison of the compliance outcomes between groups, with the significance level set at $P<.05$.

Outcomes	Overall (n=40)	Intervention group (n=20)	Control group (n=20)	Statistical test result (*df*)	*P* value
Compliance (number of exercises per week), mean (SD)					
Hip abduction (n=420)	360 (76)	323 (90)	397 (31)	3.5007 (38)[a]	.001
Leg curl (n=420)	358 (80)	318 (95)	397 (31)	3.5246 (38)[a]	.001
Knee extension (n=420)	356 (80)	322 (91)	390 (48)	2.9709 (38)[a]	.005
Semi-squat (n=420)	346 (93)	301 (109)	319 (40)	3.4567 (38)[a]	.001
All types of exercise (n=1640)	1420 (318)	1264 (373)	1575 (135)	3.5044 (38)[a]	.001
Compliance (%; observed number of exercises/prescribed number of exercises × 100), mean (SD)					
Hip abduction	86 (18)	77 (21)	95 (7)	3.5007 (38)[a]	.001
Leg curl	85 (19)	76 (23)	95 (7)	3.5426 (38)[a]	.001
Knee extension	85 (19)	77 (22)	93 (12)	2.9709 (38)[a]	.005
Semi-squat	82 (22)	72 (26)	93 (10)	3.4567 (38)[a]	.001
Overall compliance	85 (19)	75 (22)	94 (8)	3.5044 (38)[a]	.001
Compliance (number of participants who completed at least 80% of all exercises every day), n (%)					
Hip abduction	13 (33)	3 (15)	10 (50)	—[b]	.04
Leg curl	13 (33)	3 (15)	10 (50)	—[b]	.04
Knee extension	12 (30)	3 (15)	9 (45)	—[b]	.08
Semi-squat	12 (30)	3 (15)	9 (45)	—[b]	.08
All exercises	12 (30)	3 (15)	9 (45)	—[b]	.08

[a]Independent *t* test.

[b]2-tailed Fisher exact test.

Effects of the O-RA App on Correctness of Following the Exercise Program

Correctness of the exercise program was assessed using the correctness score and the number of participants who correctly exercised for each type as well as all types of exercise. We found no statistically significant differences in the correctness score and the number of participants who correctly exercised during each type of exercise as well as all types of exercise between the intervention and control groups (Table 3).

XSL•FO
RenderX

Table . Comparison of the correctness and pain outcomes between groups, with the significance level set at *P*<.05.

Outcomes	Overall (n=40)	Intervention group (n=20)	Control group (n=20)	Statistical test result (*df*)	*P* value
Correctness score (maximum=100), mean (SD)					
Hip abduction	80 (30)	83 (29)	78 (30)	3.5007 (38)[a]	.60
Leg curl	85 (26)	85 (29)	85 (24)	3.5246 (38)[a]	>.99
Knee extension	88 (27)	85 (33)	90 (21)	2.9709 (38)[a]	.59
Semi-squat	86 (25)	83 (29)	90 (21)	3.4567 (38)[a]	.36
Overall correctness	85 (21)	84 (26)	86 (15)	3.5044 (38)[a]	.78
Correctness (number of participants who correctly exercised), n (%)					
Hip abduction	26 (65)	14 (70)	12 (60)	—[b]	.74
Leg curl	29 (73)	15 (75)	14 (70)	—[b]	>.99
Knee extension	32 (80)	16 (80)	16 (80)	—[b]	>.99
Semi-squat	30 (75)	14 (70)	16 (80)	—[b]	.72
All exercises	19 (48)	11 (55)	8 (40)	—[b]	.53
Pain VAS[c], median (IQR)	0 (0-0)	0 (0-0)	0 (0-0)	192.5[d]	.72
Participants who experienced pain, n (%)	5 (13)	3 (15)	2 (10)	—[b]	>.99

[a]Independent *t* test.

[b]2-tailed Fisher exact test.

[c]VAS: visual analog scale.

[d]Mann-Whitney *U* test.

Adverse Effects, Satisfaction, and Difficulty With Using the O-RA App

Adverse effects of exercise were assessed using the pain VAS during the week of performing the exercise program and the number of participants who experienced pain. We found no differences in median pain VAS between the intervention and control groups (intervention, 0; control, 0; *U*=192.5; *P*=.73). In addition, there was no difference in the numbers of participants who experienced pain between the intervention and control groups (3/20, 15% vs 2/20, 10%; *P*>.99; Table 3).

Table . Difficulty and satisfaction with using the Osteoarthritis-Rehabilitation Assistant (O-RA) app (n=20).

Parameters	Results, mean (SD)	Results, median (IQR)	Results, range
Satisfaction	47 (32)	43 (22-77)	0-100
Difficulty	59 (32)	60 (38-84)	0-100

According to the satisfaction level and feasibility of the O-RA app for participants in the intervention group, the median satisfaction with the O-RA app was 43 (IQR 22-77), and the median difficulty using the O-RA app was 60 (IQR 38-84; Table 4). Notably, participants in the intervention group who exercised every day, at least 80% of the time (n=3), had a lower difficulty score than those who did not (n=17; median difficulty score: 30 vs 65). No statistical test was conducted due to a very small sample size (n=3) in one group.

Focusing on narrative comments on the use of the O-RA app, two participants reported that the app was not stable (could not be opened); 3 participants reported that it wrongly detected their body motions, which caused too many repetitions of the exercise; and 2 participants complained that the app requires a continuous WiFi internet connection during the entire exercise program. However, their WiFi internet connection was not stable, which made the exercise program incomplete. Overall, 8 participants (8/20, 40%) indicated that it was too difficult for them to use the O-RA app as a routine exercise assistant.

Discussion

Principal Findings

This study aimed to demonstrate the effect of the O-RA app on increasing compliance and correctness of the exercise program for knee OA. Since this is the first study using the O-RA app with humans, we conducted this assessor-blinded, randomized controlled clinical trial with older adults who had not been

XSL•FO
RenderX

diagnosed with and had no symptoms of knee OA for safety purposes. Therefore, this study should be considered as a phase 1 or early exploratory study. We hypothesized that the O-RA app can increase the compliance with and correctness of the exercise program. Contrary to our hypothesis, results from this study demonstrated that the O-RA app was not a facilitator but a barrier to the exercise program. In addition, the correctness of the exercise program was not different between the groups that used and did not use the O-RA app. These unexpected results indicate that this version of the O-RA app may not be appropriate for improving exercise adherence and an improved version should be developed. Beyond the instability and the requirement for continuous WiFi connectivity, participants also reported difficulty navigating the interface, lack of familiarity with smartphone technology, and a preference for in-person physiotherapy guidance as additional reasons for reduced app usage. A formal usability assessment using standardized instruments (eg, System Usability Scale) would be valuable for systematically characterizing these barriers in future studies.

The explanation for why participants who used the O-RA app had lower compliance with the exercise program than participants who did not use it may be associated with the difficulty using the app. Although the developers of the O-RA app tested its usability in their proof-of-concept experiment, it was never tested in a clinical setting before this study. In a real-life clinical setting, users of the app are varied in many aspects, including knowledge, skills, and experiences regarding mobile app use, which may cause difficulty when using the mobile app to facilitate exercise program compliance. This problem may even worsen exercise adherence by making the participant frustrated. Notably, optimal usability study design involves recruiting end users with variable skills and knowledge levels to obtain broad-based feedback from both tech-savvy and tech-challenged users. This explanation is confirmed by the results of this study, which demonstrated that participants in the intervention group who completely exercised at least 80% every day had a lower difficulty score than those who did not complete the exercises. The main problems that made the use of the O-RA app difficult were (1) stability of app access, (2) invalidity of movement detection, (3) requirement for a continuous WiFi internet connection, and (4) overall difficulty for Thai older adults. Therefore, the developers of this app should consider these pain points and improve the app accordingly. For example, built-in data storage in the app instead of using cloud storage should be considered to increase the stability of app access and make the app usable without a continuous WiFi internet connection. In addition, a user-friendly UX/UI system should be specifically developed to facilitate use by Thai older adults, such as large font size and easy-to-understand calibration processes and performance steps. Additionally, options should be built into the app to accommodate users with variable levels of health literacy and technical capability, such as a simplified navigation mode and multilingual support.

Focusing on the effect of each exercise type, a trend of increased efficacy of the intervention for more difficult exercise types was observed. For example, the intervention group had a significantly lower percentage of participants who exercised at least 80% of the time every day for hip abduction and leg curl but not for knee extension and semi-squat exercises, resulting in a nonsignificant difference in the overall exercise program. This discrepancy may indicate that the O-RA app should be more effective at improving compliance with the more complicated exercises. Most of the exercises used in this study are relatively easy; therefore, the participants can perform them without problems. Adding the use of the O-RA app to these easy-to-perform exercises may make the exercise activities more complicated and redundant. Therefore, it may not increase but even decrease compliance with the exercise program, causing lower compliance in the intervention group than in the control group.

According to the correctness outcomes, since the exercise program was relatively easy, no effect of the O-RA app was observed since the correctness reached the ceiling in the control group. It may be both interesting and more valuable to add only the O-RA app to complicated exercises, such as semi-squat exercises, back extension exercises, or back flexion exercises, to make the potential positive effect of the O-RA app more evident. In addition, no significant increase in pain was observed in the intervention group when compared with the pain in the control group. This result can confirm the safety of O-RA app use by older adults who have no diagnosis and no symptoms of knee OA. However, since the safety outcome was not the primary outcome, the sample size calculated from the efficacy outcome is likely inadequate, which may cause an underpowered safety result. Therefore, adverse effects of using the O-RA app should be monitored in future studies.

When compared with results from previous studies, there were some differences since most of them demonstrated significant effects of mobile apps for facilitating an exercise program. These inconsistencies may be due to the differences in study methodology as well as the mobile app used between this study and the previous ones. For example, the study by Alasfour and Almarwani [10], which demonstrated a significant improvement in adherence to the exercise program between the mobile app group and the handout group, used a mobile app that did not include motion analysis. The study by Thiengwittayaporn et al [12], which demonstrated a significant improvement in the correctness of the exercises between the mobile app group and the handout group, also used a mobile app that did not include motion analysis. We hypothesized that, although motion analysis is a powerful tool for creating a personalized exercise program, its complicated function may make the users, especially those who are not familiar with this technology, frustrated, causing unfavorable compliance and correctness outcomes. On the other hand, the study by Biebl et al [11] demonstrated a noninferior effect on the correctness of the exercises in the motion analysis–based mobile app group and physiotherapist-supervised group. This inconsistency may be due to the difference in the mobile app used between the studies. In the study by Biebl et al [11], the motion analysis mobile app was easier to use than our mobile app since it required no calibration at the beginning of each exercise session. In addition, the exercise program in the study by Biebl et al [11] was more complicated than the program in our study. Therefore, motion analysis feedback may add some value. These issues address the importance of the

XSL•FO
RenderX

feasibility of mobile apps and the selection of an exercise program where motion analysis should be added.

Strengths and Limitations

We would like to appraise our study and discuss its strengths and limitations using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluations) paradigm [16]. According to this paradigm, 5 issues relevant to this study that affect the certainty or quality of evidence should be considered, including study design, risk of bias, imprecision, inconsistency, and indirectness of the evidence. Its strength is related to the study design, which is a phase 1 randomized controlled trial providing preliminary evidence with a relatively high level of certainty and quality. As a preliminary finding, the effect size of the primary outcome has a small SD, although the precision of these estimates should be confirmed in a larger phase 2 trial. However, the results are internally consistent since the effects of the intervention on all outcomes were in the same direction (ie, favoring the control), which indicates a problem regarding the use of the O-RA app. Notably, as a phase 1 trial, these findings are exploratory and should be interpreted with caution.

According to the limitations of this study, first, potential biases in this study should be discussed since they may introduce systematic error that makes the results deviate from the truth. The first issue is selection bias (ie, unequal baseline characteristics between the comparison groups). Since this study used a valid randomization method (simple randomization) and allocation concealment (a sealed envelope) and the statistical tests demonstrated that baseline characteristics were not different between the groups, there may not be problematic selection bias in this study. The second issue is performance bias, both for participants and investigators. In this study, the investigators (the physiotherapist and rehabilitation physician) but not the participants were blinded to the study group, which may cause a risk of performance bias from the participant side. For example, participants who knew that they were in the control group tried to complete the exercise program as much as possible since they wanted to be a responsible person from the perspectives of health care providers. This behavior, known as the Hawthorne effect, may cause a higher increase in exercise compliance in the control group, resulting in performance bias that reduces the effect size. For the detection bias, since this study mostly used the objective outcome (compliance was evaluated by the family member, and correctness was assessed by the investigator), there may be a low risk of detection bias. Since this study has no missing participants and all outcomes were reported according to the protocol, it may have a low risk of attrition and reporting bias. Therefore, this study may have a risk of performance bias from a lack of participant blinding. However, it is difficult to blind the participants of an exercise intervention [17]. One potential approach to mitigate this limitation is to have the control group perform sham exercises, which is a method commonly used in exercise intervention trials to reduce performance bias. Next, according to the indirectness issue, the study population of this study was older adults without knee OA symptoms or diagnosis. Therefore, the results cannot be generalized to older adults who already have knee OA symptoms or diagnosis. Performing exercises by a population with the disease may be more challenging than by those without disease; therefore, further studies should be conducted in the target population. Additionally, this study did not use a standardized app usability evaluation tool (eg, the System Usability Scale or Technology Acceptance Model), which limits our ability to systematically characterize usability barriers. Future iterations of the study should incorporate such validated instruments.

Clinical and Research Implications

For clinical implications, we suggest not using this version of the O-RA app for increasing compliance with a lower extremity exercise program in older adults with knee OA symptoms or diagnosis. Since there is no direct evidence, we also suggest not using this version of the O-RA app for increasing compliance with the exercise program in older adults who already have knee OA symptoms or a diagnosis.

For research implications, an improved version of the O-RA app should be developed, one that is more stable, more user-friendly, and easier to use. After that, a study aiming to evaluate the usability and reliability of the new version of the O-RA app should be conducted before any further clinical trials. In addition, further studies aiming to adjust the O-RA app to assist users performing more complicated exercise programs should be considered. Further studies should also be conducted in another target population, such as those who already have knee OA symptoms or diagnosis, before using the O-RA app, preferably an updated one, with that specific population.

Conclusions

Results from this randomized controlled parallel trial demonstrate that O-RA, a motion analysis–based mobile app, was not a facilitator but a barrier for older adults without knee OA symptoms or diagnosis to complete the lower extremity exercise program. This result may be due to the instability of and difficulty using the app. An updated version, aiming to increase the stability and be more user-friendly, should be developed. More comprehensive data, including qualitative user feedback and standardized usability metrics (eg, System Usability Scale), should be collected to effectively guide the design of future versions.

Acknowledgments

We would like to thank Ms Saifon Phatpaeng and Ms. Nantawan Daolert, PT, for assisting with data collection. No generative artificial intelligence (AI) was used during the development of this manuscript.

XSL•FO
RenderX

Funding

This study was funded by the Faculty of Medicine Chiang Mai University Research Fund (Contract Number: REH- 2566-0329). The funder had no role in any process of the study.

Data Availability

The datasets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

NI, KT, KS, RW, KW, and SW developed the Osteoarthritis-Rehabilitation Assistant (O-RA) app. PS, SP, and JK designed the study. PS, NI, KT, and KS collected the data. PS and SP analyzed the data. PS and SP wrote and revised the manuscript. SP and JK critically reviewed the manuscript. All authors reviewed the final manuscript before submission.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH checklist.
[PDF File, 177 KB - rehab_v13i1e80971_app1.pdf]

References

1. Pereira D, Peleteiro B, Araújo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. Osteoarthr Cartil 2011 Nov;19(11):1270-1285. [doi: 10.1016/j.joca.2011.08.009]
2. Koedwan C, Bun in J, Teerasimbut C, Kuptniratsaikul V. The prevalence of knee OA in community-based elders [Article in Thai]. Thai Journal of Physical Therapy 2016;38(2):59-70 [FREE Full text]
3. Martel-Pelletier J, Pelletier JP. Is osteoarthritis a disease involving only cartilage or other articular tissues? Eklem Hastalik Cerrahisi 2010 Apr;21(1):2-14. [Medline: 20302555]
4. Ferreira RM, Martins PN, Gonçalves RS. Non-pharmacological and non-surgical interventions to manage patients with knee osteoarthritis: an umbrella review 5-year update. Osteoarthr Cartil Open 2024 Sep;6(3):100497. [doi: 10.1016/j.ocarto.2024.100497] [Medline: 39040626]
5. Golightly YM, Allen KD, Caine DJ. A comprehensive review of the effectiveness of different exercise programs for patients with osteoarthritis. Phys Sportsmed 2012 Nov;40(4):52-65. [doi: 10.3810/psm.2012.11.1988] [Medline: 23306415]
6. Zacharias A, Green RA, Semciw AI, Kingsley MIC, Pizzari T. Efficacy of rehabilitation programs for improving muscle strength in people with hip or knee osteoarthritis: a systematic review with meta-analysis. Osteoarthr Cartil 2014 Nov;22(11):1752-1773. [doi: 10.1016/j.joca.2014.07.005]
7. Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthr Cartil 2008 Feb;16(2):137-162. [doi: 10.1016/j.joca.2007.12.013]
8. Liaqat M, Mushtaq M, Jamil A, et al. Mobile health interventions: a frontier for mitigating the global burden of cardiovascular disease. Cureus 2024 Jun;16(6):e62157. [doi: 10.7759/cureus.62157] [Medline: 38993461]
9. Ryan S, Ní Chasaide N, O' Hanrahan S, Corcoran D, Caulfield B, Argent R. mHealth apps for musculoskeletal rehabilitation: systematic search in app stores and content analysis. JMIR Rehabil Assist Technol 2022 Aug 1;9(3):e34355. [doi: 10.2196/34355] [Medline: 35916688]
10. Alasfour M, Almarwani M. The effect of innovative smartphone application on adherence to a home-based exercise programs for female older adults with knee osteoarthritis in Saudi Arabia: a randomized controlled trial. Disabil Rehabil 2022 Jun;44(11):2420-2427. [doi: 10.1080/09638288.2020.1836268] [Medline: 33103499]
11. Biebl JT, Rykala M, Strobel M, et al. App-based feedback for rehabilitation exercise correction in patients with knee or hip osteoarthritis: prospective cohort study. J Med Internet Res 2021 Jul 13;23(7):e26658. [doi: 10.2196/26658] [Medline: 34255677]
12. Thiengwittayaporn S, Wattanapreechanon P, Sakon P, et al. Development of a mobile application to improve exercise accuracy and quality of life in knee osteoarthritis patients: a randomized controlled trial. Arch Orthop Trauma Surg 2023 Feb;143(2):729-738. [doi: 10.1007/s00402-021-04149-8] [Medline: 34453570]
13. Kudo M, Watanabe K, Otsubo H, et al. Analysis of effectiveness of therapeutic exercise for knee osteoarthritis and possible factors affecting outcome. J Orthop Sci 2013 Nov;18(6):932-939. [doi: 10.1007/s00776-013-0443-9] [Medline: 24085378]
14. Chow SC, Wang H, Shao J. Sample Size Calculations in Clinical Research, 2nd edition: Chapman and Hall/CRC; 2007. [doi: 10.1201/9781584889830]

XSL•FO
RenderX

15. Personal Data Protection Act, B.E. 2562. Kingdom of Thailand. 2019. URL: https://www.mdes.go.th/uploads/tinymce/source/%E0%B8%AA%E0%B8%84%E0%B8%AA/Personal%20Data%20Protection%20Act%202019.pdf [accessed 2026-04-08]
16. Guyatt GH, Oxman AD, Kunz R, et al. What is "quality of evidence" and why is it important to clinicians? BMJ 2008 May 3;336(7651):995-998. [doi: 10.1136/bmj.39490.551019.BE] [Medline: 18456631]
17. Thakur B, Ayers GD, Atem F, DeClercq JJ, Jain NB. Statistical and methodological considerations for randomized controlled trial design in physical medicine and rehabilitation. Am J Phys Med Rehabil 2023 Oct 1;102(10):855-860. [doi: 10.1097/PHM.0000000000002225] [Medline: 36882301]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
CPU: central processing unit
CSS: cascading style sheets
GPU: graphics processing unit
GRADE: Grading of Recommendations, Assessment, Development, and Evaluations
mHealth: mobile health
O-RA: Osteoarthritis-Rehabilitation Assistant
OA: osteoarthritis
UI: user interface
UX: user experience
VAS: visual analog scale

Edited by S Munce; submitted 20.Jul.2025; peer-reviewed by R Marshall, Z Zandesh; revised version received 17.Mar.2026; accepted 23.Mar.2026; published 15.Apr.2026.

Please cite as:
Sornmayura P, Pattanakuhar S, Intapan N, Tragoolpua K, Sroykabkaew K, Wangboon R, Wachirangkul K, Wannarat S, Klaphajone J
Effects of an Exercise-Assisting Mobile App (Osteoarthritis-Rehabilitation Assistant [O-RA]) on Rehabilitation Outcomes in Older Adults: Randomized Controlled Parallel Clinical Trial
JMIR Rehabil Assist Technol 2026;13:e80971
URL: https://rehab.jmir.org/2026/1/e80971
doi:10.2196/80971

© Pajaree Sornmayura, Sintip Pattanakuhar, Napaschol Intapan, Krittipat Tragoolpua, Kaewkla Sroykabkaew, Rungkan Wangboon, Krittipong Wachirangkul, Sartita Wannarat, Jakkrit Klaphajone. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 15.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone-Supported Vestibular Rehabilitation in Individuals With Vestibular Dysfunction: Pilot Randomized Crossover Trial Assessing Functional Clinical Outcomes and Anxiety

Azriel Kaplan[1], MSc, PT; Liran Kalderon[1], PT, PhD; Amit Wolfovitz[2], MD; Yoav Gimmon[2,3*], PT, PhD; Shelly Levy-Tzedek[1,4,5*], PhD

[1]Department of Physical Therapy, Recanati School for Community Health Professions, Ben-Gurion University of the Negev, Beersheva, Israel

[2]Department of Otolaryngology—Head and Neck Surgery, Sheba Tel-HaShomer Medical Centre, Ramat Gan, Israel

[3]Department of Physical Therapy, Faculty of Social Welfare & Health Sciences, University of Haifa, Eshkol Tower, Department of Physical Therapy, Haifa, Israel

[4]Zelman Center for Neuroscience, Ben-Gurion University of the Negev, Beersheva, Israel

[5]Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Freiburg, Germany

[*]these authors contributed equally

Corresponding Author:
Yoav Gimmon, PT, PhD
Department of Otolaryngology—Head and Neck Surgery, Sheba Tel-HaShomer Medical Centre, Ramat Gan, Israel

Related Articles:

Comment in: https://rehab.jmir.org/2026/1/e107851

Comment in: https://rehab.jmir.org/2026/1/e109396

Abstract

Background: Vestibular disorders impair balance, increase fall risk, and reduce quality of life due to dizziness and vertigo. They are frequently accompanied by heightened anxiety, which may further limit daily functioning and contribute to avoidance behaviors. Although vestibular rehabilitation has been extensively studied and shown to be effective in managing vestibular disorders, adherence to home-based exercises remains low for many dizzy patients. This is often attributed to uncertainty about correct performance, lack of feedback, or difficulty maintaining a structured routine. To help address these barriers, Vestibulon, a smartphone app co-designed with clinicians and patients, was developed to support rehabilitation practice by providing guided exercise, structured scheduling, progress monitoring, and clear instructions intended to promote confidence and engagement.

Objective: We aim to evaluate the potential contribution of a smartphone-based app to vestibular rehabilitation outcomes and to explore the relationship between dizziness-related disability and anxiety during the intervention.

Methods: This randomized 2-period crossover pilot trial included 20 adults with vestibular dysfunction (mean age 52, SD 12 y) who completed 6 weeks of rehabilitation. Participants were randomized to begin with either app-supported or conventional treatment before crossing over to the alternate condition after 3 weeks. This design enabled each participant to experience both modes of rehabilitation. Assessments were conducted at baseline (T0), mid-study (after 3 wk; T1), and end of study (T2). Outcome measures included the Dizziness Handicap Inventory (DHI) to assess dizziness-related disability, the State-Trait Anxiety Inventory to evaluate anxiety, and the Instrumented Timed Up and Go test to quantify functional mobility. The primary outcome was the change in DHI scores across time points.

Results: Significant improvements in DHI were observed between T0 and T2 in both groups (median change: app first=34, IQR 6 - 35 points, P=.006; conventional first=18, IQR 10 - 19 points, P=.009). Improvement in Instrumented Timed Up and Go performance was observed only when the app-supported phase occurred first (Z=−2.45, P=.01), suggesting a potential early benefit of structured smartphone guidance. State-Trait Anxiety Inventory scores did not change significantly in either sequence. Across all time points, dizziness-related disability and state anxiety demonstrated a consistent and significant moderate correlation (r=0.64, P<.001), emphasizing the strong interplay between physical and psychological symptoms in individuals with vestibular disorders.

Conclusions: This pilot study indicates that smartphone-supported vestibular rehabilitation has the potential to enhance functional outcomes for some patients. The consistent association between dizziness and anxiety underscores the relevance of considering

XSL•FO
RenderX

psychological factors in vestibular rehabilitation. Given the preliminary nature of this study and the small sample size, these findings should be interpreted cautiously, and further research is needed to determine the app's effectiveness in larger randomized controlled trials.

Trial Registration: ClinicalTrials.gov NCT05959278; https://clinicaltrials.gov/study/NCT05959278

(JMIR Rehabil Assist Technol 2026;13:e84207) doi:10.2196/84207

KEYWORDS

vestibular rehabilitation; dizziness; vestibular application; vestibular anxiety; vestibular technology; rehabilitation technology; mobile phone

Introduction

Central or peripheral vestibular dysfunction may cause dizziness, vertigo, and unsteadiness [1]. The peripheral injury can be classified as either unilateral or bilateral vestibular hypofunction [2]. In 2014, a total of 53 - 95 million people across the United States and Europe endured vestibular hypofunction [3]. The common symptoms are vertigo, dizziness, instability, ataxia, nystagmus, and impaired navigational abilities [1,4,5]. A decrease in balance caused by vestibular hypofunction significantly increases the risk of falling in active older adults [6]. The symptoms of vestibular hypofunction may cause impairment in participation in social gatherings and events, quality of life, and a feeling of anxiety [7,8].

Alongside its physical and functional effects, dizziness has been associated with psychological factors, particularly anxiety [7,9,10], although the mechanisms underlying this association are still being investigated. Recent evidence suggests that cerebellar involvement in vestibular processing may play a role in this relationship [11]. Accordingly, interactions between the vestibular system and the cerebellum could be key to understanding the relationship between balance issues and anxiety disorders [11]. In addition, several studies have shown that the scores of the Dizziness Handicap Inventory (DHI) questionnaire are correlated with various anxiety questionnaires, but most of these studies examined specific vestibular populations, such as people with vestibular migraine, or measured the correlation at only one point in time [12-14].

Over the years, several studies [15-18] have shown the safety and effectiveness of vestibular rehabilitation. In vestibular rehabilitation, exercises are performed to alleviate symptoms with habituation exercises, improve gaze stability with adaptation and substitution exercises, and improve balance [15]. According to the clinical practice guidelines for vestibular rehabilitation published by the American Physical Therapy Association and updated in 2022, it is recommended to practice 3 to 5 times per day for a total of 20 - 40 minutes daily over 4 to 6 weeks [15].

Despite vestibular rehabilitation's effectiveness and safety, patients report difficulty during the process, including time commitment and difficulty getting to many face-to-face meetings with the therapist [19]. Moreover, access to vestibular rehabilitation is limited (eg, in terms of distance from the clinic or clinician availability); in a European survey among vestibular therapists, 48% described accessibility as difficult or very difficult [20].

In recent years, studies have examined several ways to make vestibular practice more accessible, such as through practicing tai chi or using virtual reality in vestibular rehabilitation [21]. These methods have shown efficacy in some studies [22,23]. Despite this, these methods are rarely used in practice, possibly because it can be challenging to combine short periods of exercise several times a day [24], as recommended in clinical practice guidelines. In addition, the high cost of robotic [25] or virtual-reality [22] systems may also account for the absence of these technologies from the rehabilitation landscape [22].

Several phone applications have been developed in recent years to aid in vestibular rehabilitation [26-28]. DSilva et al [26] and Nehrujee et al [27] examined the usability of the application after a 1-time experience on individuals with vestibular disorders as well as on healthy individuals. Hovareshti et al [28] validated the performance of a tablet application (which can also be adapted for mobile phones) designed for monitoring practice while identifying head movements and gaze focus using the camera. In addition, they also presented a case series examining the effect on eye-gaze accuracy of 3 patients and 1 healthy individual who used the application [29]. However, these studies did not examine the effects of the phone applications' impact on patients' clinical outcomes over the course of the vestibular rehabilitation program.

To meet the emerging need and address the existing difficulties in the accessibility of vestibular rehabilitation [19,20,24], we developed the "Vestibulon" phone app, aimed to help patients perform their vestibular exercises; the app provides feedback for the user, gives a clear explanation of each exercise, helps manage practice time, and provides information and feedback to the therapist and patient.

The first goal of this research is to examine the effect of practice with the "Vestibulon" app on functional balance and anxiety indices. These were measured using the following parameters: (1) Instrumented Timed Up and Go Test (iTUG), (2) DHI, and (3) State-Trait Anxiety Inventory (STAI). The second goal is to assess the association between balance function (iTUG) and anxiety (STAI). We hypothesized that practice supported by the app would increase compliance with the practice, thereby improving indicators of balance and anxiety.

Methods

Study Design

This study was conducted using a randomized crossover design. Each participant was randomly assigned to 1 of 2 groups, while

XSL•FO
RenderX

gender variables were stratified to ensure an equal distribution. A computer program determined the randomization into groups. Due to the nature of the intervention, outcome assessors were not blinded; no blinding of participants or therapists was possible, and allocation was not concealed. A total of 6 weeks of vestibular rehabilitation were given to each participant. During the first period, which lasted 3 weeks, group 1 used the app for their vestibular rehabilitation, and group 2 did their vestibular rehabilitation exercises without the app. At the crossover point, they switched, such that in the second period, which also lasted 3 weeks, group 1 did their rehabilitation exercises without the app, and group 2 used the app. Measurement sessions were held at Sheba Hospital's otolaryngology clinic between June 2023 and July 2024. This study was concluded in July 2024, following completion of planned recruitment. All study participants gave their written informed consent. This trial is reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) 2010 statement: extension to randomized crossover trials (Checklist 1).

Participants

A total of 26 participants were randomly assigned to the study. Further, 6 participants dropped out during the study, resulting in 20 participants completing this pilot study. As this is a preliminary investigation, a formal power analysis was not performed. Instead, our sample size aligns with literature suggesting a minimum of 12 participants per arm to adequately gauge feasibility and variance for future calculations [30]. In this crossover design, participants received both interventions, ensuring that the final sample exceeded this minimum requirement per arm. A broad age range (18 - 75 y) was intentionally included to allow an initial feasibility assessment across diverse adult users, in line with the aims of a pilot study.

Inclusion and Exclusion Criteria

The inclusion criteria were (1) participants' age of 18 - 75 years, (2) diagnosis of vestibular dysfunction, and (3) being fluent in Hebrew. The exclusion criteria were (1) a medical condition that prevents the participant from performing home vestibular rehabilitation practice, such as orthopedic, neurological, cardiac, or visual impairment; (2) individuals with dizziness who have already been given a home vestibular rehabilitation exercise program; and (3) diagnosis of central vestibular disorder due to brain structural damage.

Recruitment

Participants were recruited through the Sheba Hospital Dizziness Clinic. Clinicians informed eligible patients about this study during routine visits and provided them with a study information sheet. Those interested provided their contact details, and a research team member subsequently contacted them to explain this study further, confirm eligibility, and schedule the baseline assessment.

Ethical Considerations

All study procedures were approved by the Helsinki Committee of Sheba Medical Center (protocol number 9576 - 22-SMC), and the trial was prospectively registered on ClinicalTrials.gov (NCT05959278; registration date: 07/06/2023). All participants provided written informed consent before enrollment and were informed of their right to withdraw from this study at any time without consequence. No secondary use of data from prior studies was involved, and thus no additional waivers or exemptions were required.

To ensure privacy and confidentiality, all data collected for this study were anonymized using coded identifiers. No personally identifying information was stored within the research database or within the investigational app. Paper-based documents (eg, consent forms) were stored in a locked cabinet inside a locked research laboratory, and all digital data were stored on secure, access-restricted institutional servers. Only authorized study personnel had access to coded datasets. Participants did not receive financial compensation for participation in this study.

Outcome Measures and Data Collection

The investigational smartphone app, Vestibulon, was designed to support the delivery of conventional vestibular rehabilitation remotely, with an emphasis on accessibility, clarity, and adherence. Its development was informed by a participatory design process, incorporating insights from patients and clinicians regarding usability, clarity of instructions, accessibility, and barriers to home-based vestibular rehabilitation [25]. These stakeholder inputs directly shaped core features of the app, including its simple interface, structured reminders, and emphasis on supporting daily adherence to standard vestibular exercises. All participants used the same version of the app during this study.

Vestibulon enables the clinician to predefine the rehabilitation parameters for each user, including the type of exercise (adaptation, habituation, and substitution), frequency, intensity, and recommended duration, based on the patient's diagnosis and clinical characteristics. The exercises included in the app correspond to standard, widely accepted components of vestibular rehabilitation. Before each exercise, the app provides step-by-step written and visual instructions to ensure correct performance.

To promote adherence, the app delivers push-notification reminders for exercise sessions. Each user selects the preferred times of day for receiving these reminders and can modify these settings at any time during the intervention period. When a reminder appears, users can choose to begin the exercise, postpone it ("snooze"), or dismiss it. After completing an exercise session, the user is prompted to provide brief feedback (eg, whether the exercise was completed and whether they wish to continue). These features are intended to help patients structure their daily practice and maintain consistency across sessions. Participants received a brief standardized onboarding session of approximately 10 minutes, during which the app was installed and demonstrated, and users were instructed on selecting reminder times, navigating the exercise interface, and submitting feedback. No advanced technical skills were required to operate the app. The app functioned reliably throughout this trial, and no major technical issues or crashes were reported.

XSL•FO
RenderX

Outcome Measures

Instrumented Timed Up and Go Test

The iTUG examines functional balance [31]. The test measures the time it takes for a participant to stand up from a seated position on a chair, walk for 3 meters at their own pace, turn around, return for 3 meters, turn again, and return to their seated position [31]. We used the iTUG for a quantitative assessment for this test [32]. For the measurement, we used Opal (APDM) sensors and the instructions in the APDM "mobility lab" application.

State-Trait Anxiety Inventory

The STAI is a self-report questionnaire designed to examine a person's general and current anxiety levels [33]. The questionnaire is divided into 2 parts, each consisting of 20 statements, which the participant should rate on a scale of 1 - 4 according to their degree of agreement [34]. This questionnaire has been used among patients who endure vestibular disorders [35]. We used this questionnaire to better distinguish between the participant's general and momentary anxiety to assess the effect of the vestibular rehabilitation on anxiety levels.

Dizziness Handicap Inventory

The DHI is a 25-item questionnaire designed to assess the impact of dizziness on everyday life in people with vestibular disorders [36]. The questionnaire has 3 parts: physical, emotional, and functional [36]. The minimal clinically important difference (MCID) of the DHI is an 18-point difference [36]. Participants with a baseline DHI score of ≤17 were excluded from the MCID analysis because, given the 18-point MCID threshold, these individuals could not mathematically achieve an MCID improvement. Including them would introduce a floor effect and underestimate the proportion achieving clinically meaningful change. In 2010, a Hebrew translation of the questionnaire was published [37], and this is the version we administered to participants. No changes to outcome measures were made after trial commencement.

Procedures

This was a 2-period, 2-sequence randomized crossover trial. A washout period was not implemented to avoid interrupting rehabilitation and potentially compromising clinical progress, although this may allow residual effects to influence the second phase. Each participant was administered home exercises for gaze stabilization and habituation for a total of 20 - 30 minutes each day, across 7 to 10 sessions (2 - 4 min per session). The exercises were administered by a vestibular physical therapist. Upon entry to this study (T0), we took the baseline measurements of the iTUG, DHI, and STAI. Participants were then randomized into 1 of 2 groups: one that practiced with the app first (AF), or one that started with conventional practice first (without an app; conventional first, CF). At the end of the first intervention period (T1), we performed the second set of measurements of the same parameters. This was the crossover point at which participants from each group switched to the alternative treatment condition. That is, during the second intervention period, the AF group received conventional treatment, and the CF group used the app to perform their rehabilitation exercises. Each intervention period lasted approximately 3 weeks, for a total of 6 weeks of rehabilitation per participant.

After the completion of both intervention periods (T2), we performed the endpoint measurement of the iTUG, DHI, and STAI. The 6-week period was chosen based on the clinical practice guidelines for vestibular rehabilitation [15]. Participants were aware of the intervention type. The app-based and conventional rehabilitation provided the same therapeutic content but differed in their mode of delivery.

Data Analysis

We used the Kolmogorov-Smirnov test to examine the normal distribution of quantitative variables on the ratio scale.

For evaluating differences in the STAI, DHI, and iTUG indices between groups, we used the Mann-Whitney *U* test. To assess changes in STAI, DHI, and iTUG across the 3 time points (T0, T1, and T2), we applied the Friedman test. When significant differences were identified, post hoc analyses were conducted using the Wilcoxon signed-rank test, with Bonferroni correction applied to account for multiple comparisons. Finally, to explore potential correlations between the iTUG indices, DHI scores, and the STAI questionnaire, we calculated Spearman correlation coefficient. Statistical analysis was performed with SPSS (version 28; IBM Corp) software, and results were considered significant when *P* values were <.05.

Results

Participant Characteristics

A total of 26 participants were recruited for this study. Further, 20 participants (mean age 52.4, SD 12.6 y; 8 men and 12 women) completed this study (Figure 1, Table 1). There was no statistically significant difference in the baseline data between the 2 groups. Participants who completed this study presented with the following vestibular disorders: bilateral vestibular hypofunction (n=6), unilateral vestibular hypofunction (n=4), Meniere disease (n=3), vestibular migraine (n=2), persistent postural perceptual dizziness (PPPD; n=3), labyrinthitis (n=1), and postconcussive syndrome (n=1). Furthermore, 6 participants dropped out during this study (mean age 46.7, SD 14.0 y; 2 men and 4 women).

XSL•FO
RenderX

Figure 1. Study overview. The flowchart illustrates the participant enrollment, intervention groups (one that started with app assistance [AF] and one that started with conventional rehabilitation [CF]), and assessment time points. The app icon indicates the app-assisted rehabilitation period. AF: app first; CF: conventional first; DHI: Dizziness Handicap Inventory; iTUG: Instrumented Timed Up and Go; STAI: State Trait Anxiety Inventory.

XSL•FO
RenderX

Table . Demographics and baseline questionnaire scores of each group. Baseline data of each group and the baseline data of all groups together. No statistically significant difference was found between the groups in any of the characteristics.

Characteristic	AF[a] (n=11)	CF[b] (n=9)	Total (N=20)	*P* value
Age (years), mean (SD)	48.5 (12.8)	57.3 (11.1)	52.4 (12.6)	.16[c]
Gender (woman/man), n (%)	7 (63.6) / 4 (36.4)	5 (55.6) / 4 (44.4)	12 (60) / 8 (40)	.53[d]
DHI[e] total score, median (IQR)	54.0 (36-64)	44.0 (32-74)	49.0 (33-68)	.88[c]
DHI physical, median (IQR)	14.0 (9-19)	14.0 (12-18)	14.0 (9-18)	.91[c]
DHI functional, median (IQR)	18.0 (8-28)	20.0 (10-30)	19.0 (9-30)	.73[c]
DHI emotional, median (IQR)	20.0 (11-29)	14.0 (10-30)	17.0 (10-30)	.73[c]
STAI[f] total, median (IQR)	83.0 (64-92)	92.0 (78-93)	87.0 (64-94)	.62[c]
STAI state, median (IQR)	42.0 (32-49)	44.0 (33-51)	43.0 (32-50)	.94[c]
STAI trait, median (IQR)	34.0 (30-50)	45.0 (34-55)	36.5 (31-54)	.67[c]
iTUG[g] total time, mean (SD)	11.4 (2.2)	11.1 (1.5)	11.1 (1.8)	.79[c]

[a]AF: app first.

[b]CF: conventional first.

[c] Mann-Whitney test.

[d]Chi-square test (Fisher Exact Test).

[e]DHI: Dizziness Handicap Inventory.

[f]STAI: State Trait Anxiety Inventory.

[g]iTUG: Instrumented Timed Up and Go.

Effects of the Phone App on Disability, Anxiety, and Functional Balance Measures

DHI Total

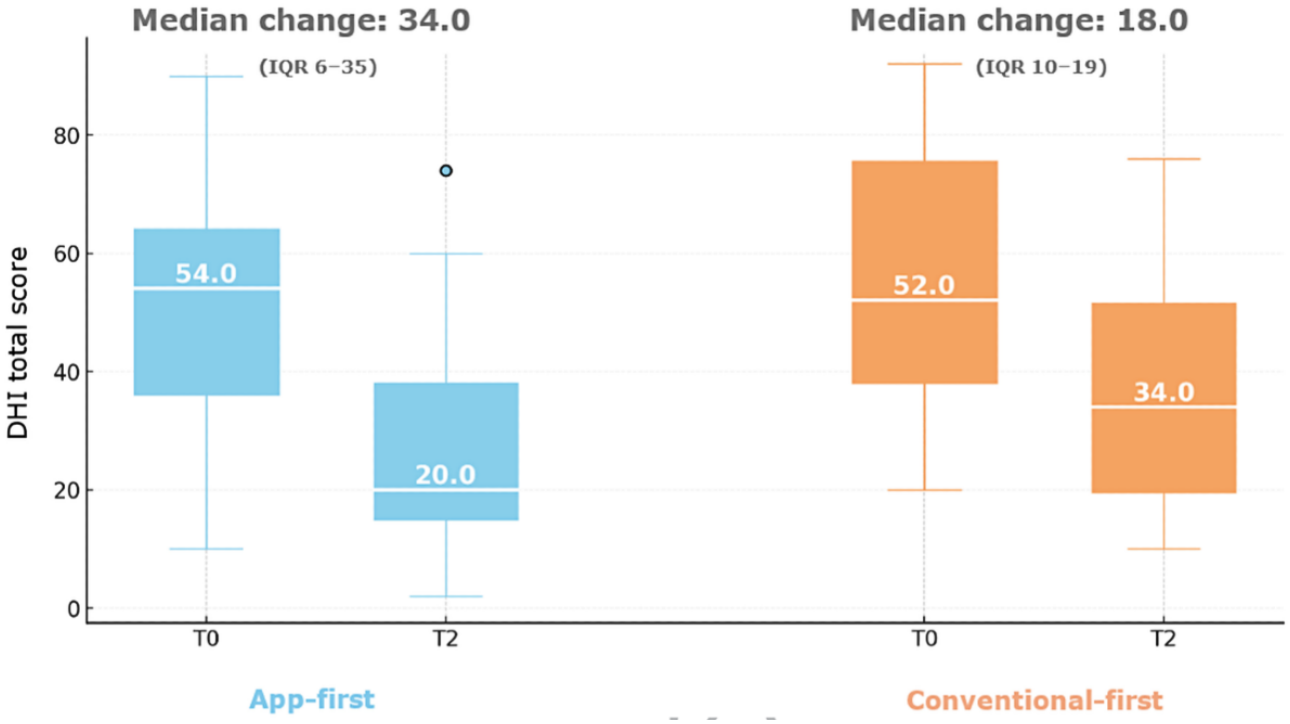
Both groups (AF and CF) show statistically significant differences between T0-T2 in the DHI total (AF: Z=−2.80, *P*=.006; CF: Z=−2.67, *P*=.009).

The total DHI score decreased by 62.9% in the AF group and by 34.6% in the CF group between baseline and end-of-study; however, the difference between the groups was not statistically significant (Z=−0.80, *P*=.42). Figure 2 summarizes the change in DHI total.

Figure 3 illustrates the changes in the DHI total score during both practice periods, as well as the number of participants who reached or exceeded the MCID.

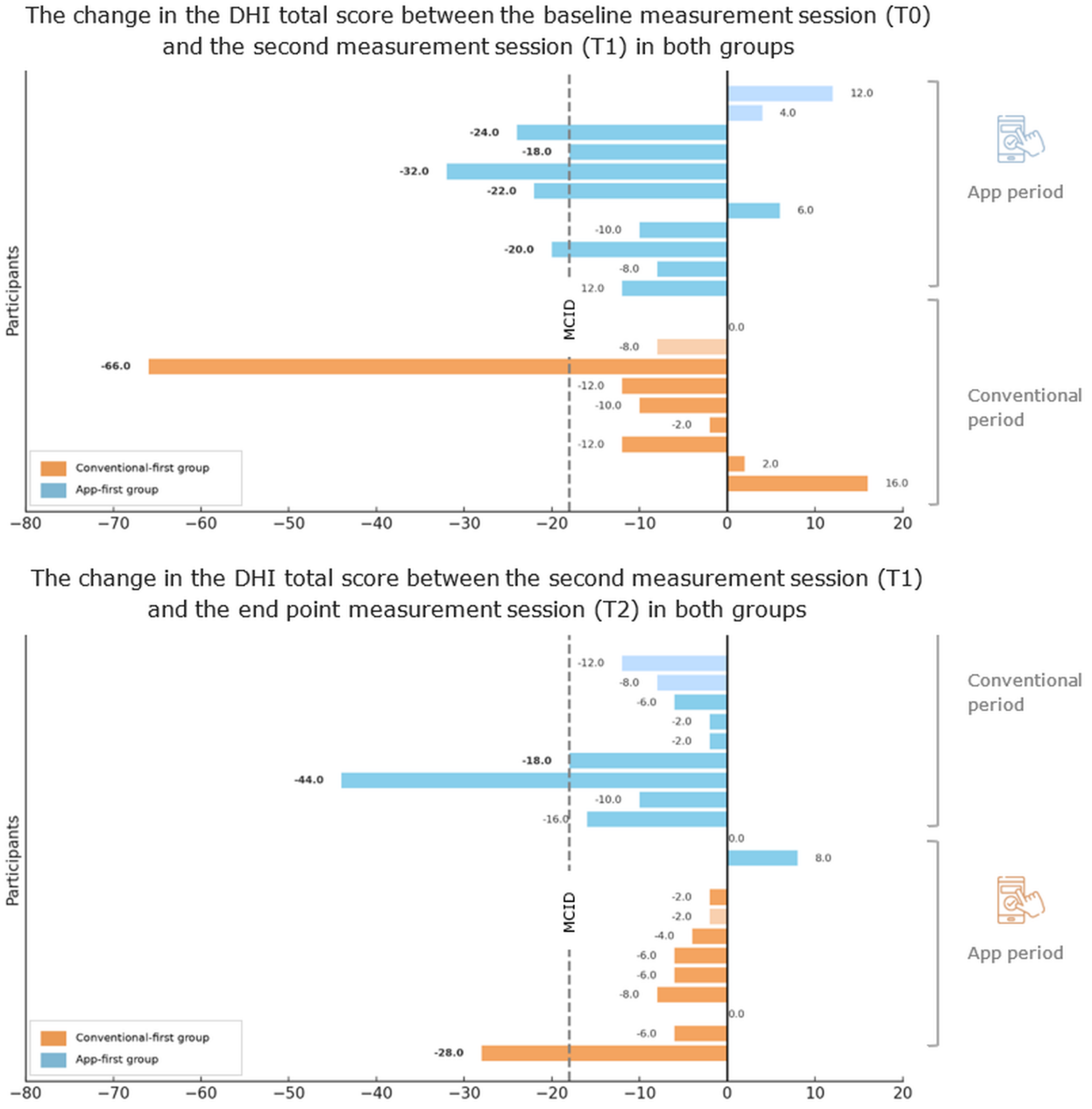
XSL•FO
RenderX

Figure 2. The change in median DHI total between T0-T2. Within each group (AF and CF), we found a statistically significant difference between T0 and T2. A blue dot indicates an outlier. Blue color represents the app-first group, and orange color represents the conventional-first group. The values in the box plot represent the median values at T0 (first evaluation when beginning the rehabilitation period) and at T2 (last evaluation after 6 weeks of rehabilitation). AF: app first; CF: conventional first; DHI: Dizziness Handicap Inventory.



XSL•FO
RenderX

Figure 3. Change in DHI total score and number of patients achieving the minimal clinically important difference score. Top: the change in the DHI total score between the baseline measurement (first evaluation before the rehabilitation period [T0]) and the second measurement session (second evaluation after 3 weeks of rehabilitation [T1]) in both groups. Bottom: the change in the DHI total score between the second measurement session (T1) and the end point measurement session (last evaluation after 6 weeks of rehabilitation [T2]) in both groups. The results of the app-first group are marked in light blue; those of the conventional-first group are marked in orange; a phone-app icon indicates that the app was used during this period. The bolded numbers indicate that these participants' score changes met or exceeded the MCID threshold. The lighter bars indicate participants with a floor effect, that is, whose initial DHI score was 17 or below. DHI: Dizziness Handicap Inventory; MCID: minimal clinically important difference.



By the end of this study, 7 of 9 (77%) participants in the AF group exceeded the MCID threshold, compared to 2 of 8 (25%) participants in the CF group. This difference between the groups did not quite reach statistical significance (χ^2_1=4.97, *P*=.06).

When comparing the with-app practice period to the no-app practice period, regardless of the group (AF/CF), we found that 6 of 17 (35.2%) participants met or superseded the MCID threshold during the period with the app (T0-T1 for the AF group and T1-T2 for the CF group), while 3 of 17 (17.6%) participants met or superseded the MCID threshold without the app (T1-T2 for the AF group and T0-T1 for the CF group).

No statistically significant change was found between the 3 measurement sessions in the STAI state (χ^2_2=0.80, *P*=.67) or in the STAI trait (χ^2_2=1.71, *P*=.42) in the AF+CF together, and in each group separately.

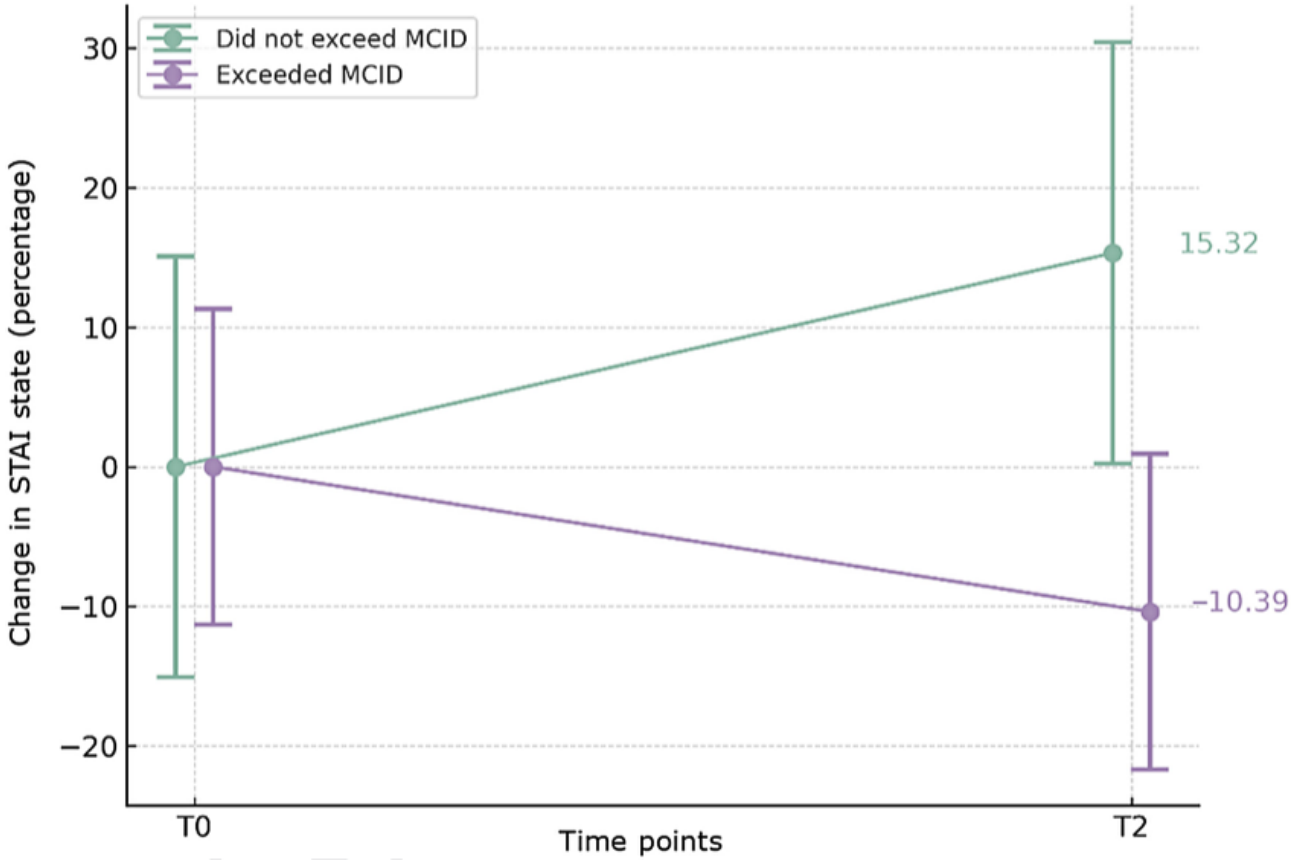
When we divided the data of all the participants into 2 groups: those who met or superseded the MCID threshold in the DHI total (a change of 18 points or more) and those who did not (a change of fewer than 18 points), we found opposing trends: the group that superseded the MCID showed a 10.3% decrease in the STAI state score between T0 and T2, while the group that did not pass the MCID showed an increase of 15.3% (Figure

XSL•FO
RenderX

4). This difference is not statistically significant (Z=−1.73, *P*=.09).

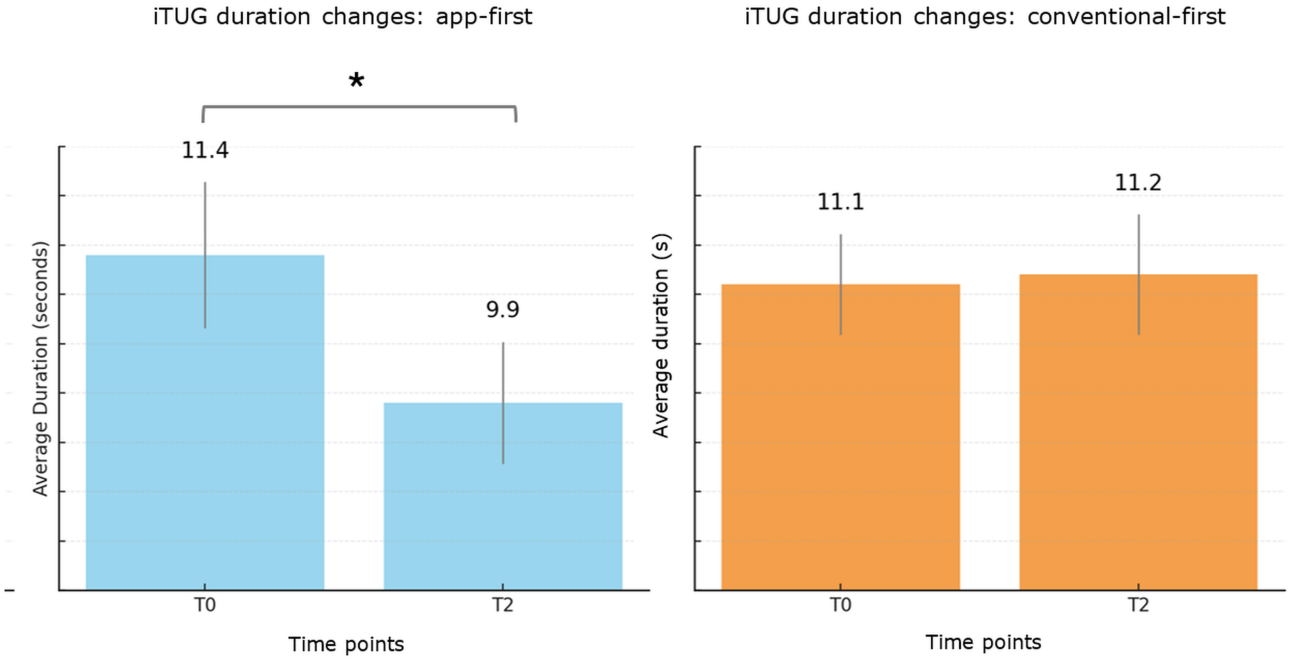
In the AF group, a statistically significant change was found across the 3 iTUG measurements (χ^2_2=9.45, *P*=.009; Figure 5). We found a statistically significant change in the post hoc test only between T0-T2 (Z=−2.45, *P*=.01). In the CF group, no statistically significant change was found across the 3 measurements (χ^2_2=1.75, *P*=.42).

Figure 4. Contrasting trends in STAI State anxiety changes relative to DHI improvement. Participants who achieved or exceeded the MCID in DHI scores between T0 and T2 showed a trend towards lower STAI State scores. Conversely, those who did not reach the MCID in DHI improvement demonstrated a trend towards higher STAI State scores. DHI: Dizziness Handicap Inventory; MCID: minimal clinically important difference; STAI: State Trait Anxiety Inventory.



XSL•FO
RenderX

Figure 5. Change in iTUG total duration between baseline (first evaluation at the beginning of the rehabilitation period [T0]) and the end of this study (last evaluation after 6 weeks of rehabilitation [T2]) * indicates a statistically significant difference between the time points. iTUG: Instrumented Timed Up and Go.



Association Between Anxiety (STAI) to Balance Function (iTUG) and Dizziness Handicap (DHI)

There was no statistically significant correlation between STAI state and iTUG duration at any time point. However, we observed a statistically significant moderate-strong positive correlation between STAI state and DHI total in all 3 measurement sessions (Table 2). Table 2 shows the Spearman correlation between the STAI state and iTUG duration, and between the STAI state and DHI total.

Table . Correlations between anxiety scores and DHI[a] and iTUG[b] for all participants.

	T0[c]	T1[d]	T2[e]
STAI[f] state and iTUG duration			
r	0.12	−0.36	−0.16
P value	.59	.19	.51
STAI state and DHI total			
r	0.59	0.65	0.68
P value	.005	.002	<.001

[a]DHI: Dizziness Handicap Inventory.

[b]iTUG: Instrumented Timed Up and Go.

[c]T0: first evaluation when beginning the rehabilitation period.

[d]T1: second evaluation after 3 weeks of rehabilitation.

[e]T2: last evaluation after 6 weeks of rehabilitation.

[f]STAI: State Trait Anxiety Inventory.

Discussion

Principal Findings

Twenty participants completed this 6-week study, with each subject undergoing 2 phases in random order: a period of conventional vestibular rehabilitation and a period of app-assisted vestibular rehabilitation. We examined how mobile-assisted vestibular rehabilitation affects functional balance measure (iTUG), dizziness disability (DHI), and anxiety (STAI). In addition, we examined the association between the measures (STAI with DHI and iTUG). We found that more individuals (7/9, 77%) in the AF group reached an improvement of 18 points or more on the DHI scale, indicating a clinically meaningful improvement, compared to the conventional-first group (2/8, 25%). Given the heterogeneity of vestibular diagnoses represented in our sample, it is possible that different conditions may respond differently to rehabilitation. For example, individuals with Ménière disease often show fluctuating symptom patterns [38,39], whereas patients with PPPD may require longer or psychologically oriented interventions [40]. As each diagnostic subgroup in our study

XSL•FO
RenderX

included only a few participants, subgroup analyses were not feasible.

The main contributions of this paper are (1) we found that using the "Vestibulon" phone app, as part of the vestibular rehabilitation process, was associated with improvements in functional clinical indices in vestibular rehabilitation (DHI and iTUG), with a trend suggesting a possible advantage of the app-assisted condition over conventional rehabilitation; (2) we found that the correlation between the DHI questionnaire and the STAI questionnaire remained relatively constant throughout the rehabilitation period and was moderate (r=0.64, P<.001), highlighting the importance of addressing both physical and psychological aspects of vestibular disorders.

The Effect of Using the Phone App on Functional Outcomes

The finding that DHI scores improved following vestibular rehabilitation in both groups is echoed in the literature [15,23,41,42]. Some of those studies showed greater improvement in DHI than shown in the current study; this difference may be due to the diverse population participating in our study, which included a relatively high percentage of patients (45%) with bilateral vestibular dysfunction, or PPPD, who sometimes require a more extended period of rehabilitation or a combination of additional treatments [15].

Our study revealed no significant changes in STAI scores in either group or in both groups together. This aligns with a report by Johansson et al [43], who found no effect on STAI score after vestibular rehabilitation based on home exercises combined with cognitive behavioral therapy, in a diverse cohort of vestibular patients. However, Meli et al [44] found an improvement in STAI score that was preserved after 2 months of follow-up; this was after a daily vestibular rehabilitation program supervised by a therapist, coupled with home practice, in diverse chronic vestibular patients. Our results may offer a potential explanation for these conflicting findings. We noted a distinction between participants who achieved the DHI MCID and those who did not. Specifically, those who met the MCID threshold demonstrated improvements in STAI scores, whereas those below the threshold showed deterioration (Figure 4). Although a higher proportion of participants reached the MCID during the app-assisted phase, this difference was not statistically significant (P=.057). Therefore, this pattern should be interpreted cautiously and viewed as preliminary, given the small sample size. Similarly, this observation is further supported by the study by Johansson et al [43], which reported no change in STAI scores and a relatively modest average change in the DHI of less than 8 points (whereas the MCID is 18). Combined, these findings suggest that vestibular rehabilitation's impact on anxiety may be closely linked to the degree of improvement in dizziness-related symptoms, highlighting the complex interplay between physical symptoms and psychological well-being in vestibular disorders.

The results of our study suggest potential advantages of the app-assisted vestibular rehabilitation over conventional methods, although the small sample size limited our ability to demonstrate statistically significant differences. Notably, the AF group showed promising trends in several key areas. First, for DHI MCID attainment, 3 times as many participants in the AF group (7/9, 77%) met or surpassed the DHI MCID threshold compared to the CF group (2/8, 25%). Although these differences in MCID attainment did not reach statistical significance when directly comparing the 2 conditions (P=.057), the pattern may still indicate a potential benefit of the app-assisted approach and warrants further investigation in a larger sample. Second, for iTUG performance, only the AF group demonstrated a statistically significant improvement in the iTUG test (Figure 5), indicating a trend toward enhanced functional mobility. Third, for DHI improvement, the AF group achieved a median change of −34 (IQR 6 - 35) points on the DHI, compared to −18 (IQR 10 - 19) points in the CF group, potentially suggesting a more substantial reduction in dizziness-related handicap.

Furthermore, twice as many participants met or surpassed the MCID threshold of DHI while using the app (6/17, 35.2%) as participants during conventional treatment (3/17, 17.6%). While this difference did not reach statistical significance, the directionality of the findings suggests a potential advantage of the app-assisted condition. These findings, which suggest a positive trend in favor of rehabilitation with the app, need further investigation. A larger-scale study with an expanded sample size and additional outcome measures is necessary to validate the observed trends and conclusively establish whether indeed the app-assisted vestibular rehabilitation is advantageous over conventional rehabilitation.

The "Vestibulon" App Compared to Other Apps

Several applications have been developed to enhance vestibular rehabilitation, each with unique features and limitations.

For example, D'Silva et al [26] and Meldrum et al [45] developed tablet-based apps integrated with an inertial measurement unit (IMU) sensor placed on the body. D'Silva et al [26] demonstrated high usability and improved accuracy in exercise execution compared to conventional exercises among individuals with vestibular disorders. However, their findings were based on a single session rather than a full vestibular rehabilitation program, and clinical changes were not assessed. In contrast, Meldrum at el [45] evaluated their app over an average of 8 weeks of vestibular rehabilitation and reported high usability, strong adherence, and reductions in dizziness, nausea, and anxiety. Nevertheless, they did not compare their treatment to conventional rehabilitation methods, and symptom improvement was assessed by subjective ratings rather than validated clinical tests or questionnaires typically used in vestibular rehabilitation. Hall et al [46], on the other hand, developed a vestibular rehabilitation training system incorporating a computer application and an external sensor, based on feedback from clinicians and patients. In a pilot study with individuals with vestibular hypofunction, their system demonstrated feasibility and effectiveness over a 4-week rehabilitation period. Unlike the systems mentioned above, the solution by Hall et al [46] also included functional outcomes, which provide more robust evidence of clinical effectiveness. However, similar to the systems in studies by D'Silva et al [26] and Meldrum et al [45], the system requires external hardware, such as a sensor and a computer or tablet, limiting its portability and accessibility.

XSL•FO
RenderX

A main strength of these systems is the ability to provide accurate real-time feedback through the sensors, increasing exercise precision. However, the reliance on multiple components, such as sensors, laptops, or tablets, constitutes a barrier to portability and accessibility.

Hovareshti et al [28] developed a tablet application that detects eye movement without using IMUs, using the tablet's camera itself, and compared its capabilities with IMUs, which showed good accuracy in head angle error and mean interpeak time errors. They presented a case series examining the impact of the practice with the app on eye-gaze accuracy (the ability to maintain one's gaze focused on a target while the head moves) in 3 patients and 1 healthy individual. They found poor eye-gaze accuracy in a participant exposed to directed energy [29]. Their significant advantage is their simplicity of use and portability, which means they do not need external sensors while providing accurate feedback based on movement detection. On the other hand, they have not yet demonstrated the app's efficacy within the long vestibular rehabilitation process or compared it with conventional rehabilitation.

Our smartphone app was developed through a participatory-design process involving focus groups with clinicians and patients [25]. They emphasized the need for ease of use, clear instructions, feedback, cost-effectiveness, and the portability of mobile technology such as phone apps [25]. These insights guided the development of our app. We then tested the app throughout 6 weeks of vestibular rehabilitation with patients, using accepted clinical questions and functional tests.

Association Between Dizziness Disability and Anxiety

Our study revealed a moderately positive correlation between the DHI and the STAI state scores across all 3 measurement points (r=0.59, r=0.65, and r=0.68 for T0, T1, and T2, respectively). These findings align with previous research on the association between DHI and STAI state. For instance, Hong et al [47] found that DHI was significantly correlated with the STAI questionnaire in a multicenter study on 407 individuals with various vestibular disorders, which examined the psychological effects of the vestibular condition. In addition, Kim et al [48] examined 544 individuals with various disorders and examined the effects of anxiety and depression on the intensity of dizziness; they found that compared to individuals with low STAI scores, individuals with high STAI scores also demonstrated higher DHI scores. Both studies examined the association between STAI and DHI at a single time point, regardless of the effect of vestibular rehabilitation. Our study extends these findings by demonstrating the correlation between DHI and STAI state at 3 points throughout the process of vestibular rehabilitation: before treatment initiation, at the midpoint, and at the conclusion of the 6-week treatment. Additionally, we observed that participants who exceeded the MCID threshold on the DHI also showed improvement on the STAI state, in contrast with those who did not reach the MCID threshold and exhibited an increase in their STAI state scores (Figure 4). These findings suggest that one possible contributor to the positive trend observed with the app-assisted condition may relate to psychological elements, such as participants' perception of support or feedback, which have been highlighted as important by patients and clinicians [24]. However, because these factors were not directly measured in the present study, this explanation should be considered a hypothesis for future investigation. The results underscore the relationship between dizziness symptoms, functional limitations, and anxiety. The incorporation of these psychological components appears to address not only the physical aspects of vestibular rehabilitation but also the psychological factors that can impact recovery and adherence to treatment.

Interestingly, we found no correlation between STAI trait and DHI. Schuhbeck et al [49] did find an association between high STAI trait and DHI scores in the acute phase after vestibular stroke and high DHI scores at long-term follow-up. The discrepancy in the findings may be attributed to differences in study populations, as we excluded individuals with vestibular strokes.

We did not observe a statistically significant correlation between the iTUG test and the DHI scores at any measurement point (Table 2). Previous studies have reported conflicting results on the correlation between TUG and DHI, possibly due to variations in study populations and test instructions (eg, walking as quickly as possible vs at a comfortable pace). Further research is needed to clarify the relationship between DHI and iTUG, with careful consideration of population characteristics and standardized instructions.

Limitations

This is a pilot study with a small sample size, which limits the external validity and generalizability of the findings. This study included participants with a diverse range of vestibular disorders and a broad adult age range, which offers the advantage of reflecting the clinical heterogeneity seen in practice. However, this diagnostic and age diversity may also influence the results, as certain vestibular conditions or age groups could have a disproportionate impact on specific outcomes, particularly in a small pilot sample.

Furthermore, crossover study designs present inherent challenges, as it can be difficult to ascertain which phase of the intervention contributed to the observed changes. In addition, the lack of a washout period may have introduced carryover effects between phases, and this should be considered when interpreting the comparative findings. The absence of blinding for participants, therapists, and assessors represents an additional limitation. This may have contributed to performance or detection bias, especially for subjective outcomes (DHI and STAI) and for functional assessments such as the iTUG. Although standardized assessment procedures were used, the influence of expectancy or assessor awareness cannot be fully excluded. To address these limitations, future research should implement a randomized controlled trial with a larger sample size to enhance the robustness and applicability of the results.

Conclusions

Our findings contribute to the understanding of the interplay between dizziness, anxiety, and functional mobility in vestibular rehabilitation. The consistent correlation between DHI and STAI state scores throughout the rehabilitation process underscores

XSL•FO
RenderX

the importance of addressing both physical and psychological aspects in vestibular disorders.

The "Vestibulon" smartphone app, when integrated into vestibular rehabilitation, showed improvements across several outcomes, with patterns suggesting a possible trend favoring the app-assisted phase. Taken together with the significant within-participant improvement in iTUG performance, the observation that more participants exceeded the DHI MCID threshold during the app-assisted phase suggests that integrating the "Vestibulon" app into vestibular rehabilitation may offer clinical benefits. However, these trends did not reach statistical significance and should therefore be interpreted as preliminary. These encouraging preliminary findings support the need for larger, adequately powered trials to determine the extent of the app's contribution to rehabilitation outcomes.

Acknowledgments

The authors would like to thank Yuliya Berdichevsky, MSc, for her assistance in data analysis and figure preparation. The authors acknowledge the use of generative AI (artificial intelligence) tools for language editing purposes only. Specifically, Grammarly (Superhuman Platform) was used to assist with grammar and style corrections. No generative AI (artificial intelligence) tools were used for data analysis, interpretation, or generation of scientific content.

Funding

The research was partially supported by the Rosetrees Trust and by the Consolidated Anti-Aging Foundation. The funders had no involvement in this study's design, data collection, analysis, interpretation, or paper writing.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: YG (lead), SL-T (equal), AW (supporting), LK (supporting)
Data curation: AK
Formal analysis: AK (lead), LK (supporting)
Funding acquisition: SL-T
Investigation: AK (lead), LK (supporting)
Methodology: YG (lead), SL-T (equal), LK (supporting), AW (supporting), AK (supporting)
Project administration: YG (lead), SL-T (equal), AW (supporting)
Resources: SL-T
Supervision: AW
Validation: YG (lead), SL-T (equal), LK (supporting)
Visualization: AK (lead), LK (supporting)
Writing – original draft: AK (lead), LK (supporting)
Writing – review & editing: YG (lead), SL-T (equal), LK (supporting), AW (supporting)

Conflicts of Interest

None declared.

Checklist 1
CONSORT checklist.
[PDF File, 155 KB - rehab_v13i1e84207_app1.pdf]

References

1. Strupp M, Dlugaiczyk J, Ertl-Wagner BB, Rujescu D, Westhofen M, Dieterich M. Vestibular disorders. Dtsch Arztebl Int 2020 Apr 24;117(17):300-310. [doi: 10.3238/arztebl.2020.0300] [Medline: 32530417]
2. Starkov D, Strupp M, Pleshkov M, Kingma H, van de Berg R. Diagnosing vestibular hypofunction: an update. J Neurol 2021 Jan;268(1):377-385. [doi: 10.1007/s00415-020-10139-4] [Medline: 32767115]
3. Grill E, Heuberger M, Strobl R, et al. Prevalence, determinants, and consequences of vestibular hypofunction. Results from the KORA-FF4 survey. Front Neurol 2018;9:1076. [doi: 10.3389/fneur.2018.01076] [Medline: 30581415]
4. Whitney SL, Alghwiri AA, Alghadir A. An overview of vestibular rehabilitation. Handb Clin Neurol 2016;137:187-205. [doi: 10.1016/B978-0-444-63437-5.00013-3] [Medline: 27638071]
5. Zingler VC, Weintz E, Jahn K, et al. Causative factors, epidemiology, and follow-up of bilateral vestibulopathy. Ann N Y Acad Sci 2009 May;1164:505-508. [doi: 10.1111/j.1749-6632.2009.03765.x] [Medline: 19645958]

XSL•FO
RenderX

6. Tuunainen E, Rasku J, Jäntti P, Pyykkö I. Risk factors of falls in community dwelling active elderly. Auris Nasus Larynx 2014 Feb;41(1):10-16. [doi: 10.1016/j.anl.2013.05.002] [Medline: 23763793]
7. Bayat A, Hoseinabadi R, Saki N, Sanayi R. Disability and anxiety in vestibular diseases: a cross-sectional study. Cureus 2020 Nov 30;12(11):e11813. [doi: 10.7759/cureus.11813] [Medline: 33409058]
8. Guinand N, Boselie F, Guyot JP, Kingma H. Quality of life of patients with bilateral vestibulopathy. Ann Otol Rhinol Laryngol 2012 Jul;121(7):471-477. [doi: 10.1177/000348941212100708] [Medline: 22844867]
9. Elyoseph Z, Geisinger D, Zaltzman R, Hartman TG, Gordon CR, Mintz M. The overarching effects of vestibular deficit: imbalance, anxiety, and spatial disorientation. J Neurol Sci 2023 Aug 15;451:120723. [doi: 10.1016/j.jns.2023.120723] [Medline: 37393737]
10. Jacob RG, Furman JM. Psychiatric consequences of vestibular dysfunction. Curr Opin Neurol 2001 Feb;14(1):41-46. [doi: 10.1097/00019052-200102000-00007] [Medline: 11176216]
11. Hilber P, Cendelin J, Le Gall A, Machado ML, Tuma J, Besnard S. Cooperation of the vestibular and cerebellar networks in anxiety disorders and depression. Prog Neuropsychopharmacol Biol Psychiatry 2019 Mar 8;89:310-321. [doi: 10.1016/j.pnpbp.2018.10.004] [Medline: 30292730]
12. Zamyslowska-Szmytke E, Politanski P, Jozefowicz-Korczynska M. Dizziness handicap inventory in clinical evaluation of dizzy patients. Int J Environ Res Public Health 2021 Feb 24;18(5):2210. [doi: 10.3390/ijerph18052210] [Medline: 33668099]
13. Kim EK, Hum M, Sharon JD. Correlating vestibular migraine patient assessment tool and handicap inventory to daily dizziness symptoms. Otol Neurotol 2023 Dec 1;44(10):1052-1056. [doi: 10.1097/MAO.0000000000004014] [Medline: 37733970]
14. Zhu C, Li Y, Ju Y, Zhao X. Dizziness handicap and anxiety depression among patients with benign paroxysmal positional vertigo and vestibular migraine. Medicine (Baltimore) 2020 Dec 24;99(52):e23752. [doi: 10.1097/MD.0000000000023752] [Medline: 33350759]
15. Hall CD, Herdman SJ, Whitney SL, et al. Vestibular rehabilitation for peripheral vestibular hypofunction: an updated clinical practice guideline from the Academy of Neurologic Physical Therapy of the American Physical Therapy Association. J Neurol Phys Ther 2022 Apr 1;46(2):118-177. [doi: 10.1097/NPT.0000000000000382] [Medline: 34864777]
16. Herdman SJ, Hall CD, Maloney B, Knight S, Ebert M, Lowe J. Variables associated with outcome in patients with bilateral vestibular hypofunction: preliminary study. J Vestib Res 2015;25(3-4):185-194. [doi: 10.3233/VES-150556] [Medline: 26756134]
17. Horak FB, Jones-Rycewicz C, Black FO, Shumway-Cook A. Effects of vestibular rehabilitation on dizziness and imbalance. Otolaryngol Head Neck Surg 1992 Feb;106(2):175-180. [doi: 10.1177/019459989210600220] [Medline: 1738550]
18. McDonnell MN, Hillier SL. Vestibular rehabilitation for unilateral peripheral vestibular dysfunction. Cochrane Database Syst Rev 2015 Jan 13;1(1):CD005397. [doi: 10.1002/14651858.CD005397.pub4] [Medline: 25581507]
19. Cohen HS. Assessment of functional outcomes in patients with vestibular disorders after rehabilitation. NeuroRehabilitation 2011;29(2):173-178. [doi: 10.3233/NRE-2011-0692] [Medline: 22027079]
20. Meldrum D, Burrows L, Cakrt O, et al. Vestibular rehabilitation in Europe: a survey of clinical and research practice. J Neurol 2020 Dec;267(Suppl 1):24-35. [doi: 10.1007/s00415-020-10228-4] [Medline: 33048219]
21. Deveze A, Bernard-Demanze L, Xavier F, Lavieille JP, Elziere M. Vestibular compensation and vestibular rehabilitation. Current concepts and new trends. Neurophysiol Clin 2014 Jan;44(1):49-57. [doi: 10.1016/j.neucli.2013.10.138] [Medline: 24502905]
22. Rosiak O, Krajewski K, Woszczak M, Jozefowicz-Korczynska M. Evaluation of the effectiveness of a virtual reality-based exercise program for unilateral peripheral vestibular deficit. J Vestib Res 2018;28(5-6):409-415. [doi: 10.3233/VES-180647] [Medline: 30714985]
23. Xie M, Zhou K, Patro N, et al. Virtual reality for vestibular rehabilitation: a systematic review. Otol Neurotol 2021 Aug 1;42(7):967-977. [doi: 10.1097/MAO.0000000000003155] [Medline: 33782257]
24. Kalderon L, Kaplan A, Wolfovitz A, Levy-Tzedek S, Gimmon Y. Barriers and facilitators of vestibular rehabilitation: patients and physiotherapists' perspectives. J Neurol Phys Ther 2024 Jul 1;48(3):140-150. [doi: 10.1097/NPT.0000000000000470] [Medline: 38426842]
25. Kalderon L, Kaplan A, Wolfovitz A, Gimmon Y, Levy-Tzedek S. Do we really need this robot? Technology requirements for vestibular rehabilitation: input from patients and clinicians. Int J Hum Comput Stud 2024 Dec;192:103356. [doi: 10.1016/j.ijhcs.2024.103356]
26. DSilva LJ, Skop KM, Pickle NT, et al. Use of stakeholder feedback to develop an app for vestibular rehabilitation-input from clinicians and healthy older adults. Front Neurol 2022;13:836571. [doi: 10.3389/fneur.2022.836571] [Medline: 35280295]
27. Nehrujee A, Vasanthan L, Lepcha A, Balasubramanian S. A smartphone-based gaming system for vestibular rehabilitation: a usability study. J Vestib Res 2019;29(2-3):147-160. [doi: 10.3233/VES-190660] [Medline: 31356221]
28. Hovareshti P, Roeder S, Holt LS, et al. VestAid: a tablet-based technology for objective exercise monitoring in vestibular rehabilitation. Sensors (Basel) 2021 Dec 15;21(24):8388. [doi: 10.3390/s21248388] [Medline: 34960480]
29. Whitney SL, Ou V, Hovareshti P, et al. Utility of VestAid to detect eye-gaze accuracy in a participant exposed to directed energy. Mil Med 2023 Jul 22;188(7-8):e1795-e1801. [doi: 10.1093/milmed/usac294] [Medline: 36208334]

XSL•FO
RenderX

30. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
31. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
32. Kim KJ, Gimmon Y, Millar J, Brewer K, Serrador J, Schubert MC. The instrumented Timed “Up & Go” test distinguishes turning characteristics in vestibular hypofunction. Phys Ther 2021 Jul 1;101(7):pzab103. [doi: 10.1093/ptj/pzab103] [Medline: 33774661]
33. Spielberger CD. State-trait anxiety inventory for adults. APA PsycTests 1983. [doi: 10.1037/t06496-000]
34. Julian LJ. Measures of anxiety. Arthritis Care Res 2011 Nov;63(S11):S467-S472. [doi: 10.1002/acr.20561] [Medline: 22588767]
35. Kalderon L, Chaimoff M, Katz-Leurer M. The distinction between state and trait anxiety levels in patients with BPPV in comparison with healthy controls. Front Psychol 2022;13:1055467. [doi: 10.3389/fpsyg.2022.1055467] [Medline: 36533063]
36. Jacobson GP, Newman CW. The development of the Dizziness Handicap Inventory. Arch Otolaryngol Head Neck Surg 1990 Apr;116(4):424-427. [doi: 10.1001/archotol.1990.01870040046011] [Medline: 2317323]
37. Kaplan DM, Friger M, Racover NK, Peleg A, Kraus M, Puterman M. The Hebrew dizziness handicap inventory. Harefuah 2010 Nov;149(11):697-700. [Medline: 21250408]
38. Wright T. Menière’s disease. BMJ Clin Evid 2015 Nov 5;2015:0505. [Medline: 26545070]
39. Hoskin JL. Ménière’s disease: new guidelines, subtypes, imaging, and more. Curr Opin Neurol 2022 Feb 1;35(1):90-97. [doi: 10.1097/WCO.0000000000001021] [Medline: 34864755]
40. Popkirov S, Staab JP, Stone J. Persistent postural-perceptual dizziness (PPPD): a common, characteristic and treatable cause of chronic dizziness. Pract Neurol 2018 Feb;18(1):5-13. [doi: 10.1136/practneurol-2017-001809] [Medline: 29208729]
41. Millar JL, Gimmon Y, Roberts D, Schubert MC. Improvement after vestibular rehabilitation not explained by improved passive VOR gain. Front Neurol 2020;11:79. [doi: 10.3389/fneur.2020.00079] [Medline: 32153490]
42. Tokle G, Mørkved S, Bråthen G, et al. Efficacy of vestibular rehabilitation following acute vestibular neuritis: a randomized controlled trial. Otol Neurotol 2020 Jan;41(1):78-85. [doi: 10.1097/MAO.0000000000002443] [Medline: 31789800]
43. Johansson M, Akerlund D, Larsen HC, Andersson G. Randomized controlled trial of vestibular rehabilitation combined with cognitive-behavioral therapy for dizziness in older people. Otolaryngol Head Neck Surg 2001 Sep;125(3):151-156. [doi: 10.1067/mhn.2001.118127] [Medline: 11555746]
44. Meli A, Zimatore G, Badaracco C, De Angelis E, Tufarelli D. Effects of vestibular rehabilitation therapy on emotional aspects in chronic vestibular patients. J Psychosom Res 2007 Aug;63(2):185-190. [doi: 10.1016/j.jpsychores.2007.02.007] [Medline: 17662755]
45. Meldrum D, Murray D, Vance R, et al. Toward a digital health intervention for vestibular rehabilitation: usability and subjective outcomes of a novel platform. Front Neurol 2022;13:836796. [doi: 10.3389/fneur.2022.836796] [Medline: 35422750]
46. Hall CD, Flynn S, Clendaniel RA, et al. Remote assessment and management of patients with dizziness: development, validation, and feasibility of a gamified vestibular rehabilitation therapy platform. Front Neurol 2024;15:1367582. [doi: 10.3389/fneur.2024.1367582] [Medline: 38872821]
47. Hong SM, Lee HJ, Lee B, et al. Influence of vestibular disease on psychological distress: a multicenter study. Otolaryngol Head Neck Surg 2013 May;148(5):810-814. [doi: 10.1177/0194599813476476] [Medline: 23380760]
48. Kim SK, Kim YB, Park IS, Hong SJ, Kim H, Hong SM. Clinical analysis of dizzy patients with high levels of depression and anxiety. J Audiol Otol 2016 Dec;20(3):174-178. [doi: 10.7874/jao.2016.20.3.174] [Medline: 27942604]
49. Schuhbeck F, Strobl R, Conrad J, et al. Determinants of functioning and health-related quality of life after vestibular stroke. Front Neurol 2022;13:957283. [doi: 10.3389/fneur.2022.957283] [Medline: 36158947]

Abbreviations

AF: app first
CF: conventional first
CONSORT: Consolidated Standards of Reporting Trials
DHI: Dizziness Handicap Inventory
IMU: inertial measurement unit
iTUG: Instrumented Timed Up and Go
MCID: minimal clinically important difference
PPPD: persistent postural perceptual dizziness
STAI: State-Trait Anxiety Inventory

XSL•FO
RenderX

Edited by S Munce; submitted 16.Sep.2025; peer-reviewed by FN Ardic, ZW Chen; revised version received 04.Jan.2026; accepted 26.Jan.2026; published 24.Mar.2026.

Please cite as:
Kaplan A, Kalderon L, Wolfovitz A, Gimmon Y, Levy-Tzedek S
Smartphone-Supported Vestibular Rehabilitation in Individuals With Vestibular Dysfunction: Pilot Randomized Crossover Trial Assessing Functional Clinical Outcomes and Anxiety
JMIR Rehabil Assist Technol 2026;13:e84207
URL: https://rehab.jmir.org/2026/1/e84207
doi:10.2196/84207

© Azriel Kaplan, Liran Kalderon, Amit Wolfovitz, Yoav Gimmon, Shelly Levy-Tzedek. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone-Based Grading and Rehabilitation in Patients With Facial Palsy Using Computer Vision: Prospective Validation Study

Franz-Tassilo Müller-Graff[1], MD, Dr Med; Fabian Essig[2], MD, Dr Med; Maximilian U Friedrich[3], MD, Dr Med; Kathrin Hoika[1], DDS; Stephan Hackenberg[1], MD, Prof Dr Med; Kristen Rak[1], MD, Prof Dr Med; Johannes Taeger[1], MD, Dr Med

[1]Department of Oto-Rhino-Laryngology, Head and Neck Surgery, University Hospital Wuerzburg, Josef-Schneider-Strasse 11, Wuerzburg, Germany

[2]Department of Neurology, University Hospital Wuerzburg, Wuerzburg, Germany

[3]Department of Neurology, University Hospital Ulm, Ulm, Germany

Corresponding Author:
Franz-Tassilo Müller-Graff, MD, Dr Med
Department of Oto-Rhino-Laryngology, Head and Neck Surgery, University Hospital Wuerzburg, Josef-Schneider-Strasse 11, Wuerzburg, Germany

Abstract

Background: Peripheral facial palsy causes significant functional and psychosocial impairments, requiring precise assessment and patient engagement for effective rehabilitation. However, conventional clinician-graded scales (eg, House-Brackmann Scale, Sunnybrook Facial Grading System, and Stennert Index) are subjective and prone to interobserver variability, limiting their reliability for tracking recovery. Smartphone-based computer vision solutions offer objective, standardized facial movement grading, and interactive home-based training to improve adherence and outcomes.

Objective: This pilot study evaluated a novel iOS smartphone app (Apple Inc.) for facial palsy management. The app uses the iPhone TrueDepth 3D camera and on-device computer vision to compute a Digital Facial Index (DFI) for objective facial movement analysis, and provides guided neuromuscular facial exercises with real-time biofeedback. The study aimed to validate DFI against standard clinical grading scales and assess patient-reported outcomes and usability.

Methods: A 4-week single-arm pilot included 21 patients with unilateral facial palsy. Participants used the app at home for daily facial exercises and periodic self-assessments with DFI. Clinicians, blinded to DFI, rated facial function from standardized video exams at baseline and 4 weeks using the House-Brackmann Scale, the Sunnybrook Facial Grading System, and the Stennert Index. DFI concurrent validity was evaluated via correlation with these clinician scores. Patient-reported outcomes included pre- and postintervention Facial Disability Index (FDI) physical and social scores, the System Usability Scale, and a poststudy user feedback questionnaire.

Results: During the study period, strong correlations were observed between DFI and conventional clinical scores. FDI physical and social showed significant functional improvement. Mean System Usability Scale was 88.3 (SD 15.4), indicating excellent usability, and participants reported high satisfaction, preferring the app over traditional paper-based exercises.

Conclusions: The app's DFI provided objective facial function grading that correlated well with standard clinical scales. Patients' FDI scores improved significantly over 4 weeks. High usability and patient preference support the app's feasibility for home-based rehabilitation. This digital approach is promising for facial palsy management, and controlled studies are needed to confirm efficacy and improve long-term engagement.

(JMIR Rehabil Assist Technol 2026;13:e81963) doi:10.2196/81963

KEYWORDS

Digital Facial Index; facial palsy; facial neuromuscular retraining; facial diagnostics; truedepth camera; smartphone sensors; mHealth; usability; telerehabilitation; computer vision; mobile health

Introduction

Facial nerve palsy (FNP) is a common neurological disorder affecting approximately 30 per 100,000 individuals annually, often leading to significant functional and aesthetic impairments [1,2]. Difficulties with facial expressions, speech articulation, and ingestion severely impact quality of life [3,4]. The treatment of FNP encompasses pharmacological, surgical, and rehabilitative strategies [5]. Among these, facial neuromuscular retraining is a widely recommended noninvasive approach aimed at restoring facial motor function [6,7]. Traditionally, such training relies on printed exercise instructions, requiring patients to perform facial exercises in front of a mirror as a biofeedback mechanism. Everyday clinical experience shows that adherence

XSL•FO
RenderX

is generally low, and the evidence base for its effectiveness remains variable [8]. Albeit more efficacious, supervised therapy by specialized therapists is limited by cost and availability, especially in rural areas and therefore underused [9,10].

Effective FNP management requires tools for precise diagnostic assessment, structured rehabilitation training, and disease monitoring. The current standard of care is semiquantitative, clinician-administered visual scales such as the House-Brackmann Scale (HBS) [11] and the Sunnybrook Facial Grading System (SFGS) [12-14]. Other established scales include the Facial Nerve Grading System 2.0 [15], movement, rest, secondary defects, and subjective scoring grading system described by de Ru et al [16], Yanagihara Scale [17], or the Stennert Index (SI) [18]. Despite their widespread use, these approaches are limited by time-intensive scoring and are affected by clinimetric limitations such as subjectivity and interobserver variability [19], prompting innovative solutions.

Efforts to improve objectivity in FNP assessment encompass standardized photo and video documentation as well as semiautomated software solutions [20], which often require specialized and costly hardware, again restricting their widespread clinical adoption [19]. Digital tools such as eFACE (a clinician-graded electronic facial paralysis assessment; Banks et al [21]) support FNP quantification, but are limited by their reliance on manual user input [21].

In recent years, digital medicine approaches have revolutionized health care by leveraging consumer technologies and artificial intelligence (AI) algorithms to enhance diagnostics and disease monitoring [22,23]. In particular, the rapid evolution of smartphone sensors—such as high-resolution cameras and Light Detection and Ranging–based depth-sensing—has opened new possibilities for medical applications [24-28]. These built-in capabilities, coupled with AI methods such as computer vision (CV), enable precise motion tracking, real-time biometric analysis, and automated assessments, reducing the need for specialized equipment. This advancement is particularly valuable in fields requiring vision-based analyses, such as facial nerve assessment and rehabilitation, and might therefore provide an ideal solution that would provide time-efficient, user-friendly, and reproducible assessments, minimizing interobserver variability while maintaining diagnostic accuracy [23]. Indeed, previous pilot studies by our group demonstrated the feasibility of objective FNP assessment as well as app-supported facial rehabilitation using the advanced facial recognition and tracking capabilities afforded by Apple's TrueDepth camera technology [29,30]. However, the clinical utility of these functionalities remained undetermined.

Here, we evaluate a novel smartphone app combining objective FNP grading and structured rehabilitation training in a clinical setting. The app leverages the iPhone's TrueDepth camera and on-device CV algorithms to generate a Digital Facial Index (DFI), enabling rapid and reproducible facial movement assessment. Additionally, it features a neuromuscular training program with real-time biofeedback and educational video content to enhance patient engagement and guide proper exercise execution. This solution addresses key limitations in FNP assessment and rehabilitation, offering a highly accessible, scalable, and objective tool.

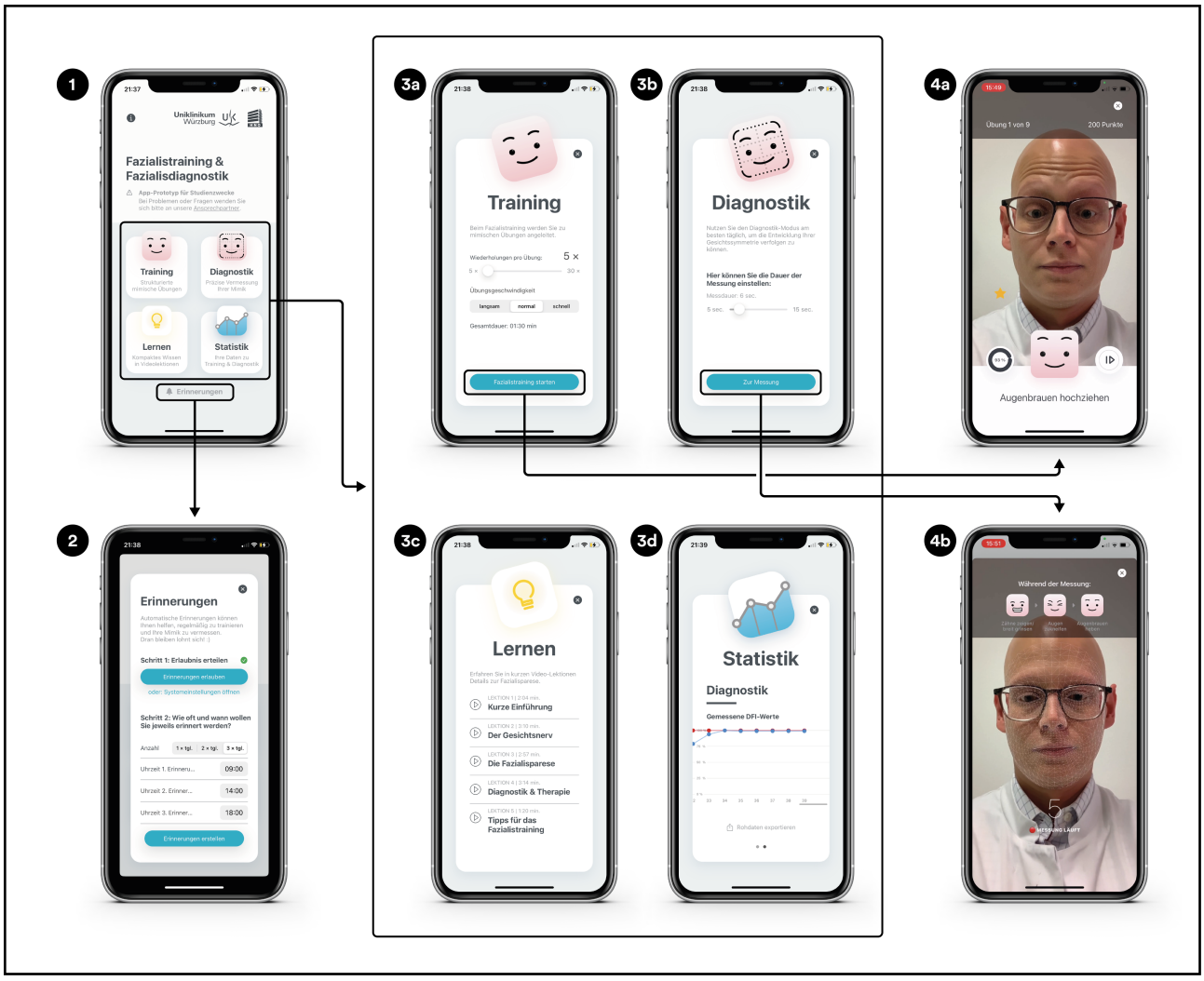
Methods

App Interface

The app's diagnostic and training capabilities have been extensively described in the literature [29,30]. In summary, in addition to its core functions, including a facial training program with real-time biofeedback, in which users view a live camera-based reflection of their own face, and facial diagnosis, as outlined in the introduction, the app also features an educational component with easy-to-understand videos explaining the clinical pathophysiology of acute facial palsy. The training module is based on a modified facial exercise protocol described by Okreu and Beckers [31] and comprises 9 standardized facial movements, including eyebrow elevation, eye closure, cheek inflation, lip pursing, lateral mouth movements to the right and left, wide mouth opening, elevation of the mouth corners, and showing the teeth. Exercises are performed in front of the smartphone's front-facing camera and are supported by real-time visual feedback. Instead of conventional paper-based instructions, exercise execution is guided by animated visual cues and accompanying text instructions. The number of repetitions and training pace can be individually adjusted, and a percentage-based feedback score indicates the accuracy of movement execution. In the current version, the training module is static and provides the same standardized, movement-based facial exercises to all users, independent of DFI values. Users can monitor their individual progress in managing facial paralysis through statistical data and set daily training reminders. An overview of the study app interface is presented in Figure 1.

XSL•FO
RenderX

Figure 1. Overview of the study app interface (German version). (1) Home screen with buttons for the 4 main components: facial training ("Training"), diagnostics ("Diagnostik"), learning ("Lernen"), and statistics ("Statistik"), as well as the reminder function. (2) Reminder screen with up to 3 daily reminders. (3a/3b) Setting screens for facial training and facial diagnosis. (3c) Learning screen with 5 video lessons on the subject of facial palsy. (4a) Facial neuromuscular retraining with real-time biofeedback using animated visual cues. (4b) Facial nerve diagnostics with a specification of 3 grimaces, which the user must perform within the previously set time window (in 3b). The app interface is shown in its original German-language version. Written informed consent for the use of the image was obtained from the individual shown.



App Development

As in the preliminary work, a native app architecture under iOS (Apple Inc.) was selected [29]. This approach allowed for the efficient transfer of existing code directly into the study app, enabling seamless integration with new modules and ensuring full compatibility with the Face ID technology underlying the DFI measurement. The use of on-device computer algorithms via Apple's ARKit enabled real-time facial analysis without requiring data to leave the device. Common graphic and animation tools, such as the Adobe Creative Suite (Photoshop, Illustrator, and After Effects; Adobe Inc.) and Sketch (Bohemian Coding), were used to design the user interface. The app was developed using the Swift programming language within the XCode development environment (Apple Inc.). Open-source code libraries were integrated via the CocoaPods (Eloy Durán and Fabio Pelosin) dependency manager [32], including "Hero" for general view controller transitions [33], "ScrollableGraphView" for displaying graphs in the statistics function [34], and "Lottie" for vector-based animations used in the animated smileys of the training function [35].

Grading of FNP

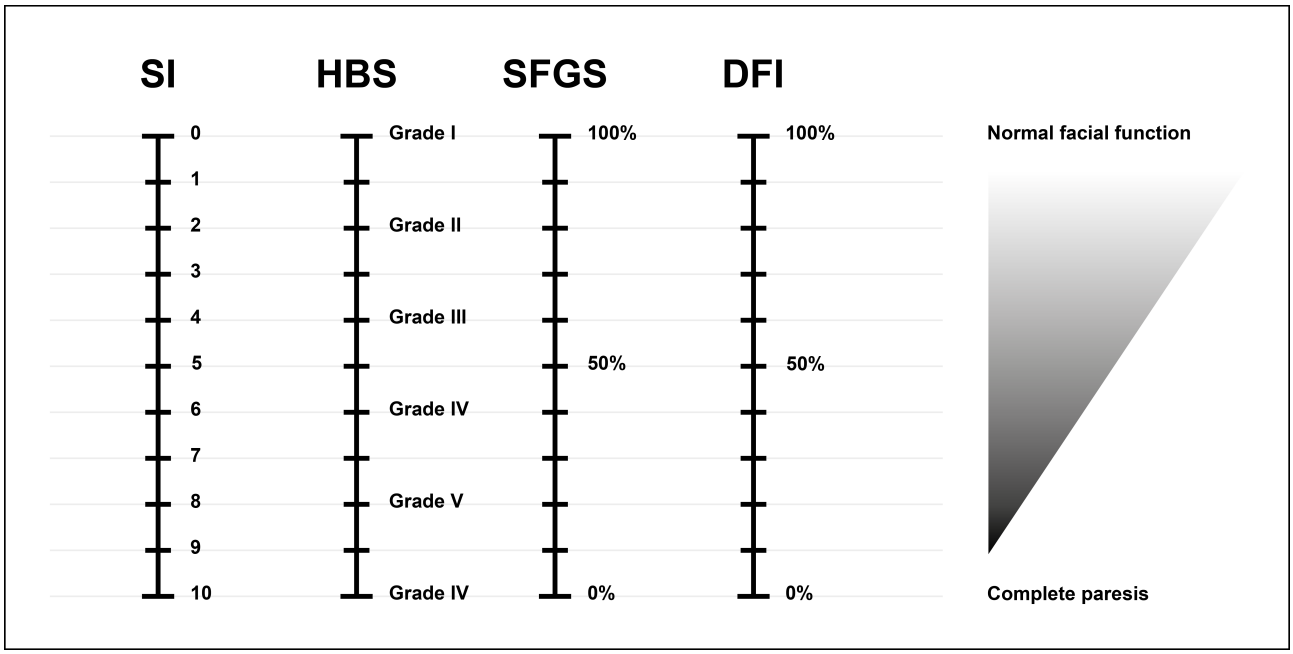
The facial grading tool of the app objectively quantifies facial movements, generating a result known as the DFI. The DFI is calculated based on region-specific facial movement amplitudes derived from 3D depth data captured by the smartphone's TrueDepth camera during standardized facial expressions. Movements of the forehead, eye region, and mouth are quantified separately, with the affected side normalized relative to the contralateral side. These values are combined into a weighted composite score (forehead 10%, eye region 40%, and mouth region 50%) to generate the final DFI expressed as a percentage. The DFI is expressed on a scale from 0% to 100%, where 0% represents complete paresis and 100% indicates normal, intact facial nerve function. Thus, a maximum value of 100% corresponds to fully preserved and normal facial nerve function.

To compare the DFI with clinically established scoring systems, the SI, HBS, and SFGS were assessed by 3 experienced raters (2 ENT specialists and 1 dental doctoral candidate) using standardized facial videos. Raters were blinded to the DFI

XSL•FO
RenderX

results, patient identity, and each other's assessments. The SI is a historically established grading system commonly used in German-speaking clinical practice and uses inverse scoring, with lower scores indicating better facial nerve function. In relation to these conventional clinical scores commonly used to evaluate facial function, the DFI corresponds directly to the SFGS (0% - 100%) [12] and inversely to the SI (0 - 10) [18] and HBS (grade I-VI) [11]. For example, a DFI of 60% is equivalent to an SI of 4, an HBS grade of III, and an SFGS score of 60%. To enhance comparability, all scores were standardized so that 100% reflects fully intact facial function. A conversion chart aligning the various clinical scores (SI, HBS, SFGS, and DFI) is provided in Figure 2.

Figure 2. Conversion scheme for the 4 different scores: Stennert Index, House-Brackmann Scale, Sunnybrook Facial Grading System, and Digital Facial Index. DFI: Digital Facial Index; HBS: House-Brackmann Scale; SFGS: Sunnybrook Facial Grading System; SI: Stennert Index.



Subjective User Assessment

The Facial Disability Index (FDI) is a disease-specific, self-reported questionnaire designed to assess the physical and psychosocial disability associated with facial neuromuscular disorders [36]. It consists of 2 subscales: physical function, which evaluates difficulties in essential activities such as eating, drinking, speaking, and oral hygiene, and social and well-being function, which assesses the emotional and social impact of facial dysfunction, including social withdrawal, irritability, and psychological distress. The FDI has been validated as a reliable tool for measuring patient-reported disability and treatment outcomes, providing a more comprehensive assessment than general health-related quality of life measures [36]. The exact items of the FDI can be found in the Results section.

The System Usability Scale (SUS) is a widely used, standardized questionnaire designed to assess the overall usability of a system [37]. It consists of 10 Likert scale items that evaluate aspects such as ease of use, complexity, consistency, and the need for technical support. The SUS provides a single composite score ranging from 0 to 100, offering a global measure of subjective usability. Originally developed for industrial usability assessments, the SUS is a quick, reliable, and cost-effective tool that has been validated across various domains, making it suitable for comparing different systems and tracking usability improvements over time [37]. A score of 85 - 100 is considered "excellent" usability, indicating that the system largely meets user expectations; a score of 70 - 84 is classified as "good"; a score of 50 - 69 is deemed "acceptable"; and a score below 50 is categorized as "poor," indicating significant usability issues that require comprehensive improvements [38]. The exact items of the SUS can be found in the Results section.

The study-specific questionnaire is a structured, self-reported assessment tool designed to evaluate user experience and engagement with the facial app used in this study. It consists of 10 Likert scale items, covering aspects such as overall satisfaction, clarity and usefulness of educational videos, ease of use of the diagnostic function, motivation for training, and preference compared to traditional paper-based exercises. Additionally, the questionnaire assesses the effectiveness of features such as progress tracking and reminders in supporting adherence to therapy. An open-ended section allows participants to provide qualitative feedback and suggestions for improvement, making the questionnaire a comprehensive tool for assessing user acceptance and potential optimizations of the app. The exact items of the study-specific questionnaire can be found in the Results section.

Study Design

The study app was made available to the participants either on loaned iPhones (XR) or on their own compatible devices.

Patients were eligible for inclusion if they had facial palsy of any etiology, provided that a routine diagnostic evaluation had been conducted in the ENT clinic or neurology department of the University Hospital of Würzburg. Only patients in the acute stage of facial palsy, defined as symptom onset within a

XSL•FO
RenderX

maximum of 7 days prior to study inclusion, were enrolled. Additional inclusion criteria required participants to be at least 18 years of age and in a clinical condition that permitted study participation, meaning they were awake, cooperative, and capable of providing informed consent. Exclusion criteria comprised patients whose general health status did not allow for informed consent or study participation. Furthermore, individuals with other diseases or prior surgeries in the head and neck region, unrelated to facial palsy but potentially influencing measurement outcomes, were excluded from the study.

Over the course of the 4-week study period, participants were encouraged to use the app regularly for training and diagnostics. At both the beginning and end of the study, standardized videos were recorded according to the protocol suggested by Schaede et al [39], from which 3 experienced raters assessed 3 clinical scores: SI [18], HBS [11], and SFGS [12]. In addition, the DFI was recorded at these time points, and disease-specific quality of life was assessed using the German version of the FDI [36,40]. At the final appointment, the SUS according to Brooke [37] was used to evaluate the usability of the app, and further questions were collected through a tailored, study-specific questionnaire consisting of 10 items. To assess therapy frequency and duration, usage data were retrieved from the app at the final appointment (the app includes a logbook feature that tracks the respective usage times and periods).

Statistical Analysis

No a priori power analysis was conducted. Due to the exploratory nature of the study, effect sizes are reported alongside *P* values to aid in interpretation. Prior to the analysis of scores and the FDI, the normality of the data was assessed using the Shapiro-Wilk test. Based on the distributional characteristics of the data, nonparametric or parametric tests were applied where appropriate. To compare the facial scores and the FDI between the beginning and the end of the study, the Wilcoxon matched-pairs signed-rank test was used for nonparametric tests and a paired *t* test was used for parametric tests. Differences were considered statistically significant at a significance level of α=.05.

Both Spearman rank correlations were used to assess the relationship between the newly developed Digital Facial Index (DFI) and existing facial palsy scores (SI, HBS, and SFGS). Spearman correlation was used to evaluate the strength and direction of monotonic relationships, making it suitable for ordinal and nonnormally distributed data.

To evaluate the degree of agreement between the DFI and the other facial palsy scores and to identify potential systematic bias, Bland-Altman plots were generated. These plots are presented with bias (mean difference), SD of the bias, and the 95% limits of agreement (LoA).

Intraclass correlation coefficients (ICCs) were calculated to compare the DFI with the SI, HBS, and SFGS, as well as to assess intrarater variability across 3 repeated measurements within each score. Based on the methodology described by Wirtz [41] and Koo and Li [42], a 2-way mixed model was applied, and average-measure ICC values were reported. ICC was tested for absolute agreement and interpreted according to the classification proposed by Cicchetti [43]: unacceptable (ICC<0.4), fair (0.4≤ ICC<0.6), good (0.6≤ ICC<0.75), and excellent (ICC≥0.75). Furthermore, to analyze the consistency of the 3 repeated measurements within each score, ICC with a 2-way mixed model and single-measure values and the associated Cronbach α were computed. Cronbach α was interpreted based on the following scale: unacceptable (α<0.7), fair (0.7≤ α<0.8), good (0.8≤ α<0.9), and excellent (α≥0.9). Only statistically significant ICC and Cronbach α values were considered for interpretation.

Statistical analyses and data visualization were conducted using GraphPad Prism (version 8.4.0; GraphPad Software) and SPSS (version 30.0.0.0; IBM Corp). Data are presented in bar charts, scatter plots, and Bland-Altman plots.

Ethical Considerations

The prospective study was approved by the Ethics Committee of the University of Würzburg (file number 96 - 20-sc), and informed consent was obtained from all participants prior to their inclusion in the research. All participants provided written informed consent prior to inclusion in the study. As facial video recordings were required for the assessment of facial function, the study involved identifiable data. These data were handled confidentially and in accordance with applicable data protection regulations, and access was restricted to authorized study personnel. No financial compensation was provided for participation in this study.

Results

Facial Function: Study's Bookends

A total of 21 patients were included in the study, comprising 11 males (52.4%) and 10 females (47.6%) with a mean age of 40.1 (SD 11.33) years. The underlying causes of unilateral FNP were idiopathic in 10 cases (47.6%) and inflammatory and neoplastic in 7 cases (33.4%), including 6 cases of zoster oticus (28.6%) and 1 case of T2 acoustic neuroma (4.8%). Additionally, 4 cases (19%) were classified as traumatic, comprising 2 cases of FNP following traumatic brain injury with petrous bone fracture (9.5%) and 2 cases of iatrogenic postoperative facial nerve injury (9.5%), which occurred following type III tympanoplasty and parotidectomy. Regarding laterality, facial palsy affected the left side in 11 cases (52.4%) and the right side in 10 cases (47.6%). All participants were included in the acute phase of peripheral facial palsy (≤7 days after symptom onset). Synkinesis, assessed using the synkinesis subscore of the SFGS, was rated as 0 by all 3 blinded raters at both baseline and follow-up.

Overall, a statistically significant improvement in facial nerve function was observed across all 4 scoring systems from baseline to study completion. At the onset of the study, the median facial nerve function levels were 43.3% (IQR 33-77) for SI, 40% (IQR 27-67) for HBS, 34.3% (IQR 21-70) for SFGS, and 58.3% (IQR 36-98) for DFI. By the end of the study, these values had all significantly improved to 96.7% (IQR 87-100) for SI (*P*<.001), 100% (IQR 73-100) for HBS (*P*<.001), 97% (76-99) for SFGS

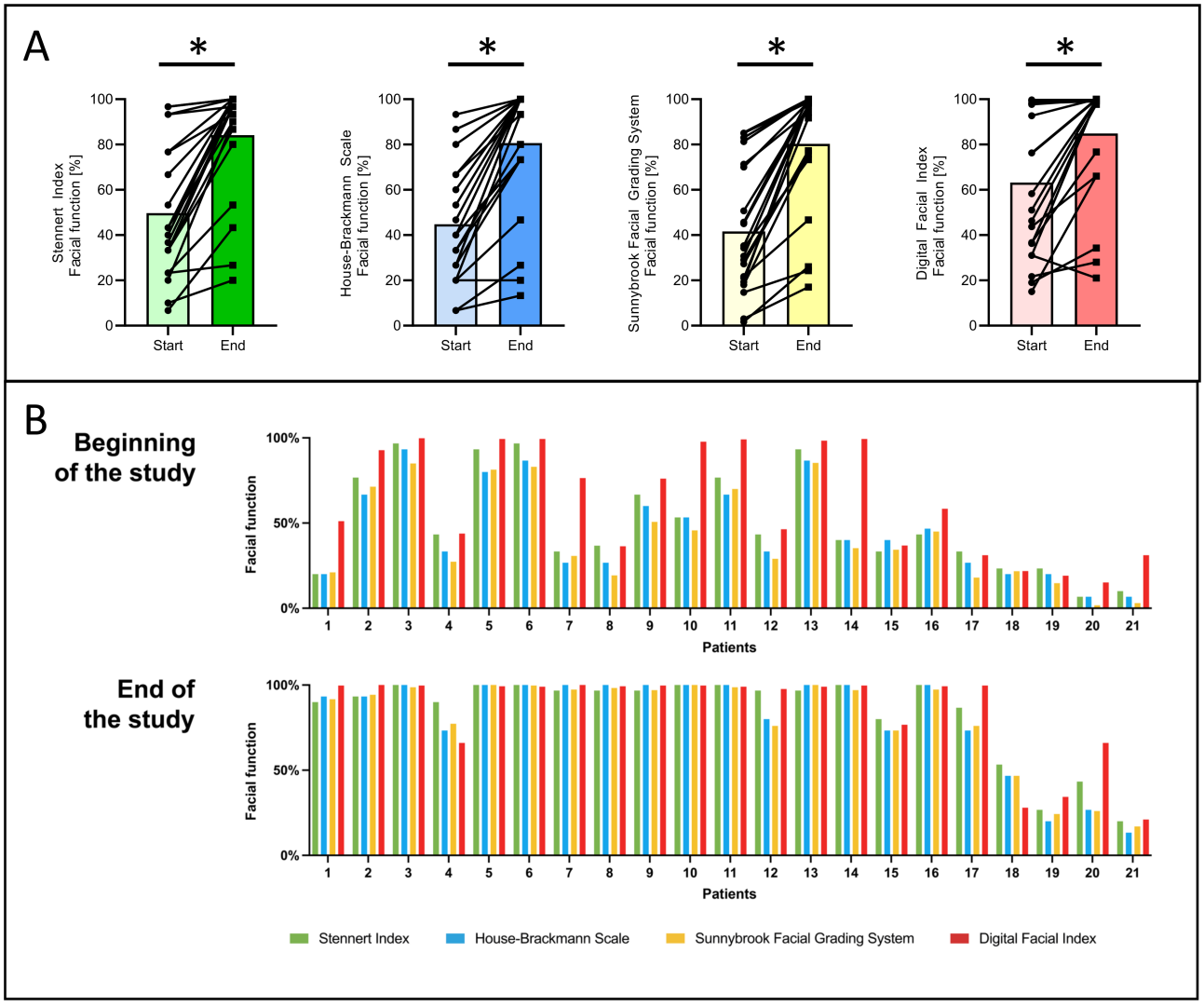
XSL•FO
RenderX

(*P*<.001), and 99.3% (IQR 77-100) for DFI (*P*<.001; Figure 3A).

Figure 3B

A more detailed examination of individual patient data shows a high degree of convergence between the DFI and established clinical scores, particularly at the end of the study. While some variability was observed at baseline, overall values aligned more closely over time. Figure 3B provides an overview of individual score trajectories, illustrating that changes observed in clinical grading scales are paralleled by corresponding changes in DFI values. These findings support the clinical validity of the DFI and its responsiveness to functional changes over time.

Figure 3. Comparison of facial function at the beginning and end of the study. Line plots illustrating facial function at the beginning and end of the study. A significant improvement is evident across all scoring systems. (B) The Digital Facial Index shows the closest correlation with the House-Brackmann Scale results. A more detailed examination of all participants can be found in (B). At the start of the study, the Digital Facial Index appears to produce overly optimistic results in certain cases compared to the other scoring systems (eg, patient 7). By the end of the study, the results are more consistent, with the Digital Facial Index largely providing an accurate representation of normal facial function. Differences between the different cohorts are indicated as significant,*P<.001 (paired t test and Wilcoxon matched-pairs signed-rank test). DFI: Digital Facial Index; HBS: House-Brackmann Scale; SFGS: Sunnybrook Facial Grading System; SI: Stennert Index.

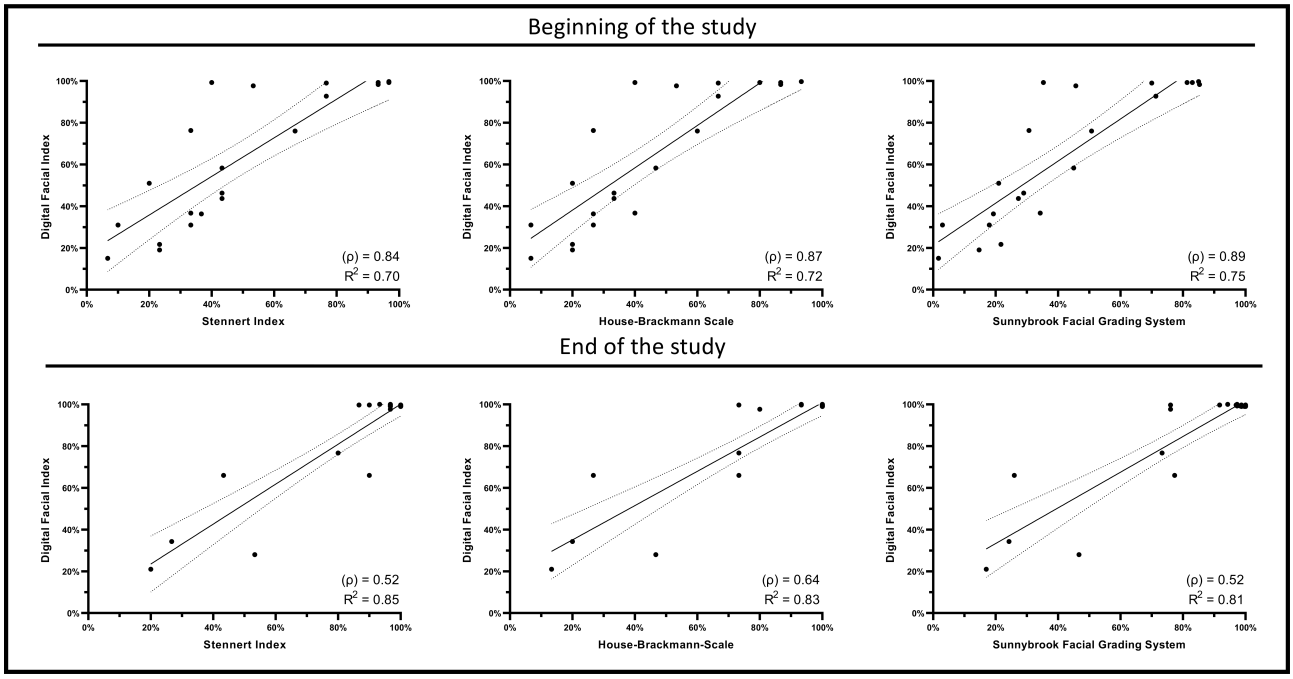


Facial Grading Scores: Comparative Analysis

In regression analyses, the 3 clinically established scoring systems were compared with the DFI. At the beginning of the study, the Spearman correlation coefficients (ρ) were 0.84 for the SI with the DFI, (ρ)=0.87 for the HBS with the DFI, and (ρ)=0.89 for the SFGS with the DFI, all indicating a strong correlation. By the end of the study, the correlation coefficient for the SI with the DFI decreased to (ρ)=0.52, while the HBS and the SFGS showed correlation coefficients of (ρ)=0.64 and (ρ)=0.52, respectively (Figure 4).

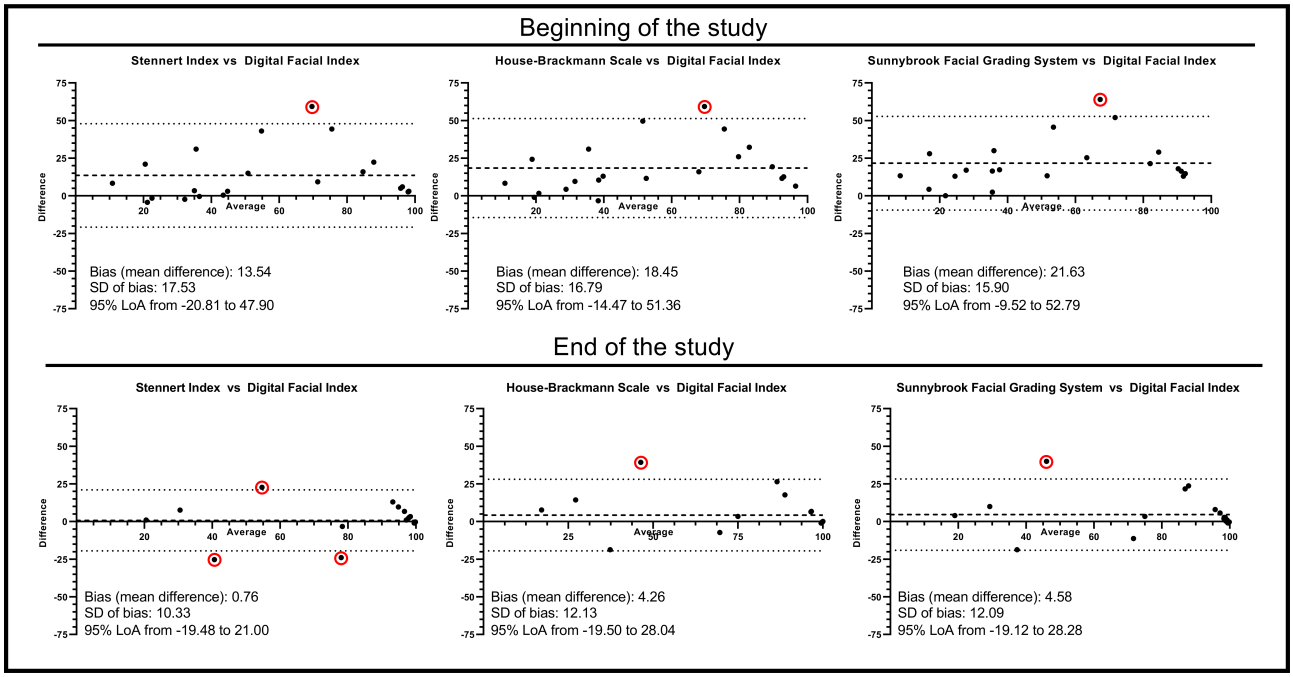
XSL•FO
RenderX

Figure 4. Correlation analysis comparing Digital Facial Index and the other 3 clinical scores shows a good to strong correlation at both the beginning and end of the study. Note that the scores were also adapted here as shown in Figure 2.



Furthermore, Bland-Altman plots were generated to assess the agreement between the clinically established scoring systems and the DFI. These plots confirmed that, at the beginning of the study, the DFI provided systemically more optimistic results regarding facial function compared to the SI (mean difference 13.54, SD 17.53), HBS (mean difference 18.45, SD 16.79), and SFGS (mean difference 21.63, SD 15.90). By the end of the study, this trend was only partially observed (SI: mean difference 0.76, SD 10.3; HBS: mean difference: 4.26, SD 12.13; SFGS: mean difference: 4.58, SD 12.09). Nevertheless, in all plots, the majority of the points fell within the LoA (items out of LoA: SI: 1 beginning, 3 end; HBS: 1 each beginning and end; SFGS: 2 beginning, 2 end), which generally indicates a satisfactory level of agreement between the scores (Figure 5).

Figure 5. Bland-Altman plots comparing the Digital Facial Index with the Stennert Index, House-Brackmann Scale, and Sunnybrook Facial Grading System at baseline and at the end of the study. LoA: limits of agreement.



To further assess the agreement between the conventional clinical scoring systems and the DFI, both ICC and Cronbach α were calculated. The results were classified as “excellent” (Table 1). Additionally, to evaluate the agreement as well as the internal consistency and reliability of the 3 measurements within the 4 scores (DFI: 3 measurements; SI, HBS, SFGS: 3 raters each providing one rating), both ICC and Cronbach α

XSL•FO
RenderX

were calculated. The results were also classified as "excellent" (Table 1).

ICC and CA were used to statistically compare DFI with SI, HBS, and SFGS (average measures) at the beginning and the end of the study. Interpreting the values as described by Cicchetti [43], there are values at both the beginning and the end of the study that can be classified as "excellent." Furthermore, ICC and CA were used to evaluate agreement within the 3 ratings of the 4 scores (DFI, SI, HBS, and SFGS). As well, all scores were classified as "excellent."

Table . Reliability analysis.

	SI[a] – DFI[b]		HBS[c] – DFI		SFGS[d] – DFI		DFI		SI		HBS		SFGS	
	ICC[e,f]	CA[g,h]	ICC	CA	ICC	CA	ICC	CA	ICC	CA	ICC	CA	ICC	CA
Between scores														
Beginning of the study	0.828[i]	0.910[i]	0.868[i]	0.910[i]	0.828[i]	0.960[i]	—[j]	—	—	—	—	—	—	—
End of the study	0.946[i]	0.958[i]	0.960[i]	0.958[i]	0.946[i]	0.982[i]	—	—	—	—	—	—	—	—
Within scores														
Beginning of the study	—	—	—	—	—	—	0.908[i]	0.968[i]	0.828[i]	0.950[i]	0.908[i]	0.968[i]	0.828[i]	0.950[i]
End of the study	—	—	—	—	—	—	0.931[i]	0.975[i]	0.914[i]	0.977[i]	0.931[i]	0.975[i]	0.914[i]	0.977[i]

[a]SI: Stennert Index.

[b]DFI: Digital Facial Index.

[c]HBS: House-Brackmann Scale.

[d]SFGS: Sunnybrook Facial Grading System.

[e]ICC: intraclass correlation coefficient.

[f]ICC: excellent (ICC≥0.75); good (0.6≤ICC<0.75); fair (0.4≤ICC<0.6); unacceptable (ICC<0.4).

[g]CA: Cronbach α.

[h]CA: excellent (α≥0.9); good (0.8≤α<0.9); fair (0.7≤α<0.8); unacceptable (α<0.7).

[i]Excellent.

[j]Not applicable.

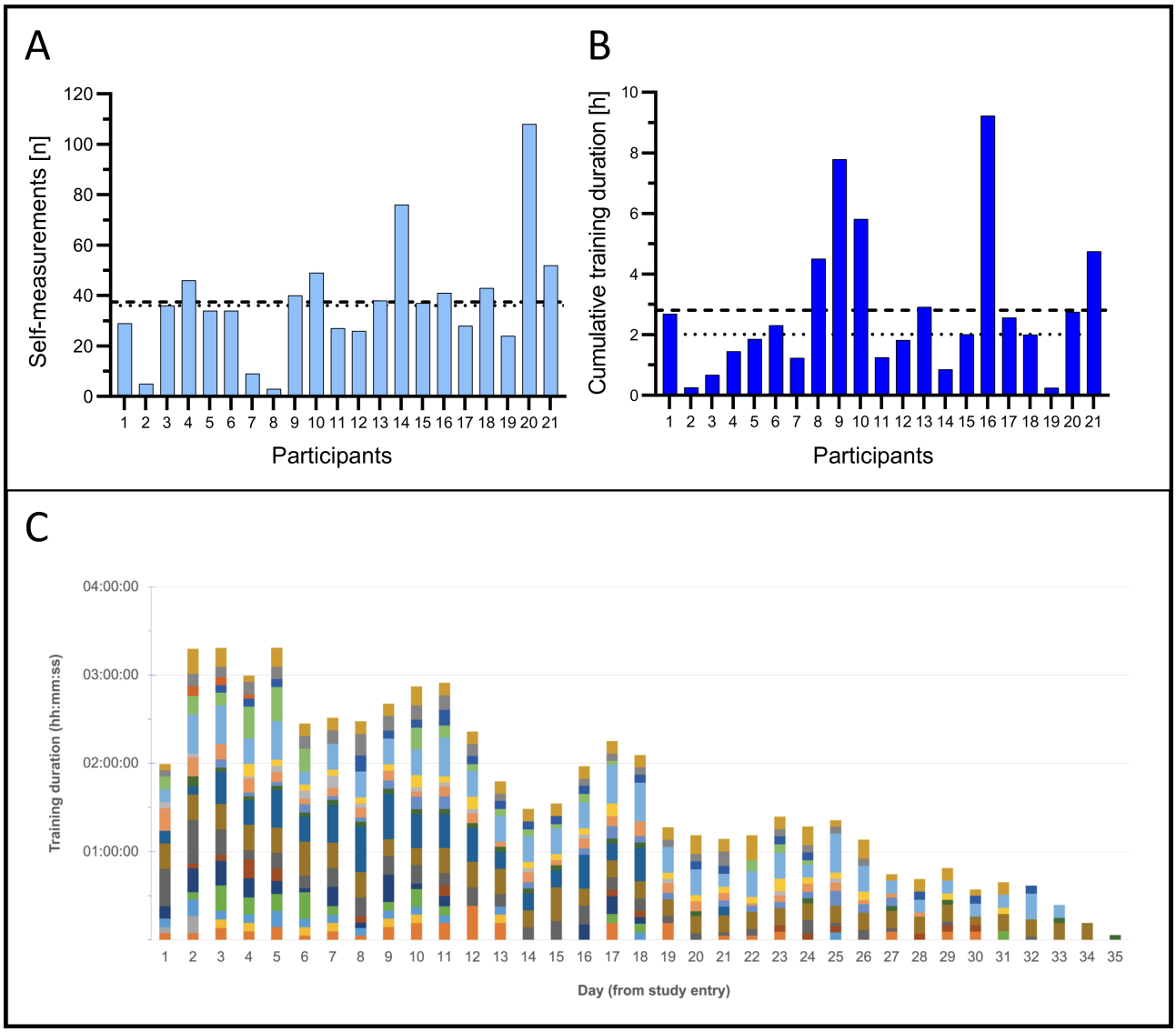
App Usage Frequency

The use of the app was analyzed through a logbook. Regarding the diagnostic function, it was observed that patients used the app in a highly variable manner. The range of usage varied from only 3 applications (case 8) to 108 uses (case 20) of the DFI self-measurements. On average, the participants performed a mean of 37.4 DFI self-measurements (SD 23.0; median 36.0; IQR 27–45) during the study period (Figure 6A).

Regarding the training function of the app, the average mean cumulative usage was 2 hours and 48 minutes (SD 2 h 22 min; median 2 h 0 min, IQR 1 h 15 min to 3 h 43 min) over the 30-day study period (Figure 6B). It should be noted that, due to the individual scheduling of the follow-up appointments, participation sometimes extended beyond the initially planned 30 days up to 35 days from study entry. The cumulative exercise time for all patients is visualized in Figure 6C. A characteristic pattern emerges, resembling a series of progressively lower peaks every sixth day (at approximately days 5, 11, 17, 23, and 29). Between these peaks, training duration exhibits a trough-like decline, creating a wave-like structure. This pattern suggests an initial phase of high engagement, followed by periodic fluctuations and an overall decline in training activity. In the last days of the study, training durations are minimal.

XSL•FO
RenderX

Figure 6. Use of the Digital Facial Index self-measurement and training function within the study app. (A) The Digital Facial Index self-measurements ranged from 3 to 108 (mean 37.4, SD 23.0, dashed line; median 36.0, IQR 27-45, dotted line), shown per participant. (B) Training duration per participant (mean 2 h 48 min, SD 2 h 22 min, dashed line; median 2 h 0 min, IQR 1 h 15 min to 3 h 43 min, dotted line) ranging from 14 min 50 s to 9 h 14 min 33 s. (C) Cumulative facial training time per participant, with each patient represented by a unique color. Due to the individualized scheduling of follow-up appointments, some participants' involvement extended beyond the initially planned 30 days.



Subjective User Assessment

The FDI demonstrated a significant improvement in both physical function and social function and well-being among the participants between the start and end of the study. Physical function improved significantly from a mean of 57.1 (SD 13.2) at the beginning of the study to 85.0 (SD 15.2) at the end of the study (P<.001). Similarly, the average scores for social function and well-being showed a significant improvement, with the score rising from a mean of 49.3 (SD 16.5) at the start of the study to 83.6 (SD 24.5) at the end (P<.001). A detailed list of the questions with the respective means, SDs, and P values can be found in Table 2. Correlations between the FDI and other scoring systems ranged from moderate to strong. For the SI, correlations with the FDI increased from r=0.41 to r=0.61 (physical) and from r=0.50 to r=0.66 (social). With the HBS, correlations rose from r=0.46 to r=0.72 (physical) and from r=0.67 to r=0.74 (social). The SFGS showed increases from r=0.53 to r=0.74 (physical) and from r=0.53 to r=0.71 (social). Correlations with the DFI remained stable for physical function (r=0.50) and increased slightly for social function (r=0.52 to r=0.56).

XSL•FO
RenderX

Table . Evaluation of the Facial Disability Index (according to VanSwearingen and Brach [36] and Volk et al [40]).

Questionnaire items	Facial Disability Index - Item	Beginning of the study, mean (SD)	End of the study, mean (SD)	*P* value
	Physical function			
1	How much difficulty did you have keeping food in your mouth, moving food around in your mouth, or getting food stuck in your cheek while eating?	3.3 (0.8)	4.4 (0.8)	<.001[a]
2	How much difficulty did you have drinking from a cup?	3.1 (0.8)	4.5 (0.7)	<.001[b]
3	How much difficulty did you have saying specific sounds while speaking?	3.8 (0.9)	4.5 (0.9)	<.001[c]
4	How much difficulty did you have with your eye tearing excessively or becoming dry?	3.0 (0.6)	4.2 (0.5)	<.001[b]
5	How much difficulty did you have with brushing your teeth or rinsing your mouth?	3.3 (0.6)	4.4 (0.7)	<.001[b]
	Overall score in %	57.1 (13.2)	85.0 (15.2)	<.001[b]
	Social and well-being function			
6	How much of the time have you felt calm and peaceful?	2.8 (1.0)	4.5 (0.7)	<.001[b]
7	How much of the time did you isolate yourself from people around you?	2.8 (1.0)	4.4 (1.0)	<.001[b]
8	How much of the time did you get irritable toward those around you?	3.3 (1.1)	4.3 (1.1)	<.001[a]
9	How often did you wake up early or wake up several times during your nighttime sleep?	3.3 (1.1)	4.3 (1.2)	<.001[c]
10	How often has your facial function kept you from going out to eat, shop, or participate in family or social activities?	2.7 (0.8)	4.2 (1.3)	<.001[a]
	Overall score in %	49.3 (16.5)	83.6 (24.5)	<.001[b]

[a]*P*<.001.

[b]*P*<.0001.

[c]*P*<.01.

The SUS indicated that the system was generally perceived as highly user-friendly. It should be noted that the SUS score ranges from 0 to 4, with 4 representing the most favorable response, and that negative items (items 2, 4, 6, 8, and 10) are reversed to ensure consistency in interpretation. Positive items, such as “I thought the system was easy to use” and “I would imagine that most people would learn to use this system very quickly,” scored an average of 3.81 (SD 0.51) and 3.62 (SD 0.50), respectively. Negative items, such as “I found the system unnecessarily complex” and “I think that I would need the support of a technical person to be able to use this system” scored 3.57 (SD 0.68) and 3.52 (1.12), suggesting that participants did not perceive these aspects negatively. The overall mean SUS score was 88.33 (SD 15.36; out of 100), which is considered excellent, indicating high user-friendliness and that the system met users’ expectations to a large extent. A detailed list of the questions with the respective means and SDs can be found in Table 3.

XSL•FO
RenderX

Table . System Usability Score (according to Brooke [37]).

Item	Item end of the study	Positive (P), Negative (N) item	Value, mean (SD)
1	I think that I would like to use this system frequently	P (0 - 4)	2.67 (1.28)
2	I found the system unnecessarily complex	N (0 - 4 reversed)	3.57 (0.68)
3	I thought the system was easy to use	P (0 - 4)	3.81 (0.51)
4	I think that I would need the support of a technical person to be able to use this system	N (0 - 4 reversed)	3.52 (1.12)
5	I found the various functions in this system were well integrated	P (0 - 4)	3.67 (0.66)
6	I thought there was too much inconsistency in this system	N (0 - 4 reversed)	3.67 (0.66)
7	I would imagine that most people would learn to use this system very quickly	P (0 - 4)	3.62 (0.50)
8	I found the system very cumbersome to use	N (0 - 4 reversed)	3.57 (0.68)
9	I felt very confident using the system	P (0 - 4)	3.67 (0.73)
10	I needed to learn a lot of things before I could get going with this system	N (0 - 4 reversed)	3.57 (0.68)
	SUS[a] (sum of all points x 2.5)	Range: 0 - 40	88.33 (15.36)

[a]SUS: System Usability Scale.

A study-specific questionnaire was used to assess participants' perceptions of the facial nerve rehabilitation app (Taeger) and its various functions. Overall, the results indicate a predominantly positive evaluation of the app, with a mean score of 4.1 (SD 0.9). Participants particularly appreciated the clarity and usefulness of the instructional videos (mean 4.9, SD 0.3) and the explanation of training exercises (mean 4.9, SD 0.3). The app was generally well-received, with high enjoyment ratings (mean 4.5, SD 0.9) and straightforward usability of the diagnostic function (mean 4.4, SD 0.7).

However, user engagement with certain features was more variable. While the diagnostic function provided motivation for regular training (mean 3.9, SD 1.1), the training results (mean 3.2, SD 1.2), statistics function (mean 3.1, SD 1.4), and reminder function (mean 3.1, SD 1.3) were rated lower. Despite these aspects, the majority of participants preferred using the app over the paper-based exercise log (mean 3.9, SD 1.3) and supported the app as a standard supplementary therapy for facial palsy (mean 4.8, SD 0.5). A detailed list of the questions with the respective means, medians, and SDs can be found in Table 4.

XSL•FO
RenderX

Table . Study-specific questionnaire. Likert scale with 1 (strongly disagree) to 5 (strongly agree).

Item end of the study	Value, mean (SD)
I enjoyed using the facial rehabilitation app.	4.5 (0.9)
The instructional videos were easy to understand and helpful.	4.9 (0.3)
Using the diagnostic function was straightforward.	4.4 (0.7)
The diagnostic function motivated me to use the training regularly.	3.9 (1.1)
The training exercises were clearly explained.	4.9 (0.3)
The training results encouraged me to continue with further training sessions.	3.2 (1.2)
I regularly used the statistics function.	3.1 (1.4)
The reminder function helped me adhere consistently to my training sessions.	3.1 (1.3)
I prefer using the facial rehabilitation app over the paper-based exercise log.	3.9 (1.3)
The facial rehabilitation app should be offered as a standard supplementary therapy option for facial palsy.	4.8 (0.5)
Overall score (mean)	4.1 (0.9)

Discussion

Facial Function

Digital apps are increasingly integral to modern medicine, leveraging smartphone sensors for diverse clinical applications. From heart rate monitoring [44-46] to retinal imaging [47-49] and sleep tracking [50-52], smartphones have demonstrated their potential across various medical specialties. In this context, the iPhone's TrueDepth camera has been explored for assessing facial nerve function, as first described by Taeger et al [29,30]. This study prospectively evaluated a corresponding app-based diagnostic and training approach in a clinical cohort. Our findings highlight 3 main strengths of the DFI: first, it demonstrated high convergent validity, with strong correlations to established clinical grading systems such as the HBS, SFGS, and SI, confirming its consistency with clinical standards. Second, the DFI may offer higher sensitivity in detecting subtle improvements over time, potentially capturing changes that conventional ordinal scales may overlook. Third, user feedback was positive, indicating good usability and patient acceptance of the app-based approach.

A cohort of 21 patients with unilateral facial palsy was assessed under standardized conditions using the newly developed DFI, alongside 3 clinician-graded scoring systems recommended by the German guideline for facial palsy management [5]: the SI, the HBS, and the SFGS. Our findings indicate a strong correlation between the DFI and conventional grading methods over a 4-week follow-up, suggesting that this app-based tool offers a standardized, digital approach to facial function assessment. By integrating objective motion analysis with smartphone accessibility, the DFI may enhance diagnostic consistency and facilitate long-term patient monitoring.

Across all scoring systems, facial nerve function improved significantly over 4 weeks. The initially higher values measured with the DFI suggest that it tends to underestimate the severity of facial palsy. Therefore, correction factors or an optimized analysis algorithm may need to be applied in the future. However, given the exploratory nature of this pilot study and the limited sample size, no statistical correction for the observed DFI overestimation was applied, as such adjustments would be unstable and potentially misleading. Robust calibration approaches will require larger, stratified studies with multiple measurement time points and in-depth raw data analysis. By study completion, these discrepancies diminished, with all measures converging toward near-complete recovery, consistent with the natural course of idiopathic facial palsy, where most cases resolve within 3 weeks [53,54]. As facial function improves across participants, variability between scoring methods naturally decreases. Given that all patients were examined in the acute phase of facial palsy, the absence of synkinesis is in line with the expected disease course and indicates that the present findings primarily apply to early-stage facial nerve dysfunction.

Regression analyses showed strong initial correlations between the DFI and traditional grading systems (ρ=0.84 - 0.89), which decreased by study end (ρ=0.52 - 0.64). This decline likely reflects reduced variance as patients approached full recovery, resulting in ceiling effects where categorical scales fail to capture subtle residual asymmetries. Similar patterns were described by Katsumi et al [55], who reported reduced score variability in late recovery stages when comparing 3D imaging-based assessments to the HBS [55]. In contrast, the DFI appears more sensitive to minor deficits, potentially indicating greater granularity or capturing distinct aspects of facial nerve recovery. Bland-Altman analyses confirmed an initial bias toward higher DFI scores, which diminished over time, supporting overall agreement with conventional assessments. Reliability metrics, including ICC and Cronbach α, demonstrated excellent reproducibility across all tools, underscoring the robustness of the DFI. However, the divergence in late-stage correlations raises the question of whether the DFI measures distinct recovery aspects or simply offers greater

XSL•FO
RenderX

granularity. Further studies with larger, more diverse populations are needed to validate its clinical utility.

It is not uncommon for different clinical grading scales to yield varying results, as each system emphasizes distinct aspects of facial nerve dysfunction. The HBS, for instance, a 6-grade categorical system, is widely used for its simplicity but exhibits substantial grading shifts when converted to other scales. In contrast, the SFGS provides finer gradations, improving precision but requiring more detailed assessment. Studies have demonstrated the superior reliability of SFGS over HBS [56-58], highlighting the advantages of nuanced grading.

While increased complexity enhances accuracy, it can also hinder routine clinical use due to time constraints [56]. The DFI, using a 0% - 100% scale within a smartphone-based app, streamlines data collection and reduces clinician workload. A key innovation is the integration of on-device CV algorithms that automatically analyze facial movements captured via the smartphone camera. This enables high-fidelity, real-time tracking without manual input, improving accessibility and efficiency. Compared to eFACE [21,59], the current digital standard for facial nerve assessment, the DFI offers a key advantage: fully automated video-based data acquisition. While eFACE is highly reliable, it requires manual input, limiting practicality in time-sensitive environments. Independent recording of the DFI by the patient is possible after just a few trials within a time window of only 5 seconds. Automated tools such as the Facogram (O'Reilly et al [60]) have been shown to enhance grading consistency and reduce subjectivity [60,61]. Similar to the Facogram, the DFI aims to improve reproducibility and minimize interobserver variability, offering a real-time, app-based solution for longitudinal monitoring.

App Usage Frequency

The analysis of app usage frequency provided insights into user engagement with the diagnostic and training functions of the digital facial app. DFI self-measurements varied widely among participants, ranging from 3 to 108 applications. This broad distribution suggests that while some patients actively integrated the app into their routine, others engaged only sporadically. The mean of 37.4 self-measurements (SD 23.0) indicates regular use, reflecting a general willingness to incorporate digital self-assessments into rehabilitation. Similar trends have been observed in other mobile health (mHealth) interventions, where engagement variability is influenced by individual motivation, technological familiarity, and perceived benefit [62-65].

The training function showed a cumulative usage of 2 hours 48 minutes, with wave-like engagement patterns similar to those seen in digital rehab tools [66,67]. While declining use over time could reflect waning motivation [68], it may also indicate functional recovery and reduced perceived need, consistent with prior findings [69-71]. Although efficacy cannot be assessed without controlled studies, usage trends suggest good acceptance of app-based training. Due to the short observation period and the heterogeneous usage patterns, app usage data were interpreted descriptively and not used to infer training efficacy.

Despite these fluctuations, usage patterns support the feasibility of app-based rehabilitation for facial nerve dysfunction. The DFI's integration into a mobile platform facilitated frequent, flexible self-assessment—an advantage over clinic-based evaluations. High variability in engagement suggests that while some patients adopt digital tools seamlessly, others require structured guidance or reminders [72,73]. Research indicates that engagement in digital health interventions can be improved through gamification, personalized reminders, or real-time feedback [74,75].

Similar variability has been observed in other mHealth apps for facial nerve rehabilitation. Machetanz et al [76] reported differing engagement levels with the FACEsemper app (Machetanz et al [76]), with higher adherence among patients with more pronounced functional deficits. Their study highlights the role of perceived need in sustaining engagement and suggests that automated reminders and tailored exercises could enhance adherence.

Incorporating individualized training plans, manual assistance prompts, or region-specific exercise recommendations may further optimize adherence and therapeutic efficacy. Future studies should explore whether such refinements sustain engagement and improve functional recovery compared to conventional rehabilitation approaches.

Subjective User Assessment

The subjective evaluation of the app-based rehabilitation approach for FNP showed promising results, with significant improvements in disease-specific quality of life (FDI). The increase in physical function (mean 57.1, SD 13.2 to mean 85.0, SD 15.2; $P<.001$) and social well-being (mean 49.3, SD 16.5 to mean 83.6, SD 24.5; $P<.001$) suggests that digital rehabilitation tools may support motor recovery and broader psychosocial well-being, consistent with previous mHealth studies [77,78]. However, these improvements likely reflect the natural recovery trajectory of FNP rather than a direct effect of the app. While the intervention facilitated structured rehabilitation, its precise impact on quality of life remains unclear. Future controlled studies are necessary to differentiate app-related benefits from spontaneous recovery. Nonetheless, these findings highlight the role of digital tools in rehabilitation and the psychosocial management of FNP. Currently, no studies have directly examined the FDI in digital rehabilitation contexts, such as mHealth solutions. However, integrating the FDI into digital programs, including the app used in this study, could provide valuable insights into their effectiveness.

Usability is critical for long-term adoption. The app achieved a SUS score of 88.3, exceeding the commonly reported average of 68 for digital health apps [37,79], which is also confirmed for digital health apps [80]. This suggests that patients found the app intuitive, well-integrated, and easy to navigate, with minimal technological barriers. Our SUS results align with findings by Machetanz et al [76], who reported a SUS score of 89.6 for a mobile facial grading solution.

Despite positive feedback on core functionalities, secondary features such as statistical feedback and reminders received lower ratings (mean 3.1, SD 1.4 and mean 3.1, SD 1.3, respectively), indicating a need for refinement. Research suggests that engagement in digital health interventions can be

XSL•FO
RenderX

enhanced through gamification, personalized progress tracking, and adaptive feedback [81,82]. Future iterations should integrate such elements to optimize user motivation and adherence. A strong preference for the app over traditional paper-based exercise logs underscores a broader shift toward digital health solutions. The ability to perform self-assessments using the DFI and engage in structured rehabilitation without continuous clinical supervision represents a significant advancement in accessibility and patient autonomy in facial palsy management.

This study has several limitations. First, the single-arm design without a control group and the short observation period preclude conclusions regarding therapeutic efficacy. Improvements in facial function are likely dominated by spontaneous recovery, particularly in idiopathic peripheral facial palsy. Second, the sample size was small and heterogeneous with respect to etiology and recovery stage, limiting generalizability. Third, the DFI tended to overestimate facial function in early recovery stages, and no correction factors were applied in this exploratory pilot study.

Limitations

This study has several limitations. The single-arm design without a control group and the short observation period preclude conclusions regarding therapeutic efficacy, as improvements in facial function are likely influenced by spontaneous recovery, particularly in acute peripheral facial palsy. The small and etiologically heterogeneous sample limits generalizability. The DFI tended to overestimate facial function in early recovery stages, and no correction factors were applied in this exploratory pilot study. Furthermore, the app does not currently detect synkinetic movements or monitor muscle overuse, and its educational content does not address synkinesis; therefore, applicability to patients with chronic facial palsy or established synkinesis cannot be inferred. Accordingly, the present findings should be interpreted as evidence of feasibility, usability, and concurrent validity rather than clinical effectiveness. Future controlled studies with larger cohorts and longer follow-up are required.

Conclusion

This study highlights the feasibility of an app-based approach for FNP assessment and rehabilitation. The DFI strongly correlated with conventional grading systems, offering a standardized and objective evaluation method. High usability (SUS=88.3) supports its integration into clinical practice, though engagement varied, emphasizing the need for personalized adherence strategies. Future developments such as gamification, adaptive feedback, and AI-driven personalization could further optimize long-term adherence and therapeutic efficacy.

Acknowledgments

We thank the entire team of the Department of Otorhinolaryngology and the Department of Neurology at the University Hospital of Würzburg for their support in patient recruitment.

Generative artificial intelligence (AI) tools (ChatGPT and OpenAI) were used to support language editing of the manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author and the last author on reasonable request. The custom swift code used in this study is also available from the corresponding and last author on reasonable request. The custom swift code is provided for research and educational purposes only, and any commercial use, including distribution, sale, or licensing to third parties, is not permitted.

Authors' Contributions

FTMG carried out patient acquisition, provided data, performed the analysis, wrote and edited the paper. FE and KH carried out patient acquisition, provided data, and edited the paper. MUF, SH, and KR contributed to the writing, interpretation of the content, and editing of the manuscript, revising it critically. JT conceived and designed the study, programmed the medical app used for the study, provided data, performed the analysis, wrote and edited the paper. All authors approved the completed version.

Conflicts of Interest

None declared.

References

1. Hauser WA, Karnes WE, Annis J, Kurland LT. Incidence and prognosis of Bell’s palsy in the population of Rochester, Minnesota. Mayo Clin Proc 1971 Apr;46(4):258-264. [Medline: 5573820]
2. Gilden DH. Clinical practice. Bell’s Palsy. N Engl J Med 2004 Sep 23;351(13):1323-1331. [doi: 10.1056/NEJMcp041120] [Medline: 15385659]

XSL•FO
RenderX

3. Luijmes RE, Pouwels S, Beurskens CHG, Kleiss IJ, Siemann I, Ingels K. Quality of life before and after different treatment modalities in peripheral facial palsy: a systematic review. Laryngoscope 2017 May;127(5):1044-1051. [doi: 10.1002/lary.26356] [Medline: 27859341]
4. Saadi R, Shokri T, Schaefer E, Hollenbeak C, Lighthall JG. Depression rates after facial paralysis. Ann Plast Surg 2019 Aug;83(2):190-194. [doi: 10.1097/SAP.0000000000001908] [Medline: 31232802]
5. Heckmann JG. Treatment of idiopathic facial palsy (Bell's palsy) [article in German].. DGNeurologie 2022 Nov;5(6):449-465. [doi: 10.1007/s42451-022-00489-w]
6. Pereira LM, Obara K, Dias JM, Menacho MO, Lavado EL, Cardoso JR. Facial exercise therapy for facial palsy: systematic review and meta-analysis. Clin Rehabil 2011 Jul;25(7):649-658. [doi: 10.1177/0269215510395634] [Medline: 21382865]
7. Baricich A, Cabrio C, Paggio R, Cisari C, Aluffi P. Peripheral facial nerve palsy: how effective is rehabilitation? Otol Neurotol 2012 Sep;33(7):1118-1126. [doi: 10.1097/MAO.0b013e318264270e] [Medline: 22872180]
8. Khan AJ, Szczepura A, Palmer S, et al. Physical therapy for facial nerve paralysis (Bell's palsy): an updated and extended systematic review of the evidence for facial exercise therapy. Clin Rehabil 2022 Nov;36(11):1424-1449. [doi: 10.1177/02692155221110727] [Medline: 35787015]
9. Brach JS, VanSwearingen JM. Physical therapy for facial paralysis: a tailored treatment approach. Phys Ther 1999 Apr;79(4):397-404. [Medline: 10201545]
10. Ross B, Nedzelski JM, McLean JA. Efficacy of feedback training in long-standing facial nerve paresis. Laryngoscope 1991 Jul;101(7 Pt 1):744-750. [doi: 10.1288/00005537-199107000-00009] [Medline: 2062155]
11. House JW, Brackmann DE. Facial nerve grading system. Otolaryngol Head Neck Surg 1985 Apr;93(2):146-147. [doi: 10.1177/019459988509300202] [Medline: 3921901]
12. Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. Otolaryngol Head Neck Surg 1996 Mar;114(3):380-386. [doi: 10.1016/S0194-59989670206-1] [Medline: 8649870]
13. Fattah AY, Gurusinghe ADR, Gavilan J, et al. Facial nerve grading instruments: systematic review of the literature and suggestion for uniformity. Plast Reconstr Surg 2015 Feb;135(2):569-579. [doi: 10.1097/PRS.0000000000000905] [Medline: 25357164]
14. Kim SJ, Lee HY. Acute peripheral facial palsy: recent guidelines and a systematic review of the literature. J Korean Med Sci 2020 Aug 3;35(30):e245. [doi: 10.3346/jkms.2020.35.e245] [Medline: 32743989]
15. Vrabec JT, Backous DD, Djalilian HR, et al. Facial Nerve Grading System 2.0. Otolaryngol Head Neck Surg 2009 Apr;140(4):445-450. [doi: 10.1016/j.otohns.2008.12.031] [Medline: 19328328]
16. de Ru JA, Braunius WW, van Benthem PPG, Busschers WB, Hordijk GJ. Grading facial nerve function: why a new grading system, the MoReSS, should be proposed. Otol Neurotol 2006 Oct;27(7):1030-1036. [doi: 10.1097/01.mao.0000227896.34915.4f] [Medline: 17006355]
17. Yanagihara N. Grading of facial palsy. In: Fisch U, editor. Facial Nerve Surgery: Proceedings of the Third International Symposium on Facial Nerve Surgery: Kugler Medical Publications; 1977.
18. Stennert E, Limberg CH, Frentrup KP. An index for paresis and defective healing--an easily applied method for objectively determining therapeutic results in facial paresis (author's transl). HNO 1977 Jul;25(7):238-245. [Medline: 893151]
19. Dusseldorp JR, van Veen MM, Mohan S, Hadlock TA. Outcome tracking in facial palsy. Otolaryngol Clin North Am 2018 Dec;51(6):1033-1050. [doi: 10.1016/j.otc.2018.07.003] [Medline: 30172554]
20. Finkensieper M, Volk GF, Guntinas-Lichius O. Facial nerve disorders. Laryngorhinootologie 2012 Feb;91(2):121-141. [doi: 10.1055/s-0031-1300965] [Medline: 22318445]
21. Banks CA, Bhama PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment: the eFACE. Plast Reconstr Surg 2015 Aug;136(2):223e-230e. [doi: 10.1097/PRS.0000000000001447] [Medline: 26218397]
22. Shajari S, Kuruvinashetti K, Komeili A, Sundararaj U. The emergence of AI-based wearable sensors for digital health technology: a review. Sensors (Basel) 2023 Nov 29;23(23):9498. [doi: 10.3390/s23239498] [Medline: 38067871]
23. Friedrich MU, Relton S, Wong D, Alty J. Computer vision in clinical neurology: a review. JAMA Neurol 2025 Feb 17. [doi: 10.1001/jamaneurol.2024.5326] [Medline: 39960732]
24. Majumder S, Deen MJ. Smartphone sensors for health monitoring and diagnosis. Sensors (Basel) 2019 May 9;19(9):2164. [doi: 10.3390/s19092164] [Medline: 31075985]
25. Shaheen A, Qazalbash AA, Noor M, Arif M, Asim T, Khan WS, et al. Chapter 24 - smartphone-based biosensors for healthcare. In: Nguyen TA, editor. Advanced Sensors for Smart Healthcare: Elsevier; 2025:387-409. [doi: 10.1016/B978-0-443-24790-3.00025-9]
26. Banik S, Melanthota SK, et al. Recent trends in smartphone-based detection for biomedical applications: a review. Anal Bioanal Chem 2021 Apr;413(9):2389-2406. [doi: 10.1007/s00216-021-03184-z] [Medline: 33586007]
27. Friedrich MU, Schneider E, Buerklein M, et al. Smartphone video nystagmography using convolutional neural networks: ConVNG. J Neurol 2023 May;270(5):2518-2530. [doi: 10.1007/s00415-022-11493-1] [Medline: 36422668]
28. Friedrich MU, Roenn AJ, Palmisano C, et al. Validation and application of computer vision algorithms for video-based tremor analysis. NPJ Digit Med 2024 Jun 21;7(1):165. [doi: 10.1038/s41746-024-01153-1] [Medline: 38906946]

XSL•FO
RenderX

29. Taeger J, Bischoff S, Hagen R, Rak K. Utilization of smartphone depth mapping cameras for app-based grading of facial movement disorders: development and feasibility study. JMIR Mhealth Uhealth 2021 Jan 26;9(1):e19346. [doi: 10.2196/19346] [Medline: 33496670]
30. Taeger J, Bischoff S, Hagen R, Rak K. Development of a smartphone app for neuromuscular facial training. HNO 2020 Aug;68(Suppl 2):79-85. [doi: 10.1007/s00106-020-00880-7] [Medline: 32638059]
31. Okreu S, Beckers M. Mundmotorik & Fazialisübungen: Ein Übungsheft Für: Hippocampus Verlag; 2002.
32. What is CocoaPods. CocoaPods. URL: https://cocoapods.org [accessed 2026-04-17]
33. HeroTransitions. GitHub. URL: https://github.com/HeroTransitions/Hero [accessed 2026-04-17]
34. Philackm/scrollablegraphview. GitHub. URL: https://github.com/philackm/ScrollableGraphView [accessed 2026-04-17]
35. airbnb/lottie-ios. GitHub. URL: https://github.com/airbnb/lottie-ios [accessed 2026-04-17]
36. VanSwearingen JM, Brach JS. The Facial Disability Index: reliability and validity of a disability assessment instrument for disorders of the facial neuromuscular system. Phys Ther 1996 Dec;76(12):1288-1298. [doi: 10.1093/ptj/76.12.1288] [Medline: 8959998]
37. Brooke J. SUS: A Quick and Dirty Usability Scale: Taylor & Francis; 1996:189-194.
38. Bangor AK. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
39. Schaede RA, Volk GF, Modersohn L, Barth JM, Denzler J, Guntinas-Lichius O. Video instruction for synchronous video recording of mimic movement of patients with facial palsy. Laryngorhinootologie 2017 Dec;96(12):844-849. [doi: 10.1055/s-0043-101699] [Medline: 28470660]
40. Volk GF, Steigerwald F, Vitek P, Finkensieper M, Kreysa H, Guntinas-Lichius O. Facial Disability Index and Facial Clinimetric Evaluation Scale: validation of the German versions. Laryngorhinootologie 2015 Mar;94(3):163-168. [doi: 10.1055/s-0034-1381999] [Medline: 25089633]
41. Wirtz M. Determining the quality of rater judgements using intraclass correlation, and enhancing rater judgements. Rehabilitation (Stuttg) 2004 Dec;43(6):384-389. [doi: 10.1055/s-2003-814935] [Medline: 15565540]
42. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
43. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6(4):284-290. [doi: 10.1037/1040-3590.6.4.284]
44. Bánhalmi A, Borbás J, Fidrich M, Bilicki V, Gingl Z, Rudas L. Analysis of a pulse rate variability measurement using a smartphone camera. J Healthc Eng 2018;2018:4038034. [doi: 10.1155/2018/4038034] [Medline: 29666670]
45. Koenig N, Seeck A, Eckstein J, et al. Validation of a new heart rate measurement algorithm for fingertip recording of video signals with smartphones. Telemed J E Health 2016 Aug;22(8):631-636. [doi: 10.1089/tmj.2015.0212] [Medline: 26938673]
46. Siddiqui SA, Zhang Y, Feng Z, Kos A. A pulse rate estimation algorithm using PPG and smartphone camera. J Med Syst 2016 May;40(5):126. [doi: 10.1007/s10916-016-0485-6] [Medline: 27067432]
47. Bastawrous A. Smartphone fundoscopy. Ophthalmology 2012 Feb;119(2):432-433. [doi: 10.1016/j.ophtha.2011.11.014] [Medline: 22305322]
48. Haddock LJ, Kim DY, Mukai S. Simple, inexpensive technique for high-quality smartphone fundus photography in human and animal eyes. J Ophthalmol 2013;2013:518479. [doi: 10.1155/2013/518479] [Medline: 24171108]
49. Lord RK, Shah VA, San Filippo AN, Krishna R. Novel uses of smartphones in ophthalmology. Ophthalmology 2010 Jun;117(6):1274-1274. [doi: 10.1016/j.ophtha.2010.01.001] [Medline: 20522335]
50. Choi YK, Demiris G, Lin SY, et al. Smartphone applications to support sleep self-management: review and evaluation. J Clin Sleep Med 2018 Oct 15;14(10):1783-1790. [doi: 10.5664/jcsm.7396] [Medline: 30353814]
51. Niemeijer K, Mestdagh M, Kuppens P. Tracking subjective sleep quality and mood with mobile sensing: multiverse study. J Med Internet Res 2022 Mar 18;24(3):e25643. [doi: 10.2196/25643] [Medline: 35302502]
52. Alamoudi D, Breeze E, Crawley E, Nabney I. The feasibility of using smartphone sensors to track insomnia, depression, and anxiety in adults and young adults: narrative review. JMIR Mhealth Uhealth 2023 Feb 17;11:e44123. [doi: 10.2196/44123] [Medline: 36800211]
53. Peitersen E. Bell's palsy: the spontaneous course of 2,500 peripheral facial nerve palsies of different etiologies. Acta Otolaryngol Suppl 2002;2002(549):4-30. [Medline: 12482166]
54. Peitersen E. The natural history of Bell's palsy. Am J Otol 1982 Oct;4(2):107-111. [Medline: 7148998]
55. Katsumi S, Esaki S, Hattori K, Yamano K, Umezaki T, Murakami S. Quantitative analysis of facial palsy using a three-dimensional facial motion measurement system. Auris Nasus Larynx 2015 Aug;42(4):275-283. [doi: 10.1016/j.anl.2015.01.002] [Medline: 25655983]
56. Mengi E, Orhan Kara C, Necdet Ardıç F, et al. Comparison of the reliability of the House- Brackmann, Facial Nerve Grading System 2.0, and Sunnybrook Facial Grading System for the evaluation of patients with peripheral facial paralysis. J Int Adv Otol 2024 Jan;20(1):14-18. [doi: 10.5152/iao.2024.231162] [Medline: 38454283]
57. Kanerva M, Poussa T, Pitkäranta A. Sunnybrook and House-Brackmann Facial Grading Systems: intrarater repeatability and interrater agreement. Otolaryngol Head Neck Surg 2006 Dec;135(6):865-871. [doi: 10.1016/j.otohns.2006.05.748] [Medline: 17141075]

XSL•FO
RenderX

58. Neely JG, Cherian NG, Dickerson CB, Nedzelski JM. Sunnybrook facial grading system: reliability and criteria for grading. Laryngoscope 2010 May;120(5):1038-1045. [doi: 10.1002/lary.20868] [Medline: 20422701]
59. Banks CA, Jowett N, Hadlock TA. Test-retest reliability and agreement between in-person and video assessment of facial mimetic function using the eFACE Facial Grading System. JAMA Facial Plast Surg 2017 May 1;19(3):206-211. [doi: 10.1001/jamafacial.2016.1620] [Medline: 28006048]
60. O'Reilly BF, Soraghan JJ, McGrenary S, He S. Objective method of assessing and presenting the House-Brackmann and regional grades of facial palsy by production of a facogram. Otol Neurotol 2010 Apr;31(3):486-491. [doi: 10.1097/MAO.0b013e3181c993dc] [Medline: 20042993]
61. Lee LN, Susarla SM, H Hohman M, Henstrom DK, Cheney ML, Hadlock TA. A comparison of facial nerve grading systems. Ann Plast Surg 2013 Mar;70(3):313-316. [doi: 10.1097/SAP.0b013e31826acb2c] [Medline: 23241802]
62. Fan S, Jain RC, Kankanhalli MS. A comprehensive picture of factors affecting user willingness to use mobile health applications. ACM Trans Comput Healthc 2024 Jan 31;5(1):1-31. [doi: 10.1145/3626962]
63. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567. [doi: 10.2196/14567] [Medline: 31573916]
64. Palos-Sanchez PR, Saura JR, Rios Martin M, Aguayo-Camacho M. Toward a better understanding of the intention to use mHealth apps: exploratory study. JMIR Mhealth Uhealth 2021 Sep 9;9(9):e27021. [doi: 10.2196/27021] [Medline: 34499044]
65. Eaton C, Vallejo N, McDonald X, et al. User engagement with mHealth interventions to promote treatment adherence and self-management in people with chronic health conditions: systematic review. J Med Internet Res 2024 Sep 24;26:e50508. [doi: 10.2196/50508] [Medline: 39316431]
66. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283. [doi: 10.2196/20283] [Medline: 32990635]
67. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
68. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11. [doi: 10.2196/jmir.7.1.e11]
69. Kernebeck S, Busse TS, Ehlers JP, Vollmar HC. Adherence to digital health interventions: definitions, methods, and open questions. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2021 Oct;64(10):1278-1284. [doi: 10.1007/s00103-021-03415-9] [Medline: 34559252]
70. Barun B, Divić Z, Martinović Kaliterna D, Poljičanin A, Benzon B, Aljinović J. Mobile app intervention increases adherence to home exercise program after whiplash injury-a randomized controlled trial (RCT). Diagnostics (Basel) 2024 Dec 4;14(23):2729. [doi: 10.3390/diagnostics14232729] [Medline: 39682637]
71. Jack K, McLean SM, Moffett JK, Gardiner E. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228. [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
72. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198. [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
73. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152. [doi: 10.2196/jmir.2104] [Medline: 23151820]
74. Floryan M, Chow PI, Schueller SM, Ritterband LM. The model of gamification principles for digital health interventions: evaluation of validity and potential utility. J Med Internet Res 2020 Jun 10;22(6):e16506. [doi: 10.2196/16506] [Medline: 32519965]
75. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86. [doi: 10.2196/jmir.2583] [Medline: 23598614]
76. Machetanz K, Lins M, Roder C, Naros G, Tatagiba M, Hurth H. Innovative mobile app solution for facial nerve rehabilitation: a usability analysis. Front Digit Health 2024;6:1471426. [doi: 10.3389/fdgth.2024.1471426] [Medline: 39498101]
77. Rintala A, Kossi O, Bonnechère B, Evers L, Printemps E, Feys P. Mobile health applications for improving physical function, physical activity, and quality of life in stroke survivors: a systematic review. Disabil Rehabil 2023 Dec;45(24):4001-4015. [doi: 10.1080/09638288.2022.2140844] [Medline: 36325613]
78. Davergne T, Meidinger P, Dechartres A, Gossec L. The effectiveness of digital apps providing personalized exercise videos: systematic review with meta-analysis. J Med Internet Res 2023 Jul 13;25:e45207. [doi: 10.2196/45207] [Medline: 37440300]
79. Sauro JL, James R. Quantifying the User Experience: Practical Statistics for User Research: Morgan Kaufmann; 2012. [doi: 10.1016/B978-0-12-384968-7.00002-3]
80. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
81. Looyestyn J, Kernot J, Boshoff K, Ryan J, Edney S, Maher C. Does gamification increase engagement with online programs? A systematic review. PLoS One 2017;12(3):e0173403. [doi: 10.1371/journal.pone.0173403] [Medline: 28362821]

XSL•FO
RenderX

82. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48. [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]

Abbreviations

AI: artificial intelligence
CV: computer vision
DFI: Digital Facial Index
FDI: Facial Disability Index
FNP: facial nerve palsy
HBS: House-Brackmann Scale
ICC: intraclass correlation coefficient
LoA: limits of agreement
mHealth: mobile health
SFGS: Sunnybrook Facial Gradings Scale
SI: Stennert Index
SUS: System Usability Scale

Edited by S Munce; submitted 07.Aug.2025; peer-reviewed by G Siciliano, K Young; revised version received 11.Feb.2026; accepted 19.Feb.2026; published 25.May.2026.

Please cite as:
Müller-Graff FT, Essig F, Friedrich MU, Hoika K, Hackenberg S, Rak K, Taeger J
Smartphone-Based Grading and Rehabilitation in Patients With Facial Palsy Using Computer Vision: Prospective Validation Study
JMIR Rehabil Assist Technol 2026;13:e81963
URL: https://rehab.jmir.org/2026/1/e81963
doi:10.2196/81963

© Franz-Tassilo Müller-Graff, Fabian Essig, Maximilian U Friedrich, Kathrin Hoika, Stephan Hackenberg, Kristen Rak, Johannes Taeger. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 25.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Reducing Educational Bias in Cognitive Assessment via Dynamic Support Vector Machine Weighting: Validation Study on an Education-Stratified Dataset

Qing Liu[1*], MMed; Chi Ma[2*], BSc; Mengyuan Liu[1], MA; Suhui Chen[3], BSc; Mengting Yu[4], BSc; Lijuan Xia[5], BSc; Qi Zhang[3], BSc; Ming Wu[3], BSc

[1]School of Humanities and Social Sciences, University of Science and Technology of China, Hefei, China

[2]School of the Gifted Young, University of Science and Technology of China, Hefei, China

[3]Department of Rehabilitation Medicine, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Tian'e Hu No.1, Hefei, China

[4]Department of Rehabilitation Medicine, The Second People's Hospital, Wuhu, China

[5]Shuguang Hospital Anhui Branch Affiliated to Shanghai University of Traditional Chinese Medicine, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Ming Wu, BSc
Department of Rehabilitation Medicine, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Tian'e Hu No.1, Hefei, China

Abstract

Background: The Mini-Mental State Examination (MMSE) remains widely used for cognitive screening, yet its performance varies substantially across educational backgrounds. Linear education corrections fail to capture the nonlinear interference patterns among subitems.

Objective: This study aimed to analyze how educational level shapes MMSE subitem contributions and to develop an education-adaptive optimization strategy using support vector machine–based weighting.

Methods: MMSE data from 812 participants were stratified into 4 education groups. Subitem deletion experiments quantified each subitem's contribution (Δ). Education-specific support vector machine models were then constructed to derive dynamic weighting coefficients. Performance improvements were assessed before and after weighting.

Results: The illiterate group relied heavily on spatial orientation and memory, whereas university-educated individuals depended more on executive and calculation functions. Several education-dependent interference items were identified (eg, visuospatial construction in the primary group and basic orientation tasks in the university group). Dynamic weighting improved accuracy in all cohorts, most notably among illiterate individuals (Δ=7.25%; *P*=.06), followed by the primary school group (Δ=3.12%; *P*=.03).

Conclusions: Education-stratified weighting enhances the fairness and interpretability of MMSE-based screening. External validation confirmed generalizability, although multicenter studies are needed.

(JMIR Rehabil Assist Technol 2026;13:e79841) doi:10.2196/79841

KEYWORDS

machine learning; Mini-Mental State Examination; MMSE; support vector machine; SVM; dynamic weighted model; educational background

Introduction

The Mini-Mental State Examination (MMSE) is one of the most widely used cognitive screening tools in clinical and community settings, designed for the rapid detection of cognitive impairment [1]. Since its development by Folstein et al in 1975 [2], the MMSE has become a core instrument for assessing five cognitive domains—orientation, immediate recall, attention or calculation, language, and visuospatial ability—and can typically be completed within 10 to 15 minutes [3]. Its broad adoption reflects its operational simplicity and solid psychometric performance, with reported sensitivity of 80% to 85% and specificity of 75% to 80% for dementia screening [4].

Alzheimer disease is characterized by progressive neurodegeneration, including amyloid-β deposition and structural decline. As Alzheimer disease progresses in a relatively predictable neuroanatomical sequence, MMSE subdomains provide clinically meaningful stage markers:

XSL•FO
RenderX

orientation and memory deficits often correspond to early hippocampal involvement, while language and visuospatial impairments reflect later temporoparietal degeneration [5,6]. This multidimensional structure allows the MMSE to map aspects of the disease trajectory beyond a single total score.

Nevertheless, the diagnostic performance of the MMSE varies considerably across educational levels [7,8]. Individuals with higher education often exhibit false-negative results due to two well-recognized mechanisms: compensatory neuroplasticity that delays the manifestation of cognitive symptoms [9] and the use of test-taking strategies that allow them to maintain near-normal scores despite underlying pathology [10,11]. In contrast, individuals with limited formal education show substantially higher false-positive rates, particularly on education-dependent items, such as literacy tasks and object naming, with error rates increasing by approximately 30% to 35% compared with education-neutral instruments [12-15]. As a result, the MMSE demonstrates reduced sensitivity (68% - 72%) and specificity (65% - 70%) when a conventional cutoff score of 24 points is applied across heterogeneous educational backgrounds [16-18]. Current clinical guidelines therefore emphasize the need to incorporate educational history into cognitive assessment to minimize diagnostic inaccuracies [19,20].

Machine learning (ML) has emerged as a transformative approach in medical diagnostics, with particular relevance to cognitive assessment optimization. Support Vector Machines (SVMs) offer several advantages for psychometric refinement, including (1) robust performance in limited-sample settings (typically n<1000) [21,22], (2) effective handling and interpretability of high-dimensional feature spaces [23,24], and (3) strong compatibility with multimodal data integration frameworks that combine neuropsychological assessments with imaging- or biomarker-derived measures [25-28]. Earlier ML classifiers trained solely on MMSE total scores or subitems—most commonly logistic regression, SVM, or random forest—have demonstrated moderate diagnostic performance, with reported accuracies of 72% to 85% and area under the curve (AUC) values of 0.75 to 0.85 [29]. In contrast, multimodal models integrating magnetic resonance imaging (MRI) radiomics, positron emission tomography (PET) signatures, or speech biomarkers consistently achieve accuracies exceeding 90% to 95% [30,31]. This performance gap highlights the need for methodological innovations that enhance unimodal MMSE-based classifiers while preserving their scalability and low implementation cost.

This study introduces an education-sensitive adaptation of the MMSE based on an SVM-guided dynamic weighting framework. We analyzed cognitive screening data (n=812) collected from Chinese tertiary hospitals and community health centers and stratified participants into four educational cohorts: illiterate (0 y), primary (≤6 y), secondary (7 - 12 y), and tertiary (≥13 y). The model uses systematic item-response analysis combined with SVM-guided weighting to adjust the relative contribution of each MMSE subitem in an education-specific manner.

This work provides 2 methodological innovations. First, it moves beyond conventional linear education corrections by using nonlinear SVM modeling to identify "cognitive interference" patterns unique to each education level. Second, it introduces a dynamic weighting strategy specifically designed to mitigate the disproportionate false-positive burden experienced by low-literacy populations. Together, these innovations aim to deliver a more equitable and clinically practical adaptation of the MMSE, particularly relevant for resource-limited settings.

Although unimodal or condensed classifiers derived solely from MMSE subitems are highly accessible, they are inherently constrained by limited feature richness and therefore tend to exhibit moderate predictive performance. In contrast, multimodal diagnostic systems—such as those leveraging MRI, PET, radiomics, or speech biomarkers—achieve substantially higher accuracies but remain impractical for widespread screening due to cost, technical requirements, and limited availability in primary care. Thus, methodological advances that enhance the diagnostic utility of unimodal MMSE-based approaches, while maintaining their affordability and scalability, are urgently needed. This study addresses this gap by proposing an education-stratified dynamic weighting method designed to improve predictive performance within the intrinsic limitations of unimodal cognitive classifiers.

Methods

Ethical Considerations

This study followed the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of the University of Science and Technology of China (approval number 2024-RE-431). All clinical and community datasets were obtained with authorization from the original data custodians. During data processing, only anonymized records were used, and no personally identifiable information (eg, names, addresses) was accessed, ensuring compliance with the Declaration of Helsinki.

Sample

Cognitive assessment data were obtained from two sources: the Rehabilitation Medicine Department of a tertiary hospital and a community-based cognitive screening program. A total of 812 valid records were included. Among these, 404 samples were collected from outpatients and inpatients with clinically diagnosed cognitive impairment. The inclusion and exclusion criteria are summarized in Table 1.

XSL•FO
RenderX

Table . Data inclusion and exclusion criteria.

Source	Number (n)	Inclusion criteria	Exclusion criteria	Data assessment requirements
Clinical dataset	404	• Outpatients and inpatients with clinically confirmed cognitive impairment • Complete clinical and assessment records	• Patients who underwent surgery or died during hospitalization • Missing MMSE[a] assessment data	• For clinical data: All MMSE assessors completed a standardized 12-h training program and passed certification. • Postcollection, automated missing value detection and manual verification were performed.
Community dataset	408	• Community residents with no medical consultation for cognitive complaints within the past 12 mo • Intact consciousness and ability to complete cognitive assessments	• Individuals with diagnosed mild cognitive impairment or suspected cognitive decline • Individuals with incomplete assessment procedures or severe organic neurological conditions	• For community data: Assessments followed standardized community screening procedures. Assessments were administered following standardized community screening procedures with verification of data completeness and adherence to protocol.

[a]MMSE: Mini-Mental State Examination.

Scale Selection and Scoring Criteria

The MMSE was used as the primary cognitive assessment tool. The scale comprises five cognitive domains: orientation, memory, attention and calculation, language, and visuospatial construction. Education-adjusted cutoff scores were based on Chinese normative standards [3]: illiterate group: ≤17, primary school group: ≤20, secondary school group: ≤22, and university group: ≤23.

Data Processing and Grouping

Overview

A total of 1000 assessments were collected initially. After quality control, 188 (18.8%) records were excluded due to missing data (n=107, 10.7%), nonstandardized test administration (n=35, 3.5%), or unqualified assessors (n=46, 4.6%). Before merging the hospital and community datasets, baseline demographic comparability was examined using an independent-samples *t* test (age) and chi-square test (gender) to minimize potential selection bias.

The final dataset (N=812) was stratified by educational attainment as follows: illiterate (n=108, 13.3%), primary school (n=105, 12.9%), secondary school (n=364, 44.8%), and university (n=235, 28.9%). Within each group, patient and control counts were recorded as illiterate (patients: n=60, 7.4%; and controls: n=48, 5.9%), primary (patients: n=53, 6.5%; and controls: n=52, 6.4%), secondary (patients: n=185, 22.8%; and controls: n=179, 22.0%), and university (patients: n=106, 13.1%; and controls: n=129, 15.9%). This preprocessing ensured a standardized data foundation for subsequent ML procedures.

Model Selection

A SVM classifier with a Radial Basis Function kernel was selected due to its suitability for medium-sized datasets and its capacity to model nonlinear decision boundaries. A stratified 5-fold cross-validation scheme was used to maximize data utilization while preserving the original case-control ratio within each educational group.

Hyperparameter Optimization

Hyperparameters were optimized through a grid search embedded within the cross-validation framework. The search grid included penalty parameter C: {0.1, 1, 5, 10, 100} and kernel coefficient γ: {0.001, 0.01, 0.1, 1}.

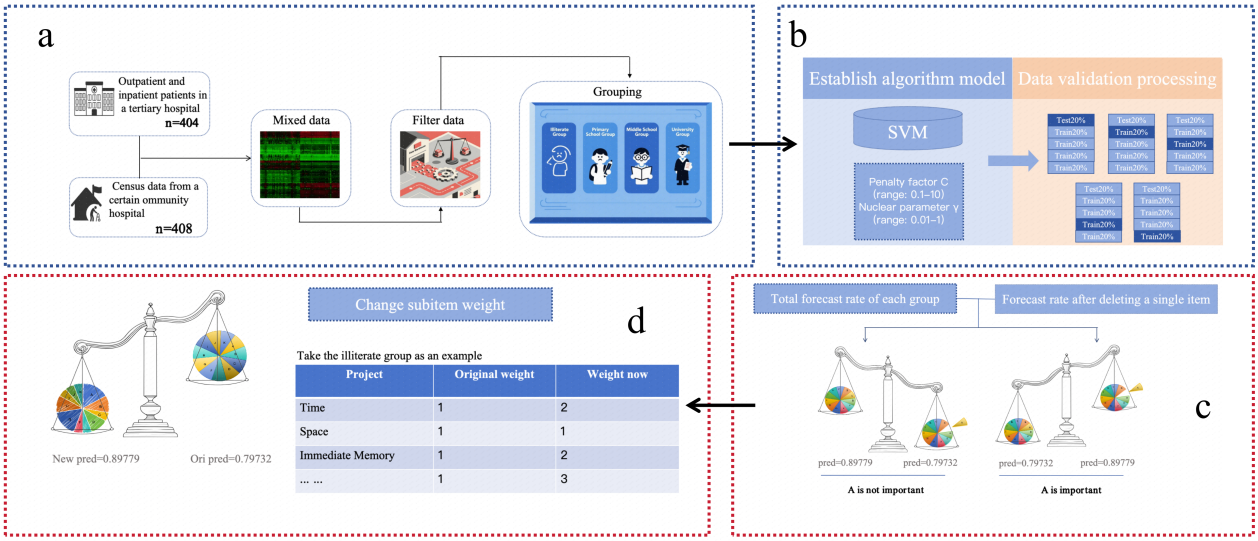
For each parameter pair, the model was trained on 4 folds and evaluated on the remaining fold, and the average accuracy across all 5 folds was used to identify the optimal configuration. Education-specific models were then retrained on the full dataset using the selected hyperparameters to ensure maximal predictive performance while avoiding information leakage.

Experimental Design

The overall experimental workflow is presented in Figure 1.

XSL•FO
RenderX

Figure 1. Contribution of individual Mini-Mental State Examination (MMSE)j items across educational strata. Feature importance was quantified using the percentage change in cross-validated accuracy (Δ_i) after deleting each subitem. We operationally defined critical factors as items whose removal caused a performance decrease of $\Delta<-1.0$ and interference factors as items whose removal increased accuracy by $\Delta>+0.5$. SVM: support vector machine.



The process began with data fusion and initial grouping: clinical and community data were merged to construct the initial dataset (n=812), and participants were stratified into 4 educational groups (ie, illiterate, primary school, secondary school, and university) to ensure subsequent analyses reflected education-specific cognitive characteristics. Next, a Support Vector Machine classifier with a radial basis function kernel was trained using a stratified 5-fold cross-validation scheme, preserving the original case-control ratio within each group. To quantify the contribution of each MMSE subitem, a systematic ablation procedure was performed. For each educational group, each subitem was removed in turn, and an SVM model was retrained. The feature contribution was quantified using the cross-validated change in prediction accuracy (Δ), calculated as:

$$\Delta = Accuracy_{postdeletion} - Accuracy_{baseline}$$

A negative Δ indicates a positive contribution (performance dropped when the item was removed), while a positive Δ suggests the item acted as noise. Finally, based on the subitem contribution profiles, an education-specific dynamic weighting scheme was constructed to adjust the relative importance of MMSE subitems before classification.

Other Statistical Techniques

Routine statistical analyses were performed using SPSS version 24.0 (IBM Corp), and ML model construction and evaluation were conducted using Python (version 3.10.10; Python Software Foundation)

For categorical variables, chi-square or Fisher exact tests were applied depending on expected frequencies. Continuous variables with a normal distribution were expressed as mean (SD) and compared using independent-samples *t* tests. Variables not following a normal distribution were summarized using median (IQR) and analyzed using nonparametric tests.

A 2-tailed $P<.05$ was considered statistically significant.

Results

Demographic Comparison

Following data cleansing, demographic characteristics of the clinical and community datasets were compared to verify baseline homogeneity before integration. As presented in Table 2, differences in age and gender distributions were not statistically significant ($P>.05$).

XSL•FO
RenderX

Table . Demographic characteristics and data composition stratified by educational background.

Characteristic	Illiterate (n=108)	Primary school (n=105)	Secondary school (n=364)	University (n=235)	*P* value
Age, mean (SD)	59.02 (8.5)	58.55 (7.8)	58.62 (8.1)	56.7 (7.4)	.05
Gender, n (%)					.37
Male	67 (62.0)	60 (57.1)	219 (60.2)	140 (59.6)	
Female	41 (38.0)	45 (42.9)	145 (39.8)	95 (40.4)	
Data source, n (%)					.12
Clinical[a] (patient)	60 (55.6)	53 (50.5)	185 (50.8)	106 (45.1)	
Community[b] (control)	48 (44.4)	52 (49.5)	179 (49.2)	129 (54.9)	

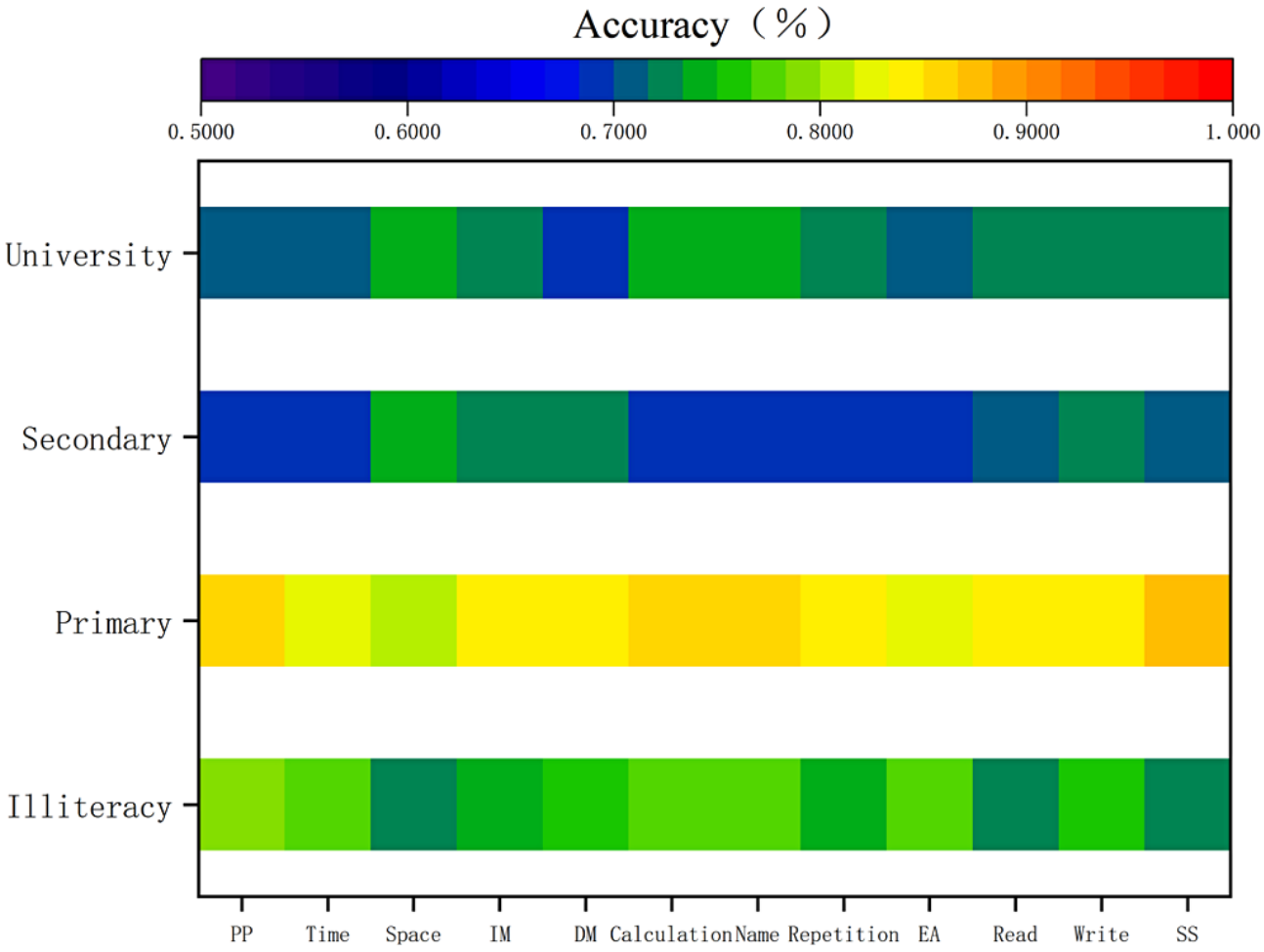
[a]Hospital dataset: patients diagnosed or evaluated in the tertiary hospital.

[b]Community dataset: participants recruited from community-based cognitive screening programs.

Item-Wise Contribution Analysis

To assess the contribution of individual MMSE items across educational groups, we conducted an item deletion experiment. The change in cross-validated accuracy (Δi) after removing each subitem was used to quantify its importance. A negative Δi (< -1.0) indicated a critical item whose removal impaired performance, while a positive Δi (> +0.5) indicated an interference item whose removal improved accuracy. The results of this analysis are visualized in Figure 2.

Figure 2. Heatmap visualization of feature contribution analysis. The color spectrum reflects prediction accuracy after item deletion. Cooler colors (blue and green) denote "critical features" (accuracy drop), while warmer colors (yellow and red) denote "neutral/interfering features" (accuracy stable/increase). The y-axis represents educational groups with the following sample sizes (N_patient/N_control): illiterate (60/48), primary school (53/52), secondary school (185/179), and university (106/129). DM: delayed memory; EA: execution ability; IM: immediate memory; PP: penmanship and praxis; SS: spatial structure.



XSL•FO
RenderX

Illiterate Group

In the illiterate cohort (baseline accuracy=78.83%), all subitems demonstrated positive contributions. The most influential items were spatial orientation (Δ=−6.58% postremoval), immediate memory (Δ=−3.85%), and writing (Δ=−2.86%). No interference items were identified.

Primary School Group

In the primary school group (baseline accuracy=85.71%), spatial construction functioned as an interference factor (Δ =+0.95%), while calculation showed a negligible effect (Δ=0). Removal of immediate memory, delayed recall, reading, or writing reduced accuracy to 83.81% (Δ=−1.90%).

Secondary School Group

In the secondary school group (baseline accuracy=69.78%), multiple interference factors were observed: spatial orientation (Δ=+3.57%), delayed recall (Δ=+3.02%), executive function (Δ=+2.19%), reading (Δ=+1.64%), immediate memory (Δ=+2.47%), writing (Δ=+0.20%), and spatial construction (Δ=+1.10%). Only temporal orientation (Δ=−0.83%) and calculation ability (Δ=−0.80%) demonstrated positive contributions.

University Group

In the university group (baseline accuracy=71.49%), five interference items were identified: spatial orientation (Δ=+2.98%), calculation (Δ=+2.98%), immediate memory (Δ=+1.70%), naming (Δ=+1.70%), and delayed recall (Δ=+1.42%). Repetition (Δ=−0.57%) and reading (Δ=−0.30%) showed small positive contributions.

Item Weight Analysis

On the basis of these contribution patterns, education-specific weighted coefficients were calculated for all MMSE subitems. The weighted diagnostic score was defined as follows:

$$S=\sum t-111\ (vi\times w1)$$

where S is the final weighted diagnostic score, v_i is the raw score of the i-th MMSE subitem, and w_i is the dynamic weighting coefficient assigned to the i-th subitem.

Classification was determined by comparing S to a group-specific classification threshold (T): S<T indicated cognitive impairment, and S≥T indicated normal cognition.

After normalization, the maximum possible weighted score was 60. Education-specific diagnostic thresholds were defined as follows: illiterate (*t*=30), primary (*t*=31), secondary (*t*=32), and university (*t*=33). As presented in Tables 3 and 4, accuracy improved across all educational strata after weight calibration.

Table . Model performance and weighting coefficients for core cognitive domains by educational group.

Group	Prediction accuracy before weighting	Prediction accuracy after weighting	Orientation (time)	Orientation (space)	Memory (immediate)	Memory (delayed)	Attention and calculation
Illiterate (%）	79.83	86.08	1	3	2	2	1
Primary school (%）	85.71	88.16	2	3	2	2	1
Secondary school (%）	69.78	69.83	2	3	2	1	2
University (%）	71.49	73.62	1	3	2	2	2

Table . Weighting coefficients for language and visuospatial domains by educational group.

Group	Naming	Repetition	Execution ability	Reading	Writing	Spatial structure
Illiterate (%)	3	3	2	3	2	3
Primary school (%)	1	2	3	2	2	1
Secondary school (%)	2	2	2	1	1	1
University (%)	2	3	1	1	2	2

Table 5 summarizes the diagnostic performance before and after dynamic weighting, including accuracy, sensitivity, specificity, AUC, and the statistical significance of performance changes. The largest improvement occurred in the illiterate group (Δ=+7.25%), whereas the secondary school group showed the smallest improvement (Δ=+0.05%).

Weight distributions differed across education levels. For basic cognitive domains, spatial orientation weights were uniformly set to 3 across all groups, and immediate memory weights remained at 2. In the illiterate group, naming, repetition, and spatial construction received the highest weights (all=3). In the primary school group, temporal orientation was weighted at 2 and spatial construction was weighted at 1. In the secondary school group, executive function and calculation were weighted at 2, while spatial construction remained at 1. In the university group, executive function received the highest weight (3), calculation was weighted at 2, and spatial construction, reading, and writing each received a weight of 1.

XSL•FO
RenderX

Table . Comparison of diagnostic performance metrics and statistical significance before and after dynamic weighting.

Education group and metric	Before weighting	After weighting	Improvement (Δ)	*P* value
Illiterate			7.25%	.06
Accuracy (%)	78.83	86.08		
Sensitivity (%)	81.45	84.92		
Specificity (%)	76.21	87.24		
AUC[a]	0.835	0.912		
Primary school			2.45%	.03
Accuracy (%)	85.71	88.16		
Sensitivity (%)	86.10	88.54		
Specificity (%)	85.32	87.78		
AUC	0.894	0.928		
Secondary school			0.05%	.46
Accuracy (%)	69.78	69.83		
Sensitivity (%)	70.25	70.31		
Specificity (%)	69.31	69.35		
AUC	0.752	0.753		
University			2.13%	.18
Accuracy (%)	71.49	73.62		
Sensitivity (%)	72.15	74.88		
Specificity (%)	70.83	72.36		
AUC	0.768	0.795		

[a]AUC: area under the curve.

Discussion

Principal Findings

This study developed an education-sensitive optimization framework for the MMSE using a dynamic SVM-based weighting strategy. Across 812 education-stratified participants, the model demonstrated that cognitive subitems contribute differently to diagnostic prediction depending on years of formal education. Dynamic weighting enhanced the performance of unimodal MMSE classifiers, particularly in low-literacy groups, and reduced education-related diagnostic bias. These findings refine the diagnostic utility of condensed cognitive screening tools and support their broader applicability in resource-limited settings.

Comparison With Prior Work

ML has been widely applied to cognitive assessment [32-35], yet existing studies typically use MMSE scores as features or labels without addressing education-driven measurement bias [36,37]. Prior condensed MMSE-only classifiers generally achieve accuracies of 72% to 85%, whereas multimodal MRI-, PET-, and biomarker-based models often exceed 90% [32,33]. Although our model remains unimodal, the dynamic weighting mechanism improved prediction by amplifying high-value signals and attenuating education-dependent interference, partially narrowing the performance gap with multimodal systems while sustaining low cost and operational simplicity.

Traditional MMSE optimization approaches—such as linear education corrections or standard ML models—lack the capacity to capture nonlinear, education-specific feature interactions [38]. In contrast, our framework integrates item deletion experiments with SVM-informed dynamic weighting to generate transparent, subgroup-adaptive coefficients. This approach offers an interpretable alternative to black box ensemble models and provides an effective mechanism for addressing item-level educational bias in cognitive assessment. A comprehensive comparison between the proposed dynamic weighted SVM model and existing cognitive assessment optimization strategies is presented in Table 6.

XSL•FO
RenderX

Table . Comparison of the proposed dynamic weighted Support Vector Machine (SVM) model with existing cognitive assessment optimization strategies.

Optimization strategy	Key methodology	Data requirements	Advantages	Limitations
Traditional linear correction (eg, MMSE[a]-E)	Linear regression and fixed point addition	MMSE scores+demographics	Simple calculation and clinically familiar	Ignored nonlinear "cognitive interference" and low precision
Standard machine learning (eg, RF[b], ANN[c])	Black box classification using raw scores	MMSE subitems or total scores	High classification accuracy	Low interpretability ("black box") and hard to explain to clinicians
Multimodal fusion models	Deep learning (CNN[d]/RNN[e]) integration	MRI[f]/PET[g] imaging+biomarkers+scale	Highest accuracy (>90%) and comprehensive pathology mapping	High cost, low accessibility in primary care, and complex deployment
This study (dynamic weighted SVM)	Nonlinear dynamic weighting via SVM	MMSE subitems only (low cost)	High interpretability (visible weights), education adaptive, and high accessibility	Accuracy is lower than multimodal models (feature limitation)

[a]MMSE: Mini-Mental State Examination.

[b]RF: random forest.

[c]ANN: artificial neural network.

[d]CNN: convolutional neural network.

[e]RNN: recurrent neural network.

[f]MRI: magnetic resonance imaging.

[g]PET: positron emission tomography.

Interpretation of Education-Specific Patterns

Education-stratified analyses revealed distinct, cognitive reliance patterns. The illiterate group depended heavily on orientation and immediate memory, consistent with basic functional domains commonly preserved in low-literacy populations. The primary school group demonstrated changes in weighting for temporal orientation, whereas secondary school participants exhibited varied interference across multiple domains and minimal model improvement (Δ=0.05), suggesting high within-group heterogeneity. University-educated participants showed stronger reliance on executive functioning and calculation. These patterns align with prior evidence that cognitive performance becomes increasingly distributed and strategy dependent with higher educational attainment [39,40].

The dynamic weighting strategy provided incremental accuracy improvements across all educational cohorts while reducing false-positive rates in lower-education groups. Subitem deletion further allowed visualization of feature importance (Figure 2), conceptually analogous to Shapley Additive Explanations–based explainability, thereby enhancing the transparency of SVM decision processes. Clinical validation by senior rehabilitation physicians confirmed that weight distributions reflected plausible neurocognitive patterns.

External validation using an independent community dataset (n=314) demonstrated stable performance (overall accuracy: 82.48% and AUC: 0.88), with the highest accuracy observed in the illiterate group (88.89%). These results support the generalizability and robustness of the proposed dynamic weighting model and mitigate concerns about overfitting.

Limitations

This study has several limitations. First, the sample was derived from a single geographic region, leading to educational imbalance across subgroups and potentially limiting generalizability. Broader multicenter sampling is needed. Second, demographic confounders beyond age and gender—such as socioeconomic status, occupational complexity, and urban-rural residence—were not available in the dataset, restricting the ability to fully disentangle education effects from related social determinants. Third, although the MMSE provides valuable screening utility, ceiling effects in highly educated individuals limit sensitivity to early cognitive decline. Integrating additional modalities, such as the Montreal Cognitive Assessment or imaging markers, may complement the proposed model. Finally, the SVM model relied on grid-search hyperparameter tuning; future research may explore AutoML-based optimization to improve efficiency and predictive performance.

Conclusions

This study proposes a dynamic SVM-based weighting framework that enhances the diagnostic fairness of MMSE-based cognitive screening across diverse educational backgrounds. By quantifying item-level contributions and adapting subitem weights for each education group, the method addresses a longstanding source of measurement bias in cognitive assessment. The approach retains the accessibility and scalability of condensed cognitive screening tools while improving prediction accuracy and interpretability. These findings provide a practical foundation for developing equitable cognitive assessment strategies, particularly in resource-limited regions.

XSL•FO
RenderX

Acknowledgments

The authors sincerely acknowledge all individuals and institutions that contributed to this work. This research was jointly supported by the Health Research Program of Anhui Province (project number AHWJ2023A10123) and the Anhui Provincial Teaching Research Project (project number 2022jyxm1851) titled "Research and Implementation of a Standardized Training Curriculum System for Rehabilitation Therapists Based on Competency-Based Education." Special thanks are extended to the funding bodies for their support.

Qi Zhang is co-corresponding author (email: ustczhangqi@163.com; phone +86 17333070277).

Funding

This work was supported by the Anhui Provincial Teaching Research Project (grant 2022jyxm1851), the Health Research Program of Anhui (grant AHWJ2023A10123), and the Anhui Provincial Graduate Academic Innovation Project “Optimization and Prediction of Alzheimer’s Disease-Related Cognitive Scales Based on Machine Learning” (grant 2024xscx009). This research received no commercial sponsorship or financial support that could compromise its objectivity. All authors declare no financial or proprietary interests relevant to this study. The funding agencies provided only financial support and had no involvement in study design, data collection and analysis, manuscript preparation, or submission for publication. The research team assumes full responsibility for the study's independence and scientific integrity.

Data Availability

The clinical dataset used in this study was obtained under ethical approval (2024-RE-431) and contains sensitive patient information; therefore, it cannot be made publicly available. Deidentified clinical records (n=404) may be shared upon reasonable request to the corresponding author under a data sharing agreement. Community screening data (n=408) are available with permission from the participating community health centers.

The complete source code for data preprocessing, Support Vector Machine modeling, dynamic weighting, and statistical analysis is openly available on GitHub [41].

Authors' Contributions

QL: Conceptualization, writing – original draft, formal analysis, data curation, methodology
CM: Conceptualization, writing – original draft, formal analysis, data curation, methodology
ML: Investigation, resources
SC: Data curation (clinical)
MY: Data curation (clinical and community)
LX: Data curation (clinical and community)
QZ: Supervision, project administration
MW: Supervision, project administration
All authors have reviewed and approved the final manuscript, agreeing to assume public responsibility for its content.
QZ is the co-corresponding author.

Conflicts of Interest

None declared.

References

1. Chun CT, Seward K, Patterson A, Melton A, MacDonald-Wicks L. Evaluation of available cognitive tools used to measure mild cognitive decline: a scoping review. Nutrients 2021 Nov 8;13(11):3974. [doi: 10.3390/nu13113974] [Medline: 34836228]
2. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
3. Jia X, Wang Z, Huang F, et al. A comparison of the Mini-Mental State Examination (MMSE) with the Montreal Cognitive Assessment (MoCA) for mild cognitive impairment screening in Chinese middle-aged and older population: a cross-sectional study. BMC Psychiatry 2021 Oct 4;21(1):485. [doi: 10.1186/s12888-021-03495-6] [Medline: 34607584]
4. Patnode CD, Perdue LA, Rossom RC, et al. Screening for cognitive impairment in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2020 Feb 25;323(8):764-785. [doi: 10.1001/jama.2019.22258] [Medline: 32096857]
5. Scheltens P, De Strooper B, Kivipelto M, et al. Alzheimer’s disease. Lancet 2021 Apr 24;397(10284):1577-1590. [doi: 10.1016/S0140-6736(20)32205-4] [Medline: 33667416]
6. Yin TT, Cao MH, Yu JC, et al. T1-weighted imaging-based hippocampal radiomics in the diagnosis of Alzheimer’s disease. Acad Radiol 2024 Dec;31(12):5183-5192. [doi: 10.1016/j.acra.2024.06.012] [Medline: 38902110]

XSL•FO
RenderX

7. Cardoso S, Barros R, Marôco J, de Mendonça A, Guerreiro M. Different MMSE domains are associated to cognitive decline and education. Appl Neuropsychol Adult 2024;31(4):533-539. [doi: 10.1080/23279095.2022.2041018] [Medline: 35234096]
8. Bird HR, Canino G, Stipec MR, Shrout P. Use of the Mini-Mental State Examination in a probability sample of a Hispanic population. J Nerv Ment Dis 1987 Dec;175(12):731-737. [doi: 10.1097/00005053-198712000-00005] [Medline: 3500273]
9. Koepsell TD, Kurland BF, Harel O, Johnson EA, Zhou XH, Kukull WA. Education, cognitive function, and severity of neuropathology in Alzheimer disease. Neurology 2008 May 6;70(19 Pt 2):1732-1739. [doi: 10.1212/01.wnl.0000284603.85621.aa] [Medline: 18160675]
10. Baker DW, Gazmararian JA, Sudano J, Patterson M. The association between age and health literacy among elderly persons. J Gerontol B Psychol Sci Soc Sci 2000 Nov;55(6):S368-S374. [doi: 10.1093/geronb/55.6.s368] [Medline: 11078114]
11. Ouvrard C, Berr C, Meillon C, et al. Norms for standard neuropsychological tests from the French CONSTANCES cohort. Eur J Neurol 2019 May;26(5):786-793. [doi: 10.1111/ene.13890] [Medline: 30575234]
12. Aprahamian I, Martinelli JE, Cecato J, Yassuda MS. Screening for Alzheimer's disease among illiterate elderly: accuracy analysis for multiple instruments. J Alzheimers Dis 2011;26(2):221-229. [doi: 10.3233/JAD-2011-110125] [Medline: 21593559]
13. Rezende GP, Cecato J, Martinelli JE. Cognitive abilities screening instrument-short form, Mini-Mental State Examination and functional activities questionnaire in the illiterate elderly. Dement Neuropsychol 2013;7(4):410-415. [doi: 10.1590/S1980-57642013DN74000009] [Medline: 29213866]
14. Katzman R, Zhang MY, et al. A Chinese version of the Mini-Mental State Examination; impact of illiteracy in a Shanghai dementia survey. J Clin Epidemiol 1988;41(10):971-978. [doi: 10.1016/0895-4356(88)90034-0] [Medline: 3193141]
15. Weiss BD, Reed R, Kligman EW, Abyad A. Literacy and performance on the Mini-Mental State Examination. J Am Geriatr Soc 1995 Jul;43(7):807-810. [doi: 10.1111/j.1532-5415.1995.tb07057.x] [Medline: 7602038]
16. Paddick SM, Gray WK, McGuire J, Richardson J, Dotchin C, Walker RW. Cognitive screening tools for identification of dementia in illiterate and low-educated older adults, a systematic review and meta-analysis. Int Psychogeriatr 2017 Jun;29(6):897-929. [doi: 10.1017/S1041610216001976] [Medline: 28274299]
17. Pellicer-Espinosa I, Díaz-Orueta U. Cognitive screening instruments for older adults with low educational and literacy levels: a systematic review. J Appl Gerontol 2022 Apr;41(4):1222-1231. [doi: 10.1177/07334648211056230] [Medline: 34856843]
18. Brucki SMD, Mansur LL, Carthery-Goulart MT, Nitrini R. Formal education, health literacy and Mini-Mental State Examination. Dement Neuropsychol 2011;5(1):26-30. [doi: 10.1590/S1980-57642011DN05010005] [Medline: 29213716]
19. Carnero-Pardo C. Should the Mini-Mental State Examination be retired? Neurologia 2014 Oct;29(8):473-481. [doi: 10.1016/j.nrl.2013.07.003] [Medline: 24140158]
20. Cai H, Shao Y, Liu XY, et al. Interpretable prognostic modeling for long-term survival of Type A aortic dissection patients using Support Vector Machine algorithm. Eur J Med Res 2025 Apr 15;30(1):277. [doi: 10.1186/s40001-025-02510-w] [Medline: 40229872]
21. Luo W. Predicting cervical cancer outcomes: statistics, images, and machine learning. Front Artif Intell 2021;4:627369. [doi: 10.3389/frai.2021.627369] [Medline: 34164615]
22. Cai F, Cherkassky V. Generalized SMO algorithm for SVM-based multitask learning. IEEE Trans Neural Netw Learn Syst 2012 Jun;23(6):997-1003. [doi: 10.1109/TNNLS.2012.2187307] [Medline: 24806769]
23. Huang C, Chung FL, Wang S. Multi-view L2-SVM and its multi-view core vector machine. Neural Netw 2016 Mar;75:110-125. [doi: 10.1016/j.neunet.2015.12.004] [Medline: 26773824]
24. Mastropietro A, Feldmann C, Bajorath J. Calculation of exact Shapley values for explaining Support Vector Machine models using the radial basis function kernel. Sci Rep 2023 Nov 10;13(1):19561. [doi: 10.1038/s41598-023-46930-2] [Medline: 37949930]
25. Jasodanand VH, Bellitti M, Kolachalama VB. An AI-first framework for multimodal data in Alzheimer's disease and related dementias. Alzheimers Dement 2025 Sep;21(9):e70719. [doi: 10.1002/alz.70719] [Medline: 40983947]
26. Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. Neuroimage 2017 Jul 15;155:530-548. [doi: 10.1016/j.neuroimage.2017.03.057] [Medline: 28414186]
27. Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage 2011 Apr 1;55(3):856-867. [doi: 10.1016/j.neuroimage.2011.01.008] [Medline: 21236349]
28. Jasodanand VH, Kowshik SS, Puducheri S, et al. AI-driven fusion of multimodal data for Alzheimer's disease biomarker assessment. Nat Commun 2025 Aug 11;16(1):7407. [doi: 10.1038/s41467-025-62590-4] [Medline: 40789853]
29. Dincer A, Gordon BA, Hari-Raj A, et al. Comparing cortical signatures of atrophy between late-onset and autosomal dominant Alzheimer disease. Neuroimage Clin 2020;28:102491. [doi: 10.1016/j.nicl.2020.102491] [Medline: 33395982]
30. Qiu S, Miller MI, Joshi PS, et al. Multimodal deep learning for Alzheimer's disease dementia assessment. Nat Commun 2022 Jun 20;13(1):3404. [doi: 10.1038/s41467-022-31037-5] [Medline: 35725739]
31. Castellano G, Esposito A, Lella E, Montanaro G, Vessio G. Automated detection of Alzheimer's disease: a multi-modal approach with 3D MRI and amyloid PET. Sci Rep 2024 Mar 3;14(1):5210. [doi: 10.1038/s41598-024-56001-9] [Medline: 38433282]

XSL•FO
RenderX

32. Vyas A, Aisopos F, Vidal ME, Garrard P, Paliouras G. Identifying the presence and severity of dementia by applying interpretable machine learning techniques on structured clinical records. BMC Med Inform Decis Mak 2022 Oct 17;22(1):271. [doi: 10.1186/s12911-022-02004-3] [Medline: 36253849]
33. Feng F, Wang P, Zhao K, et al. Radiomic features of hippocampal subregions in Alzheimer's disease and amnestic mild cognitive impairment. Front Aging Neurosci 2018;10:290. [doi: 10.3389/fnagi.2018.00290] [Medline: 30319396]
34. Battista P, Salvatore C, Castiglioni I. Optimizing neuropsychological assessments for cognitive, behavioral, and functional impairment classification: a machine learning study. Behav Neurol 2017;2017:1850909. [doi: 10.1155/2017/1850909] [Medline: 28255200]
35. Khatun S, Morshed BI, Bidelman GM. A single-channel EEG-based approach to detect mild cognitive impairment via speech-evoked brain responses. IEEE Trans Neural Syst Rehabil Eng 2019 May;27(5):1063-1070. [doi: 10.1109/TNSRE.2019.2911970] [Medline: 30998476]
36. Wu Y, Jia M, Xiang C, Lin S, Jiang Z, Fang Y. Predicting the long-term cognitive trajectories using machine learning approaches: a Chinese nationwide longitudinal database. Psychiatry Res 2022 Apr;310:114434. [doi: 10.1016/j.psychres.2022.114434] [Medline: 35172247]
37. Martin SA, Townend FJ, Barkhof F, Cole JH. Interpretable machine learning for dementia: a systematic review. Alzheimers Dement 2023 May;19(5):2135-2149. [doi: 10.1002/alz.12948] [Medline: 36735865]
38. Kantayeva G, Lima J, Pereira AI. Application of machine learning in dementia diagnosis: a systematic literature review. Heliyon 2023 Nov;9(11):e21626. [doi: 10.1016/j.heliyon.2023.e21626] [Medline: 38027622]
39. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 8;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
40. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol 2012 Nov;11(11):1006-1012. [doi: 10.1016/S1474-4422(12)70191-6] [Medline: 23079557]
41. Dynamic-SVM-weighting-for-MMSE-optimization. GitHub. URL: https://github.com/masC1962/Dynamic-SVM-Weighting-for-MMSE-Optimization [accessed 2026-01-05]

Abbreviations

AUC: area under the curve
ML: machine learning
MMSE: Mini-Mental State Examination
MRI: magnetic resonance imaging
PET: positron emission tomography
SVM: support vector machine

Edited by S Munce; submitted 30.Jun.2025; peer-reviewed by AM Hasan, H Kan, S Neufang, ZY Jiang; accepted 23.Dec.2025; published 25.Feb.2026.

Please cite as:
Liu Q, Ma C, Liu M, Chen S, Yu M, Xia L, Zhang Q, Wu M
Reducing Educational Bias in Cognitive Assessment via Dynamic Support Vector Machine Weighting: Validation Study on an Education-Stratified Dataset
JMIR Rehabil Assist Technol 2026;13:e79841
URL: https://rehab.jmir.org/2026/1/e79841
doi:10.2196/79841

© Qing Liu, Chi Ma, Mengyuan Liu, Suhui Chen, Mengting Yu, Lijuan Xia, Qi Zhang, Ming Wu. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using a Co-Designed Digital Self-Management Program to Prepare Patients for Hip or Knee Replacement Surgery: Pragmatic Pilot Study

Elizabeth Horton[1], PhD; Hayley Wright[1], PhD; Andy Turner[1], PhD; Louise Moody[2], PhD; Lucy Aphramor[3], PhD; Anna Carlson[4], PhD; Hesam Ghiasvand[5], PhD; Shea Palmer[6], PhD

[1]Centre for Intelligent Healthcare, Coventry University, Priory Street, Coventry, United Kingdom

[2]Centre for Arts, Memory and Communities, Coventry University, Coventry, United Kingdom

[3]Centre for Agroecology, Water and Resilience, Coventry University, Coventry, United Kingdom

[4]Hope For The Community Community Interest Company, Enterprise Hub, Coventry, United Kingdom

[5]Centre for Healthcare and Communities, Coventry University, Coventry, United Kingdom

[6]School of Healthcare Sciences, Cardiff University, Cardiff, United Kingdom

Corresponding Author:
Elizabeth Horton, PhD
Centre for Intelligent Healthcare, Coventry University, Priory Street, Coventry, United Kingdom

Abstract

Background: The aging population has resulted in more people living longer with musculoskeletal conditions who require hip and knee replacement surgery. Lengthening waiting lists continue to be a challenge for patients and health care services.

Objective: This pragmatic study aimed to develop and test a digital self-management intervention (the HOPE [Help Overcome Problems Effectively] program) to better prepare patients waiting for hip and knee replacement surgery.

Methods: The study used a pragmatic, pre-post with follow-up, single-arm design. All intervention and data collection components were delivered online. Patients were recruited from those on the waiting list for hip or knee surgery. Following iterative co-development of the intervention, the content was refined and optimized into a final version for testing. The resulting program was an 8-week intervention delivered via the HOPE 4 The Community (H4C) digital platform (powered by H4C). Data were collected at baseline (pre-HOPE program), 8 weeks (post-HOPE program), and 6-month follow-up. Patient-reported outcome measures related to preparation for surgery, quality of life, physical function, pain, mental well-being, self-efficacy, and physical activity. Resource usage data were collected to calculate health and social care costs. System Usability Scale data were collected post-HOPE program.

Results: One hundred participants enrolled in the HOPE program. Of these, 57 (57%) consented to take part in the evaluation and returned the baseline questionnaire. Thirty-nine participants completed ≥5 of the 8 sessions and all surveys. Among the 25 participants who had surgery at 6 months, 23 (92%) felt better prepared due to the HOPE program. Median improvements in most outcomes were observed at 8 weeks, with several continuing to improve at 6 months. The Friedman test showed significant improvements over 6 months in self-efficacy (pain: P=.002; other symptoms: P<.01), pain (P=.04), health status (P=.02), and mental well-being (P=.01). No significant changes were noted in physical activity. While the early cost analysis did not reach statistical significance, it indicated potential cost savings from reduced patient interactions with health care professionals. Sixty-four percent (25/39) of participants had surgery, and this likely contributed in part to improvements in outcomes. System usability was rated above average (mean score 70.1, SD 15.9).

Conclusions: The results are promising in relation to participants attending the HOPE program feeling better prepared for surgery. A fully powered efficacy and cost-effectiveness trial is needed to determine the contribution of the HOPE program to outcomes, over and above the contribution of surgery.

(JMIR Rehabil Assist Technol 2026;13:e68286) doi:10.2196/68286

KEYWORDS

prehabilitation; self-management; peer support; digital platform; hip and knee arthroplasty; musculoskeletal

XSL•FO
RenderX

Introduction

In the United Kingdom, an estimated 20.3 million people are affected by musculoskeletal conditions. These conditions account for 21% of the years lived with illness and disability [1]. The global prevalence of osteoarthritis is increasing, and if the trend continues, osteoarthritis will become one of the most prevalent diseases in populations from high-income countries in the coming decades [2]. The aging UK population is living longer with complex musculoskeletal conditions and comorbidities, causing increased demand on National Health Service (NHS) health and social care services [1], accounting for up to 30% of general practice consultations in England [3].

Lengthening waiting lists are particularly problematic in musculoskeletal medicine. A 2019 report found that in England alone, 726,000 people had severe hip osteoarthritis and 1.4 million people had severe knee osteoarthritis [4]. For those whose condition is severe, joint replacement surgery is the only option to alleviate pain and improve mobility and the ability to self-manage. Under the NHS constitution, 92% of patients should be treated within 18 weeks as part of the referral-to-treatment scheme. However, in 2019, nearly 4000 patients had been waiting for over 2 years for surgery [4], and more than 690,000 were on waiting lists in 2021 [5]. The COVID-19 pandemic had an unprecedented impact on secondary care orthopedic services, with a significant increase in waiting times for the majority of patients [5]. While on the waiting list, patients are likely to experience worsening pain, reduced mobility, increased anxiety, and deteriorating health, leading to greater demand for health and care services. In recognition of wait times, Versus Arthritis [5] and Arthritis Action [6] offer resources for self-management on their websites. By 2060, it is projected that the demand for hip and knee joint replacements in the United Kingdom will rise by nearly 40% from current levels, which will have significant implications for the health care system [7].

New ways of working are needed to optimize support for patients, maximize capacity, and mitigate risk. It is also important to address inequities: the COVID-19 pandemic foregrounded deep-rooted equality, diversity, and inclusion issues in relation to morbidity and mortality that are entangled with access to health care services. Inequities in treatment waiting time [8] for musculoskeletal services and in treatment outcomes [9] reflect this general picture and highlight the need for action. There is a need for holistic support among those waiting for hip and knee surgery in England. The NHS personalized care team recommends that patients on the waiting list should receive self-management support to “wait well” by undertaking prehabilitation. This support should empower patients through information, health coaching, and digital resources [10].

Prehabilitation is an effective way of improving perioperative outcomes through support to increase physical and mental resilience for surgery. Systematic reviews and meta-analyses have found some, generally low-quality evidence that prehabilitation improves a range of postoperative outcomes for patients undergoing hip and knee surgery, including function, pain, strength, and quality of life [11-13]. A more recent systematic review and meta-analysis specifically focused on the effects of digital prehabilitation in a range of musculoskeletal conditions awaiting surgery, including knee and hip replacements [14]. They found evidence that advanced technologies supported greater improvements in function pre- and post-operatively than standard care for knee and hip replacements. Greater improvements were also seen in preoperative pain, preoperative risk of falling, and postoperative stiffness. There was no evidence for spinal surgery or other conditions. However, few orthopedic prehabilitation interventions are digitally delivered, nor do they provide peer or emotional support, which is highly valued by many patients living with long-term conditions [11-13]. Indeed, a recent survey conducted in the United Kingdom [15] found that, although the vast majority of hospitals (97%) offered preoperative education, only 59% and 48% offered prehabilitation for knee and hip arthroplasty, respectively. Education was mainly delivered as a single talk supported by a booklet, and prehabilitation mainly as strengthening exercise, advice, and written information. Reported barriers included lack of facilities, funding, and staff. There was also a reported lack of robust evidence to support practice [15]. Across various surgical specialties, multimodal prehabilitation includes nutrition and psychological support alongside exercise training. There is some evidence of psychological factors improving postsurgical outcomes [13]. A systematic review and meta-analysis found low-quality evidence that psychological interventions have a positive effect on postsurgical anxiety and on mental components of quality of life [16].

In a review [17] of over 30 prehabilitation surgery schools in the United Kingdom and Ireland (these schools inform patients about what to expect and guide them on how to prepare physically and mentally to reduce postoperative risks of surgery), only 40% contained content to manage emotional well-being, and only 13% used digital apps. Further, many interventions were not underpinned by behavior change theory and techniques.

In 2022, Coventry University and its university spin-out social enterprise, H4C (HOPE [Help Overcome Problems Effectively] for The Community) interest company, developed a proof-of-concept digital intervention, called the Help Overcome Problems Effectively (HOPE) program, to help patients prepare for hip and knee surgery. The HOPE program for hip and knee patients shares the same underlying theoretical framework as other HOPE programs for long-term conditions offered by H4C, which have been taxonomized using the taxonomy of self-management support [18] and are described in detail in published papers [19-21]. All 14 digital versions of the HOPE program have been approved by the Quality Institute for Self-Management Education and Training for the provision of self-management structured education (QIS2020 and QIS2023 [22]) and certified by the Organisation for the Review of Care and Health Apps (ORCHA [23]), scoring 88% for Android and iOS (Apple Inc), and 86% for WebApp, indicating compliance with best practice in data security, professional assurance, usability, and accessibility.

XSL•FO
RenderX

The HOPE program for hip and knee patients combines evidence-based self-management content with a validated exercise program, incorporating a home exercise component tailored to individual needs and abilities, drawing from the work of Ageberg et al [24].

In 2023 H4C was awarded funding through the UK Research and Innovation Healthy Ageing Challenge Scaling Social Ventures competition to co-design and evaluate the HOPE program for hip and knee surgery patients. The funding competition was to support social enterprises in scaling products and services to support healthy aging and deliver social value.

The pragmatic, multimethod study aimed to optimize and evaluate the HOPE program to determine whether patients were better prepared for surgery. The study objectives included optimizing the HOPE program through co-design with stakeholders, implementing and testing the program with patients waiting for a joint replacement, and assessing their preparedness for surgery.

Methods

Study Design

This study used a co-design phase followed by a pragmatic, pre-post, with follow-up, single-arm intervention study. All intervention components and data collection were delivered online. This study is reported according to the CONSORT (Consolidated Standards of Reporting Trials) 2016 statement: extension for nonrandomized pilot trials [25]. CHERRIES (Checklist for Reporting Results of Internet E-Surveys) was used to guide the survey report [26]. All intervention and data collection activities took place online. All study data were collected online via questionnaires administered through Qualtrics Survey Software (Qualtrics).

Co-Design Phase to Optimize the HOPE Program

Ten participants took part in the development activities, which included 3 online workshops. One workshop was undertaken with 6 patient participants waiting for a hip (n=3) or knee replacement surgery (n=3), who had completed an earlier proof-of-concept HOPE program (5 female participants, aged 60 - 80 years). The purpose of the workshop was to explore their experiences and generate feedback on the HOPE program.

Two health professional workshops involving 4 NHS staff from our partner organizations were held to discuss referral pathways and useful resources for patients awaiting surgery. The roles of the professionals were Elective Recovery Lead, Team Lead Physiotherapist in Elective Orthopedics, Project Manager of a Musculoskeletal Clinical Program, and Senior Primary and Community Care Lead. Workshops and interviews were conducted online via Zoom (Zoom Video Communications, Inc) and MS Teams (Microsoft Corp) to allow for geographically dispersed participation.

Development of the Exercise Program

The exercise program central to the intervention was based on the The Neuromuscular Exercise training program for patients with knee or hip osteoarthritis assigned for total joint replacement the neuromuscular exercise training program for patients with knee or hip osteoarthritis assigned for total joint replacement program [24,27], which was specifically developed for older patients with severe knee and hip osteoarthritis before having total joint replacement surgery. Only the exercises from the neuromuscular exercise training program for patients with knee or hip osteoarthritis assigned for total joint replacement program were adopted within the HOPE program. Those exercises have also been incorporated into the Good Life with osteoArthritis: Denmark program [28-30]. The exercises have been demonstrated to be safe, patients can successfully progress them, and they contribute to improvements in a range of outcomes, including symptoms, function, medication use, and sick leave. A range of video and visual resources had previously been developed to support the exercise components [31]. Following feedback from the co-design phase, new video resources were developed to illustrate how the exercises could be adapted within the home environment. Forty-three videos were filmed in a home setting (living room, bedroom, and kitchen), using home furniture (sofa, chair, and bed) and both exercise equipment and everyday household items as exercise props, with volunteers representing different ages and genders, and incorporating visual prompts and voiceover instructions. The exercises target major lower limb muscle groups and can be adapted to individual capabilities, with 3 difficulty levels and encouragement to alter repetitions and sets. Participants could build their own home-based exercise program by answering 6 questions about their ability (eg, if they can easily get on and off the floor) and equipment (eg, if they have a step they can use at home). An algorithm was then built to create their personalized exercise program from the 43 videos. Participants progressed up and down levels of difficulty at their own pace, monitored progress, and set exercise reminders. Participants could download their exercise record in PDF format to keep or share with a health care professional. Tips on creating a safe exercise space, as well as important information to mitigate any worries or injuries, were part of the program.

The HOPE program: Intervention Content and Structure

The resulting program comprised 8 modules and was hosted on H4C's digital platform, powered by H4C. The content comprised text, images, videos, downloadable documents, interactive activities (eg, quizzes, self-monitoring tools, and diaries), and discussion forums and messaging facilities. The digital content was released at set times over the 8 weeks but could be accessed at any time (asynchronous). Participants had the option to "fast-track" the content if they were due to have surgery during the 8 weeks.

Once accessed and viewed, the app content could be viewed offline, reducing the need for a data plan or high-quality internet connection. An analog print booklet was produced, containing the same content as the digital version of the HOPE program, for those who were digitally excluded and/or experienced low digital literacy.

XSL•FO
RenderX

Pre-Post With Follow-Up Study

Participants

Broad eligibility criteria were used to ensure the study was as inclusive as possible and to provide ample opportunity for participation. Individuals were eligible if they were adults aged 18 years or older, lived in the South West of England in the United Kingdom, were currently on a waiting list for hip or knee replacement surgery, had access to the internet and a suitable device to engage with the intervention, and were able to interact with all materials provided as part of the intervention.

Patients interested in attending the HOPE program were referred to the study sign-up webpage through several routes. NHS South West referral sources included secondary care, primary care, and musculoskeletal clinics. Eligible participants were referred directly to H4C to enroll in the HOPE program and given the option to take part in the research study. Patients who chose to take part in the study were directed to the participant information sheet and consent form in Qualtrics Survey Software. Patients were informed that participation was voluntary and that their decision would not affect their quality of care.

We collected the following sociodemographic information: name, email address, gender, age, postcode, occupation and employment, and some details about their emotional health and their illness diagnosis, level of physical activity, health care visits, time on the waiting list, and date of surgery. Postcode data were used to calculate the English index of multiple deprivation (IMD [32]). IMD is an official measure of deprivation ranked from 1 (most deprived) to 10 (least deprived).

The questionnaire was administered through Qualtrics, using responsive and mobile-ready question formats. Adaptive questioning was used to conditionally display questions based on previous responses to reduce the number and complexity of the questions. Most pages contained between 1 and 6 items. Excluding the introduction, participant information sheet, and consent form, the survey was distributed over 14 pages. The responses were made mandatory to avoid missing data. The survey was not set up to allow participants to change their responses. The procedure, as outlined in the participant information sheet and survey structure, involved collecting identifiable information at registration—specifically, name and email address (rather than via technical means such as cookies or IP addresses)—which was then used by the research team to ensure each individual only completed the survey once per time point. Pre-HOPE program (baseline) questionnaires were completed during the period of July 6-13, 2023, for the first HOPE program and July 20-31, 2023, for the second HOPE program. Participants received a £60 (approximately US $80) electronic gift voucher for completion of all pre- and postprogram questionnaires. Participants were informed in the Patient Information Sheet how their data would be processed in accordance with the Data Protection Act 2018. Participation in the study was optional for patients who accessed the HOPE program.

The HOPE Program: Accessing and Completing the Program

Following completion of the pre-HOPE program survey, participants were given access to the HOPE program (start dates: July 13 or 27, 2023) through a personalized log-in link.

Throughout the program, participants were supported by 2 facilitators who were trained in line with Quality Institute for Self-Management Education and Training standards. The program content was organized into themed sessions across the 8 weeks, with an integrated tailored exercise program (described in the “Development of the Exercise Program” section above; Table 1 lists session content; refer to Multimedia Appendix 1 for a brief description of each session and screenshots of the intervention).

Table . Session content of the HOPE[a] program.

Session	Session content
1	Instilling HOPE
2	Managing pain and fatigue
3	Stress and shifting your thinking
4	Communication
5	Sleep and mindfulness
6	Setbacks and hospital stay
7	Happiness and strengths
8	Moving on with HOPE

[a]HOPE: Help Overcome Problems Effectively.

Patient-Reported Outcome Measures

Surgery Preparation

At 6-month follow-up, participants were asked if they felt better prepared for surgery using the following question from the Patient Preparedness for Surgery questionnaire [33]: “Overall, I feel or felt (if I had surgery) prepared for my upcoming surgery.” There were 6 response options: strongly agree, agree, somewhat agree, somewhat disagree, disagree, and strongly disagree. Participants were also asked to provide reasons for their answers. Those who had surgery indicated whether they felt the HOPE program helped them prepare before surgery, after surgery, or both. Participants provided textual responses

XSL•FO
RenderX

to explain why they agreed or disagreed that the HOPE program helped them prepare for surgery.

The following validated patient-reported outcome measures (PROMs) were collected at baseline, post-HOPE program (8 weeks), and 6-month follow-up via Qualtrics.

Short Warwick-Edinburgh Mental Well-Being Scale

The Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS [34]) is a short version that assesses mental well-being within the adult population. The SWEMWBS uses 7 items from the full WEMWBS [35], which relate more to mental functioning than feelings. The 7 statements are positively worded, with 5 response categories ranging from "none of the time" to "all of the time." Total scores range from 7 to 35, with higher scores indicating higher mental well-being. A change of one point or more on the SWEMWBS total score represents a minimally important level of change.

The EQ-5D Index and EQ-Visual Analogue Scale

The EQ-5D index [36] and the EQ-Visual Analogue Scale (EQ-VAS) are widely used measures of health status and health-related quality of life, respectively. The EQ-5D index assesses patients' health state across 5 dimensions (self-care, mobility, anxiety and depression, usual activities, and pain and discomfort) that are weighted to provide a utility value based on a population tariff. Scores range from 0 (death) to 100 (perfect health). The EQ-VAS is a vertical rating scale for health, scored between 0 (worst imaginable health) and 100 (best imaginable health).

Western Ontario and McMaster Universities Arthritis Index

The Western Ontario and McMaster Universities Arthritis Index (WOMAC [37]) consists of 24 items divided into 3 subscales: Pain (5 items), Stiffness (2 items), and Physical Function (17 items). Items are scored on a scale of 0 - 4, which corresponds to: None (0), Mild (1), Moderate (2), Severe (3), and Extreme (4). The scores for each subscale are summed, with possible score ranges of 0 - 20 for Pain, 0 - 8 for Stiffness, and 0 - 68 for Physical Function. A sum of the scores for all 3 subscales gives a total WOMAC score (maximum 96). Higher scores indicate worse pain, stiffness, and functional limitations.

Arthritis Self-Efficacy Scale

The Arthritis Self-Efficacy Scale (ASES [38]) measures a person's confidence to self-manage their arthritis symptoms and consists of 2 subscales: Pain (5 items) and Other Symptoms (6 items). Items are scored from 1 (very uncertain) to 10 (very certain). The scores for each subscale are summed, with a possible score range of 10 - 50 for Pain and 10 - 60 for Other Symptoms. Higher scores indicate higher self-efficacy.

International Physical Activity Questionnaire–Short Form

The International Physical Activity Questionnaire–Short Form (IPAQ-SF [39]) assesses physical activity undertaken across a comprehensive set of domains including: (1) leisure-time physical activity, (2) domestic and gardening (yard) activities, (3) work-related physical activity, and (4) transport-related physical activity. The items are structured to provide separate scores on walking, moderate-intensity, and vigorous-intensity activity, as well as a combined total score to describe the overall level of activity. Computation of the total score requires summation of the duration (in minutes) and frequency (days) of walking, moderate-intensity, and vigorous-intensity activity. The IPAQ-SF scoring protocol assigns the following metabolic equivalent of task (MET) values to walking, moderate, and vigorous-intensity activity: 3.3 METs, 4.0 METs, and 8.0 METs, respectively. Participants are considered to have met Centers for Disease Control and Prevention and American College of Sports Medicine physical activity recommendations if they reported at least 150 minutes per week of walking, moderate, or vigorous intensity physical activity.

Numerical Pain Rating Scale

The Numerical Rating Scale (NPRS)-11 [40] is an 11-point scale for self-report of pain. It is the most commonly used unidimensional pain scale. The respondent selects a whole number (integers 0 - 10) that best reflects the intensity (or other quality, if requested) of their pain. The anchors are 0=no pain and 10=worst possible pain (there are various wordings of the upper anchor).

HOPE Program Usability and User Engagement

Usability

The usability of the system was assessed by the System Usability Scale (SUS [41]), which was embedded in the last session of the HOPE program. It was optional for participants to complete. The SUS uses a 5-point Likert scale ranging from 1=strongly disagree to 5=strongly agree across 10 items. Odd-numbered questions (1, 3, 5, 7, and 9) generate a positive response. Even-numbered questions (2, 4, 6, 8, and 10) generate a negative response, which must be inverted. All the points added up together could gain a maximum of 40, thus the multiplication by 2.5 to make the scale out of 100. A total score of ≥68 is considered above-average usability.

User Engagement

The intervention platform collected user engagement data. For this study, we report the number of sessions completed, the number of participants who used the personalized exercise program, and the most commonly bookmarked content or activities.

Sample Size

This pragmatic study enrolled participants from an opportunity sample (n=100) comprising eligible candidates. Potential participants received an email containing a link to the study website hosted by Qualtrics. Here, participants were required to review the digital Participant Information Sheet, provide digital consent, and complete the digital questionnaires.

Analytical Methods

Data relating to sociodemographic characteristics and outcome measures were collated and presented descriptively at the group level. Outcome data were mostly ordinal and nonnormally distributed, so descriptive data were limited to frequencies (and proportions) and medians (and IQRs). While the study was not powered to detect statistically significant changes in outcomes between time points, nonparametric Friedman and post hoc Wilcoxon signed-rank tests were used to explore changes over

XSL•FO
RenderX

time between baseline, post-HOPE program (8 weeks), and 6-month follow-up. All analyses were performed using IBM SPSS (version 28). The level of statistical significance was set at $P<.05$. Textual responses to the question about surgery preparedness at the 6-month follow-up survey were summarized to illustrate the quantitative findings.

Given this was a feasibility study with complete-case analysis as the prespecified approach, we focused on participants who engaged with ≥5 sessions and completed follow-ups. This decision was made because (1) the primary aim was assessing intervention feasibility and acceptability under optimal conditions, (2) minimal data were available from noncompleters (only 4/15 provided follow-ups), and (3) high follow-up rates among completers (98% at 8 weeks and 93% at 6 months) reduced concerns about attrition bias. Future efficacy trials will use intention-to-treat (ITT) analysis.

Resource Usage

An early cost-impact analysis evaluated the change in costs associated with patients' appointments and visits with NHS England to understand the potential cost impact of the program and assess whether it could be expanded into a broader study. These data were captured via the Qualtrics survey at baseline, post-HOPE program (8 weeks), and 6-month follow-up.

The economic analysis focused on changes in the number of interactions patients had with NHS health and social care staff, measured by appointments and visits. A decision model was developed using parameters from a before-and-after analysis, literature review, and incorporating assumptions. The mean values, associated SEs, and assumptions populated the model, detailed in Multimedia Appendix 2. The total cost impact was calculated from the NHS personal and social care perspective, both per patient and per patient per week.

Costs associated with interaction changes were evaluated at 8 weeks and 6 months compared to baseline using unit costs from the Unit Costs of Health and Social Care report by the Personal Social Services Research Unit at the University of Kent [42] and the NHS National Tariff [43]. A probabilistic sensitivity analysis explored uncertainty around the results.

Ethical Considerations

The user requirements research undertaken by Coventry University received ethical approval from the Coventry University Research Ethics Committee (P151751). The research and evaluation activity has also received approval from Coventry University (P106036) and, as an amendment to a preexisting HOPE evaluation, from the Health Research Authority and Health and Care Research Wales (Integrated Research Applications System, project ID 283172).

Results

Co-Design Phase Adaptations

Adaptations to the intervention, as an outcome of patient and health professionals' feedback, were as follows: Adaptations suggested by patients were (1) guidance on how to adjust the exercises to meet individual needs and capabilities; (2) a broader range of additional activities to try, for example, pool-based exercises; (3) reassurance for people who may struggle to keep up with the program; (4) information to challenge misinformation, controversies, and conflicting advice; and (5) clearer guidance on how to access some features, for example, messaging functions.

Adaptations suggested by health professionals were (1) reminders and nudges to make healthy changes and prevent deconditioning, (2) long-term access to information for use postoperatively, (3) adjustment of exercises to cater to different abilities and comorbidities, (4) program certification to demonstrate credibility, and (5) reference to the expert input that informed the program content.

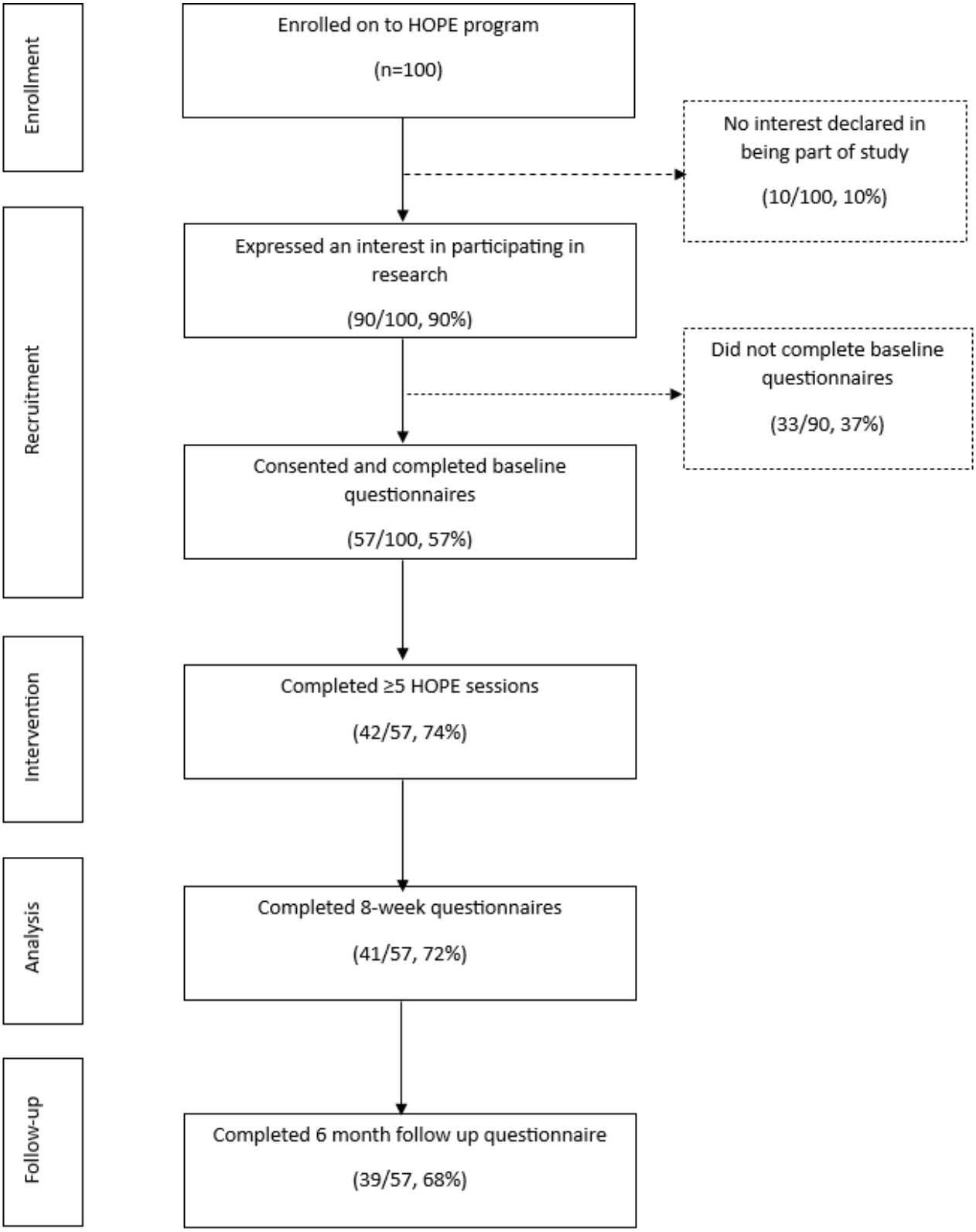
Participants

One hundred participants enrolled in the 2 HOPE programs (HOPE 1: n=59 and HOPE 2: n=41). Of these, 57 (57%) consented to take part in the evaluation and returned the baseline questionnaire (n=39, HOPE program 1 and n=18, HOPE program 2). Forty-one participants returned follow-up questionnaires at 8 weeks (41/57, 71.9%), and 39 participants returned questionnaires at 6-month follow-up (39/57, 68.4%). Forty-two participants (42/57, 73.7%) accessed ≥5 of the 8 sessions and were considered program completers.

Almost all of the HOPE program completers (41/42, 98%) returned follow-up questionnaires at 8 weeks, and 39 (93%) returned questionnaires at 6-month follow-up (Figure 1). HOPE program completers who returned both questionnaires (39/42, 93%) were included in the primary analysis. There was no missing outcome data, as these fields were required during questionnaire completion.

XSL•FO
RenderX

Figure 1. Flow of participants through the study. HOPE: Help Overcome Problems Effectively.



Participant characteristics are presented in Table 2 and are similar in the total sample (n=57) and completers (n=39). All completer participants (n=39) identified as White-English, Welsh, Scottish, Northern Irish, or British ethnicity and described English as their first language (all 39/39, 100%). One-third of participants were male (13/39, 33%) and two-thirds were female (26/39, 67%). Age was only reported by 21 (54%) participants, with a median age of 66.0 (IQR 63.0-69.5) years. The majority of participants were retired (23/39, 59%). A third of participants (13/39, 33%) were listed for hip replacement surgery, and two-thirds (26/39, 67%) for knee replacement surgery. The median IMD was 7.00 (IQR 2.5-13) and the median time on the waiting list for surgery was 6.00 (IQR 2-12) months.

XSL•FO
RenderX

Table . Participant baseline characteristics of completers (n=39) and total sample (N=57).

Characteristic	Completers (n=39)	Total sample (N=57)
Gender, n (%)		
Male	13 (33)	20 (35)
Female	26 (67)	36 (63)
Not specified	0 (0)	1 (2)
Age (years), median (IQR)	66 (63-69.5)	66 (63-69.5)
Ethnicity, n (%)		
White-English, Welsh, Scottish, Northern Irish, or British	39 (100)	56 (98)
Black, African, or Caribbean	0 (0)	1 (2)
Disability, n (%)		
Mental health condition (long-term)	5 (13)	7 (12)
Blind or partially sighted	1 (3)	1 (2)
Hard of hearing or deaf	0 (0)	1 (2)
Long-term illness or health condition (lasting more than 12 months or terminal)	4 (10)	7 (12)
Mobility impairment	24 (62)	32 (56)
Employment, n (%)		
In paid work: full-time	4 (10)	8 (14)
In paid work: part-time	4 (10)	8 (14)
Retired	23 (59)	31 (54)
Not in paid work	8 (21)	10 (18)
Not in paid work due to hip or knee condition?	8 (21)	10 (18)
Index of multiple deprivation, median (IQR)	7 (2.25)	7 (2)
Joint replacement, n (%)		
Hip	13 (33)	22 (39)
Knee	26 (67)	35 (61)
Waiting time (months), median (IQR)	6 (2-12)	7 (2.5-13)

Twenty-four out of 39 (62%) participants considered themselves to have a disability, with 9 (23%) participants reporting that day-to-day activities were “limited a little,” and 15 (39%) reporting that activities were “limited a lot.” Seven (18%) participants reported more than one specific type of disability (refer to Table 2).

User Engagement

Just over half of all participants completed all 8 sessions (30/57, 53%), with 6 participants completing 7 sessions (6/57, 11%), 1 completing 6 sessions (1/57, 2%), and a further 5 participants completing 5 sessions (5/57, 9%). Forty-nine out of 57 (86%) participants used the personalized exercise program. The top 5 bookmarked content or activities were (1) exercise program, (2) relaxed breathing, (3) mindfulness meditation, (4) compassionate approach to pain, and (5) cognitive diffusion activity.

Patient-Reported Outcomes and Estimations

By the time of the 6-month follow-up, 25 out of 39 (64%) participants had already received their surgery. Of those who had their surgery, the majority (23/25; 92%) agreed with the statement: “As a result of attending the HOPE program, overall, I felt better prepared for my surgery.” Eight out of 23 (35%) participants selected “strongly agree,” 10 (44%) selected “agree,” and 5 (22%) selected “somewhat agree.” Of the 23 participants who agreed that they were better prepared, 16 (70%) felt better prepared in the presurgery period, 3 (13%) felt better prepared postsurgery, and 5 (17%) felt better prepared pre- and postsurgery.

Of those who had not yet had surgery, the majority (13/14, 93%) agreed with the statement: “As a result of attending the HOPE program, overall, I feel better prepared for my surgery.” Of these, 1 participant (1/14, 7%) selected “strongly agree,” 7 (n=7/14, 50%) selected “agree,” and 5 (5/23, 30%) selected “somewhat agree.”

XSL•FO
RenderX

All 39 participants who completed the 6-month follow-up questionnaire provided reasons why they agreed or disagreed that the HOPE program helped them prepare for surgery. The findings are presented under 4 headings: personalized exercise, physical and mental preparation, peer support, and nothing new. Participant ID numbers 1 - 14 are those that were still waiting for surgery at 6-month follow-up, and IDs 15-39 are participants who had undergone surgery. No harms or unintended consequences were reported during the study.

Personalized Exercise

The program offered exercises that helped patients improve their physical condition and overall preparedness for surgery.

> *Better exercised and with better muscle definition.* [ID19]
>
> *It gave me some exercises to prepare for surgery.* [ID29]
>
> *It encouraged me to do the preparation exercises and helped lift my mood when needed.* [ID20]

Physical and Mental Preparation

Patients found the program beneficial for preparing both physically and mentally for surgery. It provided information and insight about what to expect before and after surgery, helping to manage pain, reduce anxiety, increase hope, and plan for the future.

> *The program gave me an insight into what to expect during and after the procedure.* [ID18]
>
> *I found the information useful and the relaxation techniques particularly helpful.* [ID38]
>
> *The information given was clear about the future after the operation.* [ID7]
>
> *I feel I manage pain better even if it becomes more painful.* [ID8]
>
> *I knew so much about what to expect, and I learned techniques to calm any anxiety.* [ID33]

Peer Support

Connecting with others who have arthritis and are waiting for surgery made patients feel less isolated. The program offered a platform for discussing shared challenges, such as surgery delays and recovery expectations, fostering a sense of community among participants. Participants valued the emotional support they received through the program. Sharing experiences with others who were undergoing similar surgeries provided comfort, while insights into the surgical process helped ease fears.

> *Hearing what other arthritis sufferers are going through made you feel that you are not alone in dealing with the pain.* [ID1]
>
> *The HOPE program gave me the opportunity to share my thoughts/fears with others who had either had their joint surgery or were waiting to undergo it.* [ID17]

Nothing New

A few participants found that the program covered what they already knew or that they already had a positive mindset.

> *I haven't found out anything new.* [ID9]
>
> *I already had a very positive view of how to deal with the issues arising from my arthritis.* [ID10]

Usability

Only 16 participants completed the optional SUS. Participants reported a mean SUS score of 70.1 (SD 15.9; range 50 - 95). The 10-item frequency response data are provided in Multimedia Appendix 3. Compared with the 23 participants who did not complete the SUS, the 16 SUS completers were younger (median age of 64, IQR 8 vs 67, IQR 6 years; sample size n=8 and n=13, respectively), included a higher proportion of males (44% vs 26%), and more knee surgery patients (75% vs 60%) with a mobility impairment (69% vs 50%). Other patient characteristics were broadly similar. On average, the 16 completers had slightly lower disease severity: total WOMAC median 49 (IQR 23) versus 53 (IQR 17).

Table 3 summarizes the patient-reported outcomes at baseline, 8 weeks (post-HOPE program), and 6-month follow-up. Median values suggested potential improvements in many outcome measures at the end of the HOPE program (8 weeks). There were sustained improvements in median values for several outcomes at 6 months.

XSL•FO
RenderX

Table . Summary of baseline, post–Help Overcome Problems Effectively (HOPE) program (8 weeks), and 6-month follow-up outcomes (n=39).

Outcome variable	Baseline, median (IQR)	8 weeks, median (IQR)	*P* value, Wilcoxon test (baseline to 8 week)	6 months, median (IQR)	*P* value, Friedman test
Arthritis Self-Efficacy Scale (ASES)					
Confidence to manage pain (1 - 10, ↑=better)	3.8 (2.0-5.6)	4.6 (3.6-5.4)	.07	5.6 (3.8-8.2)	.002[a]
Confidence to manage other symptoms (1 - 10, ↑=better)	4.5 (2.5-5.5)	5.0 (4.2-6.5)	.001[a]	6.5 (4.2-8.3)	<.001[a]
Western Ontario and McMaster Osteoarthritis Index (WOMAC)					
Pain (0 - 20, ↑=worse)	10.0 (8.0-12.0)	10.0 (7.0-13.0)	.19	8.0 (5.0-12.0)	.0[a]4
Stiffness (0 - 8, ↑=worse)	4.0 (4.0-5.0)	4.0 (4.0-5.0)	.92	4.0 (2.0-5.0)	.11
Physical functioning (0 - 68, ↑=worse)	35.0 (26.0-41.0)	34.0 (24.0-42.0)	.44	29.0 (10.0-42.0)	.07
Total (0 - 96, ↑=worse)	49.0 (40.0-58.0)	48.0 (33.0-59.0)	.38	39.0 (19.0-60.0)	.09
Numerical Pain Rating Scale (NPRS)					
Pain (0 - 10, ↑=worse)	6.0 (5.0-7.0)	5.0 (4.0-7.0)	.19	5.0 (2.0-6.0)	.002[a]
EQ-5D					
Quality of Life (EQ-VAS[c]; 0 - 100, ↑=better)	58.0 (35.0-80.0)	60.0 (35.0-80.0)	.37	70.0 (45.0-85.0)	.05
Health status (EQ-Index; 0 - 1, ↑=better)	0.62 (0.30-0.74)	0.60 (0.23-0.78)	.54	0.75 (0.30-0.83)	.02[a]
Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS)					
Mental well-being (SWEMWBS; 5 - 35, ↑=better)	25.0 (21.0-28.0)	25.0 (22.0-28.0)	.86	27.0 (23.0-29.0)	.01[a]
International Physical Activity Questionnaire–Short Form (IPAQ-SF)					
Total (MET-min/wk, ↑=better)	2340 (393-7464)	2628 (480-6152)	.98	2079 (306-5988)	.55
Sitting time (min/d, ↑=worse)	360 (285-480)	360 (181-540)	.41	300 (240-480)	.15
Inactive (<600 MET[b]-min/wk), n (%)	12 (30.8)	10 (25.6)	—[d]	12 (30.8)	—[d]

[a]Statistically significant *P*<05. Data is for participants who participated in ≥5 sessions and completed both follow-up questionnaires (n=39).

[b]MET: metabolic equivalent.

[c]EQ-VAS: EQ-Visual Analogue Scale.

[d]Not applicable.

The Friedman test indicated several potential improvements across the 6 months. Median scores for ASES pain were 3.8 (IQR 2.0-5.6), 4.6 (IQR 3.6-5.4), and 5.6 (IQR 3.8-8.2) at baseline, 8 weeks, and 6-month follow-up, respectively (*P*=.002); ASES other symptoms: 4.5 (IQR 2.5-5.5), 5.0 (IQR 4.2-6.5), and 6.5 (IQR 4.2-8.3*; P*<.001); WOMAC pain: 10.0 (IQR 8.0-12.0), 10.0 (IQR 7.0-13.0), and 8.0 (IQR 5.0-12.0*; P*=.04); NRPS: 6.0 (IQR 5.0-7.0), 5.0 (IQR 4.0-7.0), and 5.0 (IQR 2.0-6.0*; P*=.002); EQ-index: 0.62 (IQR 0.3-0.74), 0.60 (IQR 0.23-0.78), and 0.75 (IQR 0.30-0.83*; P*=.02); and SWEMWBS: 25.0 (IQR 21.0-28.0), 25 (IQR 22.0-28.0), and 27.0 (IQR 23.0-29.0*; P*=.01). Separate Wilcoxon tests at 8 weeks (immediately following the end of the HOPE program) found that only ASES other symptoms was statistically significant (*P*=.001; refer to Table 3).

Ancillary Analyses

Assessment of Bias: Program Completers Versus Program Noncompleters at Baseline

A total of 15 participants were categorized as noncompleters of the Hope program (ie, completing <5 of 8 sessions). Only 4 (27%) of these participants returned both follow-up questionnaires, which was insufficient for meaningful analysis.

XSL•FO
RenderX

Therefore, bias assessment was conducted using baseline data only. Potential differences at baseline between noncompleters (<5 sessions; n=15) and completers (≥5 sessions; n=42) were explored using descriptive statistics. There were no obvious differences between noncompleters and completers in age (median 65.50, IQR 64-72.75 vs median 66.00, IQR 63.0-69.0 years, respectively) or IMD (median 7.00, IQR 5.0-8.0 vs median 7.00, IQR 6.0-8.25). There were slight differences between noncompleters and completers in gender (male: 50% vs 31%), ethnicity (White: 93% vs 100%), disability (yes: 53% vs 62%), employment (in paid work—full time and part time: 54% vs 20%), joint replacement (knee: 47% vs 67%), and waiting time (7.00, IQR 4.0-16.0 vs 6.00, IQR 2.0-13.0 months). On average, noncompleters had slightly greater disease severity. For example, noncompleters reported more pain (NPRS: median 7.00, IQR 6.0-8.0 vs median 6.00, IQR 5.0-7.0, respectively), had a higher total WOMAC score (median 60.00, IQR 40.0-71.0 vs median 49.00, IQR 38.75-58.25), and a lower EQ-5D index value (median 0.45, IQR 0.18-0.73 and median 0.636, IQR 0.29-0.74).

Impact of Surgery on Outcomes

To understand the potential impact of surgery on outcomes, an exploratory descriptive comparison between those who had and had not received surgery at 6 months was made (data presented in Multimedia Appendix 4). This comparison was only based on the ASES data, since improvements in this outcome were statistically significant at both 8 weeks and 6 months. The results show that those who had received surgery at 6 months had larger median improvements in self-efficacy (for both pain and other symptoms). Those who had not had surgery showed marginal improvements in self-efficacy for pain and for other symptoms at 8 weeks. These were maintained at 6 months for pain self-efficacy but not for other symptoms.

Health Care Resource Usage

The component of the study that focused on resource use for this early cost-impact analysis had a sample size of 39 patients, who completed the web-based questionnaire at all 3 time points: baseline, after 8 weeks, and after 6 months. Of these, 25 patients had their surgical intervention within the period covered (ie, within 6 months) and were therefore excluded from the analysis, leaving a total sample size of 14 analyzed. Results are provided in Table 4 (total cost-impact per patient) and Table 5 (cost-impact per patient per week). Cost-impact per patient per week evaluation revealed overall cost savings over 8 weeks as well as over 6 months, but this failed to reach statistical significance. Face-to-face general practitioner interactions at the 6-month interval showed a statistically significant change. Further details of the economic analysis are provided in Multimedia Appendix 2.

Table . Total cost impact per patient (£per week).

Cost category[a]	Cost change from baseline to 8 weeks, mean (95% CI)	Cost change from baseline to 6 months, mean (95% CI)
Face-to-face visit with a physiotherapist	12.85 (–3.80 to 36.02)	–8.64 (–76.34 to 27.29)
Remote visit with GP[b]	–7.72 (–19.55 to 0.36)	0.22 (–3.38 to 4.07)
Face-to-face visit with GP	9.66 (–12.57 to 34.00)	14.09 (–5.39 to 36.29)
Face-to-face hospital visit	–7.46 (–29.65 to 10.52)	–8.91 (–38.35 to 10.84)
Total	7.34 (46.49 to –27.61)	–3.25 (46.10 to –87.02)

[a]Positive values correspond to cost savings.

[b]GP: general practitioner.

Table . Cost impact per patient (£ per week).

Cost category[a]	Cost change from baseline to 8 weeks, mean (95% CI)	Cost change from baseline to 6 months, mean (95% CI)
Face-to-face visit with physiotherapist	1.61 (–0.48 to 4.5)	0.69 (–3.34 to 3.94)
Remote visit with [c]GP	–0.96 (–2.44 to 0.04)	0.20 (–0.1 to 0.63)
Face-to-face visit with GP	1.21 (–1.57 to 4.25)	2.45 (0.50 to 5.08)[b]
Face-to-face hospital visit	–0.93 (–3.71 to 1.31)	–0.03 (–1.85 to 1.82)
Total	0.92 (–3.45 to 5.81)	3.31 (–1.93 to 8.12)

[a]Positive values correspond to cost savings.

[b]Statistically significant (P<.05).

[c]GP: general pratitioner.

Sample Size Calculation for Future Trial

Data collected as part of the current evaluation were used to inform likely sample sizes for future studies in this area. This sample size calculation was based on ASES-8 data. Unfortunately, the minimum clinically important difference of the ASES-8 is unknown [44]. However, it is sensitive to change, with an effect size of 0.31 previously reported for the ASES-8

XSL•FO
RenderX

following interdisciplinary group therapy for fibromyalgia [45]. Moderate effect sizes of this magnitude are common for conservative interventions in musculoskeletal conditions. In this pilot study, the mean and SD values for ASES pain and ASES other symptoms at baseline were 3.98 (SD 1.93) and 4.29 (SD 2.08), respectively. Assuming a 1-tailed hypothesis, an effect size of 0.3, α=.05, 90% power, and a 1:1 allocation ratio, 191 participants would be required in each group (N=382) to detect a ≥0.58-point difference in ASES pain and ≥0.62-point difference in ASES other symptoms.

Discussion

Principal Findings

This study evaluated the HOPE program, a digital self-management intervention designed to support patients awaiting hip and knee replacement surgery. Results from 39 completers suggested potential improvements in self-efficacy, pain, health status, and mental well-being over 6 months. Most participants felt better prepared for surgery, and the program was rated above average for usability (mean SUS score 70.1).

Participant feedback revealed some key areas that underscore the program's potential usefulness. Some participants appreciated the targeted exercises that improved their physical and mental readiness for surgery. The program provided comprehensive information about the surgical process, helping patients manage pain, reduce anxiety, and plan for the future. Studies have shown that patients have difficulties remembering information immediately after deciding to undergo surgery [46]. Having access to digital information, which can be regularly and quickly updated with evidence-based information, is a useful resource for patients. By fostering a sense of community, the program helped some participants connect with others facing similar challenges. However, some participants noted that the program offered nothing new, as they already enjoyed a positive mindset or previous knowledge.

The demographic profile of completers (median age 66, IQR 63-69.5 years; 100% White; and 66.7% female) was almost identical to a recent UK study, which found that digital health coaching delivered to patients waiting for lower limb arthroplasty improved patient activation and reduced length of hospital stay [47]. It should be noted that noncompleters of the program were more likely to be male, in paid employment, and awaiting a hip replacement.

Engagement with the HOPE program was high, with 73.7% (42/57) of participants attending ≥5 of 8 sessions. Follow-up and engagement rates were lower when based on the 100 participants who enrolled: 39% (39/100) completed the 6-month follow-up questionnaires, and 42% (42/100) who completed ≥5 sessions. Among those who completed all study procedures, 93% (39/42) engaged with the program.

A recent national digital attitudes and behavior survey conducted in the United Kingdom by ORCHA in 2023 described the willingness of older respondents to use digital apps for self-monitoring, symptom tracking, and managing recovery [48].

At the 6-month follow-up, nearly two-thirds (25/39, 64%) of participants had undergone surgery. More than 90% (23/25) of these participants agreed that the program helped them prepare better for surgery. Statistically significant median improvements in most PROMs were evident at the end of the HOPE program, and several scores continued to improve at 6-month follow-up, including self-efficacy, pain, health status, and mental well-being. The exercise program was the most bookmarked page, and despite the majority of participants (49/57, 86%) starting the personalized exercise program, there were no improvements in time spent sitting or in the proportion of participants classified as inactive. The exercise program may require greater input from facilitators to encourage optimum engagement. Research shows that exercise supervision involving trained physical therapists improves compliance with exercises, especially in older adults [13,49]. Alternatively, it may be that the IPAQ-SF lacks sensitivity to adequately assess physical activity [39]. More objective measures of physical activity, such as accelerometry, could be considered in future research.

The high number of participants undergoing surgery makes it challenging to attribute potential improvements in PROMs to either intervention. In their systematic review and meta-analysis, Punnoose et al [13] showed that variability in surgical procedures can influence postoperative recovery; therefore, postsurgical improvements cannot be attributable solely to prehabilitation. Owing to the often degenerative nature of musculoskeletal conditions, potential improvements in PROMs in this study were not anticipated a priori. Rather, it was hypothesized that attending the HOPE program would slow the rate of decline through the acquisition of effective self-management and coping strategies. Thus, the observed trend for median improvements across the majority of PROMs is encouraging.

Resource Usage

This early cost analysis suggests that the HOPE program may lead to a reduction in patient interactions with care professionals at both 8 weeks and 6 months. However, the small sample size results in wide CIs, which limits the reliability of these findings and affirms the need for further studies to assess the cost-effectiveness of the program. Despite this limitation, the initial results highlight the potential for the HOPE program to offer cost-saving benefits at a societal level.

Strengths and Limitations

A strength of this real-world study was the inclusion of participants with lived experience at all stages of the project, providing input into the HOPE program intervention development process and follow-up feedback to optimize it for further studies. The majority of participants started the exercise program, which is a cornerstone of prehabilitation. Other strengths include the use of validated PROMs, high levels of engagement with the intervention, and good survey completion rates at 6 months. This version of the HOPE program was rapidly developed and deployed by adding new musculoskeletal content to an existing taxonomized evidence-based intervention. Some of the health professionals involved in the co-design workshops suggested that patients needing only conservative management and not requiring surgery would also benefit from

XSL•FO
RenderX

the program. Our co-creation and intervention development process could develop and test a program for these patients and for other groups of nonorthopedic presurgery patients. The powered by H4C platform currently hosts more than 15 digital self-management and health interventions. Using a single platform to deliver multiple interventions and modules offers several advantages for funders, researchers, health care providers, and patients. Many patients live with comorbid conditions requiring diverse information and self-management techniques. Platform delivery can incorporate and streamline self-management support. Torous and Vaidyam [50] asserted that "instead of a plethora of apps, there is a need for a few that meet the needs of many." Drawing on successful examples from the automobile, space, and clean energy sector, Ansar and Flyvbjerg [51] outline the benefits of platforms over one-off designs, such as repeatability, extendibility, absorptive and adaptive capacity, resulting in "faster, better, cheaper" services and products. They concluded that sectors such as health "are ripe for a platform rethink." Another strength of this application is the partnership between a social enterprise company and an academic institution. A recent Wellcome report [52] recommended that companies, including nonprofits, can be better at developing and scaling digital health solutions than university research groups.

The limitations of the study include the small number (16/39) of participants who completed the SUS. It is possible that these participants had a more positive user experience compared to those who did not complete the scale. Among the 39 study completers, most participants (>90%) agreed that the program helped them prepare better for surgery, and the textual responses supporting this question provided limited feedback. A broader set of feedback questions and/or postprogram qualitative interviews or focus groups analyzed using rigorous and transparent methods—with participants who did not complete the program—could elicit more critical or negative experiences.

The self-selecting nature of recruitment may have resulted in participants who were inherently more inclined to seek assistance or engage in self-help efforts.

Without a control group comparator, it is not possible to directly attribute any change in the PROMs to the HOPE program. It is important to note that many improvements were not statistically significant, and the statistical analyses performed were likely to be underpowered. Furthermore, a recent systematic review of hip arthroplasty prehabilitation interventions suggested that measures such as the WOMAC may not be the most appropriate measure to detect differences and suggest alternative objective measures such as the chair rise test, gait speed, or stair climbing [53]. That review also found that more than 8 weeks of prehabilitation was associated with improved outcomes, suggesting that future trials of the HOPE program should consider extending the length of the intervention. While our completer analysis provides valuable proof-of-concept data, it limits generalizability to real-world implementation, where attrition is typically higher. The baseline differences between completers and noncompleters suggest our effect estimates may be optimistic. Future randomized controlled trials (RCTs) should combine ITT and per-protocol analyses to distinguish efficacy from effectiveness.

Separating the effects of the intervention from the effects of surgery is problematic. The ancillary analysis of the ASES data (refer to Table 4) suggests that surgery was probably a major contributor to improvements in self-efficacy at 6 months. This is not surprising, given that the excellent outcomes of hip replacement surgery have led to the procedure being described by The Lancet as the "operation of the century" [54]. Approximately 96.2% and 90.8% of patients have previously reported satisfaction with their hip and knee replacement surgery, respectively [55].

A future definitive RCT should be appropriately powered to directly compare an intervention group (ie, the HOPE program) against an appropriate control group (ie, treatment as usual). Subgroup analysis should compare PROMs in those who have had, or are still awaiting, surgery at 6-month follow-up. Such a design would help to distinguish the effects of the intervention from the effects of surgery.

The baseline data show that program noncompleters (ie, those who completed <5 sessions) had slightly greater disease severity at baseline than program completers. Owing to limited follow-up data, it is not known whether these participants could not complete the program due to factors relating to their musculoskeletal condition, their experience of the program, or random intervening factors. Nonresponders were also more likely to be male, in paid employment, and awaiting a hip replacement. Such findings raise questions about how to engage people with greater disease severity and these sociodemographic characteristics in future support programs. Further research is needed to understand individual needs and how they change as disease and pain progress, and to determine how best to support individuals through targeted interventions.

In line with the wealth of other UK health care research studies, the participant sample in this study lacked diversity in terms of ethnicity and socioeconomic characteristics. The study sample reflects the demographics of NHS waiting lists and can be understood as a manifestation of structural inequalities. People living in the most deprived areas of the United Kingdom are more likely to require replacement surgery but less likely to receive it [56] and less likely to have good outcomes [57,58] compared with those living in the least deprived areas. This recurring finding underscores the need for research into the impact of structural barriers to self-management, which may, in turn, suggest the need for more options or a new paradigm approach. Health care interventions that disproportionately meet the needs of nonmarginalized groups embed injustice by widening health inequity. The earlier statement that no harm was reported during the study holds when "harm" is understood within the parameters of evidence-based medicine and its associated framework of biomedical ethics. However, when a framing such as distributive justice is applied, the intervention may be associated with unintended adverse consequences that emerge from and perpetuate ideologies such as structural racism and classism. Lack of attention to unintended harm linked to the lack of diversity in self-management research highlights the need for an expanded ethical framework informed by disability justice scholarship [59]. Recommendations from a recent report into musculoskeletal health inequalities in the United Kingdom included prioritizing surgery and self-management support for

XSL•FO
RenderX

patients living in the most deprived areas [60]. More effort is required to understand the needs of and actively recruit these groups of participants in future self-management trials. A national digital attitudes and behavior survey conducted in the United Kingdom by ORCHA in 2021 [61] found that advocacy for digital health apps was highest among people of Black African heritage (89%), followed by Asian (80%), and then White (64%) respondents. Studies from the United States highlight the importance of recruiting low-income and ethnic minority participants, showing that these groups are more willing to attend [62] and engage more [63] with health interventions compared with White participants in higher-income groups. However, data from this study show that deprivation levels were similar between HOPE program completers and noncompleters.

Conclusion

The results are promising in relation to the acceptability of a peer-supported self-management program for people awaiting hip or knee surgery. Overall, participants felt better prepared for surgery. Textual feedback was generally positive, and participants attributed improvements in their mental and physical well-being to techniques they learned in the HOPE program. However, comparing self-efficacy in those who had and had not received surgery suggests that surgery might have been a more important agent of change than the HOPE program. Overall, the study has demonstrated potential benefit and no evidence of harm or unintended consequences. A randomized controlled efficacy and cost-effectiveness trial design, involving a socioeconomically and ethnically representative sample, is required to delineate the effects attributable to the HOPE program, as opposed to effects of having surgery or natural variation in PROMs. While these preliminary results are promising, they require confirmation in a fully powered RCT using ITT analysis to account for real-world attrition patterns.

Acknowledgments

The authors would like to thank the participants involved in the co-creation workshops, stakeholders, partners, and course participants. The HOPE program is delivered by the Coventry University spin-out social enterprise, Hope 4 The Community (H4C) CIC. We would also like to thank Health Tech Enterprise for contributing to the health economic analysis. The project, Hope and Prepare Effectively for Surgery (HoPES; 10053263), was funded by UK Research and Innovation as part of the Healthy Ageing Challenge Scaling Social Ventures competition, delivered by Innovate UK. Funding was awarded to H4C through the Small Business Research Initiative Social Ventures competition, run by UK Research and Innovation as part of the Healthy Ageing Challenge, delivered by Innovate UK. This pilot study was not registered with a clinical trials registry. There was no published protocol for this pilot study.

Funding

No external financial support or grants were received for this work.

Conflicts of Interest

AT is a co-founder and director of H4C and the co-inventor of the original HOPE program. AC is the head of Partnerships and Projects at H4C.

Multimedia Appendix 1
Intervention screenshot and session content.
[DOCX File, 560 KB - rehab_v13i1e68286_app1.docx]

Multimedia Appendix 2
Health economic evaluation analysis.
[DOCX File, 40 KB - rehab_v13i1e68286_app2.docx]

Multimedia Appendix 3
System Usability Scale scores.
[DOCX File, 22 KB - rehab_v13i1e68286_app3.docx]

Multimedia Appendix 4
Changes in Arthritis Self-Efficacy Scale.
[DOCX File, 29 KB - rehab_v13i1e68286_app4.docx]

Checklist 1
CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist.
[DOCX File, 28 KB - rehab_v13i1e68286_app5.docx]

XSL•FO
RenderX

Checklist 2
CONSORT-EHEALTH checklist.
[PDF File, 683 KB - rehab_v13i1e68286_app6.pdf]

References

1. The state of musculoskeletal health 2021: arthritis and other musculoskeletal conditions in numbers. Arthritis UK. URL: https://www.versusarthritis.org/media/24238/state-of-msk-health-2021.pdf [accessed 2025-12-12]
2. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. The Lancet 2019 Apr;393(10182):1745-1759. [doi: 10.1016/S0140-6736(19)30417-9]
3. Musculoskeletal health. NHS England. 2020. URL: https://www.england.nhs.uk/elective-care-transformation/best-practice-solutions/musculoskeletal/ [accessed 2024-08-07]
4. Versus arthritis position statement on the clinically-led review of NHS access standards. Versus Arthritis. 2019. URL: https://www.versusarthritis.org/media/21674/clinical-standards-review-nov19.pdf [accessed 2024-08-07]
5. What we're doing to tackle waiting times for people with arthritis. Versus Arthritis. 2023 Dec. URL: https://www.versusarthritis.org/news/2023/december/what-we-re-doing-to-tackle-waiting-times-for-people-with-arthritis/ [accessed 2024-08-07]
6. Treatment waiting times. Arthritis Action. 2024. URL: http://www.arthritisaction.org.uk/living-with-arthritis/resource-centre/waiting-times/ [accessed 2024-08-07]
7. Matharu G, Culliford D, Blom A, Judge A. Projections for primary hip and knee replacement surgery up to the year 2060: an analysis based on data from The National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. Ann R Coll Physicians Surg Can 2022 Jun;104(6):443-448. [doi: 10.1308/rcsann.2021.0206]
8. People living in the poorest areas waiting longer for hospital treatment: The King's Fund and Healthwatch England share new analysis. The King's Fund. 2021 Sep. URL: https://www.kingsfund.org.uk/insight-and-analysis/press-releases/poorest-areas-waiting-longer-hospital-treatment#:~:text=New%20analysis%20from%20The%20King's,in%20the%20least%2Ddeprived%20areas [accessed 2024-08-07]
9. Bergstra SA. Health inequalities across patients with early inflammatory arthritis of different ethnicities: what could be the driving factors? Rheumatology (Oxford) 2022 Dec 23;62(1):7-8. [doi: 10.1093/rheumatology/keac383] [Medline: 35786709]
10. Elective recovery planning supporting guidance. NHS England. 2022 Apr. URL: https://www.england.nhs.uk/wp-content/uploads/2021/12/B1269-elective-recovery-planning-supporting-guidance.pdf [accessed 2024-08-07]
11. Vasta S, Papalia R, Torre G, et al. The influence of preoperative physical activity on postoperative outcomes of knee and hip arthroplasty surgery in the elderly: a systematic review. J Clin Med 2020 Mar 31;9(4):969. [doi: 10.3390/jcm9040969] [Medline: 32244426]
12. Wang L, Lee M, Zhang Z, Moodie J, Cheng D, Martin J. Does preoperative rehabilitation for patients planning to undergo joint replacement surgery improve outcomes? A systematic review and meta-analysis of randomised controlled trials. BMJ Open 2016 Feb;6(2):e009857. [doi: 10.1136/bmjopen-2015-009857]
13. Punnoose A, Claydon-Mueller LS, Weiss O, Zhang J, Rushton A, Khanduja V. Prehabilitation for patients undergoing orthopedic surgery: a systematic review and meta-analysis. JAMA Netw Open 2023 Apr 3;6(4):e238050. [doi: 10.1001/jamanetworkopen.2023.8050] [Medline: 37052919]
14. Guida S, Vitale JA, Swinnen E, et al. Effects of prehabilitation with advanced technologies in patients with musculoskeletal diseases waiting for surgery: systematic review and meta-analysis. J Med Internet Res 2024 Dec 12;26:e52943. [doi: 10.2196/52943] [Medline: 39666971]
15. Omar I, Wylde V, Fogg J, Whitehouse M, Bertram W. Pre-operative education and prehabilitation provision for patients undergoing hip and knee replacement: a national survey of current NHS practice. BMC Musculoskelet Disord 2025 Apr 29;26(1):421. [doi: 10.1186/s12891-025-08637-5] [Medline: 40301783]
16. Perry R, Herbert G, Atkinson C, et al. Pre-admission interventions (prehabilitation) to improve outcome after major elective surgery: a systematic review and meta-analysis. BMJ Open 2021 Sep 30;11(9):e050806. [doi: 10.1136/bmjopen-2021-050806] [Medline: 34593498]
17. Fecher-Jones I, Grimmett C, Carter FJ, Conway DH, Levett DZH, Moore JA. Surgery school—who, what, when, and how: results of a national survey of multidisciplinary teams delivering group preoperative education. Perioper Med 2021 Dec;10(1):20. [doi: 10.1186/s13741-021-00188-2]
18. Pearce G, Parke HL, Pinnock H, et al. The PRISMS taxonomy of self-management support: derivation of a novel taxonomy and initial testing of its utility. J Health Serv Res Policy 2016 Apr;21(2):73-82. [doi: 10.1177/1355819615602725] [Medline: 26377727]
19. Martin F, Wright H, Moody L, et al. Help to overcome problems effectively for cancer survivors: development and evaluation of a digital self-management program. J Med Internet Res 2020 May 19;22(5):e17824. [doi: 10.2196/17824] [Medline: 32209529]
20. Wright H, Martin F, Clyne W, et al. A digital self-management program (help to overcome problems effectively) for people living with cancer: feasibility randomized controlled trial. J Med Internet Res 2021 Nov 5;23(11):e28322. [doi: 10.2196/28322] [Medline: 34738912]

XSL•FO
RenderX

21. Wright H, Turner A, Ennis S, et al. Digital peer-supported self-management intervention codesigned by people with long COVID: mixed methods proof-of-concept study. JMIR Form Res 2022 Oct 14;6(10):e41410. [doi: 10.2196/41410] [Medline: 36166651]
22. Accreditation register. Quality Institute for Self Management Education and Training (QISMET). 2024. URL: https://www.qismet.org.uk/accreditation/accreditation-register/ [accessed 2024-08-07]
23. Delivering safe digital health. The Organisation for the Review of Care and Health Apps (ORCHA). 2024. URL: https://orchahealth.com [accessed 2024-08-07]
24. Ageberg E, Link A, Roos EM. Feasibility of neuromuscular training in patients with severe hip or knee OA: the individualized goal-based NEMEX-TJR training program. BMC Musculoskelet Disord 2010 Jun 17;11(1):126. [doi: 10.1186/1471-2474-11-126] [Medline: 20565735]
25. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5(1):114. [doi: 10.1186/s40814-019-0499-1] [Medline: 31608150]
26. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
27. Ageberg E, Nilsdotter A, Kosek E, Roos EM. Effects of neuromuscular training (NEMEX-TJR) on patient-reported outcomes and physical function in severe primary hip or knee osteoarthritis: a controlled before-and-after study. BMC Musculoskelet Disord 2013 Dec;14(1):232. [doi: 10.1186/1471-2474-14-232]
28. Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA:D™): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. BMC Musculoskelet Disord 2017 Feb 7;18(1):72. [doi: 10.1186/s12891-017-1439-y] [Medline: 28173795]
29. Roos EM, Barton CJ, Davis AM, et al. GLA:D to have a high-value option for patients with knee and hip arthritis across four continents: Good Life with osteoArthritis from Denmark. Br J Sports Med 2018 Dec;52(24):1544-1545. [doi: 10.1136/bjsports-2017-098904]
30. Barton CJ, Kemp JL, Roos EM, et al. Program evaluation of GLA:D® Australia: physiotherapist training outcomes and effectiveness of implementation for people with knee osteoarthritis. Osteoarthr Cartil Open 2021 Sep;3(3):100175. [doi: 10.1016/j.ocarto.2021.100175] [Medline: 36474815]
31. Goff AJ, De Oliveira Silva D, Ezzat AM, et al. Co-design of the web-based “My Knee” education and self-management toolkit for people with knee osteoarthritis. Digit HEALTH 2023;9:20552076231163810. [doi: 10.1177/20552076231163810] [Medline: 37009308]
32. English indices of deprivation 2019. GOV.UK. 2019 Sep. URL: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 [accessed 2025-12-12]
33. Kenton K, Pham T, Mueller E, Brubaker L. Patient preparedness: an important predictor of surgical outcome. Am J Obstet Gynecol 2007 Dec;197(6):654. [doi: 10.1016/j.ajog.2007.08.059] [Medline: 18060968]
34. Ng Fat L, Scholes S, Boniface S, Mindell J, Stewart-Brown S. Evaluating and establishing national norms for mental wellbeing using the short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): findings from the Health Survey for England. Qual Life Res 2017 May;26(5):1129-1144. [doi: 10.1007/s11136-016-1454-8] [Medline: 27853963]
35. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. Health Qual Life Outcomes 2009 Feb 19;7(1):15. [doi: 10.1186/1477-7525-7-15] [Medline: 19228398]
36. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
37. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988 Dec;15(12):1833-1840. [Medline: 3068365]
38. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self‐efficacy in people with arthritis. Arthritis & Rheumatism 1989 Jan;32(1):37-44. [doi: 10.1002/anr.1780320107]
39. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the international physical activity questionnaire short form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8(1):115. [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
40. Stratford PW, Spadoni G. The reliability, consistency, and clinical application of a numeric pain rating scale. Physiotherapy Canada 2001;53:88-91.
41. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
42. Unit costs 2022-2027. University of Kent. URL: https://www.pssru.ac.uk/unitcostsreport/ [accessed 2024-08-07]
43. National tariff payment system documents, annexes and supporting documents. NHS England. 2023. URL: https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/ [accessed 2023-08-07]
44. Brady TJ. Measures of self-efficacy: Arthritis Self-Efficacy Scale (ASES), Arthritis Self-Efficacy Scale-8 Item (ASES-8), Children’s Arthritis Self-Efficacy Scale (CASE), Chronic Disease Self-Efficacy Scale (CDSES), Parent’s Arthritis

XSL•FO
RenderX

Self-Efficacy Scale (PASE), and Rheumatoid Arthritis Self-Efficacy Scale (RASE). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S473-S485. [doi: 10.1002/acr.20567] [Medline: 22588769]
45. Mueller A, Hartmann M, Mueller K, Eich W. Validation of the arthritis self-efficacy short-form scale in German fibromyalgia patients. Eur J Pain 2003;7(2):163-171. [doi: 10.1016/S1090-3801(02)00097-6] [Medline: 12600798]
46. Specht K, Kjaersgaard-Andersen P, Pedersen BD. Patient experience in fast-track hip and knee arthroplasty--a qualitative study. J Clin Nurs 2016 Mar;25(5-6):836-845. [doi: 10.1111/jocn.13121] [Medline: 26708610]
47. Powley N, Tew GA, Durrand J, et al. Digital health coaching to improve patient preparedness for elective lower limb arthroplasty: a quality improvement project. BMJ Open Qual 2023 Dec 7;12(4):e002244. [doi: 10.1136/bmjoq-2022-002244] [Medline: 38061840]
48. Digital health in the UK today, annual national attitudes and behaviour research. ORCA. 2023. URL: https://info.orchahealth.com/digital-health-attitudes-behaviour-2023-report [accessed 2025-12-12]
49. Shubert TE. Evidence-based exercise prescription for balance and falls prevention: a current review of the literature. J Geriatr Phys Ther 2011;34(3):100-108. [doi: 10.1519/JPT.0b013e31822938ac] [Medline: 22267151]
50. Torous J, Vaidyam A. Multiple uses of app instead of using multiple apps - a case for rethinking the digital health technology toolbox. Epidemiol Psychiatr Sci 2020 Jan 31;29:e100. [doi: 10.1017/S2045796020000013] [Medline: 32000876]
51. Ansar A, Flyvbjerg B. How to solve big problems: bespoke versus platform strategies. Oxford Review of Economic Policy 2022 May 18;38(2):338-368. [doi: 10.1093/oxrep/grac009]
52. The Wellcome Trust. Catalytic environments to drive progress in digital mental health research. Wellcome Open Res 2024;9:301. [doi: 10.21955/wellcomeopenres.1115389.1]
53. Widmer P, Oesch P, Bachmann S. Effect of prehabilitation in form of exercise and/or education in patients undergoing total hip arthroplasty on postoperative outcomes-a systematic review. Medicina (Kaunas) 2022 May 30;58(6):742. [doi: 10.3390/medicina58060742] [Medline: 35744005]
54. Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. The Lancet 2007 Oct;370(9597):1508-1519. [doi: 10.1016/S0140-6736(07)60457-7]
55. Appiah KOB, Khunti K, Kelly BM, et al. Patient‐rated satisfaction and improvement following hip and knee replacements: development of prediction models. Evaluation Clinical Practice 2023 Mar;29(2):300-311. [doi: 10.1111/jep.13767]
56. Cook MJ, Lunt M, Ashcroft DM, Board T, O'Neill TW. The impact of frailty and deprivation on the likelihood of receiving primary total hip and knee arthroplasty among people with hip and knee osteoarthritis. J Frailty Aging 2023;12(4):298-304. [doi: 10.14283/jfa.2023.36] [Medline: 38008980]
57. Clement ND, Muzammil A, MacDonald D, Howie CR, Biant LC. Socioeconomic status affects the early outcome of total hip replacement. J Bone Joint Surg Br 2011 Apr;93-B(4):464-469. [doi: 10.1302/0301-620X.93B4.25717]
58. Clement ND, Jenkins PJ, MacDonald D, et al. Socioeconomic status affects the Oxford knee score and short-form 12 score following total knee replacement. Bone Joint J 2013 Jan;95-B(1):52-58. [doi: 10.1302/0301-620X.95B1.29749] [Medline: 23307673]
59. Piepzna-Samarasinha LL. Care Work: Dreaming Disability Justice: Arsenal Pulp Press; 2018:263.
60. Act now: musculoskeletal health inequalities and deprivation. : Arthritis and Musculoskeletal Alliance (ARMA); 2024 URL: https://arma.uk.net/wp-content/uploads/2024/03/Musculoskeletal-Health-Inequalities-and-Deprivation-report_v07.pdf [accessed 2024-08-07]
61. Digital health in the UK: national attitudes and behaviours. ORCHA. 2021. URL: https://us.orchahealth.com/wp-content/uploads/2021/07/2107_ICS_Research_Report_2021_National_final.pdf [accessed 2025-12-12]
62. Lorenzo-Luaces L, Wasil A, Kacmarek CN, DeRubeis R. Race and socioeconomic status as predictors of willingness to use digital mental health interventions or one-on-one psychotherapy: national survey study. JMIR Form Res 2024 Apr 11;8:e49780. [doi: 10.2196/49780] [Medline: 38602769]
63. Horrell LN, Kneipp SM, Ahn S, et al. Chronic disease self-management education courses: utilization by low-income, middle-aged participants. Int J Equity Health 2017 Jun 27;16(1):114. [doi: 10.1186/s12939-017-0604-0] [Medline: 28655319]

Abbreviations

ASES: Arthritis Self-Efficacy Scale
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
CONSORT: Consolidated Standards of Reporting Trials
EQ-VAS: EQ-Visual Analogue Scale
H4C: HOPE 4 The Community Interest Company
HOPE: Help Overcome Problems Effectively
IMD: index of multiple deprivation
IPAQ-SF: International Physical Activity Questionnaire–Short Form
ITT: intention-to-treat
MET: metabolic equivalent of task
NHS: National Health Service

XSL•FO
RenderX

NPRS: Numerical Pain Rating Scale
ORCHA: Organisation for the Review of Care and Health Apps
PROM: patient-reported outcome measure
RCT: randomized controlled trial
SUS: System Usability Scale
SWEMWBS: Short Warwick-Edinburgh Mental Well-Being Scale
WOMAC: Western Ontario and McMaster Universities Arthritis Index

Edited by S Munce; submitted 01.Nov.2024; peer-reviewed by G Martin-Alemany, J Bailey, P Marier-Deschenes; revised version received 23.Oct.2025; accepted 04.Nov.2025; published 07.Jan.2026.

Please cite as:
Horton E, Wright H, Turner A, Moody L, Aphramor L, Carlson A, Ghiasvand H, Palmer S
Using a Co-Designed Digital Self-Management Program to Prepare Patients for Hip or Knee Replacement Surgery: Pragmatic Pilot Study
JMIR Rehabil Assist Technol 2026;13:e68286
URL: https://rehab.jmir.org/2026/1/e68286
doi:10.2196/68286

© Elizabeth Horton, Hayley Wright, Andy Turner, Louise Moody, Lucy Aphramor, Anna Carlson, Hesam Ghiasvand, Shea Palmer. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 7.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Stakeholder Perspectives on Facilitators of and Barriers to the Implementation of Internet of Things–Equipped Powered Wheelchairs in Sweden: Qualitative Case Study

Johan Borg[1*], PhD; Marie Elf[1*], PhD; Anders Brantnell[2,3*], PhD; Klas Palm[2*], PhD

[1]School of Health and Welfare, Dalarna University, Högskolegatan 2, Falun, Sweden

[2]Department of Civil and Industrial Engineering, Division of Industrial Engineering and Management, Uppsala University, Uppsala, Sweden

[3]Department of Women's and Children's Health, CIRCLE - Complex Intervention Research in Health and Care, Uppsala University, Uppsala, Sweden

[*]all authors contributed equally

Corresponding Author:
Johan Borg, PhD
School of Health and Welfare, Dalarna University, Högskolegatan 2, Falun, Sweden

Abstract

Background: The implementation of Internet of Things (IoT) technologies to proactively monitor the functioning of assistive products remains underexplored, particularly in public assistive technology services.

Objective: This qualitative single-case study investigated how a public assistive technology provider in Sweden implemented IoT-equipped assistive products.

Methods: The implementation process was mapped, and a thematic analysis of interviews with 6 participants representing key organizational roles was conducted.

Results: The process comprised 9 major events involving 4 main stakeholders. The findings included 4 themes of facilitators and 5 themes of barriers, which were synthesized into three meta-themes: (1) system-level structures that enable and constrain implementation, (2) tensions between technological potential and the challenges of realizing it, and (3) relational and cultural dynamics that balance trust with caution.

Conclusions: The findings underscore the importance of using a sociotechnical framework to guide the implementation of IoT in assistive technology services. Such a framework supports a holistic approach that integrates user needs, privacy considerations, and regulatory requirements, while drawing on leadership, technical expertise, and collaboration to address legal, organizational, technical, and user-related challenges. The implications for health care technology implementation are that organizations should adopt sociotechnical implementation strategies that account for legal, organizational, technical, and user-related considerations.

(JMIR Rehabil Assist Technol 2026;13:e92414) doi:10.2196/92414

KEYWORDS

assistive products; assistive technology; implementation; Internet of Things; IoT; powered wheelchairs

Introduction

Assistive products that function poorly can limit users' participation at home and in the community and may lead to reduced quality of life, injury, or pain [1-4]. Maintenance and repair services are therefore essential components of the provision of assistive products [5]. Using technologies such as the Internet of Things (IoT) to monitor use, condition, and status offers a proactive approach to identifying maintenance needs, preventing users from being stranded, and providing insights that can inform design improvements.

IoT refers to a network of uniquely identifiable objects connected to the internet, often equipped with sensing, actuation, and programming capabilities [6]. These capabilities allow data collection and remote adjustment of device states. In health care, IoT has been recognized as a tool for improving quality and efficiency, enabling real-time monitoring and enhanced coordination among devices, service users, and providers. Such connectivity can support more proactive and integrated care, potentially reducing system strain by lowering resource demands and workloads [7,8].

Breakdowns and repairs of powered wheelchairs are relatively common [2-4]. Applying IoT to assistive products like these enables remote monitoring (eg, overheating and location), optimization of individual settings (eg, speed, footrest angle), troubleshooting (eg, identifying damaged components), and maintenance planning (eg, detecting battery degradation). Thus, the IoT has the potential to improve the management and use of assistive products, and consequently, enhance users' quality of life. However, evidence regarding the successful

XSL•FO
RenderX

implementation of IoT in health care remains limited [9], and to our knowledge, no studies have examined the implementation of IoT-equipped assistive products in practice.

Successful IoT adoption in health care is challenging because it involves more than introducing new devices. It requires integrating new technologies into an existing network of actors, routines, and governance structures [10,11]. A sociotechnical perspective can help explain why this integration succeeds or fails by treating implementation as a dynamic interaction of factors across social, organizational, and technical dimensions that influence each other [9]. From this perspective, leadership is central for providing direction, legitimacy, and resources [12]. The involvement of critical actors within the organization can support local adaptation, workflow fit, and trust among users. Understanding how these dimensions influence each other is important for translating IoT's potential into sustained use in health services [13].

The sociotechnical perspective has been applied in health care research such as the adoption of health IT [14], the usability of electronic health records [15], and the use of AI in health care [16]. Although these studies demonstrate the importance of studying health care technologies as sociotechnical systems, in which the implementation depends on interactions between technical and organizational factors [14-16], few studies have examined the actual implementation of IoT solutions in health care practice [9]. Moreover, a recent review of IoT adoption in health care did not identify any studies exploring the implementation process from a sociotechnical perspective, particularly regarding IoT-equipped assistive products within assistive technology services [9].

Furthermore, to our knowledge, no previous studies have specifically investigated the implementation of IoT-equipped assistive products in real-world health care settings. Consequently, there is insufficient understanding of how social, organizational, and technical factors interact during the implementation of IoT-equipped assistive products.

The objective of this study was to explore the implementation process of IoT-equipped powered wheelchairs within a public assistive technology provider in Sweden from a sociotechnical perspective. The study focused on the implementation journey from initiation through to the organizational decision to deploy the technology to end users.

Specifically, it addressed the following research questions:

1. What were the major events that shaped the implementation of IoT-equipped powered wheelchairs?
2. Who were the stakeholders involved in this process and what roles did they play?
3. What facilitators and barriers did stakeholders perceive as influencing the implementation of IoT-equipped powered wheelchairs?
4. How can the implementation of IoT-equipped powered wheelchairs be understood from a sociotechnical perspective?

Methods

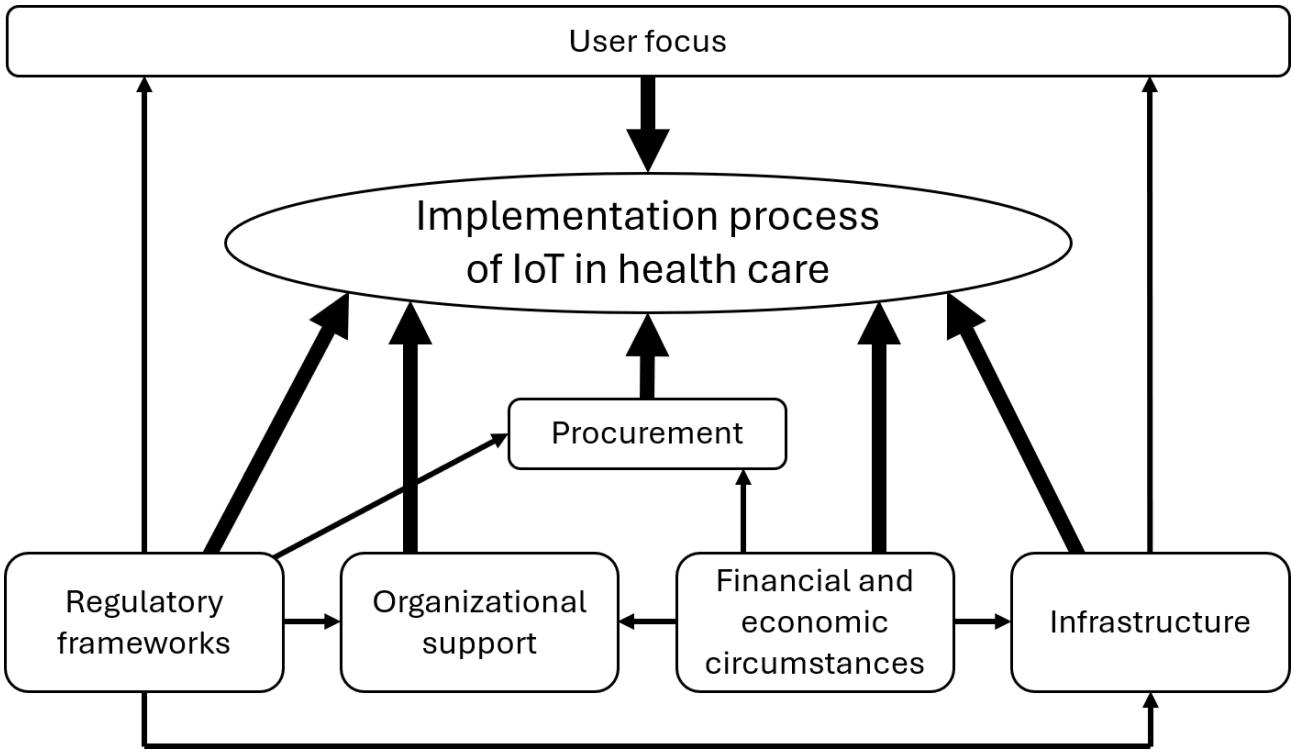
Qualitative Approach and Theory

This study employed a qualitative single-case study design [17,18], using thematic analysis of semistructured interviews. The reporting is structured in compliance with the Standards for Reporting Qualitative Research [19].

To provide a theoretical perspective and enhance the understanding of the adoption process, we drew on the sociotechnical framework developed by Davis et al [20] and later enhanced by Palm et al [21], as the original framework was found not to fully cover important dimensions in the implementation of IoT in health care [9]. The enhanced sociotechnical systems model is tailored to IoT implementation in health care and includes the most important factors that seem to influence the implementation process (Figure 1) [21].

XSL•FO
RenderX

Figure 1. A sociotechnical systems model tailored to Internet of Things (IoT) implementation in health care, showing the most important factors that seem to influence the IoT implementation process in health care. The broader arrows represent a direct impact on the IoT implementation process in health care, while the narrower arrows represent an indirect impact, from one factor to another factor, which, in turn, affects the ability to implement IoT in health care (adapted from Palm et al [21]).



Researcher Characteristics and Reflexivity

The research team comprised 4 investigators with complementary expertise: an engineer specializing in assistive technology with long-standing practice and research experience, a registered nurse contributing clinical insights and implementation competence, an implementation researcher with a track record in qualitative and quantitative studies of health care innovation and implementation, and a social psychologist with extensive practitioner- and research-based experience in organizational development processes.

All authors shared an interest in exploring the implementation of digital technologies in the health sector, which informed the study design and interpretation. None of the investigators had prior personal or professional relationships with the participants. To mitigate bias, they discussed their assumptions and interpretations during data analysis. This reflexive process helped ensure that findings were grounded in participants' experiences rather than researchers' preconceptions.

Context

The study was conducted at an Assistive Technology Center serving approximately 300,000 residents in a mostly sparsely populated county in Central Sweden. The center is responsible for assessing needs, providing, maintaining, and repairing assistive products, including IoT-equipped powered wheelchairs. It operates under the Regional Authority, which oversees health care and assistive technology provision across the county. In addition, the Regional Authority is responsible for dental care, public health, public transport, research, culture, and, to some extent, education. To support these operations, the authority maintains a central IT department and an application board, which prepares and manages matters related to digital applications and their use within the organization.

At the time of the study, the Assistive Technology Center had recently integrated IoT-equipped powered wheelchairs into routine practice to support usage monitoring, simplify troubleshooting, and improve maintenance and service processes. Through mobile network connectivity, the IoT system enabled remote monitoring of parameters such as battery condition, system errors (eg, joystick-related faults), motor current, and the position and inclination of the seat and footrest. The system also provided access to geographic location in cases of theft or other emergencies. Automatic alerts supported early fault detection, diagnostics, service planning, and preventive maintenance by technicians. Approximately 60 IoT-equipped wheelchairs were in active use during the interview period.

Recruitment and Participants

Participants were purposively selected to capture diverse experiences of how the IoT-equipped powered wheelchairs were implemented. First, the Assistive Technology Center provided the names of 2 staff members who had been engaged in the implementation process and who worked with IoT-equipped powered wheelchairs daily. During a process-mapping interview, they provided the names of 4 people who had played key roles during the implementation, representing different stakeholders. These 6 individuals were included based on their involvement in the process under study. The stakeholders and the participants' roles are presented in Table 1.

XSL•FO
RenderX

Table . Stakeholders, participants' roles and numbers, and interview durations.

Stakeholders	Participants' roles and numbers	Interview durations (min)
Assistive Technology Center, Regional Authority	Technician, 2 males	51
Assistive Technology Center, Regional Authority	Head of Department, Mobility, 1 male	34
IT Department, Regional Authority	Head of Administration, 1 male	31
Application Board, Regional Authority	Administrative Manager, 1 female	26
Supplier of IoT[a]-equipped wheelchairs	Sales Development Manager, 1 male	27

[a]IoT: internet of things.

The participants were asked to recommend additional people to talk to, ensuring that we captured relevant people involved in the implementation process. One additional person was suggested, but when contacted, he said that he had not been involved in the process. Thus, we concluded that we had interviewed all key people involved in the implementation process. The sampling strategy was guided by the study's focus on a specific implementation process rather than capturing the perspectives of all stakeholder groups affected by the technology. The participants were selected because they had direct responsibilities and active roles in the implementation of IoT-equipped powered wheelchairs. They were therefore considered information-rich sources regarding the events, decisions, facilitators, and barriers that shaped the process.

Ethical Considerations

Based on their assessment that no sensitive personal data would be collected, the Swedish Ethics Review Authority dismissed the application for ethical review of the study, as the study does not involve sensitive personal data or physical interventions. The interviews were conducted after obtaining informed consent from the participants, who could withdraw from the study at any time without providing a reason. Data protection regulations were followed.

Data Collection

Data were collected in 2 steps. In step 1, JB interviewed 2 participants in February 2024 to map the process and identify key stakeholders, with all information recorded in writing. In step 2, KP, ME, and JB collected data from 6 participants (including a second interview with the 2 participants from step 1) through semistructured interviews conducted between August 2024 and January 2025. Four individual online interviews and one pair-wise face-to-face interview were conducted and audio-recorded. The interview durations ranged from 26 to 51 minutes (Table 1), with a median time of 31 (IQR 27-34) minutes.

Based on the purpose and research questions of this study, the authors developed an interview guide for step 2, which included 5 main questions and 20 follow-up questions (Multimedia Appendix 1). The questions addressed issues such as the process, the stakeholders involved, and the barriers and facilitators to the implementation of IoT technology.

Data Processing

Manually recorded data from step 1 were compiled to map the implementation process and prepare for step 2. The audio-recorded data from step 2 were transcribed using TurboScribe (Leif Erikson Ventures, LLC) [22]. The individual transcriptions were reviewed, corrected manually, and anonymized.

Data Analysis

To address research questions 1 and 2, data from step 1 were compiled, and a preliminary process map indicating key stakeholders and major events was developed.

In the inductive, reflexive thematic analysis [23,24] of data from step 2, conducted to address research question 3, a hierarchical structure (subthemes→themes→meta-themes) was employed to represent progressively higher-order interpretations of patterned meaning across the dataset. This approach is consistent with multilevel thematic organization as described by Attride-Stirling [25]. The analysis followed six phases: (1) familiarization with the data, (2) generation of initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. An example of the analysis process from meaning units (quotes) to meta-themes is provided in Multimedia Appendix 2.

In line with the opportunities for using ChatGPT in thematic analysis identified by Lee et al [26], GPT-5 was employed in the analysis. The analysis was conducted in Swedish to ensure that all interpretations were grounded in the interview data and that the extracted quotes could be verified against the source material. In phase 2, ChatGPT was used to generate codes representing facilitators and barriers from individual interviews and to illustrate each code with a corresponding quote. In phases 3 and 5, based on the codes from all interviews, ChatGPT generated themes and subthemes addressing research question 3. Each subtheme was supported by its underlying codes to allow verification. When presenting the themes, ChatGPT provided concise descriptions of their core ideas. Given the relatively large number of themes, ChatGPT was subsequently instructed to synthesize them at a higher level, resulting in three meta-themes. In phase 6, ChatGPT was used to translate illustrative quotes, summarize the meta-themes, and review the language in selected parts of the paper. The translated quotes were compared with the original quotations to ensure accuracy and were revised where necessary to better reflect their intended meaning. Details of the use of ChatGPT are provided in Multimedia Appendix 3.

As part of the inductive, reflective thematic analysis, the authors familiarized themselves with the interview transcripts to gain an understanding of the data. They reviewed the generated codes

XSL•FO
RenderX

to ensure that they captured essential data. The authors reviewed and refined the subthemes, themes, and meta-themes to ensure comprehensive coverage of the data and to ensure that barriers and facilitators were thoroughly addressed. Quotations were used to illustrate each theme.

Finally, to address research question 4, the resulting meta-themes, themes, and subthemes were mapped onto the sociotechnical framework. Thus, the framework was applied analytically after the inductive thematic analysis and was not used deductively to generate the themes.

Techniques to Enhance Trustworthiness

The authors engaged with the data to develop informed insights and employed rich, illustrative quotations to support the identified subthemes and themes, thereby enhancing the credibility of the findings. Although the study used a case study design, descriptions of the participant sample, research setting, and methodological procedures were provided to support transferability. Given the case study design and the lesser number of participants, the findings are not intended to be statistically generalizable. Rather, the study provides contextually grounded insights into one implementation process, which may offer analytical transferability to similar health care technology implementation contexts. The use of ChatGPT contributed to the trustworthiness of the analysis, particularly in terms of credibility and dependability, by offering objectivity during data familiarization, coding, theme refinement, and the selection of representative quotes. The coding process and theme development were systematically documented, and the coding was reviewed by the authors to enhance dependability.

Results

Events and Stakeholders in the Implementation Process

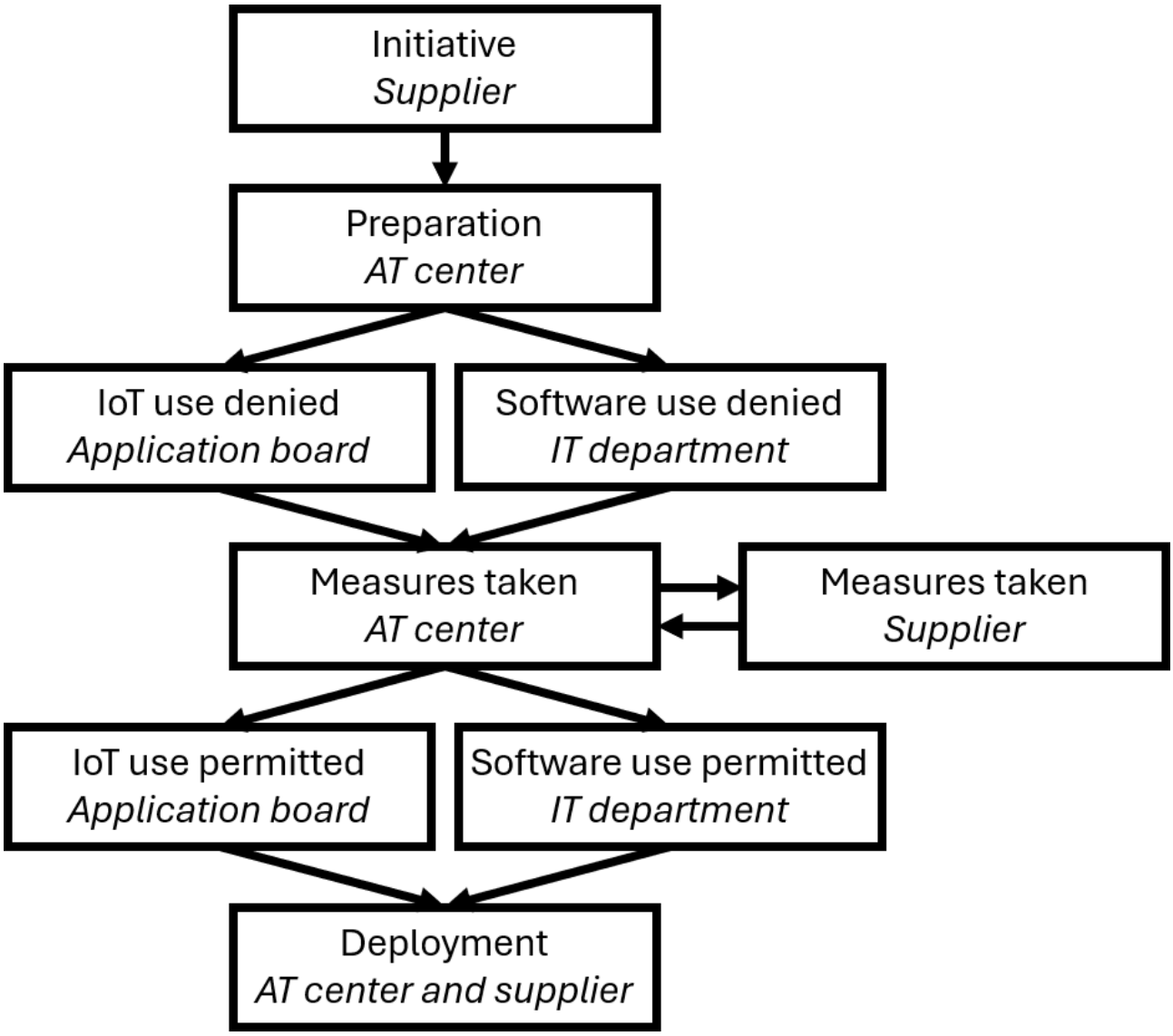
In the case studied, the implementation of IoT-equipped powered wheelchairs was initiated by the supplier. Although the technology had been integrated into their wheelchairs as early as 2019, it had not yet been used in Sweden. The supplier encouraged the Assistive Technology Center to begin using the technology. Some of the technicians at the Center realized that the technology could be useful, as they were responsible for the maintenance and repair of powered wheelchairs in a geographic area of nearly 30,000 square kilometers. The Assistive Technology Center brought the matter to the regional application board, which initially rejected the use of the technology due to concerns related to the General Data Protection Regulation (GDPR), including the ownership of the server where wheelchair data were stored, which was located outside the European Union and GPS data being classified as personal data. The regional application board also requested that 2-factor authentication be implemented. Following repeated communications between the Assistive Technology Center, the supplier, and the application board, the application board approved the use of the IoT technology. This decision was based on several changes: the data server was relocated to Sweden, the GPS functionality was disabled, and a form of 2-factor authentication was established. The latter was achieved by linking each user to their respective wheelchair using the serial number and information stored in the Assistive Technology Center's digital assistive product management system, to which the supplier did not have access. While the supplier continued to collect GPS data, it did not share it with the Center except under specific circumstances, such as theft.

In parallel with discussions involving the regional application board, the Assistive Technology Center also engaged with the IT department of the regional authority responsible for health care. The use of the supplier's software for managing IoT-equipped powered wheelchairs required technicians to have local administrator rights. These rights were initially restricted but were later permitted. After obtaining the necessary permissions, the Assistive Technology Center's management team decided to deploy the IoT technology in 2023.

Based on the initial interviews with the technicians, a preliminary reconstruction of the sequence of the 9 major events and the 4 stakeholders involved in the process is outlined schematically in Figure 2. The boxes in the figure do not represent the exact number of events or interactions between stakeholders.

XSL•FO
RenderX

Figure 2. A preliminary process map of major events and stakeholders involved in the implementation process of the Internet of Things (IoT)–equipped powered wheelchairs. AT: assistive technology.

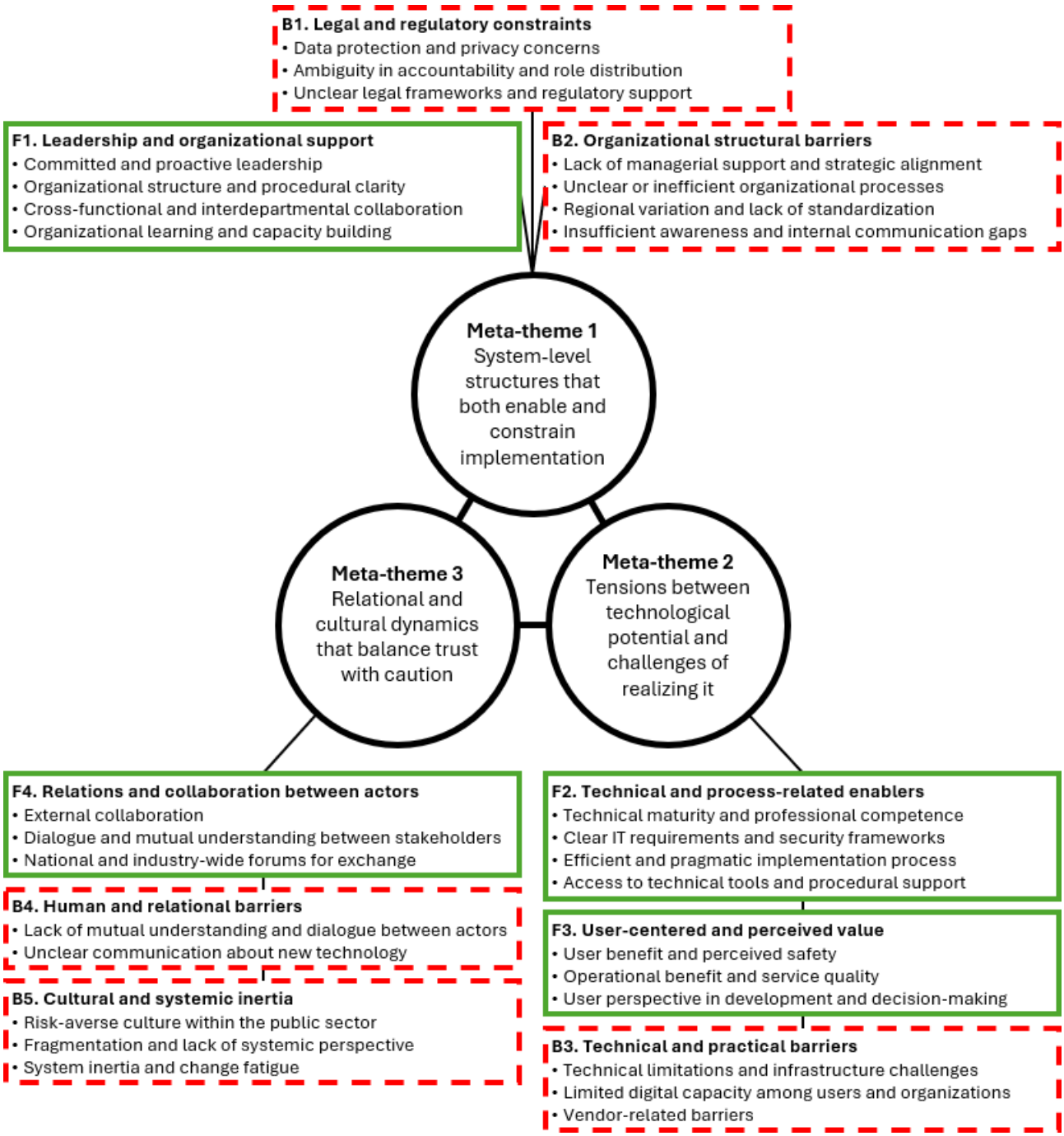


Facilitators and Barriers in the Implementation Process

The thematic analysis of the interview transcripts yielded 9 themes and 29 subthemes, which were organized into 3 meta-themes (Figure 3). These meta-themes represent overarching tensions and mechanisms that span both facilitators and barriers related to the implementation of IoT-equipped powered wheelchairs.

XSL•FO
RenderX

Figure 3. Meta-themes, themes, and subthemes of facilitators (F, solid green line) and barriers (B, dashed red line) in the implementation process of Internet of Things (IoT)–equipped powered wheelchairs.



System-Level Structures That Both Enable and Constrain Implementation

The implementation process was largely shaped by the design of organizational and legal systems, including their processes, roles, and decision-making pathways.

Leadership and Organizational Support

The implementation of IoT-equipped powered wheelchairs was facilitated by committed and proactive leadership that demonstrated engagement and clear managerial anchoring. Supportive and competent colleagues within the organization further reinforced this process. In the words of 1 participant:

> *We got a new manager who was interested in this. Because it's important to have a manager who is interested... He said, yes, go ahead and keep this going.* [Interview with Technicians, Assistive Technology Center]

Clear organizational structures and procedures also contributed to the implementation. Collaboration and structured decision-making processes involving multiple functions and departments ensured alignment across the organization. One participant described this approach as follows:

> *We have a process for new and changed needs, where we look at the entire chain and involve operations,*

XSL•FO
RenderX

> *IT, and legal before making a decision.* [Interview with Head of Administration, IT Department]

Over the course of the process, the stakeholder organizations developed and strengthened their capacity, which not only supported the implementation of the IoT-equipped wheelchairs but also generated valuable learning for future technology implementations.

Legal and Regulatory Constraints

Uncertainty about data protection, the division of responsibilities, and legislation created ambiguity and delays in the implementation process. Questions arose about whether the service complied with requirements under the GDPR and whether cloud solutions could be used for data storage. Concerns were also raised about privacy and the potential risk of data being accessed by actors in other countries.

Stakeholders experienced ambiguity in accountability and role distribution. The unclear allocation of responsibilities among stakeholders, including users themselves, was perceived as a barrier that contributed to delays. As 1 participant reflected:

> *When are we, as municipality and region, responsible for the operation, and when is it the user themselves? It's not always crystal clear.* [Interview with Head of Administration, IT Department]

Unclear legal frameworks and insufficient regulatory support further complicated the process. Changes in the classification of health data, together with the absence of legislation enabling data exchange between stakeholders such as municipalities and regions, made implementation challenging.

Organizational and Structural Barriers

A lack of coordination and clear processes made the implementation effort cumbersome and vulnerable. Limited managerial engagement and strategic alignment were evident in the initial lack of driving leadership, the absence of a joint working group within the Assistive Technology Center, and insufficient national coordination. As one of the participants described:

> *We were probably a bit resigned for a while because we didn't think our manager drove this the way they should have.* [Interview with Technicians, Assistive Technology Center]

Unclear or inefficient organizational processes further impeded progress. Participants described burdensome administrative routines, complex decision-making structures, and bureaucratic obstacles. At the same time, there was no established process for implementation, which contributed to uncertainty and inconsistency.

Regional variation and the absence of standardization added to the complexity. The supplier observed notable differences between Swedish regions, necessitating adaptations to the technical solution. Even within a single region, perceptions of the technology varied. Insufficient awareness and internal communication gaps also hindered implementation. This was reflected in the limited knowledge of the Application Board and in unclear communication channels within the region. This was illustrated by one participant, who explained:

> *Partly, I can imagine that the operations don't know that ... It may be somewhat unknown that there is an Application Board.* [Interview with Administrative Manager, Application Board]

Tensions Between Technological Potential and the Challenges of Realizing It

The studied implementation process revealed that the value of technology was realized only when stakeholders shared a common understanding of its benefits.

Technical and Process-Related Enablers

Practical, technical, and process-related factors created favorable conditions for the implementation and effective use of the IoT solution. Technical maturity and professional competence among stakeholders facilitated progress. This included the supplier's improved ability to demonstrate the technology's benefits, combined with key individuals' understanding of the underlying problems. In addition, the local technicians were enabled to act as system administrators for the required software, which was not the case earlier.

Clear IT requirements and established security frameworks at both the organizational and international levels further supported the implementation. Updates to the Network and Information Systems Directive enabled compliance, while the region's IT policies and information security classification guided the process. Additionally, the option to disable the GPS-tracking function of the IoT-equipped powered wheelchairs allowed adherence to security requirements.

An efficient and pragmatic implementation process was reflected in regular meetings, which strengthened communication between operations and IT, and a streamlined internal decision-making structure within the Assistive Technology Center. One participant put it in this way:

> *We have just over a hundred employees, and with that also a management group that keeps track of most things and can steer and make it work as it should.* [Interview with Head of Department, Mobility, Assistive Technology Center]

User-Centered and Perceived Value

Access to technical tools and procedural support contributed to a user-centered and perceived value for both the Assistive Technology Center and wheelchair users, which were recognized from the early stages of implementation. For users, perceived value was reflected in positive reactions to having their wheelchairs connected, including faster repairs and increased perceived safety, such as the ability to locate the wheelchair in the event of theft. From an operational perspective, the technology enabled earlier fault detection and planning of repairs, reduced the need to travel to investigate malfunctions, and allowed technicians to bring appropriate tools and parts, thereby saving a considerable number of working hours. According to 1 participant:

> *We have found quite a lot of benefit with it when we've learned to handle error codes. And that we don't need to travel long distances to investigate a*

XSL•FO
RenderX

> *fault.* [Interview with Head of Department, Mobility, Assistive Technology Center]

These advantages were viewed as providing not only economic but also environmental value.

Through their close contact with wheelchair users, technicians brought forward users' experiences and perspectives believed to influence the development and decision-making regarding IoT-equipped powered wheelchairs.

Technical and Practical Barriers

Technical barriers, insufficient adaptation, and low levels of digital maturity delayed or complicated the implementation process. Technical limitations and infrastructure challenges included the requirement to deactivate the GPS tracking function in the basic configuration and the inability to integrate cloud services. Additional issues involved firewalls that blocked software updates and uncertainty regarding data ownership.

Limited digital capacity among both users and organizations further constrained progress. Challenges included wheelchair users' limited ability to manage digital technology, inadequate authentication procedures in applications, and a generally slow digital transition within the region. As 1 participant remarked:

> *What you notice is this sluggish transformation into a digital society within the regions in general... they're perhaps not always the quickest to act.* [Interview with Sales Development Manager, Supplier]

Regarding the supplier, participants perceived a lack of understanding of data protection and highlighted the limited number of suppliers, which resulted in weak competition.

Relational and Cultural Dynamics That Balance Trust With Caution

The studied implementation process was influenced by relational dynamics, such as trust, communication, and collaboration, both internally and externally.

Relations and Collaboration Among Stakeholders

Successful implementation depended on effective relationships and communication between regions, suppliers, and supporting functions. External collaboration was facilitated by good coordination between the Assistive Technology Center and the supplier. Participants also suggested that an industry association could play a more active role as a driving force in fostering collaboration between suppliers and regions. As one participant pointed out:

> *I think an organization like [industry association] could be more of a driving force ... a forum where we suppliers, together with the regions, can discuss.* [Interview with Sales Development Manager, Supplier]

External factors also influenced the implementation process. For example, a new agreement between the European Union and the United States eased approval procedures for certain cloud services.

Within the regions, implementation was facilitated by dialogue and mutual understanding among stakeholders, particularly among operational, legal, and IT functions. One participant illustrated this by saying:

> *We started talking with legal: how do we solve this, then? If you say no, can we not do this? What is your thinking?* [Interview with Administrative Manager, Application Board]

National and industry-wide forums for exchange were also mentioned as facilitators. These included seminars and networks where participants could learn how other regions addressed similar challenges. The industry association was likewise seen as a potential platform for dialogue and collaboration between suppliers and regions.

Human and Relational Barriers

Uncertainty and communication difficulties across organizational boundaries negatively affected the implementation. The lack of mutual understanding and dialogue among stakeholders was reflected in the limited involvement of wheelchair users, differing perspectives between IT and operational staff, and challenges in conveying needs to decision-makers. This was illustrated by 1 participant:

> *The process itself has not really been difficult. The problem is getting in touch with all of these key stakeholders...[and] how to phrase the problem [for the Application Board].* [Interview with Technicians, Assistive Technology Center]

The limited user involvement also contributed to unclear communication about the new technology. As 1 participant reflected:

> *Was there an opportunity for the users to participate? No, it was probably mostly us who decided that.* [Interview with Head of Administration, IT Department]

Cultural and Systemic Inertia

Caution and fragmented systems made it difficult to introduce the new digital solution. A risk-averse culture within the public sector was reflected in the region's slow transition toward digitalization and a prevailing tendency to focus on problems rather than possibilities.

> *People on the IT security side more often tend to see this as a problem… Operations see more possibilities.* [Interview with Administrative Manager, Application Board]

System inertia was further evident in the complex and time-consuming relationships among the region, the supplier, and the users. In addition, legal uncertainties made it challenging to obtain approval for digital solutions.

> *[T]here have been uncertainties around GDPR and that whole area, which the service actually meets all the requirements for ... but toward the region, municipality, governing body, it isn't exactly easy to get application solutions approved.* [Interview with Sales Development Manager, Supplier]

XSL•FO
RenderX

At the national level, fragmentation and the absence of a systemic perspective resulted in regions procuring different solutions, which required suppliers to adapt to varying local conditions. At the organizational level, the process was perceived to have been easier if there had been a joint working group within the Assistive Technology Center.

A Sociotechnical Perspective on the Implementation Process

The results concerning events and stakeholders in the implementation process illustrate the complex interplay of sociotechnical factors that affect IoT implementation processes in health care. The range of factors also highlights the importance of simultaneously addressing both the technical and social factors. From a sociotechnical systems perspective, the 5 key dimensions identified by Kronlid et al [9]: laws and regulations, organizational support, user focus, resources, and infrastructure, emerged as critical to both enabling and constraining the process of implementing IoT-equipped powered wheelchairs. The results underscore the importance of aligning several aspects of those dimensions. That is, it is not enough to address a single factor; only when multiple factors are tackled in parallel, and the solutions are viewed in relation to one another, does a systemic perspective emerge, making implementation possible.

Rather than being driven by isolated technical decisions, the implementation process was shaped by interdependent cultural, organizational, and infrastructural dynamics. The case reveals how implementation was made possible through negotiation among stakeholders embedded in complex structures—where legal ambiguity, fragmented governance, and uneven digital readiness interact with leadership, operational routines, and relational trust. Progress depended not only on resolving discrete challenges but also on aligning multiple layers of the system, from regulatory compliance and organizational capacity to infrastructural adaptability and user legitimacy. The uneven involvement of end users and the tension between operational pragmatism and systemic inertia further illustrate how sociotechnical change is contingent on the ability to coordinate across domains, bridge institutional silos, and cultivate shared meaning around technological value.

The 3 meta-themes can be understood in relation to the 5 dimensions of the sociotechnical framework proposed by Kronlid et al [9] and Palm et al [21]. Meta-theme 1 primarily reflects the dimensions of laws and regulations and organizational support, illustrating how regulatory requirements and organizational structures jointly enabled and constrained implementation. Meta-theme 2 mainly relates to infrastructure and resources, highlighting how technical capabilities depended on organizational competencies, support functions, and available resources. Meta-theme 3 cuts across all dimensions of the framework by emphasizing how trust, communication, collaboration, and cultural norms influenced how regulatory, organizational, technical, and resource-related conditions were interpreted and enacted in practice.

While the first 2 meta-themes broadly correspond to dimensions identified in the sociotechnical framework, the third meta-theme emerged inductively from the data. Rather than representing a separate system dimension, it may be understood as a cross-cutting mechanism that shapes interactions among existing sociotechnical dimensions during implementation.

Discussion

Summary of Findings

This study investigated the implementation of IoT-equipped powered wheelchairs by a Swedish public assistive technology provider from a sociotechnical perspective. The findings show that implementation is not a linear or purely technical process but a negotiated and iterative alignment process in which organizational structures, technological capabilities, and relational dynamics co-shaped what was possible in practice. The 3 meta-themes capture recurring tensions within and between these dimensions that require active coordination rather than one-time decisions.

Tensions Within System-Level Structures

The first meta-theme underscores the dual role of system-level structures in shaping the implementation process. Formal organizational arrangements, legal frameworks, and decision-making pathways were central to both enabling and impeding the adoption of IoT-equipped powered wheelchairs. Consistent with sociotechnical theory, these structures functioned not merely as background conditions but as active components that interacted with human agency and technological capabilities [11]. Importantly, legal and organizational dimensions did not operate independently. Regulatory uncertainty increased the need for organizational coordination, leadership, and cross-functional dialogue, while stronger organizational support helped transform legal constraints from implementation barriers into manageable design requirements. The interaction between these dimensions, therefore, shaped not only whether implementation progressed but also how the technology was ultimately configured and used.

On the enabling side, committed leadership, managerial anchoring, and structured decision-making processes created legitimacy and momentum for implementation. Supportive leadership enabled collaboration across departments and helped translate operational needs into decisions at higher administrative levels. This finding aligns with previous research emphasizing leadership and organizational support as critical for successful digital health implementation [14].

At the same time, legal and regulatory frameworks, particularly data protection legislation, introduced uncertainty and delays. This has been observed in other studies, where regulatory ambiguity and fragmented governance complicate implementation [9]. Initial interpretations of GDPR-related requirements led to early rejection of the IoT solution by the regional application board. Importantly, however, the findings indicate that these regulatory challenges were not static barriers but evolved. A key development in the process was the shift in how legal expertise was engaged. In the early phase, legal stakeholders were approached primarily as gatekeepers tasked with answering whether the solution was permissible. In this

XSL•FO
RenderX

configuration, a single negative assessment effectively halted progress.

Later in the process, this interactional pattern changed. Legal stakeholders were invited into dialogue and encouraged to engage in problem-solving discussions focused on how the technology could be implemented in a legally compliant manner. This shift allowed regulatory requirements to be translated into concrete technical and organizational adaptations, such as relocating data servers, disabling GPS functionality, and developing alternative authentication solutions. From a sociotechnical perspective, this illustrates how legal frameworks constrain or enable implementation depending on how they are operationalized within organizational processes. Beyond legal compliance, the findings highlight the importance of robust patient data protection and cybersecurity measures when implementing IoT-enabled health care technologies. Connected assistive products generate and transmit data that may reveal sensitive information about users' health status, location, and daily activities, making secure data storage, access control, authentication, and continuous risk assessment essential. The adaptations made in this case, including the relocation of data servers, restrictions on GPS data access, and strengthened authentication procedures, illustrate how compliance with health data regulations and cybersecurity requirements can shape implementation decisions and influence stakeholder trust.

Organizational and structural barriers further amplified these tensions. The lack of standardized implementation processes and bureaucratic complexity made progress dependent on individual initiative rather than systemic support. This misalignment between organizational structures and operational practices highlights how system-level arrangements can undermine the realization of technological benefits if they are not adapted to support innovation.

Tensions Between Technological Potential and Its Realization

The second meta-theme captures the tension between the recognized potential of IoT technology and the practical challenges of integrating it into everyday operations. Stakeholders perceived clear benefits, including proactive maintenance, improved troubleshooting, reduced travel, and increased safety for wheelchair users. These anticipated gains are consistent with the broader literature on IoT in health care [7].

Realizing these benefits, however, depended on several enabling conditions. Technical maturity, professional competence, and the ability to adapt systems pragmatically were crucial. Allowing technicians to function as local system administrators and selectively disabling certain functionalities exemplify how adaptations helped bridge the gap between system requirements and operational needs. These findings reinforce sociotechnical research, showing that successful implementation often relies on local workarounds and adaptive practices rather than strict adherence to predefined technical models [15].

At the same time, technical and practical barriers persisted. Infrastructure constraints, limited system integration, and uneven levels of digital maturity among organizations and users delayed implementation and limited functionality. These challenges illustrate that technological readiness alone is insufficient; infrastructural compatibility, organizational capacity, and user capabilities are equally important. This is in line with previous studies, which indicate that technological readiness alone is insufficient and that infrastructural compatibility and user capabilities are equally important [16].

Perceived value played a critical role in sustaining engagement with the technology. User-centered benefits, such as faster repairs and increased perceived safety, along with operational efficiencies, contributed to acceptance and commitment. This highlights the importance of aligning technological solutions with the values and needs of both users and service providers.

Similarly, infrastructure and resources were closely intertwined. Technical infrastructure alone was insufficient to support implementation; its successful use depended on staff competence, available time, managerial support, and access to appropriate support structures. Resource limitations, therefore, influenced how infrastructural capabilities could be used in practice.

Tensions in Relational and Cultural Dynamics

The third meta-theme emphasizes the role of relational and cultural dynamics in shaping the implementation process. Successful implementation depended on effective collaboration and communication among multiple stakeholders, including operational staff, legal experts, IT functions, and suppliers. Dialogue and mutual understanding were essential for addressing uncertainty and reconciling differing perspectives on risk and opportunity. Collaboration with suppliers and participation in national or industry-wide forums facilitated knowledge exchange and reduced isolation, supporting findings from prior research that emphasize the role of networks in innovation diffusion [11].

A particularly important relational dynamic concerned the interaction between operational units and legal functions. When legal stakeholders were positioned primarily as external reviewers, implementation stalled. When they became collaborative partners, implementation progressed. This transition reflects a broader cultural shift from a risk-averse, compliance-focused mindset toward a more solution-oriented approach, where legal expertise was mobilized to enable innovation within regulatory boundaries rather than to block it [27].

Despite these advances, relational barriers remained. Limited user involvement, challenges in communicating operational needs to decision-makers, and differences in professional perspectives hindered a shared understanding [28]. This finding resonates with research describing a risk-averse culture in public-sector organizations, where concerns about accountability, regulatory compliance, information security, and potential negative consequences can lead decision-makers to prioritize caution over experimentation [29]. In our case, these dynamics were reflected in lengthy approval processes related to GDPR compliance, differing risk perceptions between operational and IT/security actors, and the slow pace of digital transformation within the region. This caution should not be interpreted simply as resistance to innovation; rather, it can be

XSL•FO
RenderX

understood as an attempt to balance technological opportunities with responsibilities related to safety, privacy, legal compliance, and public accountability. This suggests that the successful implementation of IoT-equipped assistive products depends not only on demonstrating technical value but also on establishing organizational legitimacy and trust across stakeholder groups.

From a sociotechnical perspective, these findings illustrate that trust, dialogue, and shared sense-making are not peripheral but central to technological implementation. The ability to move from gatekeeping to cocreation depended on leadership support, relational trust, and organizational willingness to tolerate uncertainty during the implementation process [9,30]. The findings suggest that overcoming such inertia requires cultural change, supported by leadership, shared learning, and trust-building across and within organizations.

Rather than constituting an independent sociotechnical dimension, the relational and cultural dynamics identified in this study appeared to influence how the other dimensions were enacted. Trust, communication, and shared understanding shaped how stakeholders interpreted regulations, mobilized organizational support, used resources, and engaged with technological infrastructure.

Implications From a Sociotechnical Perspective

Taken together, the findings underscore that the implementation of IoT-equipped powered wheelchairs constitutes a sociotechnical transformation rather than a discrete technological intervention. Progress depended on aligning legal, organizational, technical, and relational factors simultaneously. Legal stakeholders emerged not merely as part of the institutional context but as active participants within the sociotechnical system, and their role and positioning significantly influenced implementation outcomes [9].

The study contributes to the limited empirical literature on IoT implementation in the assistive technology sector by demonstrating how facilitators and barriers are interdependent and evolve over time. The findings suggest that future implementations should adopt holistic strategies that integrate regulatory interpretation, organizational capacity-building, technological adaptation, and relational coordination. Such approaches may help organizations navigate tensions between innovation and accountability while enabling sustainable adoption of IoT-enabled assistive products.

These findings suggest that implementation capacity in public assistive technology services depends on how governance and regulatory requirements are translated into everyday work, how technical adaptations are negotiated, and how collaboration and trust are built across functions and levels.

Policy Implications

The findings suggest 3 priority areas for policy and practice. First, uncertainty regarding data protection requirements, responsibility distribution, and regulatory compliance was a major implementation barrier. Policymakers and health care organizations should therefore develop clearer guidance regarding data governance, responsibility allocation, and GDPR compliance for IoT-enabled assistive products. Second, successful implementation depended on collaboration between operational staff, legal experts, and IT functions. Health care organizations should establish formal cross-functional implementation teams early in the implementation process to support joint problem-solving and reduce delays caused by fragmented decision-making. Third, regional variation and the lack of standardized approaches created additional complexity for both health care organizations and suppliers. Greater coordination across regions, including shared implementation procedures and common assessment criteria for connected assistive products, could reduce duplication of effort and facilitate wider adoption.

Limitations

The study has some limitations that should be considered when interpreting the findings. First, the findings rely on participants' retrospective accounts, which may have been influenced by recall bias or social desirability. To mitigate this risk, the analysis was grounded in verbatim interview transcripts, and illustrative quotations were used to support the identified themes.

Second, the study focused on stakeholders directly involved in the implementation process, from initiation to the deployment decision, and did not include wheelchair users, frontline clinicians such as occupational therapists and physiotherapists, or policymakers beyond those participating in the study. These groups may hold perspectives that differ from those reported here, particularly regarding user experience, clinical integration, and broader policy considerations. Therefore, the findings should be interpreted as an account of the implementation process from the perspective of actors involved in its planning, governance, and execution, rather than as a comprehensive representation of all stakeholders affected by the technology.

Third, as the study was conducted within a single Assistive Technology Center in Sweden, direct transferability to other health care systems or countries may be limited, particularly where organizational structures, regulatory frameworks, reimbursement systems, technical infrastructures, or roles in assistive technology provision differ. However, some identified implementation factors may be relevant to similar contexts involving connected assistive products and digitally enabled services. To support assessments of transferability, descriptions of the study context, participants, and methods have been provided.

Conclusions

This study contributes to the emerging literature on the implementation of IoT-enabled assistive products by demonstrating how implementation unfolds as a sociotechnical process shaped by interactions among regulatory, organizational, technical, and relational factors. Rather than acting independently, these factors influenced one another throughout the implementation process, highlighting the importance of understanding technology adoption as a process of system alignment rather than technology deployment alone.

A key contribution of the study is the identification of relational and cultural dynamics, including trust, communication, and cross-functional collaboration, as mechanisms that shape how existing sociotechnical dimensions are interpreted and enacted

XSL•FO
RenderX

in practice. This suggests that successful implementation depends not only on technological readiness and regulatory compliance but also on the ability of stakeholders to negotiate shared understandings and coordinate action across organizational boundaries.

Although the study was conducted in a single Swedish assistive technology service context, the findings provide transferable lessons for other public organizations implementing connected assistive products and digital health solutions. In particular, the study highlights the importance of early engagement between operational, legal, and IT functions, clear approaches to data governance, and organizational structures that support collaboration across professional and institutional boundaries.

Acknowledgments

The authors are grateful to the interviewees. Generative AI (ChatGPT, GPT-5) was used as a supportive tool during thematic analysis and to assist with the translation of quotes, the summarization of meta-themes, and the language review of selected parts of the paper. All analytical decisions remained the responsibility of the authors.

Funding

This work was supported by Vinnova, Sweden (grant 2021 - 04581).

Data Availability

Pseudonymized data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: JB, ME, AB, KP
Data curation: JB
Formal analysis: JB, ME, AB, KP
Funding acquisition: JB, ME, AB, KP (lead)
Investigation: JB, ME, KP
Methodology: JB, ME, AB, KP
Project administration: KP
Visualization: JB
Writing – original draft: JB (lead), ME, AB, KP
Writing – review and editing: JB, ME, AB, KP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 17 KB - rehab_v13i1e92414_app1.docx]

Multimedia Appendix 2
An example of the analysis process.
[DOCX File, 16 KB - rehab_v13i1e92414_app2.docx]

Multimedia Appendix 3
Details of the use of ChatGPT in the analysis.
[DOCX File, 22 KB - rehab_v13i1e92414_app3.docx]

References

1. Hogaboom NS, Worobey LA, Houlihan BV, Heinemann AW, Boninger ML. Wheelchair breakdowns are associated with pain, pressure injuries, rehospitalization, and self-perceived health in full-time wheelchair users with spinal cord injury. Arch Phys Med Rehabil 2018 Oct;99(10):1949-1956. [doi: 10.1016/j.apmr.2018.04.002] [Medline: 29698640]
2. McClure LA, Boninger ML, Oyster ML, et al. Wheelchair repairs, breakdown, and adverse consequences for people with traumatic spinal cord injury. Arch Phys Med Rehabil 2009 Dec;90(12):2034-2038. [doi: 10.1016/j.apmr.2009.07.020] [Medline: 19969165]

XSL•FO
RenderX

3. Toro ML, Worobey L, Boninger ML, Cooper RA, Pearlman J. Type and frequency of reported wheelchair repairs and related adverse consequences among people with spinal cord injury. Arch Phys Med Rehabil 2016 Oct;97(10):1753-1760. [doi: 10.1016/j.apmr.2016.03.032] [Medline: 27153763]
4. Worobey LA, Heinemann AW, Anderson KD, et al. Factors influencing incidence of wheelchair repairs and consequences among individuals with spinal cord injury. Arch Phys Med Rehabil 2022 Apr;103(4):779-789. [doi: 10.1016/j.apmr.2021.01.094] [Medline: 33845000]
5. Global report on assistive technology. : World Health Organization; 2022 URL: https://iris.who.int/server/api/core/bitstreams/17a1cbf6-be26-466c-9fca-ab80eeb2d192/content [accessed 2026-08-08]
6. Minerva R, Biru A, Rotondi D. Towards a definition of the internet of things (IoT). : IEEE Internet Initiative; 2015 URL: https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf [accessed 2026-08-08]
7. Kelly JT, Campbell KL, Gong E, Scuffham P. The internet of things: impact and implications for health care delivery. J Med Internet Res 2020 Nov 10;22(11):e20135. [doi: 10.2196/20135] [Medline: 33170132]
8. Korzun DG. Internet of things meets mobile health systems in smart spaces: an overview. In: Bhatt C, Dey N, Ashour A, editors. Internet of Things and Big Data Technologies for Next Generation Healthcare: Springer; 2017:111-129. [doi: 10.1007/978-3-319-49736-5_6]
9. Kronlid C, Brantnell A, Elf M, Borg J, Palm K. Sociotechnical analysis of factors influencing IoT adoption in healthcare: a systematic review. Technol Soc 2024 Sep;78:102675. [doi: 10.1016/j.techsoc.2024.102675]
10. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4(1):50. [doi: 10.1186/1748-5908-4-50]
11. de Bruijn H, Herder PM. System and actor perspectives on sociotechnical systems. IEEE Trans Syst, Man, Cybern A 2009;39(5):981-992. [doi: 10.1109/TSMCA.2009.2025452]
12. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
13. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. Sociology 2009 Jun;43(3):535-554. [doi: 10.1177/0038038509103208]
14. Sittig DF, Singh H. A new socio-technical model for studying health information technology in complex adaptive healthcare systems. In: Patel VL, Kannampallil TG, Kaufman DR, editors. Cognitive Informatics for Biomedicine: Human Computer Interaction in Healthcare: Springer; 2015:59-80. [doi: 10.1007/978-3-319-17272-9_4]
15. Carayon P, Salwei ME. Moving toward a sociotechnical systems approach to continuous health information technology design: the path forward for improving electronic health record usability and reducing clinician burnout. J Am Med Inform Assoc 2021 Apr 23;28(5):1026-1028. [doi: 10.1093/jamia/ocab002] [Medline: 33537756]
16. Salwei ME, Carayon P. A sociotechnical systems framework for the application of artificial intelligence in health care delivery. J Cogn Eng Decis Mak 2022 Dec;16(4):194-206. [doi: 10.1177/15553434221097357] [Medline: 36704421]
17. Priya A. Case study methodology of qualitative research: key attributes and navigating the conundrums in its application. Sociol Bull 2021 Jan;70(1):94-110. [doi: 10.1177/0038022920970318]
18. Yin RK. Case Study Research: Design and Methods, 5th edition: Sage Publications; 2013.
19. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for Reporting Qualitative Research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
20. Davis MC, Challenger R, Jayewardene DNW, Clegg CW. Advancing socio-technical systems thinking: a call for bravery. Appl Ergon 2014 Mar;45(2):171-180. [doi: 10.1016/j.apergo.2013.02.009] [Medline: 23664481]
21. Palm K, Kronlid C, Elf M, Brantnell A. Key factors shaping successful implementation of the internet of things (IoT) in health care: qualitative study. JMIR Hum Factors 2025 Jun 27;12:e71546. [doi: 10.2196/71546] [Medline: 40704745]
22. TurboScribe. URL: https://turboscribe.ai/ [accessed 2025-11-14]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
25. Attride-Stirling J. Thematic networks: an analytic tool for qualitative research. Qual Res 2001 Dec;1(3):385-405. [doi: 10.1177/146879410100100307]
26. Lee VV, van der Lubbe SCC, Goh LH, Valderas JM. Harnessing ChatGPT for thematic analysis: are we ready? J Med Internet Res 2024 May 31;26:e54974. [doi: 10.2196/54974] [Medline: 38819896]
27. Lepore D, Dolui K, Tomashchuk O, et al. Interdisciplinary research unlocking innovative solutions in healthcare. Technovation 2023 Feb;120:102511. [doi: 10.1016/j.technovation.2022.102511]
28. Tang T, Lim ME, Mansfield E, McLachlan A, Quan SD. Clinician user involvement in the real world: designing an electronic tool to improve interprofessional communication and collaboration in a hospital setting. Int J Med Inform 2018 Feb;110:90-97. [doi: 10.1016/j.ijmedinf.2017.11.011] [Medline: 29331258]

XSL•FO
RenderX

29. Cinar E, Trott P, Simms C. A systematic review of barriers to public sector innovation process. Public Manag Rev 2019 Feb;21(2):264-290. [doi: 10.1080/14719037.2018.1473477]
30. Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: scoping review. JMIR Hum Factors 2024 Aug 29;11:e48633. [doi: 10.2196/48633] [Medline: 39207831]

Abbreviations

GDPR : General Data Protection Regulation
IoT: Internet of Things

Edited by JJ Zhang; submitted 29.Jan.2026; peer-reviewed by HC Pham, Y Han; revised version received 12.Jun.2026; accepted 27.Jul.2026; published 28.Aug.2026.

Please cite as:
Borg J, Elf M, Brantnell A, Palm K
Stakeholder Perspectives on Facilitators of and Barriers to the Implementation of Internet of Things–Equipped Powered Wheelchairs in Sweden: Qualitative Case Study
JMIR Rehabil Assist Technol 2026;13:e92414
URL: https://rehab.jmir.org/2026/1/e92414
doi:10.2196/92414

© Johan Borg, Marie Elf, Anders Brantnell, Klas Palm. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 28.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Participant-Reported Reductions in Falls and Fall-Related Injuries Associated With Variable-Friction Footwear: Retrospective Pre-Post Survey Study

Sara Nataletti[1,2], PhD; Jennifer Bartloff[1,2], PhD; Sara Prokup[1,3], DPT; Joseph Harris[1], PT; Anushka Larson[4], BS; Tyler Susko[4,5], PhD; Arun Jayaraman[1,2,3,6], PhD

[1]Max Näder Center for Rehabilitation Technologies and Outcomes Research, Shirley Ryan AbilityLab, 355 E. Erie St, Chicago, IL, United States

[2]Department of Physical Medicine & Rehabilitation, Feinberg School of Medicine, Northwestern University, 710 N. Lake Shore Dr, Chicago, IL, United States

[3]Technology and Innovation Hub, Shirley Ryan AbilityLab, Chicago, IL, United States

[4]Cadense, Goleta, CA, United States

[5]Department of Mechanical Engineering, University of California, Santa Barbara, Santa Barbara, CA, United States

[6]Department of Physical Therapy & Human Movement Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author:
Arun Jayaraman, PhD
Max Näder Center for Rehabilitation Technologies and Outcomes Research, Shirley Ryan AbilityLab, 355 E. Erie St, Chicago, IL, United States

Abstract

Background: Falls represent a critical public health burden in older adults and individuals with neurological disabilities, but evidence for passive footwear-based fall prevention strategies remains limited.

Objective: This study aimed to evaluate changes in fall frequency and fall-related outcomes following adoption of variable-friction (VF) footwear.

Methods: This retrospective pre-post observational study used self-reported survey data collected from community-dwelling adults aged 18 years or older who had worn VF footwear for 3 months or longer. Participants were recruited from a convenience sample of 14,570 invited customers who purchased the device through an electronic survey platform. Outcomes included self-reported fall frequency, injurious falls, fall mechanisms, fall-related health care use (including medical visits, urgent care, and hospitalization), and assistive device use, assessed before and during VF footwear use.

Results: A total of 706 participants completed the survey, yielding 705 participants with usable data for analysis (4.8% response rate). Mean monthly participant-reported falls decreased from 2.4 (SD 4.3) before VF use to 0.3 (SD 1.9) during VF use (87.1% reduction; Wilcoxon signed-rank test, $P<.001$). The proportion reporting zero falls increased from 22.7% to 84.4%. Among participants with at least 1 fall before VF use (n=545), 95.4% (n=520) reported a reduction, including 80.7% (n=440) reporting zero falls during VF use. Injurious falls decreased by 93.2% ($P<.001$). Participants requiring medical attention decreased from 192 (27.2%) to 12 (1.7%; 93.8% reduction; $P<.001$), and hospital or urgent care visits decreased from 131 (18.6%) to 10 (1.4%; 92.4% reduction; $P<.001$). Trip-related falls decreased from 52.3% to 31.5% of reported mechanisms among participants who reported a fall mechanism in both periods ($P<.001$). Assistive device use decreased from 70.8% to 16.6% ($P<.001$). Mean annualized insurance-related fall costs decreased from US $5946.2 to US $45.7 per person-year (99.2% reduction) across the full cohort.

Conclusions: VF footwear use was associated with substantial reductions in self-reported falls, injurious falls, and fall-related health care use. Reductions in trip-related falls are consistent with the proposed biomechanical mechanism of the device. Given the low survey response rate and potential for self-selection bias, these findings should be considered preliminary. Prospective controlled studies are needed to confirm these findings.

(JMIR Rehabil Assist Technol 2026;13:e104177) doi:10.2196/104177

KEYWORDS

falls; fall prevention; rehabilitation; mobility impairment; assistive devices

XSL•FO
RenderX

Introduction

Falls are among the most consequential and costly public health events affecting older adults [1-3]. Approximately 25% to 33% of adults aged 65 years and older experience at least one fall annually, with rates approaching 50% among those over 75 [4]. More than one-third of these falls result in injuries, including fractures and traumatic brain injuries that can precipitate functional decline, institutionalization, and increased mortality [5-8]. Health care spending attributable to nonfatal falls among older adults in the United States grew from approximately US $49.5 billion in 2015 to US $80 billion in 2020, with costs projected to exceed US $100 billion by 2030 [9,10]. These figures reflect direct medical expenditures alone and do not capture the broader costs of lost independence, caregiver burden, and reduced quality of life.

Beyond physical injury, falls contribute to fear of falling [11,12], affecting an estimated 21% to 85% of older adults [13,14], which itself doubles future fall risk and creates a self-reinforcing cycle of declining activity and increasing vulnerability [15-17]. This heightened vulnerability also reflects age-related declines in postural stability. Progressive deterioration in balance control, particularly under conditions of reduced sensory input, impairs the ability to maintain stability and recover from balance disturbances, further increasing susceptibility to falls in older adults [18]. This burden is further compounded in individuals with neurological conditions such as multiple sclerosis, stroke, and Parkinson disease, where disability-related mobility impairments add to the already elevated risk associated with aging [19-23]. In a national sample of adults with multiple sclerosis aged 55 to 94 years, 64% reported at least 2 falls per year, and more than half experienced injurious falls [24].

Current evidence-based strategies for fall prevention include multicomponent exercise and balance training programs, which can reduce fall rates by 15% to 30% [25-30]. However, their population-level impact remains constrained by poor long-term adherence, with fewer than 1 in 4 participants continuing prescribed programs at 12 months, and supervised in-person programs face persistent barriers of cost, travel, and accessibility [31-34]. Footwear represents a complementary strategy of a fundamentally different kind: worn daily as part of normal routine, it continuously modifies the mechanical interface between the person and their environment without requiring any behavioral commitment. Prior work has established that design features including low heel height, slip-resistant outsoles, proper fit, and secure fastening are associated with improved stability and reduced fall risk [35-37]. Beyond these structural features, perceived footwear comfort and mediolateral support have also been associated with fear of falling and perceived stability in older adults, suggesting that both the physical design and user perception of footwear may influence factors related to fall risk [38]. Rocker-sole and balance-training shoes have also been investigated to enhance neuromuscular control, though evidence remains mixed [39,40].

Among the most common precipitating causes of falls are forward trips, occurring when toe clearance fails during the swing phase, producing abrupt foot deceleration while the center of mass continues forward [41,42]. Variable-friction (VF) footwear has been proposed as a novel strategy to mitigate this mechanism by allowing controlled sliding of the toe during a failed clearance event, thereby reducing the abrupt stopping force that contributes to forward falls. Cadense Original shoes (Cadense Inc) incorporate this system, offering a targeted, mechanism-specific, and adherence-free approach to fall prevention during normal walking. However, evidence evaluating friction-modifying footwear as a real-world fall prevention strategy is limited.

Evaluating such interventions at scale requires reliable yet feasible methods of falls ascertainment. While prospective falls calendars are considered the gold standard, their logistical burden limits applicability in large community-based studies [43,44]. Self-reported surveys offer a practical and well-validated alternative, with prior studies demonstrating accurate recall of falls over 12 months, particularly for injurious falls, and good agreement with prospective methods [44-46].

This study examined changes in fall frequency and fall-related consequences following the adoption of VF footwear, assessed via self-reported survey. Outcomes included fall frequency, injurious falls, fall mechanisms, and fall-related health care use. We hypothesized that participants wearing VF shoes would experience a reduction in fall prevalence, driven primarily by fewer forward trip-related events.

Methods

Study Design

This was a retrospective pre-post observational study using self-reported survey data to evaluate fall frequency and fall-related outcomes before and during the use of a VF footwear device. Exposure and outcomes were assessed through a single survey administration, with participants retrospectively reporting outcomes for 2 periods: before and during VF footwear use. The study was conducted anonymously via an electronic survey platform (REDCap). This study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1).

Participants

VF footwear was not provided or administered as part of this study. Rather, participants were existing Cadense customers who had independently purchased and worn the Cadense Original VF shoe for at least 3 months before survey completion. Cadense identified customers who met this minimum use criterion and provided the corresponding contact information to the research team at Shirley Ryan AbilityLab, thereby establishing the pool of potentially eligible participants. All recruitment invitations were subsequently distributed directly by the Shirley Ryan AbilityLab team. Eligible participants were adults aged 18 years or older and able to complete the survey in English. Individuals with severe cognitive or communication deficits or who were prisoners were excluded. Eligibility criteria were described in the recruitment email, and eligibility was self-evaluated by respondents prior to survey initiation. The survey link was distributed via email to all customers who met

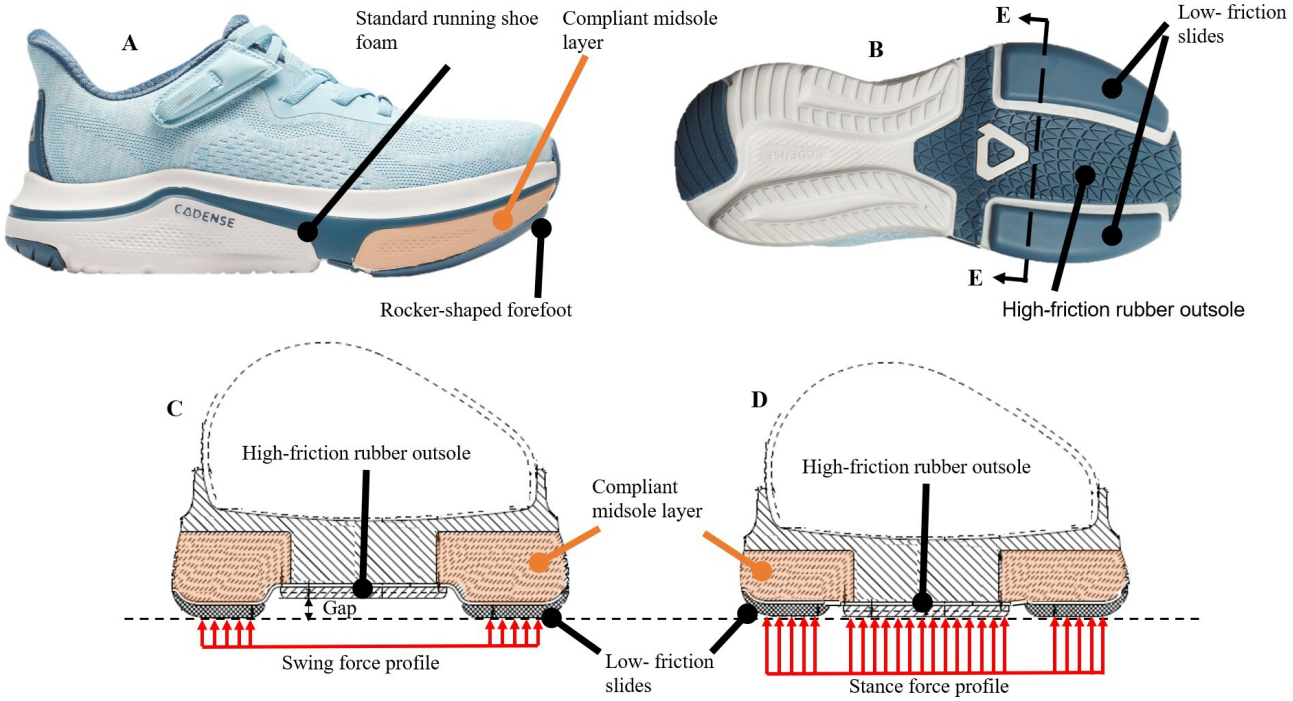
XSL•FO
RenderX

inclusion criteria. A total of 14,570 eligible customers were invited to participate in the survey. Recipients were given a response window of up to 4 weeks to complete the survey following receipt of the recruitment email.

VF Footwear

The Cadense Original VF shoe (Figure 1) is a US Food and Drug Administration–registered Class I medical device. The device operates passively and alters frictional properties of the outsole based on loading conditions during gait. During the swing phase or low ground reaction force conditions, the shoe provides reduced friction at the forefoot region through low-friction polymer slides positioned slightly proud of the primary high-friction outsole (see the gap in Figure 1C), allowing controlled sliding during incidental toe contact. During stance, increased loading compresses a compliant midsole layer set beneath the low-friction slides, retracting the low-friction slides and bringing the high-friction rubber outsole into contact with the walking surface (see Figure 1D) to support stability during weight acceptance and push-off. The design incorporates a rocker-shaped forefoot intended to reduce toe scuff during gait much like a ski tip sliding over small imperfections in the walking surface. The transition between low- and high-friction states occurs automatically based on normal force during gait and does not require active components. The device contains no electronics or batteries and is intended to mitigate forward trip-related falls by modifying toe-ground interaction mechanics. The device also provides auditory feedback during toe contact events over hard surfaces, which may increase user awareness of foot clearance during walking.

Figure 1. Variable-friction (VF) footwear design and functional concept. (A) Side view of the VF shoe showing the rocker-shaped forefoot intended to facilitate toe clearance during gait, the midsole made from standard running shoe foam, and the compliant midsole layer, which is critical to the functional concept. (B) VF outsole design showing the high-friction rubber outsole positioned between low-friction slides. (C) Section view through E-E illustrating the low-friction state seen during inadvertent swing-phase toe contact, where a small gap ensures contact force with only the low-friction slides, reducing orthogonal frictional forces that cause trips. (D) Section view through E-E illustrating the high-friction state during weight bearing, where compression of the compliant midsole layer extends the force profile to the high-friction rubber outsole, promoting high-grip stance and push-off.



Data Collection

Data were collected using a custom electronic survey administered through REDCap, a secure web-based data management platform hosted by Shirley Ryan AbilityLab. This was a closed survey, accessible only via a unique link distributed directly to identified eligible customers; it was not publicly posted or advertised. The survey instrument was reviewed internally by the research team for usability and technical functionality prior to distribution. The survey was distributed via a unique REDCap link included in the recruitment email. The survey link led respondents through 2 screens: an initial screen presenting the consent form, with an option to download a PDF copy, followed by a second screen presenting the full survey (Multimedia Appendix 1). The survey comprised 10 primary items, several with conditionally displayed follow-up questions based on branching logic; items were presented in a fixed order and were not randomized, consistent with this branching logic. Survey responses were considered complete only after respondents reached the final page and submitted the survey. Partial responses were not retained, and respondents who did not complete submission needed to restart the survey. Once a survey had been completed and submitted, REDCap's built-in mechanism prevented the same respondent from submitting the survey more than once. The survey (Multimedia Appendix 1) was designed to characterize fall frequency and fall-related outcomes before and after use of the Cadense VF shoe. The instrument collected demographic information and clinical characteristics, including age, body weight, underlying medical condition associated with fall risk, and assistive device (AD) use, as well as retrospective information on reported falls during the year prior to initiating VF shoe use and during

XSL•FO
RenderX

sustained use of the VF shoe (≥3 mo). Specifically, participants reported fall frequency, fall mechanism (eg, trip-related vs balance-related), occurrence of injurious falls, need for medical attention, health care use setting, and estimated out-of-pocket and insurance-related costs associated with fall events. Cost estimates were self-reported by participants as open numeric responses and were not verified against medical records or insurance claims. All questions were structured in closed-ended categorical or numeric formats to facilitate completion and minimize missing data. Skip logic was incorporated to display follow-up questions only when applicable (eg, injury-related items presented only to respondents reporting a fall). The survey required approximately 10 minutes to complete and was completed anonymously. Identifiable information provided for optional raffle entry was collected separately and was not linked to survey responses.

Outcome Measures

The primary outcome of this study was the change in self-reported fall frequency before VF shoe use and during sustained VF shoe use (≥3 mo). The pre-VF reporting period covered the 12 months preceding VF shoe use, whereas the VF-use reporting period reflected each participant's duration of VF shoe use and therefore varied across participants, with a minimum of 3 months. To account for these differing reporting periods, fall frequency was assessed as participants' self-reported average number of falls per month during each period. Fall frequency was analyzed both as a continuous variable and as categorical outcomes, including the proportion of participants reporting at least one fall and those reporting zero falls.

Secondary outcomes included injurious falls and other fall-related consequences and characteristics. Given the lower frequency of injurious falls relative to total falls, the number of injurious falls and their consequences were reported over the respective reporting periods, rather than as a monthly rate. Fall-related health care use included whether participants required medical attention following a fall and whether care occurred in a hospital or urgent care setting. Economic outcomes included self-reported out-of-pocket and insurance-related costs associated with fall events.

Fall mechanisms were also collected and categorized as trip-related or nontrip-related events to examine potential changes in fall etiology. AD use was also evaluated as a mobility-related outcome. Participants reported use of ADs before and during VF shoe use, including canes, walkers, ankle-foot orthoses, and neuromuscular stimulation devices. Changes in overall AD use and device-specific use were examined.

Additional descriptive outcomes included the proportion of participants reporting fewer falls, no change, or increased falls during VF shoe use.

Statistical Analysis

Descriptive statistics were used to characterize the study sample, including demographic variables and medical conditions associated with fall risk. Continuous outcomes are presented as both mean (SD) and median (IQR). Because fall frequency, injurious falls, and cost variables demonstrated nonnormal distributions, within-participant comparisons of continuous outcomes between the pre-VF and VF periods were performed using Wilcoxon signed-rank tests. Paired categorical outcomes were compared using McNemar tests, as appropriate. Effect sizes were reported as the matched-pairs rank-biserial correlation and mean difference with 95% CI (t-distribution-based) for Wilcoxon-tested outcomes, and as risk differences with 95% CI (Wald method) for McNemar-tested outcomes.

Changes in paired categorical outcomes, including fall occurrence, injurious falls, health care use, fall mechanisms, and AD use, were evaluated using McNemar tests or related paired chi-square procedures, as appropriate. A consistent paired analytic framework was applied across outcome domains to reflect the repeated-measures structure of the data.

Because the pre-VF reporting period was fixed at 12 months, whereas the VF-use reporting period varied across participants, differences in observation periods were addressed according to the method of outcome collection. Fall frequency, the primary outcome, was reported as an average monthly rate for both periods and therefore required no adjustment for exposure duration. In contrast, injurious falls and fall-related costs (out-of-pocket and insurance) were reported as totals over the respective reporting periods. To account for the variable VF-use observation window, during-VF totals were annualized for each participant by dividing the reported value by their individual VF-use duration in years before comparison with the corresponding 12-month pre-VF value. If a respondent indicated no injurious falls during a period, the out-of-pocket and insurance cost values were assigned a value of zero, consistent with the survey skip logic (Multimedia Appendix 1). Otherwise, missing cost values were retained as missing and excluded from paired analyses.

Health care use outcomes (need for medical attention, hospital or urgent care use) and the proportion of participants reporting at least one injurious fall were based on whether each event occurred during the respective reporting period. Unlike the corresponding continuous outcomes, these measures cannot be meaningfully annualized; these were therefore compared as raw proportions over the unequal exposure windows described above.

Fall mechanisms were analyzed among participants who reported mechanism information in both the pre-VF and VF-use periods. Paired changes in the occurrence of trip-related falls were evaluated using a McNemar-type test based on discordant pairs.

All analyses were performed using Python (version 3.11). Statistical significance was set at α=.05 (2-tailed).

Ethical Considerations

Ethics approval for this study was granted by the Institutional Review Board of Northwestern University (STU00224455). Participants provided passive consent by reviewing an electronic consent form and proceeding to the survey, under an institutional review board–approved waiver of documentation of informed consent given the anonymous, minimal-risk nature of the study. The study procedures adhered to the principles of the

XSL•FO
RenderX

Declaration of Helsinki and institutional ethical guidelines. All data collected were deidentified prior to analysis to protect participant privacy and confidentiality. At the conclusion of the study, all participants who completed the survey were entered into a random drawing to receive 1 of 5 free pairs of Cadense VF shoes (retail value US $199 each).

Results

Participant Characteristics

A total of 14,570 eligible customers were invited to participate. Of these, 706 (4.8%) completed the survey. One response was excluded due to a typographical data entry error in fall frequency, resulting in a final analytic sample of 705 participants. Demographic and clinical characteristics are summarized in Table 1.

Table . Participant characteristics (N=705).

Characteristic	Value
Age (y), mean (SD)	59.5 (13.8)
Weight (kg), mean (SD)	81.5 (20.4)
Duration of VF[a] shoe use (mo), mean (SD)	12.3 (8.4)
Shoe style, n (%)	
Men’s	274 (38.9)
Women’s	431 (61.1)
Reported medical condition[b], n (%)	
Multiple sclerosis	391 (55.5)
Peripheral neuropathy	127 (18.0)
Stroke	49 (7.0)
Spinal cord injury	42 (6.0)
Age-related condition	68 (9.6)
Traumatic brain injury	38 (5.4)
Parkinson disease	29 (4.1)
Charcot-Marie-Tooth disease	14 (2.0)
Other condition	131 (18.6)

[a]VF: variable-friction.

[b]Participants were allowed to select multiple conditions; therefore, percentages do not sum to 100%. Overall, 148 (21%) participants reported more than one condition.

Fall Frequency

The median self-reported monthly falls per participant decreased from 1.0 (IQR 0.5 - 3.0) prior to VF shoe use to 0.0 (IQR 0.0 - 0.0) during VF shoe use (Wilcoxon signed-rank P<.001). The corresponding mean decreased from 2.4 (SD 4.3) to 0.3 (SD 1.9), representing an absolute reduction of 2.1 falls per month and an 87.1% relative reduction (Table 2). Overall, 73.8% (520/705) of participants reported a reduction in fall frequency, 24.5% (173/705) reported no change, and 1.7% (12/705) reported an increase. Reductions in participant-reported falls were observed across all major condition groups (see Table 3). The proportion of participants reporting zero falls increased from 22.7% (160/705) prior to VF use to 84.4% (595/705) during VF use. Among participants who reported at least one fall prior to VF shoe use (n=545), 95.4% (n=520) reported a reduction in falls, including 80.7% (n=440) who reported zero falls during VF use, and 1.3% (n=7) reported an increase in falls.

A small subgroup of participants (n=12) reported increased fall frequency during VF shoe use, with median (IQR) monthly falls increasing from 0.7 (IQR 0.0 - 1.2) to 1.5 (IQR 1.0 - 3.0). The corresponding mean increased from 2.1 (SD 4.3) to 3.8 (SD 5.7). Despite this increase, injurious falls did not worsen for most individuals: only 1 participant reported an increase in injurious falls, while 3 reported fewer and 8 remained unchanged. Examination of fall mechanisms showed a shift from primarily trip-related falls prior to VF use (5/9 with available data) to predominantly balance-related falls during VF use (10/12), including both loss of balance while walking and standing. This pattern suggests that increases in falls within this small subgroup were more likely related to underlying balance impairments rather than trip-related mechanisms targeted by the VF shoes.

XSL•FO
RenderX

Table . Outcomes[a].

Outcome	Pre-VF[b] shoe	Post-VF shoe	Change (%)	*P* value	Effect size, *r*	Mean Δ (95% CI)	RD[c] percentage points (95% CI)
Falls/month, mean (SD)/median (IQR)	2.4 (4.3)/1.0 (0.5 - 3.0)	0.3 (1.9)/0.0 (0.0 - 0.0)	−87.1	<.001[d]	−0.97	−2.1 (−2.4 to −1.8)	—[e]
People with ≥1 fall, n (%)	545 (77.3)	110 (15.6)	−79.8	<.001[f]	—	—	−61.7 (−65.4 to −58.0)
People with 0 falls, n (%)	160 (22.7)	595 (84.4)	271.9	<.001[f]	—	—	+61.7 (+58.0 to +65.4)
Injurious falls/year, mean (SD)/median (IQR)	1.8 (4.2)/1.0 (0.0 - 2.0)	0.1 (0.7)/0.0 (0.0 - 0.0)	−93.2	<.001[d]	−0.95	−1.7 (−2.0 to −1.4)	—
People with ≥1 injurious fall, n (%)	381 (54.0)	43 (6.1)	−88.7	<.001[f]	—	—	−47.9 (−51.7 to −44.2)
People requiring medical attention, n (%)	192 (27.2)	12 (1.7)	−93.8	<.001[f]	—	—	−25.5 (−28.8 to −22.3)
Participants requiring hospital or urgent care, n (%)	131 (18.6)	10 (1.4)	−92.4	<.001[f]	—	—	−17.2 (−20.0 to −14.3)
Out-of-pocket fall cost, $/person-year, mean (SD)/median (IQR)	281.2 (2069.8)/0.0 (0.0 - 0.0)	17.2 (266.8)/0.0 (0.0 - 0.0)	−93.9	<.001[d]	−0.91	−264.1 (−417.7 to −110.5)	—
Insurance fall cost, $/person-year, mean (SD)/median (IQR)	5946.2 (94,754.1)/0.0 (0.0 - 0.0)	45.7 (505.0)/0.0 (0.0 - 0.0)	−99.2	<.001[d]	−0.95	−5900.6 (−12,906.5 to 1105.4)	—
People using assistive device, n (%)	499 (70.8)	117 (16.6)	−76.6	<.001[f]	—	—	−54.0 (−57.9 to −50.2)

[a]Continuous outcomes are presented as mean (SD) and median (IQR). Percentages based on the 705 sample size. Changes are calculated from the mean values as % changes [(post−pre)/pre] × 100.

[b]VF: variable-friction.

[c]RD: risk difference.

[d]Wilcoxon signed-rank test; *r*: matched-pairs rank-biserial correlation.

[e]Not applicable.

[f]McNemar test.

Table . Proportion of participants reporting reductions in falls and injurious falls by medical condition.

Condition	Participants (N)	Fewer falls, n (%)	Fewer injurious falls, n (%)
Multiple sclerosis	391	305 (78)	219 (56)
Peripheral neuropathy	127	96 (75.6)	76 (59.8)
Age-related conditions	68	49 (72.1)	32 (47.1)
Parkinson disease	29	18 (62.1)	15 (51.7)
Spinal cord injury	42	28 (66.7)	23 (54.8)
Traumatic brain injury	38	25 (65.8)	16 (42.1)
Stroke	49	32 (65.3)	20 (40.8)

XSL•FO
RenderX

Fall Mechanism

Trip-related events accounted for 67.7% (397/586) of reported fall mechanisms prior to VF shoe use, followed by loss of balance while walking (138/586, 23.5%) and loss of balance while standing (51/586, 8.7%). During VF shoe use, trip-related events accounted for 30.7% (42/137) of reported fall mechanisms, while loss of balance while walking became the most frequently reported mechanism (61/137, 44.5%), followed by loss of balance while standing (34/137, 24.8%). Of the 586 reported fall mechanisms prior to VF use, 130 also were reported a mechanism during VF use. To provide a paired within-participant comparison, we examined fall mechanisms among these 130 participants who had reported a mechanism in both periods. In the period before VF use, trip-related events accounted for 52.3% (68/130) of reported mechanisms in this subset, followed by loss of balance while walking (44/130, 33.8%) and loss of balance while standing (18/130, 13.8%). During VF use, trip-related events decreased to 31.5% (41/130), while loss of balance while walking became the most frequently reported mechanism (55/130, 42.3%), followed by loss of balance while standing (34/130, 26.2%). This paired reduction in trip-related falls was statistically significant (McNemar-type test on discordant pairs, $P<.001$). These findings are consistent with a shift in the predominant mechanism of falls from primarily trip-related events before VF use to primarily balance-related mechanisms during VF use.

Injurious Falls and Medical Attention

The median annual self-reported injurious falls decreased from 1.0 (IQR 0.0 - 2.0) prior to VF shoe use to 0.0 (IQR 0.0 - 0.0) during VF use (Wilcoxon signed-rank test, $P<.001$). The corresponding mean decreased from 1.8 (SD 4.2) to 0.1 (SD 0.7), representing a 93.2% reduction. Consistent with this reduction, 52.6% (371/705) of participants reported fewer injurious falls, 46.5% (328/705) reported no change, and 0.9% (6/705) reported an increase. The number of participants reporting at least one injurious fall per reporting period decreased from 54% (381/705) prior to VF shoe use to 6.1% (43/705) during VF use (McNemar test, $P<.001$). The number of participants requiring medical attention for falls decreased from 27.2% (192/705) prior to VF shoe use to 1.7% (12/705) during VF use, representing a 93.8% reduction (McNemar test, $P<.001$). Similarly, hospital or urgent care visits decreased from 18.6% (131/705) to 1.4% (10/705), corresponding to a 92.4% reduction (McNemar test, $P<.001$).

Fall-Related Costs

Annualized mean out-of-pocket fall-related costs decreased from $281.2 (SD $2069.8) per person-year prior to VF shoe use to $17.2 (SD $266.8) per person-year during VF shoe use, a 93.9% reduction (Wilcoxon signed-rank test, $P<.001$). Median out-of-pocket costs were $0 (IQR $0-$0) in both periods, reflecting that the majority of participants reported no fall-related out-of-pocket cost in either period. Before VF shoe use, 21.8% (154/705) of participants reported a nonzero cost, compared with 1.6% (11/705) during VF shoe use.

Similarly, annualized mean insurance-related fall costs decreased from $5946.2 (SD $94,754.1) per person-year to $45.7 (SD $505.0) per person-year, a 99.2% reduction (Wilcoxon signed-rank test, $P<.001$). As with out-of-pocket costs, median insurance costs were $0 (IQR $0-$0) in both periods. Before VF shoe use, 24.3% (171/705) of participants reported a nonzero insurance cost, versus 1.4% (10/705) during VF shoe use.

AD Use

The proportion of participants reporting use of any AD decreased from 70.8% (499/705) prior to VF shoe use to 16.6% (117/705) during VF use (McNemar test, $P<.001$). Overall, 55.2% (389/705) of participants discontinued AD use while using VF shoes, whereas only 1.0% (7/705) initiated AD use. Similar reductions were observed across individual device categories, including walker use (34.4% to 8.9%), cane use (49.7% to 11.8%), ankle-foot orthosis use (29.2% to 3.0%), and neuromuscular electrical stimulation device use (12.7% to 3.0%), with all changes significant based on McNemar tests for paired proportions (all $P<.001$).

Discussion

Principal Findings

The primary objective of this study was to evaluate the association between the use of VF footwear and participant-reported changes in fall frequency, fall-related injuries, and health care use. These findings should be interpreted in the context of the low survey response rate (706/14,570, 4.8%) and the requirement for participants to have used VF footwear for at least 3 months, as respondents may not be representative of the broader population of VF footwear users. Individuals who perceived greater benefit or were able to continue using the footwear may have been more likely to be represented in the study. Among respondents, substantial and statistically significant reductions were reported across all primary and secondary outcomes following the adoption of VF footwear. During VF footwear use, participants reported an 87.1% reduction in mean monthly fall frequency (2.4 to 0.3 falls/month) and a 93.2% reduction in injurious falls (1.8 to 0.1 falls/year). Notably, 95.4% (520/545) of participants who reported at least one fall in the previous year also reported a reduction in fall frequency during VF shoe use. Although these findings require confirmation in prospective controlled studies, they are particularly relevant given the study population, which comprised individuals with significant neurological and age-related conditions, such as multiple sclerosis, neuropathy, stroke, and Parkinson disease, who are historically at a higher risk for falls [20].

The magnitude of the reported reduction exceeds that typically observed in randomized trials of exercise and multifactorial fall-prevention interventions. A recent meta-analysis of 108 randomized controlled trials involving 23,407 participants reported an average 23% reduction in fall rates following exercise interventions [47]. Similarly, systematic reviews of multifactorial interventions have reported reductions of approximately 23% to 24% [48,49]. Although direct comparison with the present findings is limited by differences in participant populations, baseline fall risk, outcome ascertainment, follow-up duration, and study design, several factors may therefore explain the larger associations observed in the present study. Participants

XSL•FO
RenderX

self-reported neurological and age-related conditions associated with high risk of recurrent falls, potentially allowing greater opportunity for improvement than participants enrolled in many previous fall-prevention trials. Furthermore, VF footwear is designed to address a specific mechanical contributor to trip-related falls, whereas most established interventions primarily target strength, balance, or physical function. At the same time, the retrospective observational design, absence of a concurrent control group, and the potential influence of selection bias, nonresponse bias, recall bias, expectation bias, regression to the mean, and uncontrolled changes in health status or treatment may also have contributed to the observed effect size. Self-selection may be particularly relevant when interpreting the magnitude of these reductions, as customers who perceived greater benefit may have been more likely to participate than those with less favorable experiences.

A key finding of this study is the significant shift in reported fall mechanisms. Prior to using the VF footwear, for participants reporting trip-related events before and after VF use, 52.3% (68/130) of all falls were due to trips, a figure consistent with existing literature identifying forward trips as a leading cause of balance loss in older adults and those with physical disabilities [50,51]. During VF footwear use, for these users, the reported trip-related falls dropped to 31.5% (41/130). This finding is particularly relevant within the broader context of fall prevention, where age-related declines in postural stability and neurological impairments increase vulnerability to balance disturbances and make recovery from trip-related events more challenging. Previous studies have also reported that footwear comfort and perceived mediolateral stability are associated with fear of falling and perceived stability in older adults, further highlighting the potential role of footwear as one component of comprehensive fall prevention strategies [18,38]. Although the present study did not directly evaluate postural control or footwear comfort, the observed reduction in trip-related falls is consistent with the proposed biomechanical mechanism of VF footwear, whereby low-friction elements at the toe allow controlled sliding during incidental toe-ground contact rather than the abrupt deceleration that precipitates forward falls.

Reported reductions were observed consistently across diagnostic groups. Participants with multiple sclerosis reported the largest improvements, with 78% (305/391) reporting fewer falls and 56% (219/391) reporting fewer injurious falls. Similar patterns were observed in peripheral neuropathy, Parkinson disease, spinal cord injury, traumatic brain injury, stroke, and age-related conditions. Although these subgroup findings are descriptive and should not be interpreted as evidence of differential treatment effects, the consistency of the reported associations across clinically heterogeneous populations supports further prospective investigation of VF footwear in individuals at elevated risk of falling.

Similarly, self-reported fall-related health care use decreased substantially. Participants requiring medical attention after a fall decreased by 93.8% (192 to 12 participants), and reported hospital or urgent care visits decreased by 92.4% (131 to 10 participants). Although these reductions were larger than those reported in previous randomized trials, the direction of the findings is consistent with evidence showing that effective fall-prevention interventions can reduce subsequent health care use. For example, the GAPcare randomized controlled trial demonstrated fewer subsequent emergency department visits following an emergency department–initiated multidisciplinary intervention [52,53]. Given the well-established association between falls, downstream disability [10], and the substantial health care and economic burden attributable to falls [9,54], the observed participant-reported reductions in falls may have important clinical and economic implications if confirmed in prospective controlled studies. Because health care use and associated costs were self-reported and not independently verified through medical records, insurance claims, or other objective sources, these findings should be interpreted cautiously as preliminary estimates of potential impact rather than evidence of confirmed reductions attributable to VF footwear.

Participants also reported substantially reduced use of AD. Discontinuation of walkers, canes, and ankle-foot orthoses may reflect greater perceived stability, confidence, and functional independence. Given the known relationship between fear of falling, activity restriction, and functional decline, interventions that improve perceived safety during ambulation may have broader implications for participation and independence [11,55]. Previous randomized controlled trials have shown that reducing fear of falling can decrease activity restriction and improve confidence among older adults [56]. However, because balance, walking performance, or the clinical appropriateness of AD discontinuation were not objectively assessed, reduced AD use should not be interpreted as evidence of improved independence or as a recommendation to discontinue prescribed devices. Future prospective studies should evaluate changes in device use together with objective mobility outcomes and clinician assessment.

Study Limitations

Several limitations must be considered when interpreting these results. First, this was a retrospective pre-post observational study without a concurrent control group; therefore, the results demonstrate an association between VF footwear use and reported outcomes but do not establish causality. Changes in falls may have been influenced by unmeasured confounders, including rehabilitation, medication changes, disease progression or recovery, environmental modifications, altered activity levels, or increased caution. Regression to the mean may also have contributed, as participants may have sought the footwear following a period of unusually frequent falls.

Second, the low response rate is a major limitation and restricts confidence that the respondents were representative of all invited customers. Individuals who perceived substantial benefit may have been more motivated to complete the survey, potentially inflating the observed association. Conversely, participants with less favorable experiences, who continued to fall, experienced discomfort or safety concerns, or were dissatisfied with the footwear may have been less likely to respond. Because outcomes among the large majority of invited customers who did not participate are unknown, the direction and magnitude of this potential nonresponse bias cannot be determined. The low response rate therefore limits generalizability and may have affected the observed relationships.

XSL•FO
RenderX

Third, eligibility required participants to have worn the footwear for at least 3 months. Although this ensured meaningful exposure, it may have excluded individuals who discontinued use because of discomfort, lack of perceived benefit, or other reasons, introducing survivorship bias. This may have preferentially selected individuals who tolerated and continued using the footwear, potentially contributing to an overestimation of the observed benefit. Future studies should prospectively enroll participants at the time the footwear is dispensed and document adherence, discontinuation, adverse experiences, and reasons for stopping use.

Fourth, eligibility criteria, participants' characteristics, and all outcomes were self-reported and retrospectively recalled. Consequently, the findings are susceptible to misclassification, reporting, expectation, and recall bias. Because participants were aware of the intended purpose of VF footwear, their responses may have been influenced by expectations regarding its effectiveness. In addition, differences in the recall periods before and during VF footwear use may have affected the accuracy of reported fall frequencies. This was particularly relevant for binary outcomes, including the need for medical attention, hospital or urgent care use, and the occurrence of at least one injurious fall, which could not be meaningfully annualized and were therefore compared as raw proportions. The absence of objective outcome verification further limits the interpretation of these findings, as falls, injuries, diagnoses, and health care use were not confirmed through prospective monitoring, medical records, or insurance claims.

Finally, the survey collected only limited demographic and clinical information. Variables such as sex, comorbidities, disease severity, and functional status were not available, and the collected participant characteristics and diagnoses were based solely on self-report rather than independent clinical verification. Consequently, we were unable to evaluate whether the observed associations differed across important patient subgroups or account for demographic and clinical factors that may influence gait, balance, footwear perception, and fall risk. Previous work has shown that footwear comfort and foot structure may differ between women and men, particularly in older adults, suggesting that responses to footwear interventions may also vary [57]. Future prospective studies should include balanced recruitment and prespecified subgroup analyses to evaluate potential differences across demographic and clinical populations.

Despite these limitations, the consistency of the reported improvements across multiple outcomes and diagnostic groups supports further investigation. These findings should be considered preliminary real-world evidence that participants reported fewer falls and fall-related consequences during VF footwear use.

Conclusions

In this retrospective survey, participants reported fewer falls, fewer injurious falls, and lower fall-related health care use while using VF footwear. Similar improvements were observed across multiple outcomes, including a shift away from trip-related falls. Although gait, balance, and footwear biomechanics were not directly evaluated, these findings are consistent with the proposed biomechanical rationale of the device. Likewise, the reported reduction in AD use may reflect greater perceived stability or confidence, but this was not directly assessed and should not be interpreted as evidence of improved functional mobility.

Given the study's retrospective design, absence of a concurrent control group, and reliance on self-reported data, these findings should be considered preliminary associations rather than evidence of a causal treatment effect. Importantly, the low survey response rate raises substantial potential for nonresponse and self-selection bias. Participants who chose to respond may have differed systematically from the broader population of VF footwear users, including the possibility that those who experienced greater perceived benefit were more likely to participate. The magnitude of the observed associations may therefore not be representative of the full population of VF footwear users, limiting generalizability of these findings.

Prospective, controlled studies with objective outcome verification are needed to determine whether VF footwear causally reduces fall risk, to quantify its effect relative to existing fall-prevention interventions, and to identify subpopulations most likely to benefit; such studies should also collect a more comprehensive set of demographic and clinical covariates to enable multivariable analysis.

Acknowledgments

The authors would like to thank the Shirley Ryan AbilityLab REDCap administration team for their assistance in setting up the survey infrastructure, and all participants who generously contributed their time to complete the survey. JB was affiliated with the Max Näder Center for Rehabilitation Technologies and Outcomes Research at Shirley Ryan AbilityLab at the time of the study and is currently affiliated with the Physical Therapy Program in the College of Health at Oregon State University-Cascades.

ChatGPT (GPT-5.5; OpenAI) was used to assist with reformatting and rewriting select sentences for clarity during manuscript preparation. It was not used to generate study data, perform statistical analyses, or interpret results. All AI-assisted content was critically reviewed, verified, and approved by the authors.

Funding

This work was supported by the Charitable Arm of Rise Together Ventures. The funder had no involvement in the study design, survey administration, data collection, data analysis, interpretation of the results, manuscript preparation, or the decision to submit the manuscript for publication.

XSL•FO
RenderX

Data Availability

Data are available upon reasonable request to the corresponding author.

Authors' Contributions

SN, JB, SP, TS, and AJ contributed to conceptualization and methodology. SN, SP, JH, and AJ contributed to data curation, acquisition, and interpretation. SN contributed to formal analysis and statistical analysis. SN wrote the original draft. SN, JB, SP, JH, AL, TS, and AJ contributed to review and editing of the manuscript. TS and AJ contributed to funding acquisition. SN, SP, and AJ provided administrative and technical support. TS and AJ supervised the study. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

TS is the founder and chief technology officer of Cadense Inc, the manufacturer of the variable-friction footwear evaluated in this study. He is also a professor of teaching at the University of California, Santa Barbara. AL is an employee of Cadense Incorporated. These relationships have been disclosed and managed in accordance with University of California, Santa Barbara and Northwestern University policies. Cadense secured funding for this study through the charitable arm of Rise Together Ventures. TS contributed to the study design. Cadense assisted in identifying customers who had received variable-friction footwear at least 3 months before the survey to establish the pool of eligible participants and provided the corresponding contact list to the research team at Shirley Ryan AbilityLab. Recruitment emails were distributed directly by the Shirley Ryan AbilityLab team. Cadense had no role in sending recruitment invitations or any other aspect of participant recruitment, survey administration, data collection, or statistical analysis. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Survey contents.
[DOCX File, 20 KB - rehab_v13i1e104177_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 86 KB - rehab_v13i1e104177_app2.pdf]

References

1. Houry D, Florence C, Baldwin G, Stevens J, McClure R. The CDC Injury Center's response to the growing public health problem of falls among older adults. Am J Lifestyle Med 2016;10(1):74-77. [doi: 10.1177/1559827615600137] [Medline: 26688674]
2. Older adult falls data. Centers for Disease Control and Prevention (CDC). 2026. URL: https://www.cdc.gov/falls/data-research/index.html [accessed 2026-03-25]
3. Facts about falls. Centers for Disease Control and Prevention (CDC). 2026. URL: https://www.cdc.gov/falls/data-research/facts-stats/index.html [accessed 2026-03-25]
4. Finlayson ML, Peterson EW. Falls, aging, and disability. Phys Med Rehabil Clin N Am 2010 May;21(2):357-373. [doi: 10.1016/j.pmr.2009.12.003] [Medline: 20494282]
5. Moreland BL, Legha JK, Thomas KE, Burns ER. Hip fracture-related emergency department visits, hospitalizations and deaths by mechanism of injury among adults aged 65 and older, United States 2019. J Aging Health 2023 Jun;35(5-6):345-355. [doi: 10.1177/08982643221132450] [Medline: 36210739]
6. Vaishya R, Vaish A. Falls in older adults are serious. Indian J Orthop 2020 Feb;54(1):69-74. [doi: 10.1007/s43465-019-00037-x] [Medline: 32257019]
7. Falls. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2026-03-25]
8. Stevens JA, Mack KA, Paulozzi LJ, Ballesteros MF. Self-reported falls and fall-related injuries among persons aged>or=65 years--United States, 2006. J Safety Res 2008;39(3):345-349. [doi: 10.1016/j.jsr.2008.05.002] [Medline: 18571577]
9. Haddad YK, Miller GF, Kakara R, et al. Healthcare spending for non-fatal falls among older adults, USA. Inj Prev 2024 Jul 19;30(4):272-276. [doi: 10.1136/ip-2023-045023] [Medline: 39029927]
10. Florence CS, Bergen G, Atherly A, Burns E, Stevens J, Drake C. Medical costs of fatal and nonfatal falls in older adults. J Am Geriatr Soc 2018 Apr;66(4):693-698. [doi: 10.1111/jgs.15304] [Medline: 29512120]
11. Schoene D, Heller C, Aung YN, Sieber CC, Kemmler W, Freiberger E. A systematic review on the influence of fear of falling on quality of life in older people: is there a role for falls? Clin Interv Aging 2019;14:701-719. [doi: 10.2147/CIA.S197857] [Medline: 31190764]
12. MacKay S, Ebert P, Harbidge C, Hogan DB. Fear of falling in older adults: a scoping review of recent literature. Can Geriatr J 2021 Dec;24(4):379-394. [doi: 10.5770/cgj.24.521] [Medline: 34912493]

XSL•FO
RenderX

13. Scheffer AC, Schuurmans MJ, van Dijk N, van der Hooft T, de Rooij SE. Fear of falling: measurement strategy, prevalence, risk factors and consequences among older persons. Age Ageing 2008 Jan;37(1):19-24. [doi: 10.1093/ageing/afm169] [Medline: 18194967]
14. Suzuki M, Ohyama N, Yamada K, Kanamori M. The relationship between fear of falling, activities of daily living and quality of life among elderly individuals. Nurs Health Sci 2002 Dec;4(4):155-161. [doi: 10.1046/j.1442-2018.2002.00123.x] [Medline: 12406202]
15. Lavedán A, Viladrosa M, Jürschik P, et al. Fear of falling in community-dwelling older adults: a cause of falls, a consequence, or both? PLoS One 2018;13(3):e0194967. [doi: 10.1371/journal.pone.0194967] [Medline: 29596521]
16. Ellmers TJ, Ventre JP, Freiberger E, et al. Does concern about falling predict future falls in older adults? A systematic review and meta-analysis. Age Ageing 2025 Mar 28;54(4):afaf089. [doi: 10.1093/ageing/afaf089] [Medline: 40197783]
17. Young WR, Mark Williams A. How fear of falling can increase fall-risk in older adults: applying psychological theory to practical observations. Gait Posture 2015 Jan;41(1):7-12. [doi: 10.1016/j.gaitpost.2014.09.006] [Medline: 25278464]
18. Puszczałowska-Lizis E, Bujas P, Omorczyk J. Postural stability in women in the eighth and ninth decades of life. Acta Bioeng Biomech 2016;18(3):115-121. [doi: 10.5277/ABB-00475-2015-02] [Medline: 27840433]
19. Verghese J, Ambrose AF, Lipton RB, Wang C. Neurological gait abnormalities and risk of falls in older adults. J Neurol 2010 Mar;257(3):392-398. [doi: 10.1007/s00415-009-5332-y] [Medline: 19784714]
20. Homann B, Plaschg A, Grundner M, et al. The impact of neurological disorders on the risk for falls in the community dwelling elderly: a case-controlled study. BMJ Open 2013 Nov 25;3(11):e003367. [doi: 10.1136/bmjopen-2013-003367] [Medline: 24282241]
21. Matsuda PN, Verrall AM, Finlayson ML, Molton IR, Jensen MP. Falls among adults aging with disability. Arch Phys Med Rehabil 2015 Mar;96(3):464-471. [doi: 10.1016/j.apmr.2014.09.034] [Medline: 25449191]
22. Banks SA, Howe CL, Mandrekar J, et al. Assessing fall risk in multiple sclerosis using patient-reported outcomes and wearable gait metrics. Mult Scler J Exp Transl Clin 2025;11(2):20552173251329825. [doi: 10.1177/20552173251329825] [Medline: 40292035]
23. Sosnoff JJ, Socie MJ, Boes MK, et al. Mobility, balance and falls in persons with multiple sclerosis. PLoS One 2011;6(11):e28021. [doi: 10.1371/journal.pone.0028021] [Medline: 22132196]
24. Peterson EW, Cho CC, von Koch L, Finlayson ML. Injurious falls among middle aged and older adults with multiple sclerosis. Arch Phys Med Rehabil 2008 Jun;89(6):1031-1037. [doi: 10.1016/j.apmr.2007.10.043] [Medline: 18503796]
25. Guirguis-Blake JM, Perdue LA, Coppola EL, Bean SI. Interventions to prevent falls in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2024 Jul 2;332(1):58-69. [doi: 10.1001/jama.2024.4166] [Medline: 38833257]
26. Sadaqa M, Németh Z, Makai A, Prémusz V, Hock M. Effectiveness of exercise interventions on fall prevention in ambulatory community-dwelling older adults: a systematic review with narrative synthesis. Front Public Health 2023;11:1209319. [doi: 10.3389/fpubh.2023.1209319] [Medline: 37601180]
27. Caristia S, Campani D, Cannici C, et al. Physical exercise and fall prevention: a systematic review and meta-analysis of experimental studies included in Cochrane reviews. Geriatr Nurs 2021;42(6):1275-1286. [doi: 10.1016/j.gerinurse.2021.06.001] [Medline: 34555570]
28. Grabiner MD, Crenshaw JR, Hurt CP, Rosenblatt NJ, Troy KL. Exercise-based fall prevention: can you be a bit more specific? Exerc Sport Sci Rev 2014 Oct;42(4):161-168. [doi: 10.1249/JES.0000000000000023] [Medline: 25062002]
29. Sosnoff JJ, Finlayson M, McAuley E, Morrison S, Motl RW. Home-based exercise program and fall-risk reduction in older adults with multiple sclerosis: phase 1 randomized controlled trial. Clin Rehabil 2014 Mar;28(3):254-263. [doi: 10.1177/0269215513501092] [Medline: 23983091]
30. McCrum C, Bhatt TS, Gerards MHG, et al. Perturbation-based balance training: principles, mechanisms and implementation in clinical practice. Front Sports Act Living 2022;4:1015394. [doi: 10.3389/fspor.2022.1015394] [Medline: 36275443]
31. Cadmus-Bertram L, Irwin M, Alfano C, et al. Predicting adherence of adults to a 12-month exercise intervention. J Phys Act Health 2014 Sep;11(7):1304-1312. [doi: 10.1123/jpah.2012-0258] [Medline: 24176780]
32. Rivera-Torres S, Fahey TD, Rivera MA. Adherence to exercise programs in older adults: informative report. Gerontol Geriatr Med 2019;5:2333721418823604. [doi: 10.1177/2333721418823604] [Medline: 30733977]
33. Haas R, Haines TP. Twelve month follow up of a falls prevention program in older adults from diverse populations in Australia: a qualitative study. Arch Gerontol Geriatr 2014;58(2):283-292. [doi: 10.1016/j.archger.2013.10.010] [Medline: 24268766]
34. Bethancourt HJ, Rosenberg DE, Beatty T, Arterburn DE. Barriers to and facilitators of physical activity program use among older adults. Clin Med Res 2014 Sep;12(1-2):10-20. [doi: 10.3121/cmr.2013.1171] [Medline: 24415748]
35. Norouzi S, Błachnio A, Zeynalzadeh Ghoochani B, Asadollahi A, Kaveh MH. Effectiveness of shoes with non-slip insole on balance, fear of falling, and fall prevention among older women: a parallel RCT. Sage Open Aging 2025;11:30495334251358528. [doi: 10.1177/30495334251358528] [Medline: 40697361]
36. Kim IJ, Hegazy F. Enhancing footwear safety for fall prevention in older adults: a comprehensive review of design features. Ann Geriatr Med Res 2024 Jun;28(2):121-133. [doi: 10.4235/agmr.23.0167] [Medline: 38246912]

XSL•FO
RenderX

37. Feet and footwear for older adults. : Centers for Disease Control and Prevention (CDC); 2024 URL: https://www.cdc.gov/steadi/media/pdfs/2024/08/STEADI_Feet_Footwear_Guide_O.pdf [accessed 2026-09-04]
38. Puszczalowska-Lizis E, Mikulakova W, Lizis S, Koziol K, Omorczyk J. Relationships between the perception of footwear comfort and the fear of falls in people at the early period of old age. Int J Environ Res Public Health 2022 May 21;19(10):6267. [doi: 10.3390/ijerph19106267] [Medline: 35627804]
39. Chen CY, Yeh YD, Chen YC, Chuang PH, Lin HT. The rocker-soled shoes change the kinematics and muscle contractions of the lower extremity during various functional movement. Sci Rep 2022 Nov 28;12(1):20523. [doi: 10.1038/s41598-022-25116-2] [Medline: 36443456]
40. MacRae CS, Critchley D, Morrissey M, Shortland A, Lewis JS. Do rocker-sole shoes influence postural stability in chronic low back pain? A randomised trial. BMJ Open Sport Exerc Med 2016;2(1):e000170. [doi: 10.1136/bmjsem-2016-000170] [Medline: 27900198]
41. Stevens JA, Mahoney JE, Ehrenreich H. Circumstances and outcomes of falls among high risk community-dwelling older adults. Inj Epidemiol 2014 Dec;1(5):5. [doi: 10.1186/2197-1714-1-5] [Medline: 26744637]
42. Blake AJ, Morgan K, Bendall MJ, et al. Falls by elderly people at home: prevalence and associated factors. Age Ageing 1988 Nov;17(6):365-372. [doi: 10.1093/ageing/17.6.365] [Medline: 3266440]
43. Stark SL, Silianoff TJ, Kim HL, Conte JW, Morris JC. Tailored calendar journals to ascertain falls among older adults. OTJR (Thorofare N J) 2015 Jan;35(1):53-59. [doi: 10.1177/1539449214561764] [Medline: 25866488]
44. Ganz DA, Higashi T, Rubenstein LZ. Monitoring falls in cohort studies of community-dwelling older people: effect of the recall interval. J Am Geriatr Soc 2005 Dec;53(12):2190-2194. [doi: 10.1111/j.1532-5415.2005.00509.x] [Medline: 16398908]
45. Romli MH, Mackenzie L, Tan PJ, Chiew RO, Tan SH, Tan MP. Comparison of retrospective and prospective falls reporting among community-dwelling older people: findings from two cohort studies. Front Public Health 2021;9:612663. [doi: 10.3389/fpubh.2021.612663] [Medline: 33777881]
46. Mackenzie L, Byles J, D'Este C. Validation of self-reported fall events in intervention studies. Clin Rehabil 2006 Apr;20(4):331-339. [doi: 10.1191/0269215506cr947oa] [Medline: 16719031]
47. Sherrington C, Fairhall NJ, Wallbank GK, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD012424. [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
48. Gillespie LD, Robertson MC, Gillespie WJ, et al. Interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev 2012 Sep 12;2012(9):CD007146. [doi: 10.1002/14651858.CD007146.pub3] [Medline: 22972103]
49. Hopewell S, Adedire O, Copsey BJ, et al. Multifactorial and multiple component interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev 2018 Jul 23;7(7):CD012221. [doi: 10.1002/14651858.CD012221.pub2] [Medline: 30035305]
50. Crenshaw JR, Bernhardt KA, Achenbach SJ, et al. The circumstances, orientations, and impact locations of falls in community-dwelling older women. Arch Gerontol Geriatr 2017 Nov;73:240-247. [doi: 10.1016/j.archger.2017.07.011] [Medline: 28863352]
51. Pavol MJ, Owings TM, Foley KT, Grabiner MD. Mechanisms leading to a fall from an induced trip in healthy older adults. J Gerontol A Biol Sci Med Sci 2001 Jul;56(7):M428-M437. [doi: 10.1093/gerona/56.7.m428] [Medline: 11445602]
52. Goldberg EM, Marks SJ, Resnik LJ, Long S, Mellott H, Merchant RC. Can an emergency department-initiated intervention prevent subsequent falls and health care use in older adults? A randomized controlled trial. Ann Emerg Med 2020 Dec;76(6):739-750. [doi: 10.1016/j.annemergmed.2020.07.025] [Medline: 32854965]
53. Goldberg EM, Resnik L, Marks SJ, Merchant RC. GAPcare: the Geriatric Acute and Post-acute Fall Prevention Intervention-a pilot investigation of an emergency department-based fall prevention program for community-dwelling older adults. Pilot Feasibility Stud 2019;5(1):106. [doi: 10.1186/s40814-019-0491-9] [Medline: 31463079]
54. Dykes PC, Curtin-Bowen M, Lipsitz S, et al. Cost of inpatient falls and cost-benefit analysis of implementation of an evidence-based fall prevention program. JAMA Health Forum 2023 Jan 6;4(1):e225125. [doi: 10.1001/jamahealthforum.2022.5125] [Medline: 36662505]
55. Friedman SM, Munoz B, West SK, Rubin GS, Fried LP. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. J Am Geriatr Soc 2002 Aug;50(8):1329-1335. [doi: 10.1046/j.1532-5415.2002.50352.x] [Medline: 12164987]
56. Tennstedt S, Howland J, Lachman M, Peterson E, Kasten L, Jette A. A randomized, controlled trial of a group intervention to reduce fear of falling and associated activity restriction in older adults. J Gerontol B Psychol Sci Soc Sci 1998 Nov;53(6):384-392. [doi: 10.1093/geronb/53b.6.p384] [Medline: 9826971]
57. Puszczalowska-Lizis E, Koziol K, Omorczyk J. Perception of footwear comfort and its relationship with the foot structure among youngest-old women and men. PeerJ 2021;9:e12385. [doi: 10.7717/peerj.12385] [Medline: 34722004]

Abbreviations

AD: assistive device
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

VF: variable-friction

Edited by J Sarvestan; submitted 09.Jun.2026; peer-reviewed by E Puszczalowska-Lizis, M Yilmaz, O Awe, W Mikulakova; revised version received 27.Aug.2026; accepted 31.Aug.2026; published 22.Sep.2026.

Please cite as:
Nataletti S, Bartloff J, Prokup S, Harris J, Larson A, Susko T, Jayaraman A
Participant-Reported Reductions in Falls and Fall-Related Injuries Associated With Variable-Friction Footwear: Retrospective Pre-Post Survey Study
JMIR Rehabil Assist Technol 2026;13:e104177
URL: https://rehab.jmir.org/2026/1/e104177
doi:10.2196/104177

© Sara Nataletti, Jennifer Bartloff, Sara Prokup, Joseph Harris, Anushka Larson, Tyler Susko, Arun Jayaraman. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 22.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging AI and Machine Vision in Telephysiotherapy for Promoting Physical Activity and Reducing Sedentary Behavior: Scoping Review

Liva Redliha[1], BSc; Guna Semjonova[1], PhD; Liva Araka[2], BSc

[1]Department of Rehabilitation, Faculty of Health and Sports Sciences, Riga Stradiņš University, Dzirciema Street 16, Riga, Latvia

[2]Faculty of Medicine, Riga Stradiņš University, Riga, Latvia

Corresponding Author:
Liva Redliha, BSc
Department of Rehabilitation, Faculty of Health and Sports Sciences, Riga Stradiņš University, Dzirciema Street 16, Riga, Latvia

Abstract

Background: Globally, sedentary behavior is a leading risk factor for noncommunicable diseases and premature death. Although telerehabilitation has emerged as a promising method to increase access to physical therapy, the potential of artificial intelligence (AI) and machine vision technologies to objectively monitor and promote physical activity (PA) in home settings remains underexplored.

Objective: This scoping review aimed to map and systematize current evidence on the use of AI and machine vision in telephysiotherapy to promote PA and reduce sedentary behavior among healthy adults.

Methods: The review followed the Joanna Briggs Institute scoping review methodology and was reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines. The PCC (population, concepts, and context) framework defined participants as healthy adults, the concept as AI or machine learning (ML) and machine vision technologies, and the context as remote, home-based telephysiotherapy for objective PA monitoring and/or outcome measurement in interventions aimed at promoting daily activity. Searches were conducted in PubMed, ProQuest, Science Direct, and Wiley Online Library for peer-reviewed primary studies published between 2016 and 2024. Inclusion and exclusion criteria were applied systematically using the Rayyan platform.

Results: Of 7253 initial records, 4 studies met the inclusion criteria. Identified digital solutions comprised ML algorithms (eg, extreme gradient boosting and hybrid classifiers), deep learning models such as graph convolutional networks, video-based center-of-pressure assessments using RGB cameras, and ontology-driven semantic data structuring for personalization. The most frequently studied technologies combined wearable sensors with ML-based behavior change frameworks to assess usability and prototype feasibility. However, none of the studies demonstrated statistically significant improvements in reducing sedentary time or increasing PA levels among participants. The studies predominantly used exploratory or pilot designs, with small sample sizes (range: 10 - 32 participants) and short follow-up periods (≤12 wk), limiting the strength of the conclusions.

Conclusions: Current evidence on the application of AI and machine vision in telephysiotherapy remains scarce and largely exploratory. Although the integration of wearable technology and ML algorithms shows technical feasibility and personalization potential, there is insufficient evidence to confirm improvements in PA outcomes. Future research should prioritize longitudinal clinical trials, ensure the inclusion of diverse populations, and evaluate behavioral outcomes to determine the real-world impact of these technologies in telerehabilitation.

Trial Registration: OSF Registries OSF.IO/SFC97; https://doi.org/10.17605/OSF.IO/SFC97

(JMIR Rehabil Assist Technol 2026;13:e80459) doi:10.2196/80459

KEYWORDS

healthy adults; artificial intelligence; AI; machine vision; telephysiotherapy; machine learning

Introduction

Sedentary behavior is a major global health challenge. According to the World Health Organization (WHO), individuals who do not engage in adequate physical activity (PA) face a 20% to 30% higher risk of mortality compared to those who are regularly active [1]. Despite the well-documented health benefits of moderate-to-vigorous PA, large segments of the global population remain physically inactive [2]. In 1990, approximately 200 million adults worldwide lived with obesity; by 2022, this number had surged to 890 million, with the obesity rate more than doubling in every WHO region [3]. This rapid increase is closely linked to insufficient PA and sedentary

XSL•FO
RenderX

lifestyle patterns [3]. Health guidelines recommend that adults engage in at least 150 minutes of moderate-intensity aerobic PA per week and perform strength-based exercises at least 2 times weekly [4]. Even light PA offers health benefits compared to none, and interruptions of sedentary behavior—such as taking breaks from sitting or incorporating movement into leisure time—are associated with improved health outcomes [5]. In Latvia, the situation mirrors global trends: only 9.5% of adults report being physically active for at least 30 minutes four or more times per week, and over 58% are classified as overweight or obese [4]. The growing burden of noncommunicable diseases, many of which are preventable through lifestyle changes, underscores the urgent need to develop effective, scalable, and accessible PA interventions [4]. Telerehabilitation has emerged as a promising strategy for delivering physiotherapy services remotely, thereby overcoming geographical and logistical barriers to care. Real-time telephysiotherapy has been shown to be as effective as traditional face-to-face rehabilitation in improving physical function and reducing musculoskeletal pain [2]. It also offers convenience, flexibility, and the potential for long-term engagement in health-promoting behaviors [2]. Simultaneously, digital health innovations such as artificial intelligence (AI), machine learning (ML), and machine vision technologies are gaining traction in health care and rehabilitation [2]. In this review, telephysiotherapy is defined as the remote delivery of physiotherapy-related assessment, monitoring, and intervention using digital technologies outside traditional clinical settings, most commonly in the home environment. Examples include video-based exercise supervision, wearable sensor–based activity monitoring, and AI-supported feedback systems designed to guide movement and promote PA.

These technologies enable personalized, real-time monitoring and feedback; improve data collection and analysis; and support the development of adaptive, behavior-informed interventions [2]. AI-driven tools are already being used for motion analysis, diagnostic support, and virtual rehabilitation, often in combination with wearable devices, mobile applications, or virtual reality systems [6]. However, despite the growing interest and potential, the specific application of AI and machine vision technologies in telephysiotherapy—particularly for the purpose of objectively monitoring PA and promoting daily movement in home environments—remains poorly defined. There is a lack of consolidated evidence on which digital tools are currently in use, what data they collect, how reliable and user-friendly they are, and whether they effectively change behavior or health outcomes in adult users [7]. To address existing research gaps resulting from limited consolidated knowledge in this field, this scoping review aimed to collect and systematize current evidence on the use of AI and machine vision in telephysiotherapy to promote PA and reduce sedentary behavior in healthy adult populations. While AI, ML, and machine vision technologies for PA monitoring are widely developed across multiple domains, including wearable computing and human-computer interaction, this review intentionally focuses on their application within telephysiotherapy contexts. This distinction is particularly relevant for intervention-focused research, where clinical relevance, remote supervision, and integration into physiotherapy workflows are essential. Consequently, the scope of evidence relevant to technology development (research questions 1 and 2) is broader than that of intervention studies (research question 3), which remain limited within telephysiotherapy settings.

To guide this objective, the following research questions were formulated:

- What AI and machine vision solutions are currently identified in telephysiotherapy?
- What are the characteristics of these technologies, including the types of data they monitor and their reliability?
- What AI and machine vision–based interventions have been used to promote daily PA, what are their features, which target populations have been studied, and what results have been reported?

Methods

Protocol and Registration

The review study was conducted from August 2024 to May 2025. This review was conducted in accordance with the Joanna Briggs Institute methodology for scoping reviews [8] and reported in accordance with PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [9]. We followed the 6-stage framework for a scoping review proposed by Arksey and O'Malley [10]. The protocol was registered in the Open Science Framework [11].

The identified AI applications primarily focused on activity detection, movement quality assessment, and performance classification, with reported metrics such as classification accuracy and model precision. Only one study incorporated AI-driven personalized recommendations, highlighting that decision-support and behavior guidance systems remain underrepresented compared to monitoring-focused approaches. Only studies published in English were included due to resource constraints related to translation and to ensure accurate interpretation of methodological and technical details. While this approach may have excluded relevant studies published in other languages, English remains the dominant language of publication in AI, digital health, and rehabilitation research.

Eligibility Criteria

Based on recommendations from the Joanna Briggs Institute [8] and the research questions to refine the focus of our scoping study and develop an effective search strategy, we defined the eligibility criteria for the PCC (population, concepts, and context) framework [8].

Population: Healthy Adults

This review focused on healthy adults aged 18 years and older who do not present with acute or chronic medical conditions. The decision to include this population was based on the need to evaluate preventive strategies aimed at promoting PA and reducing sedentary behavior in individuals without current health limitations. Healthy adults represent a critical target group for secondary prevention efforts, particularly as sedentary lifestyles are increasingly common in working-age populations and contribute to the long-term development of noncommunicable diseases [4]. In this review, "healthy adults"

XSL•FO
RenderX

were defined as individuals without diagnosed acute or chronic physical diseases requiring medical treatment or structured rehabilitation. Individuals with lifestyle-related risk factors, such as overweight or obesity, were not excluded if no clinical diagnosis was reported. Mental health conditions were excluded only if they significantly affected physical functioning or participation in PA interventions.

Concept: AI, ML, and Machine Vision

The concept explored in this review includes AI, ML, and machine vision technologies used within telephysiotherapy contexts. AI refers to computational methods that enable systems to simulate human decision-making, adapt to data patterns, and provide real-time support [12]. ML, a subset of AI, allows for continuous system learning from large datasets, supporting personalized exercise recommendations and feedback loops [12]. Machine vision enables the interpretation of visual data for motion tracking, posture analysis, and automated exercise monitoring [12].

Context: Objective Remote Monitoring of Physical Activity (Telephysiotherapy)

The context of this review is remote physiotherapy—specifically, telephysiotherapy conducted outside formal health care settings, such as in the home environment. This includes digital interventions designed to objectively monitor or assess PA through sensors, video analysis, or algorithm-based platforms. These technologies are applied either for clinical assessment or as outcome measures in targeted interventions that aim to encourage sustained PA in daily life as part of secondary prevention strategies [8].

Selection of Sources of Evidence

Individual studies published from January 2016 to April 2024 were considered for inclusion. The selection of sources for this scoping review was guided by clearly defined inclusion and exclusion criteria aligned with the study objectives and the PCC framework (Textbox 1). These criteria were established to ensure that the review captured relevant evidence on the application of AI, ML, and machine vision in telephysiotherapy aimed at promoting PA and reducing sedentary behavior in healthy adult populations. Given the rapid advancement of AI-based solutions in health care, a broad range of study types was included, provided they met clinical relevance and methodological transparency standards.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Population: healthy adults aged 18 years and older, without acute or chronic medical diagnoses;
- Concept: digital technologies incorporating artificial intelligence (AI), machine learning (ML), or machine vision applied in physiotherapy or physical activity (PA) monitoring
- Context: objective remote monitoring of PA conducted outside of health care institutions (eg, in the home environment), either for assessment or analysis or as an outcome measure in interventions
- Study type: primary research studies of any design (quantitative, qualitative, or mixed methods), including pilot and feasibility studies
- Publication status: published in a peer-reviewed journal
- Publication language: English
- Full-text availability
- Included keywords: artificial intelligence, machine learning, machine vision, telephysiotherapy, telerehabilitation, physical activity, exercise, wearable sensors, home-based rehabilitation, digital health, and remote monitoring

Exclusion criteria

- Population: individuals with acute or chronic medical conditions, pediatric or adolescent populations, and studies involving animal models
- Concept: digital technologies not related to AI, ML, or machine vision (eg, telemonitoring without intelligent data processing)
- Context: studies not involving PA as an outcome, or interventions not conducted remotely
- Setting: institutional or inpatient rehabilitation settings
- Study type: protocols, expert opinions, narrative or systematic reviews, editorials, letters, dissertations, or conference abstracts
- Publication status: non–peer-reviewed sources
- Publication language: languages other than English
- Full-text not available
- Excluded keywords: stroke, cancer, diabetes, depression, surgery, mice, rat, animal, in vitro, cognitive training, pharmacology, neurological disorder, and rehabilitation robotics

Studies focusing primarily on rehabilitation robotics were excluded a priori, as the review focused on software-driven AI and machine vision solutions applicable within telephysiotherapy contexts. Robotic systems often require

XSL•FO
RenderX

specialized hardware, supervised environments, or clinical infrastructure, which limits their applicability to home-based, unsupervised telephysiotherapy for healthy adults.

Information Sources

For this scoping review, peer-reviewed articles were searched in the following electronic databases: PubMed, ProQuest, Science Direct, and Wiley Online Library. These databases were selected to ensure broad coverage of biomedical, rehabilitation, and digital health literature relevant to the study's PCC framework. The search results were imported into the Rayyan intelligent systematic review platform, where duplicates were identified and removed by multiple reviewers. A preliminary search was conducted between November 2024 and April 2025. The combined database search initially yielded 7253 references before deduplication. Only articles written in English were considered, and gray literature, including conference proceedings, dissertations, and preprints, was excluded from the final selection.

Search Strategy and Search

Keywords were selected based on the defined PCC framework, as well as commonly used English terminology and synonyms. The keywords were grouped according to the PCC components and are outlined in Textbox 2.

Textbox 2. Keyword combinations.

- Population: healthy adults
- Concept: artificial intelligence, machine learning, and telephysiotherapy
- Context: objective remote monitoring of physical activity (telephysiotherapy) outside the health facility (home environment) can be used for assessment and analysis or as an outcome measure in a targeted intervention to promote daily physical activity as part of secondary prevention

The initial search strategy (Multimedia Appendix 1) was developed for PubMed in consultation with the Rīga Stradiņš University librarian and was later adapted for other databases. The Boolean logic used included:

AND between PCC components OR within each component's synonyms

The following filters were used: language of publication (English), research on humans (English: human species), primary research, and publication date (from January 1, 2016). A search strategy was developed in the PubMed database using the following keywords:

Machine [All Fields] AND ("vision, ocular" [MeSH Terms]) OR "Machine Vision" [Text Word]) OR ("Artificial Intelligence/supply and distribution" [Majr] OR "Artificial Intelligence"[Text Word]) AND ("physical therapy modalities"[MeSH Terms] OR physiotherapy [Text Word]) OR ("Telerehabilitation" [Mesh] OR "Telerehabilitation" [Text Word]) OR ("Exercise" [Majr] OR "Physical Activity" [Text Word]) OR "sedentary behavior" [MeSH Terms] OR "Sedentary Behavior" [Text Word]

The search strategy used in the other databases included the following keywords:

(machine learning or artificial intelligence) AND (physical activity or exercise) OR sedentary behavior AND (telerehabilitation).

Database searches were conducted in PubMed, ProQuest, Science Direct, and Wiley Online Library from November 2024 to April 2025. The search results were imported into Rayyan, an intelligent systematic review platform. The study selection proceeded in 3 stages. First, titles and abstracts were screened to remove clearly irrelevant records. In the second phase, full-text studies were reviewed to assess eligibility based on inclusion and exclusion criteria and relevance to the primary and secondary research questions. Finally, a structured data extraction template was developed in Rayyan to support the systematic charting of evidence across studies. The template captured key details including participant characteristics (age, health status), the type of AI or machine vision technology used, PA monitoring methods, intervention design (if applicable), outcome variables, and main findings. Each study was reviewed independently by 2 team members. If Rayyan flagged a conflict between reviewers on inclusion or exclusion, a third reviewer made the final decision.

Data Charting Process

This scoping review followed a structured approach to extract and analyze data from studies examining the application of AI, ML, and machine vision technologies in telephysiotherapy for promoting PA and reducing sedentary behavior among healthy adults. Data from the included studies were systematically charted using a standardized template to ensure consistency and relevance to the research objectives and the PCC framework.

Key variables extracted for each study included: paper ID, first author (name and surname), journal name, and title of the study to establish a bibliographic reference base. Geographic and demographic context was provided by documenting the country of the study, study aim, and participant characteristics (age, sex, and health status). To understand the methodological rigor and relevance of each study, information on study design (eg, feasibility study, pilot, experimental, or observational), applied methods, and analytical approaches was collected. Special emphasis was placed on detailing the digital technology used, specifically whether it involved AI, ML, or machine vision components, and how these technologies were implemented in the remote physiotherapy context.

Furthermore, studies were reviewed for their use of PA monitoring instruments (eg, wearables, video-based systems, and sensor technologies) and whether PA was measured as an intervention outcome, behavioral indicator, or clinical metric. For studies involving interventions, data were also collected on the structure and duration of the intervention, feedback

XSL•FO
RenderX

mechanisms, and whether the systems were tailored to behavior change principles.

Given the broad diversity in how AI and digital physiotherapy interventions are designed and evaluated, studies varied in their terminology and framing of participants and settings (eg, "healthy volunteers," "nonclinical adults," and "remote physiotherapy users"). This heterogeneity was noted to better understand the generalizability and scalability of current AI-enabled telephysiotherapy solutions for healthy adult populations. The data charting process facilitated the synthesis of current practices, innovation trends, and evidence gaps in using AI-based technologies to support remote PA interventions. All data were independently extracted by 2 reviewers and cross-checked for accuracy and completeness. Discrepancies were resolved through team discussion to maintain methodological rigor and transparency.

Synthesis of Results

Following the final stages of the 6-stage framework for scoping reviews proposed by Arksey and O'Malley [10], after identifying the research questions, locating relevant studies, selecting eligible sources, and charting the data, the findings were collated, summarized, and reported. First, a *descriptive numerical summary* was conducted to outline the number of included studies, publication years, countries of origin, study designs, and types of digital technologies investigated. This provided a broad overview of the research landscape and highlighted the diversity of approaches across different geographic and methodological contexts. Second, a *qualitative thematic analysis* was applied to extract key themes related to the use of AI, ML, and machine vision technologies in remote PA monitoring and intervention delivery. This thematic synthesis focused on technology types, behavioral targets, implementation contexts, feedback mechanisms, and outcome measures. Third, *research findings were narratively evaluated* to identify results aligned with the study's core research questions. These included the types and features of digital solutions used, the contexts in which they were implemented, and their impact on promoting PA and reducing sedentary behavior. Finally, the *interpretation of findings* was contextualized in relation to practical implications for telephysiotherapy practice, limitations in current evidence, and recommendations for future research and policy development.

Results were organized descriptively and are presented using structured tables, accompanied by narrative explanations to clarify their relevance and significance.

Results

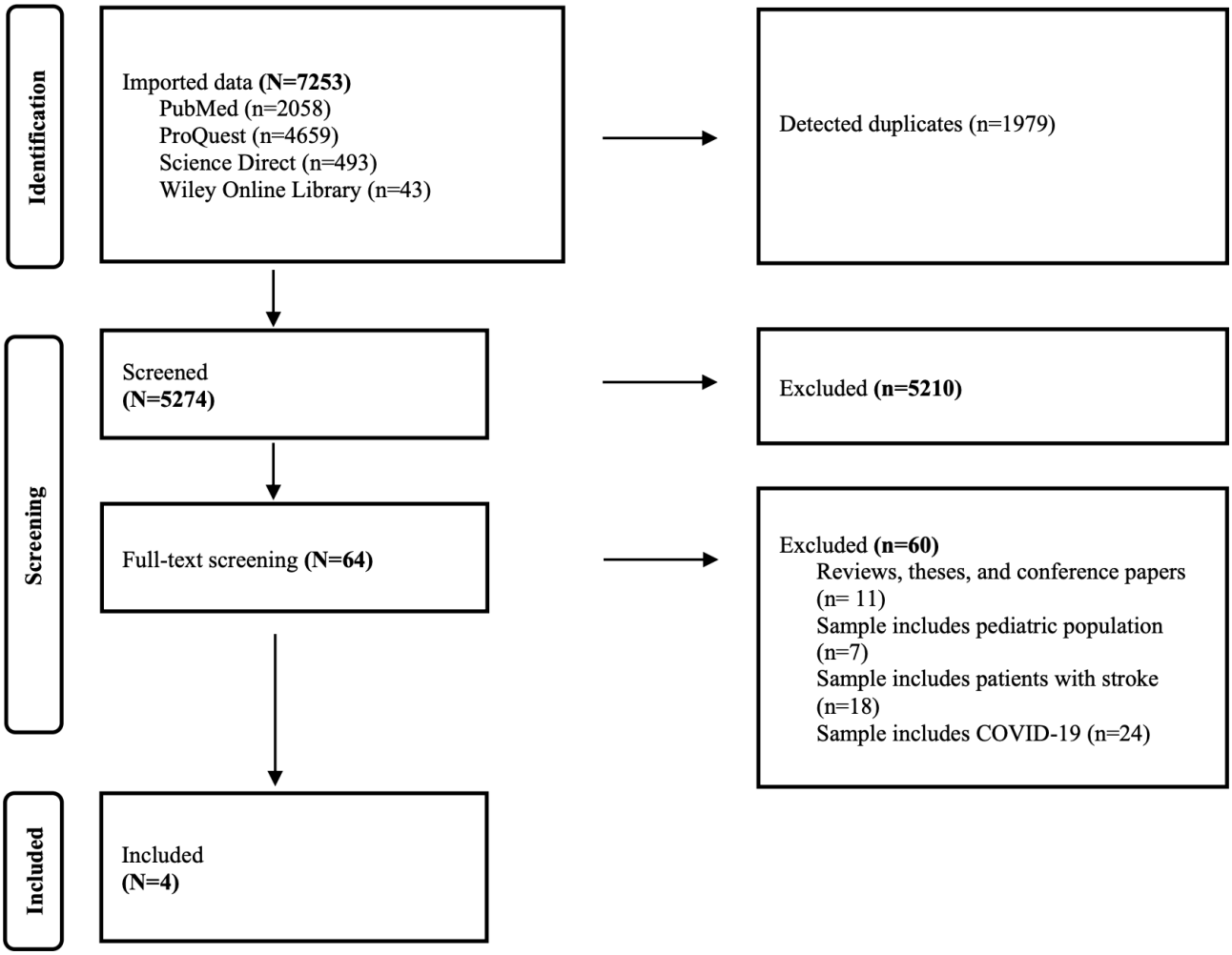
Selection of Sources of Evidence

To improve clarity and synthesis, the results are organized thematically rather than as isolated study summaries. The findings are presented according to three overarching themes: (1) types of AI and machine vision technologies used, (2) data sources and performance characteristics, and (3) intervention-related features and outcomes.

Our search identified a total of 7253 records from 4 electronic databases: PubMed, ProQuest, Science Direct, and Wiley Online Library. Among these, 1979 duplicate records were identified and removed, resulting in 5274 records eligible for title and abstract screening, indicating the effectiveness and feasibility of the search strategy. Following the initial screening based on predefined inclusion and exclusion criteria, 5210 records were excluded due to reasons such as irrelevant population (eg, clinical samples), concept mismatch (eg, digital tools without AI or machine vision), or nontelephysiotherapy contexts. Sixty-four full-text studies were assessed for eligibility. Of the 64 full-text studies reviewed, 60 were excluded for the following reasons: inclusion of clinical populations with chronic or acute diagnoses, absence of AI or machine vision components, focus on pediatric or adolescent populations, use of robotics outside telephysiotherapy, and nonprimary study designs such as reviews, conference abstracts, or protocols. A total of 4 studies met all inclusion criteria and were included in this scoping review. The selection process is illustrated in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the selection of sources of evidence.



Characteristics of Included Studies

The characteristics of the 4 selected studies in this review are summarized in Figure 1. Given that AI, ML, and machine vision were among the inclusion criteria, all included studies incorporated at least one method for the objective assessment of PA.

The identified technologies—their types and the monitored data in the analyzed studies—are summarized in Table 1.

XSL•FO
RenderX

Table . Characteristics of included studies for analysis.

Author, year (country)	Research design	Participants	Intervention	Artificial intelligence and machine vision technologies	Summary of main results
Du et al [13], 2022 (United States)	Experimental comparative validation with a preprepared public dataset UI-PRMD[a]	• n=21 healthy adults (age and gender details not specified)	• Developed a deep learning framework with GCNs[b] for physical activity assessment. • Includes video-level and frame-level CoP[c] parameter assessment.	• GCN, ST-GCN[d], and adaptive GCN • 3D body landmarks from video (with OpenPose) • Personalized CoP indicator assessment from video, also using only an RGB camera (eg, smartphone)	• The proposed method with GCN effectively estimates CoP path length, sway area, and position (as balance indicators). • The multitask model showed higher accuracy compared to single-task models. • A new dataset with synchronized 3D pose and CoP data was created and published. • The study confirms that video analysis can replace expensive laboratory equipment in everyday balance and rehabilitation monitoring.
Chatterjee et al, 2022 [12] (Norway)	Experimental study with MM algorithm testing on public and private data	• Public dataset (Fitbit): n=33 adults • Private dataset (MOX2-5 sensor): n=16 adults (aged 18 - 64 y), 30 days of observation in healthy adults	• Created a prototype of the eCoach system for generating personalized activity recommendations. • Combined ML[e] algorithms (eg, decision tree, SVM[f], and random forest) with a semantic ontology. • Used transfer learning and incremental learning approaches to improve and personalize models.	• ML classifiers: decision tree, random forest, SVM, and KNN[g]. • Use of ontologies for semantic data structuring and personalization. • Data from Fitbit and medical-grade sensor MOX2-5.	• Decision tree classifier with the “entropy” indicator showed the highest accuracy: • Fitbit data: 97.5% accuracy • MOX2-5 data: 96.1% accuracy. • After transfer learning: 97.99% accuracy. • The eCoach system can provide personalized recommendations based on activity data and user preferences. • The ontology model improves the explainability of recommendation logic (through SPARQL[h], OWL[i], and RDF[j]).

XSL•FO
RenderX

Author, year (country)	Research design	Participants	Intervention	Artificial intelligence and machine vision technologies	Summary of main results
Boyer et al [14], 2023 (Canada)	Experimental research with sensor data collection and MM model development	• Twenty one participants, age: 18 - 65 years • Health status: healthy volunteers without musculoskeletal disorders • Gender: 8 men and 12 women	• Participants performed 4 different physical therapy exercises (eg, straight leg raises and knee bends) while wearing an IMU[k] sensor (MPU-9250) attached to their ankle. • The sensor collected accelerometer, gyroscope, and magnetometer data during the exercises.	• IMU sensors as a data source (acceleration, rotation, and magnetic field measurements). • ML algorithms for exercise quality recognition and monitoring. • Classification of exercises as "correct" versus "incorrect." • The approach allows remote monitoring of patient self-discipline and technique.	• ML models can accurately determine whether a participant is performing the intended physical activity or not. • The best models achieved up to 95% accuracy, indicating that this approach is promising for remote patient monitoring and motivation during the rehabilitation process. • Thus, such systems can improve patient compliance and provide objective monitoring without the direct presence of a therapist.
Mennella et al [15], 2023 (Italy)	Experimental development with algorithm development and testing	• Seventeen healthy participants (9 women and 8 men) with an average age of 25.6 (SD 6.4) years took part in the experiment	• Motion data were obtained using MoveNet, which generates 3D skeleton coordinates. • A lightweight deep neural network with a single hidden layer was developed to classify both ROM[l] (3 classes: low, normal, and high) and compensation (correct or compensatory execution). • ROM values calculated as angles between joints were used for classification.	• ML algorithm XGBoost[m] was used to classify whether a participant performed a movement correctly or with compensation. The data were obtained using IMU sensors that recorded the angular velocity of 3 body segments: the thoracic spine, lumbar spine, and pelvis. Time-domain features were extracted from the data, such as mean value, maximum, RMS[n], and so on.	• ROM classification: average accuracy of 89% achieved (ranging from 81%-95% depending on the exercise). • Compensatory movement classification: average accuracy of 98%. • Compared to manual classification by clinicians, the system's performance is very similar or even more accurate, thus improving objectivity and repeatability.

[a]UI-PRMD: University of Idaho – Physical Rehabilitation Movement Data.

[b]GCN: graph convolutional network.

[c]CoP: center of pressure.

[d]ST-GCN: spatial-temporal graph convolutional networks.

[e]ML: machine learning.

[f]SVM: support vector machine.

[g]KNN: k-nearest neighbor.

[h]SPARQL: is a query language used to search and change data stored as graphs on the web.

[i]OWL: object windows library.

[j]RDF: resource description framework.

XSL•FO
RenderX

[k]IMU: inertial measurement unit.

[l]ROM: range of motion.

[m]XGBoost: extreme gradient boosting.

[n]RMS: root mean square.

In 2 of the 4 experimental studies, a body-mounted inertial measurement unit sensor was identified as the data source, and ML algorithms were used to recognize and monitor exercise quality [14].

AI and Machine Vision Solutions Currently Identified in Telephysiotherapy

A deep learning framework was developed with graph convolutional networks for PA assessment [13]. Video-based methods were used to assess center-of-pressure (CoP) parameters at both the level of the whole body and individual body segments [13]. An eCoach prototype system was designed to generate personalized activity recommendations [12]. ML algorithms (eg, decision tree, support vector machine, and random forest) were integrated with semantic ontology [14]. ML algorithms were applied for exercise quality recognition and monitoring [13]. The extreme gradient boosting ML algorithm was used [16].

AI and Machine Vision–Based Interventions

One of the studies included in this review introduces a system that allows real-time quantitative assessment of motor patterns in rehabilitation exercises using 2D video analysis. The main objective is to discreetly observe the exercises performed by a person during a therapy session and assess whether they are being performed correctly or incorrectly according to the instructions. The assessment corresponds to the assessment of range of motion and compensatory movement patterns that the participant may perform during a rehabilitation session [15].

To assess a patient's physical performance using 2D video, the methodology consists of the following steps: body reference point extraction, data normalization, and classification. A network approach was investigated to classify correctly performed exercises. The classifier must be sufficiently stable to assign a label to each frame that indicates incorrect execution and good movement quality (normal movement patterns without erroneous movement categories, such as limited range of motion or compensatory movements) [15].

MoveNet is an upward-looking estimation model that uses heat maps to locate key points on a person's body. The model is designed to run in a browser using TensorFlow.js or on devices using TensorFlow Lite in real time. The MoveNet Thunder variant was chosen, which provides greater power and better prediction quality while maintaining high real-time speed [16]. The architecture of the MoveNet Thunder version uses an RGB video frame or image with a resolution of 256×256 as input [15].

The research team compiled a dataset of shoulder physiotherapy exercises. It includes both concentric and isometric exercises, 9-axis data (accelerometer, gyroscope, and magnetometer data), as well as data from nonexercise activities (ie, out-of-distribution data) [14]. The researchers used a combination of ML algorithms to analyze the data and identify the most effective shoulder exercises [14].

A semantic ontology model was created to annotate ML prediction results and personal preferences, conceptualizing personalized recommendation generation using a hybrid approach. The study used a transfer learning approach to improve ML model training and performance, as well as a gradual learning approach to process the daily increasing volume of data and adapt it to ML models [12].

"Deep learning–based video analysis for human performance evaluation" investigated 2 typical tasks: a clinical biometric evaluation task—the assessment of the CoP metric—and a performance evaluation task—the assessment of physical rehabilitation exercises. The CoP metric and physical rehabilitation exercise assessment aimed to develop a new video analysis system for human performance assessment. Two gold-standard CoP indicators—CoP path length, which reflects the total distance or length of the CoP trajectory while standing, and CoP oscillation area, an ellipse that covers a certain percentage of the CoP trajectory—are widely used to quantitatively assess posture control. A shorter CoP path length and oscillation area usually indicate better postural control, while a longer length is associated with aging, diseases, or injuries that impair balance control or mobility [13].

Discussion

Principal Results

This scoping review aimed to collect and systematize current evidence on the use of AI and machine vision in telephysiotherapy to promote PA and reduce sedentary behavior in healthy adult populations. Only 4 studies met the eligibility criteria, highlighting the limited but growing body of research in this emerging field. The findings provide insight into the types of digital technologies currently in use, their core features, application settings, and the methodological gaps that need to be addressed in future research. The identified studies used a variety of digital approaches to support remote physiotherapy interventions. These included ML algorithms such as extreme gradient boosting and hybrid ensemble classifiers used to analyze sensor-derived movement data [13], as well as deep learning models like graph convolutional networks designed to interpret biomechanical features [13]. One study used a semantic ontology to structure and personalize PA recommendations [12], while another applied video-based systems to calculate the CoP using a standard RGB camera in a home environment [16]. These solutions aimed to support PA monitoring and personalized feedback delivery, aligning with the broader goal of increasing movement and reducing sedentary behavior outside traditional clinical settings [16]. In relation to the review's research questions, the first objective was to identify AI and machine vision solutions currently applied in telephysiotherapy. The review revealed that these technologies are predominantly

XSL•FO
RenderX

in the developmental or pilot-testing stages and are often integrated into wearable sensor platforms or computer vision–based applications [17]. Their use focused on either real-time exercise monitoring or movement analysis, reflecting a technical interest in enhancing the precision of remote physiotherapy systems [17]. The second objective addressed the characteristics of these technologies, including the type of data monitored and the reliability of the systems. While most studies demonstrated that it is technically feasible to implement AI and machine vision in home-based physiotherapy contexts, they rarely included robust evaluations of system accuracy, user engagement, or long-term adherence [17]. In many cases, technical functionality was prioritized over behavioral or clinical outcomes, indicating a gap between innovation and implementation readiness [16,18]. The third research question focused on identifying specific interventions and their outcomes. Although each study aimed to encourage PA, none reported statistically significant improvements in PA levels or reductions in sedentary behavior [12-15]. This may be explained by several factors, including small sample sizes, short intervention durations, and the exploratory nature of the study designs [12-15]. These limitations underscore the need for more rigorous evaluation methods and stronger behavioral outcome tracking in future research [19]. It is important to note that the limited number of included studies does not reflect the maturity of PA recognition technologies themselves, but rather the early stage of their translation into telephysiotherapy interventions for healthy adults. While AI-based activity monitoring methods are well established in broader digital health and human movement research, their integration into structured, behavior-oriented telephysiotherapy programs remains sparse.

Implications for Practice

The findings of this review emphasize the potential of integrating innovative digital solutions—specifically those using AI and machine vision—into telephysiotherapy practice aimed at healthy adults. These technologies offer opportunities to enhance remote PA monitoring through automated feedback, personalized recommendations, and real-time performance analysis [14,15]. For physiotherapists, this means the possibility of expanding their role beyond in-clinic care to continuous, preventive support in home and community settings. By leveraging AI-enabled systems, clinicians could more effectively promote PA adherence, reduce sedentary behavior, and support early intervention for individuals at risk of inactivity-related health issues. However, the limited number of studies and variation in approaches also highlight the need for a more standardized framework in designing and implementing AI-based PA monitoring. There is currently no consistent methodology for selecting sensors, interpreting data, or integrating intelligent systems into physiotherapy protocols. This inconsistency hinders comparability between studies and limits the transferability of findings to real-world practice [14]. Just as prior reviews in clinical populations have called for harmonized guidelines on accelerometry and self-reported PA [17], similar efforts are needed in the context of healthy adult telephysiotherapy. Additionally, most systems evaluated in the included studies were still in early development stages, lacking validation in routine practice. For digital health technologies to be integrated meaningfully into physiotherapy, future efforts must ensure not only technical feasibility but also clinical relevance, user-centered design, and regulatory compliance. This includes establishing protocols for data accuracy, privacy, and therapist oversight when using automated assessment and feedback tools [7].

Limitations

While the review highlighted technical innovation and promising concepts, it also exposed several limitations in the current literature. The scope of most studies was narrow, with limited demographic diversity and small participant numbers [12-15]. Longitudinal data were lacking, and many systems were only tested in highly controlled or simulated environments [14,15]. As a result, generalizability remains limited, and the effectiveness of these technologies in real-world settings is still largely unknown [13].

Another important limitation of this review is the exclusion of computer science–oriented publication venues, particularly ACM journals and conference proceedings (eg, Interactive, Mobile, Wearable and Ubiquitous Technologies), which frequently report state-of-the-art advances in PA recognition and wearable sensing. As this review prioritized clinically oriented, peer-reviewed health databases, some advanced algorithmic developments may not be represented. Future interdisciplinary reviews should integrate biomedical and human-computer interaction literature to provide a more comprehensive overview of technological innovation and readiness.

The selection of databases represents another limitation of this review. While PubMed, ProQuest, ScienceDirect, and Wiley Online Library were chosen to ensure coverage of biomedical and rehabilitation-focused literature, additional databases such as CINAHL, Embase, IEEE Xplore, Web of Science, and Scopus may have captured a broader range of AI and machine vision–related studies. The restricted database selection, combined with other eligibility criteria, likely contributed to the small number of included studies and should be considered when interpreting the scope of the findings.

Comparison With Prior Work

The findings of this scoping review partially align with previous research in adjacent areas such as telerehabilitation, digital physiotherapy, and AI-supported movement analysis. Prior systematic reviews have highlighted the increasing integration of AI and ML into rehabilitation technologies, particularly for clinical populations such as patients recovering from stroke, orthopedic surgery, or neurological conditions [18]. However, in contrast to those studies, this review focused specifically on healthy adults as a target group for secondary prevention and behavior change, rather than posttreatment rehabilitation. Previous literature has also noted the effectiveness of telehealth platforms in improving access to care, enabling remote monitoring, and supporting adherence to rehabilitation protocols [20]. However, most of these platforms rely on conventional video calls or basic sensor data, without the incorporation of intelligent systems capable of processing complex movement patterns in real time. This review identified a small number of

XSL•FO
RenderX

studies that went beyond traditional telephysiotherapy, using AI or machine vision to automate analysis, personalize feedback, or predict PA levels. These applications, though promising, are still relatively novel and not yet widely implemented or evaluated at scale. Moreover, the lack of statistically significant outcomes in the reviewed studies contrasts with findings from other digital health research, where AI-supported interventions have demonstrated improvements in chronic disease management or physical function [21]. This discrepancy may be explained by differences in study design, participant populations, and the early-stage development of many AI and machine vision tools in the PA domain. The review also supports earlier conclusions that human-technology interaction remains a critical factor in the success of digital interventions [21]. While some prior work has emphasized the potential of gamification or virtual coaching to boost engagement, the studies included in this review rarely addressed motivational strategies or user experience design, underscoring a persistent gap between technical capabilities and behavioral outcomes [12].

Despite the high reported classification accuracy in controlled or semicontrolled settings, the reviewed studies provide limited evidence regarding the robustness of AI and machine vision systems in real-world home environments. Practical challenges, such as variable lighting conditions, occlusion, sensor placement inconsistency, limited camera viewpoints, and user adherence, were rarely addressed. Moreover, high technical performance did not translate into measurable reductions in sedentary behavior or sustained increases in PA, indicating a gap between algorithmic feasibility and behavioral effectiveness.

Future Directions

Despite the high reported classification accuracy in controlled or semicontrolled settings, the reviewed studies provide limited evidence regarding the robustness of AI and machine vision systems in real-world home environments. Practical challenges such as variable lighting conditions, occlusion, sensor placement inconsistency, limited camera viewpoints, and user adherence were rarely addressed. Moreover, high technical performance did not translate into measurable reductions in sedentary behavior or sustained increases in PA, indicating a gap between algorithmic feasibility and behavioral effectiveness.

The findings of this scoping review suggest that future research should prioritize investigating how AI and machine vision technologies can not only monitor but also actively promote PA participation in healthy adult populations. While several studies have demonstrated the technical feasibility of these tools, limited research has examined their direct impact on actual behavior change, such as increased daily PA or reduced sedentary time [5]. Further exploration is needed into the integration of intelligent, interactive features—such as real-time feedback, adaptive difficulty, or personalized goal setting—within telephysiotherapy platforms. These elements have the potential to increase user engagement, support long-term adherence, and enhance motivation through continuous remote guidance [2]. As seen in adjacent fields, systems incorporating behavioral science and digital interaction strategies, including gamification or computer vision–based exercise feedback, may be more effective than passive monitoring alone [7]. It is also important that future studies evaluate user perspectives, including acceptance, usability, and trust in AI-driven physiotherapy tools. Understanding how individuals engage with and respond to these technologies in real-world settings is essential for designing interventions that are not only technically robust but also meaningful and sustainable [22]. In parallel, research should address clinical integration pathways, particularly the role of physiotherapists in overseeing or comanaging AI-supported remote programs. Moreover, the review identified a lack of standardized protocols for digital PA assessment in the context of AI and telephysiotherapy. Developing unified guidelines for system design, data interpretation, and outcome evaluation would improve comparability across studies and support evidence-based implementation into clinical and preventive care frameworks [15].

Conclusions

This scoping review examined the current evidence on the use of AI and machine vision in telephysiotherapy to promote PA and reduce sedentary behavior among healthy adults. Although the findings demonstrated the technical feasibility of integrating intelligent systems into remote physiotherapy, the number of eligible studies was limited, and most were early-stage feasibility or pilot trials. While these technologies showed potential for personalized feedback, automated monitoring, and behavior support, there was insufficient evidence on their long-term effectiveness or impact on PA outcomes. The review highlights the need for more rigorous research involving diverse populations, standardized assessment protocols, and real-world implementation. With further development and validation, AI and machine vision could offer valuable tools to enhance preventive physiotherapy, support active lifestyles, and improve accessibility to personalized care beyond traditional clinical settings.

Acknowledgments

The authors extend their gratitude to the Rīga Stradiņš University librarian.

Funding

The authors declared no financial support was received for this work.

XSL•FO
RenderX

Authors' Contributions

All authors, during the brainstorming session, conceptualized and formulated the research questions. GS and LR led the data collection methods, but all authors were involved in data collection and analysis. LR and GS conducted a 3-step search process to identify relevant peer-reviewed literature studies. To ensure that the search yielded relevant studies, variations of the search strategy were pilot-tested by authors LR and LA and refined before the final search was conducted. All authors, during the brainstorming session, conceptualized the eligibility criteria. LR developed the data charting form. All authors conducted data charting of the sample of studies selected for inclusion. LR and GS synthesized the data, and LA performed quality checks on the reported results. LR led the manuscript development. All coauthors had the opportunity to read, edit, and approve the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full search strategies.
[DOCX File, 14 KB - rehab_v13i1e80459_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 181 KB - rehab_v13i1e80459_app2.pdf]

References

1. Adhikari S, Sigdel D. Prevalence of overweight and obesity and its associated factors among adult people at Chitwan, Nepal. IOSR J Nurs Health Sci 2020;9:30-41. [doi: 10.9790/1959-0905013041]
2. Cottrell MA, Galea OA, O'Leary SP, Hill AJ, Russell TG. Real-time telerehabilitation for the treatment of musculoskeletal conditions is effective and comparable to standard practice: a systematic review and meta-analysis. Clin Rehabil 2017 May;31(5):625-638. [doi: 10.1177/0269215516645148] [Medline: 27141087]
3. World health statistics 2024: monitoring health for the SDGs, sustainable development goals. : World Health Organization; 2024 URL: https://iris.who.int/server/api/core/bitstreams/74b12494-f213-4b5b-9533-18442147e1fb/content [accessed 2026-04-29]
4. Rekomendācijas fizisko aktivitāšu veicināšanai latvijas iedzīvotājiem (politikas veidotājiem un īstenotājiem) [Report in Latvian]. : Latvijas Republikas Veselības ministrija; 2019 URL: https://www.vm.gov.lv/lv/media/6408/download [accessed 2026-04-29]
5. Park JH, Moon JH, Kim HJ, Kong MH, Oh YH. Sedentary lifestyle: overview of updated evidence of potential health risks. Korean J Fam Med 2020 Nov;41(6):365-373. [doi: 10.4082/kjfm.20.0165] [Medline: 33242381]
6. Brennan DM, Mawson S, Brownsell S. Telerehabilitation: enabling the remote delivery of healthcare, rehabilitation, and self management. Stud Health Technol Inform 2009;145:231-248. [Medline: 19592797]
7. Calabrò RS, Mojdehdehbaher S. AI-driven telerehabilitation: benefits and challenges of a transformative healthcare approach. AI 2025;6(3):62. [doi: 10.3390/ai6030062]
8. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
9. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
10. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
11. The use of machine vision and artificial intelligence in tele-physiotherapy to reduce sedentary lifestyles and promote physical activity: a scoping review. OSF. URL: https://doi.org/10.17605/OSF.IO/SFC97 [accessed 2026-04-29]
12. Chatterjee A, Pahari N, Prinz A, Riegler M. Machine learning and ontology in eCoaching for personalized activity level monitoring and recommendation generation. Sci Rep 2022 Nov 18;12(1):19825. [doi: 10.1038/s41598-022-24118-4] [Medline: 36400793]
13. Du C. Deep-learning–based video analysis for human action evaluation [Dissertation]. : University of California; 2022 URL: https://db.rsu.lv/login?url=https://www.proquest.com/dissertations-theses/deep-learning-based-video-analysis-human-action/docview/2761184410/se-2?accountid=32994 [accessed 2026-04-29]
14. Boyer P. Machine learning at-home physiotherapy adherence [Dissertation]. : University of Toronto; 2023 URL: https://db.rsu.lv/login?url=https://www.proquest.com/dissertations-theses/machine-learning-at-home-physiotherapy-adherence/docview/2889795235/se-2?accountid=32994 [accessed 2026-04-29]

XSL•FO
RenderX

15. Mennella C, Maniscalco U, De Pietro G, Esposito M. A deep learning system to monitor and assess rehabilitation exercises in home-based remote and unsupervised conditions. Comput Biol Med 2023 Nov;166:107485. [doi: 10.1016/j.compbiomed.2023.107485] [Medline: 37742419]
16. Chung JL, Ong LY, Leow MC. Comparative analysis of skeleton-based human pose estimation. Future Internet 2022;14(12):380. [doi: 10.3390/fi14120380]
17. Jolly A, Pandey V, Sahni M, Leon-Castro E, Perez-Arellano LA. Modern smart gadgets and wearables for diagnosis and management of stress, wellness, and anxiety: a comprehensive review. Healthcare (Basel) 2025 Feb 14;13(4):411. [doi: 10.3390/healthcare13040411] [Medline: 39997286]
18. Yeh SJ, Wang YC, Fang WC, Huang SC, Yang YS. Effectiveness of powered hand exoskeleton on upper extremity function in people with chronic stroke. Actuators 2025;14(2):67. [doi: 10.3390/act14020067]
19. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019 Jun;6(2):94-98. [doi: 10.7861/futurehosp.6-2-94] [Medline: 31363513]
20. Snoswell CL, Chelberg G, De Guzman KR, et al. The clinical effectiveness of telehealth: a systematic review of meta-analyses from 2010 to 2019. J Telemed Telecare 2023 Oct;29(9):669-684. [doi: 10.1177/1357633X211022907] [Medline: 34184580]
21. Capecci M, Cima R, Barbini FA, et al. Telerehabilitation with ARC intellicare to cope with motor and respiratory disabilities: results about the process, usability, and clinical effect of the “Ricominciare” pilot study. Sensors (Basel) 2023 Aug 17;23(16):7238. [doi: 10.3390/s23167238] [Medline: 37631774]
22. Burdea G, Kim N, Polistico K, et al. Assistive game controller for artificial intelligence–enhanced telerehabilitation post-stroke. Assist Technol 2021 May 4;33(3):117-128. [doi: 10.1080/10400435.2019.1593260] [Medline: 31180276]

Abbreviations

AI: artificial intelligence
CoP: center of pressure
ML: machine learning
PA: physical activity
PCC: population, concepts, and context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
WHO: World Health Organization

Edited by S Munce; submitted 11.Jul.2025; peer-reviewed by X Su, Y Xiao; revised version received 04.Mar.2026; accepted 23.Mar.2026; published 15.Oct.2026.

Please cite as:
Redliha L, Semjonova G, Araka L
Leveraging AI and Machine Vision in Telephysiotherapy for Promoting Physical Activity and Reducing Sedentary Behavior: Scoping Review
JMIR Rehabil Assist Technol 2026;13:e80459
URL: https://rehab.jmir.org/2026/1/e80459
doi:10.2196/80459

© Liva Redliha, Guna Semjonova, Liva Araka. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 15.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Indoor Navigation for People With Visual Impairment in Canada: Participatory Co-Design and Interdisciplinary Study of the Edge A-Eye Platform

Soutongnoma Safiata Kaboré[1], MSc; Deborah Annan[1], OD; Prajjol Raj Puri[1], BOptom; Sergio Mejia Romero[1], PhD; Nathalie Gingras-Royer[1], MSc; Justin Gauthier[2], DEP; H ng Nhung[2], DEP; Phuong Anh Nguyen[2], PhD; Benoit Pelletier[2], PEng; Kim Khoa Nguyen[2], PhD; Hassane Alami[3], PhD; Joseph Paul Nemargut[1], PhD

[1]School of Optomtery, University of Montreal, 3744 Jean-Brillant Street, Montreal, QC, Canada

[2]Department of Electrical Engineering, School of Higher Technology, Montreal, QC, Canada

[3]Public Health School, University of Montreal, Montreal, QC, Canada

Corresponding Author:
Soutongnoma Safiata Kaboré, MSc
School of Optomtery, University of Montreal, 3744 Jean-Brillant Street, Montreal, QC, Canada

Abstract

Background: Indoor navigation remains a major challenge for people with visual impairments, affecting autonomy, safety, and quality of life. While navigation-based assistive technologies leveraging artificial intelligence and mobile computing show promise, their adoption remains limited due to insufficient user-centered and context-aware design.

Objective: This study aimed to co-design Edge A-Eye, an inclusive artificial intelligence–powered indoor navigation platform for people with visual impairment in Canada, using a participatory and interdisciplinary approach grounded in accessibility, rehabilitation, and human-centered design principles.

Methods: From May 2024 to December 2025, we conducted research using an 8-step iterative co-design process. A total of 12 people with visual impairment and 2 specialists in vision rehabilitation participated in bilingual focus groups and semistructured interviews. A co-creation workshop gathered 21 stakeholders (users, clinicians, researchers, and industry partners). Successive prototypes were evaluated in controlled environments and real-world public settings. Data were analyzed thematically using an inductive-deductive approach informed by human-centered and inclusive design frameworks.

Results: Continuous stakeholder engagement reinforced the relevance and usability of the evolving prototypes. Participant feedback guided major design decisions, such as optimizing iPhone (Apple Inc) compatibility, excluding the gyroscope due to usability and battery constraints, and integrating hands-free navigation. The process also addressed cross-cutting issues such as digital accessibility, cybersecurity, universal design, and equitable access for users with varying technological literacy.

Conclusions: The Edge A-Eye project demonstrates how a sustained participatory co-design approach can meaningfully shape assistive technology innovation. By integrating the expertise of people with visual impairment, clinicians, researchers, and industry partners at every stage, the resulting platform is better aligned with real-world needs, safety considerations, and user expectations. This work provides a replicable framework for developing inclusive, context-aware indoor navigation technologies.

(JMIR Rehabil Assist Technol 2026;13:e81347) doi:10.2196/81347

KEYWORDS

visual impairment; indoor navigation; AI-based assistive technology; co-design; accessibility; human-centered design; universal design; Canada; artificial intelligence; mobile phone

Introduction

Global Context

More than 2.2 billion people have impaired near or distance vision around the globe, and this number is projected to double by 2050 [1]. Visual impairment affects people of all ages, but most of those affected are older than 50 years [2]. In 2017, over 1,519,000 Canadians reported having a visual disability, accounting for 5.4% of the Canadian population [3]. These individuals encounter multiple daily barriers, including physical and functional obstacles that hinder their full participation in society [4]. The most frequently mentioned obstacles include reading and accessing written information [5], engaging in leisure activities and shopping [5], as well as using public transport [6]. Visual impairment imposes a significant burden globally, not only in human and social terms but also in relation to health and economic costs, with annual productivity losses estimated at US $411 billion [7]. As the global population

XSL•FO
RenderX

continues to age and the prevalence of noncommunicable diseases rises, it is estimated that by 2050, a total of 3.5 billion people will require access to assistive technologies (ATs) [8].

Importance of Navigation-Based AT

Navigating within an environment is a fundamental aspect of the human experience, enabling individuals to carry out daily activities and engage with their surroundings [9]. Navigation-based assistive technologies (NATs) are crucial for enhancing the functional abilities of people with visual impairments [10]. NATs play a vital role in maintaining or improving the functional capabilities of individuals with disabilities and significantly impact their quality of life [11]. NATs have proliferated in recent decades to enhance orientation and mobility (O&M) for individuals with blindness and low-vision [12]. These technologies are incredibly diverse, ranging from prototypes to commercially available solutions, including electronic travel aids, smartphone apps, sensing systems, wearables, Internet of Things devices, and artificial intelligence (AI)–powered assistants [12]. O&M specialists typically provide vision rehabilitation services, instructing individuals with little to no vision on the use of NATs to support independent travel [13]. These technologies can function independently or in conjunction with traditional aids such as white canes and guide dogs [10], providing spatial awareness, obstacle detection, situational guidance, and real-time information for both indoor and outdoor navigation [14]. Ideally, NATs are designed to support and enhance existing adaptive skills, including cognitive mapping, the interpretation of nonvisual cues, and familiar travel strategies [15]. As highlighted by Nemargut, O&M training equips people with visual impairments with the foundational skills needed to interpret spatial cues, build cognitive maps, and navigate safely and independently across diverse environments [16]. This underscores that the effectiveness of NATs relies not only on technological design but also on users' existing O&M competencies, which fundamentally shape how assistive tools are interpreted and applied in real-world contexts. Although important progress has been made in the development of AT for people with visual impairment, many devices continue to endure design approaches that do not adequately reflect the physical and social realities of users' daily lives [17]. This disconnect often results in technologies that fail to meet users' actual needs, leading to limited effectiveness and high rates of abandonment [15]. These shortcomings highlight the importance of grounding technological innovation in the lived experiences of those concerned. The principle of "nothing about us without us" should guide every stage of design, development, and implementation when technologies are intended for people with disabilities [17]. In addition to user involvement, contextual factors such as environmental, social, and cultural conditions strongly influence usability and adoption [6]. Variability in lighting, acoustic conditions, spatial organization, and social norms around receiving assistance can significantly shape navigation experiences [6]. These context-dependent factors justify the need for a flexible, adaptive, and holistic user-centered design (UCD) approach that accounts for diverse real-world conditions [6]. Importantly, this study represents the first multiphase, interdisciplinary co-design process for indoor navigation technology for people with visual impairment conducted across Canadian rehabilitation centers. By situating the research within Canada's technological and clinical landscape, we demonstrate its novelty and contribution beyond reiterating the importance of co-design. This work provides a replicable framework that integrates participatory design, inclusive technology development, and cross-sector collaboration to advance the field of assistive navigation.

Limitations of Existing Navigation-Based AT

Despite technological progress, many NATs remain underused due to poor alignment with users' lived contexts [17], leading to high abandonment rates. Solutions often fail because they do not adequately consider dynamic indoor environments, usability simplicity, or the importance of multisensory cues. Moreover, trust is quickly eroded when devices produce errors or require complex interactions, reinforcing the principle "nothing about us without us" as essential for relevance, usability, and acceptability. To reduce the frequent nonadoption and abandonment of NATs, three main approaches have recently gained prominence in the development process: UCD, participatory co-design, and transdisciplinary collaboration [17].

UCD Approach

The UCD approach in the development process of AT has grown in recent years [17]. UCD, and more specifically inclusive design, is an approach aimed at developing interfaces, artifacts, products, and services that are usable, appropriate, and accessible to the widest possible range of users, while respecting the constraints of the design specifications [18]. UCD is a method for creating usable products and systems, based on a set of techniques, processes, and methods that all place the user at the center of each stage [17]. Chavarria et al [19] highlight the importance of human-centered design (HCD) in AT development. UCD plays a key role in the development of ATs, as it prioritizes users' needs and contexts through continuous engagement. For instance, Tang et al [20] underscore the role of wearable multimodal systems as a key component of inclusive design strategies. This iterative approach ensures that ATs are practical, accessible, and effective [21]. Improving the implementation of UCD in AT development and strengthening the evidence base regarding its implementation and outcomes are crucial [17]. The ISO (International Organization for Standardization) 9241-210:2019 standard provides internationally recognized principles for UCD [22]. To better engage with the realities of people with visual impairment, it is important to enhance the capacity of development teams in UCD, its principles, and components, and to improve the planning of UCD implementation [17]. UCD is guided by six key principles: (1) an explicit understanding of users, their tasks, and environments; (2) continuous user involvement throughout the design and development process; (3) design decisions informed and refined by user-centered evaluation; (4) an iterative development process; (5) consideration of the overall user experience; and (6) a multidisciplinary design team bringing diverse skills and perspectives [17].

XSL•FO
RenderX

Participatory Co-Design Approach

Co-design is a meaningful engagement of the end user at all stages of the process [18]. Participatory co-design approaches are crucial for a thorough understanding and engagement of end users, ensuring more appropriate and user-friendly products and services [19]. Gathering user information and feedback at each stage of the design and development process is essential, involving all stakeholders: users, designers, and other relevant parties [21]. A participatory co-design team comprises a range of multidisciplinary skills and viewpoints [19]. The significance of individual contributions, as highlighted in the participatory design framework by Sanders [23], lies in the valuable input each stakeholder provides at every phase of the design process. These key players are vital sources of creativity and inspiration, generating relevant and innovative ideas. Considering the diversity of the target population should not be an afterthought, but rather a fundamental prerequisite for any design, integrated from the outset in the needs specifications [21]. This approach recognizes each user as a unique individual with their own abilities, experiences, and characteristics [21]. It considers the capabilities and limitations of all potential users. Technologies designed in this way will be usable by the widest possible range of individuals, across a variety of situations, maximizing reach and impact [21]. Beyond conventional participatory approaches, sociomaterial perspectives show that facilitators and constraints in co-design emerge from dynamic configurations between actors, artifacts, material environments, and organizational routines [24]. Rong and Hansopaheluwakan-Edward [24] demonstrate that in co-design projects involving people with visual impairments, certain constraints such as material properties, organizational rules, or interaction modalities can also act as facilitators depending on how they are configured. This sociomaterial perspective informs our own co-design mechanisms, including the selection of design artifacts, the structuring of workshop activities, and the governance model across interdisciplinary teams. It also justifies key methodological trade-offs implemented to ensure effective participation and sustained alignment between technical, clinical, and experiential considerations [24].

Transdisciplinary Approach

Transdisciplinarity opens the possibility for cross-fertilization among different fields of knowledge and collaboration with people with visual impairment as potential end users and with their stakeholders [17]. Cross-fertilization across engineering, social, and clinical sciences is necessary [17]. A multidisciplinary approach, integrating diverse expertise, ensures that user needs are met [17]. Considering universal design is of great significance as it aims to optimize product design for maximum accessibility and to ensure that mainstream design is accessible to everyone [17]. The goal is for this approach to provide outstanding solutions that function equally well for users with disabilities as for those with higher abilities [22]. Adherence to the principles of nonstigmatization and nondiscrimination is essential because users are unlikely to adopt, appreciate, or purchase products that stigmatize them or emphasize their disability; however, they possess the capability and expertise to suggest ways to mitigate such stigmatizing effects [21]. Multidisciplinary teams would not only address technology-related issues but also those pertaining to user functionality. Although AI has made significant advancements in indoor navigation for people with visual impairments, its integration into an intersectional and participatory approach remains largely unexplored. These approaches are often fragmented, lacking systematic integration of people with visual impairment's lived experiences and cross-disciplinary expertise, which reduces the impact and sustainability of innovations [25]. This study aims to address this gap by considering the diverse experiences and lived realities of users while bringing together a broad range of expertise, including researchers, rehabilitation professionals, and industry stakeholders, to ensure a truly inclusive technology that meets varied needs. Hence, the interest in conducting this study.

Research Goal and Objectives

Our research aims to develop an inclusive and user-centered indoor navigation platform that effectively addresses the needs of people with visual impairment in Canada. Specifically, it is about (1) conducting qualitative research to understand the specific navigation challenges faced by people with visual impairment in various indoor environments; (2) implementing a UCD approach that ensures continuous user involvement at all stages of the platform's development, allowing for iterative feedback and refinement; (3) engaging stakeholders, including users, designers, and academics, in a collaborative workshop to gather insights and ideas that drive the design process, ensuring that the platform is user-friendly and accessible; (4) fostering cross-disciplinary partnerships among a range of fields (engineering, social sciences, and clinical sciences) to ensure the design process incorporates diverse perspectives and expertise; (5) testing and evaluating the platform in controlled and real-world environments with real users, collecting feedback to assess usability, functionality, and overall user experience; and (6) sharing the research outcomes and best practices for developing NATs with the broader community, including industry partners and policymakers, to promote wider implementation. By achieving these objectives, the research aims to deliver a robust indoor navigation solution that enhances the independence and quality of life of people with visual impairment.

Methods

Overview

The Edge A-Eye project used a participatory co-design methodology in which people with visual impairment, specialists in vision rehabilitation, engineers, human-centered computing experts, AI, computer science researchers, as well as community partners collaborated throughout all stages of development. This approach ensured that experiential knowledge, scientific evidence, and technological expertise were fully integrated into an inclusive and ethically grounded innovation process, consistent with the mechanisms of mutual learning and shared decision-making described in participatory co-design literature [26]. The sustainability of digital health innovations depends on the continuous engagement of stakeholders, the active involvement of users at every stage, and stable collaborative governance [27]. The participatory structure prioritized

XSL•FO
RenderX

transparency, responsibility, interdisciplinarity, and direct responsiveness to user-identified needs, in alignment with best practices in equitable, human-centered, and interdisciplinary design processes [28].

Ethical Considerations

Our study adheres to the ten guiding principles outlined in the Montreal Declaration on the Ethics of AI, namely: well-being, respect for autonomy, protection of privacy and personal data, solidarity, democratic participation, equity, inclusion of diversity, caution, responsibility, and sustainable development [29]. We adhered to ethical principles, in accordance with the Declaration of Helsinki [30], including approvals from the CERC (Clinical Research Ethics Committee) of the University of Montreal (CERC-2022 - 1800; January 13, 2023), CERC-2024 - 2051 (May 3, 2024), and the Centre de Recherche Interdisciplinaire en Réadaptation du Montréal Métropolitain (CRIR MP-50-2024-51; May 3, 2024). Participants provided free, voluntary, and informed consent; anonymity and data confidentiality were ensured. Source documents were accessible via VoiceOver (ResonanceVista) on adapted computers or smartphones. All data collected were anonymized and documented in a Microsoft Excel (Microsoft Corp) file coded for deidentification, while signed consent forms were securely stored in a password-protected folder, thereby guaranteeing confidentiality and adherence to ethical requirements. These measures are essential to ensure the integrity of the research and the respect for ethical values [31]. Participants received compensation of CAD $15 per hour (a currency exchange rate of CAD $1=US $0.79 was applicable).

Research Team Composition and Participation Structure

The Edge A-Eye project was spearheaded by a multidisciplinary consortium bringing together four academic institutions and several industry partners: the School of Optometry and the School of Public Health at Université de Montréal, École de Technologie Supérieure (ÉTS), in collaboration with Mathematics of Information Technology and Complex System, IVADO (Institut de Valorisation des Données), and Broadcom. This alliance integrated expertise in vision science, public health, AI, human-centered computing, computer science, and community engagement, ensuring both methodological rigor and technical relevance throughout all stages of development [26,32]. To translate participatory design principles into practice, we developed a participation matrix mapping the roles and contributions of each stakeholder group (people with visual impairment, vision rehabilitation specialists, researchers, and industry partners) across the eight co-design steps (Table 1). This structured approach demonstrates that people with visual impairment were not passive informants but active co-designers, contributing experiential knowledge to feature prioritization, prototype testing, and iterative refinements. Rehabilitation specialists ensured clinical integration, while industry partners provided insights on technical feasibility and scalability. Researchers coordinated activities, managed data analysis, and upheld ethical standards [32,33]. The leadership team comprised a professor of vision science (JPN), a professor of public health (HA), an industry expert (BP), and a professor of technological engineering (KKN). They were supported by five vision science researchers (SSK, PRP, DA, SMR, and NG-R) and three engineering researchers (PAN, JG, and HN), who contributed to participant recruitment, study coordination, and prototype development. Weekly coordination meetings fostered collaborative and iterative decision-making, reinforcing the participatory nature of the current project [34]. Partnerships with community organizations and rehabilitation centers, including the CRIR, CNIB (Canadian National Institute for the Blind) Foundation, ASAMM (Association des Sports Pour Aveugles du Montréal Métropolitain), Institut Nazareth et Louis-Braille (INLB), Lethbridge-Layton-Mackay Rehabilitation Center (LLMRC), and RAAMM (Regroupement des Aveugles et Amblyopes du Montréal Métropolitain), ensured the integration of lived experiences and promoted equity and inclusion. This framework guaranteed that empowerment was not symbolic but embedded in the design process, fully aligned with the principle of “nothing about us without us” [32,33]. Importantly, people with visual impairment were not only consulted as informants but actively engaged as co-designers of the Edge A-Eye platform, in line with established participatory design principles emphasizing end users’ involvement and empowerment [32,33]. This continuity enriched the depth and consistency of experiential insights, reduced variability, and fostered trust among stakeholders. It also enabled cumulative learning, where feedback from earlier stages informed subsequent design decisions, thereby strengthening ecological validity and usability of the Edge A-Eye platform [18,21]. Implementing a participatory co-design methodology within a large interdisciplinary consortium required active coordination to manage expectations and ensure effective communication across diverse teams. Differences in disciplinary vocabularies and priorities between technical, clinical, and user partners occasionally led to misalignment regarding feasibility and development timelines. These challenges were addressed through structured coordination mechanisms, including regular interdisciplinary meetings and iterative validation checkpoints, which facilitated clarification of terminology, alignment of objectives, and transparent decision-making throughout the development process. Table 1 presents the participation matrix.

XSL•FO
RenderX

Table . Participation matrix.

Co-design stage	People with visual impairment	Rehabilitation specialists	Researchers	Industry partners
Needs assessment (focus groups and interviews)	Shared lived experiences; identified barriers and facilitators	Validated clinical relevance	Facilitated sessions; analyzed data	Not available
Prioritization of features (workshop)	Ranked priorities; proposed functionalities	Linked priorities to rehabilitation practices	Synthesized themes; documented decisions	Provided feasibility insights
Preliminary design	Reviewed wireframes; suggested accessibility improvements	Ensured ergonomic and clinical integration	Developed prototypes	Assessed technical constraints
Iterative development	Tested prototypes; provided usability feedback	Recommended adjustments for safety	Implemented refinements	Optimized performance
Controlled testing	Evaluated navigation tasks in laboratory settings	Observed and advised on mobility strategies	Collected and analyzed data	Supported technical troubleshooting
Real-world testing	Simulated navigation in authentic environments	Monitored functional outcomes	Coordinated logistics; analyzed results	Assisted with deployment
Dissemination of results	Shared experiential insights at events	Advocated for clinical adoption	Presented findings at conferences	Explored commercialization pathways
Ongoing evaluation	Continuous feedback for improvements	Ensured alignment with rehabilitation standards	Maintained iterative updates	Planned scalability

Study Design

This study (May 15, 2024, to December 5, 2025) adopted a participatory, intersectoral co-design approach to develop a NAT tailored to the needs of people with visual impairment in Canada. People with visual impairment were actively engaged alongside vision rehabilitation specialists, academics, and industry stakeholders to ensure that the technology was both functional and responsive to real-world challenges. The current project followed the principles of transdisciplinary AT development [35] and UCD [17], which emphasize users' needs, motivations, and behaviors throughout the design process [36]. To minimize bias, no specific technology was introduced during the current study, allowing participants to consider a range of solutions, including human assistance or physical accessibility measures. The reference framework used is Bird et al [37] because it proposes a generative co-design approach centered on users, which is essential for integrating the real needs of people with visual impairment into technological development. It is structured into three phases: predesign (mobilization and understanding of needs), co-design (collaborative creation and uncovering of latent needs), and postdesign (validation and recommendations). This model promotes the acceptability and adoption of health innovations through the active participation of stakeholders [37]. The current study's methodology and findings were thoroughly documented in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [38], presented in Checklist 1. This framework ensures comprehensive reporting of essential elements, including details about the research team, study design and procedures, setting, results, data analysis, and interpretation [38]. Structured into eight steps outlined below, the process ensured that the perspectives of people with visual impairment remained central throughout, while promoting strong collaboration and sustained engagement among all stakeholders [33,36].

Integrated Scientific Reference Frameworks

The Edge A-Eye project is built on three complementary methodological pillars to ensure inclusive and sustainable innovation:

1. Technology acceptance model (Davis [39]): this model was applied to examine factors influencing technology adoption by identifying both barriers and facilitators. This approach enabled a clear understanding of actual user needs and their translation into design requirements for an adapted indoor navigation system, thereby enhancing acceptability and reducing the risk of abandonment. The model emphasizes two key determinants of technology acceptance: perceived usefulness and ease of use, which strongly shape the integration of ATs into everyday life.
2. Generative co-design framework (Bird et al [37]): this framework structured the participatory and cross-sector process, ensuring continuous involvement of end users and stakeholders. It guarantees that design decisions reflect real needs and lived contexts through prototypes tested in both laboratory and real-world conditions.
3. Technology readiness level (TRL; KTH Innovation [40]): the TRL framework was explicitly used as a development governance tool to structure decision-making, progression criteria, and validation milestones throughout the Edge A-Eye project. Rather than serving solely as a post hoc indicator of technological maturity, the TRL framework was integrated prospectively to guide the sequencing of design, testing, and evaluation activities. Each transition between TRL stages was informed by predefined feasibility criteria, user feedback, and multidisciplinary consensus involving researchers, clinicians, industry partners, and people with visual impairment. This governance approach ensured that development decisions were grounded in empirical evidence, usability outcomes, and contextual feasibility, while preventing premature claims of maturity

XSL•FO
RenderX

or scalability. The TRL framework thus functioned as a structuring and regulatory mechanism supporting iterative refinement, risk management, and transparency across this project's lifecycle. The operationalization of this governance framework and its application across successive TRL stages are detailed in the ongoing monitoring and evaluation section and summarized visually in Figures 1 and 2.

Figure 1. Ongoing monitoring and evaluation of the Edge A-Eye project. :

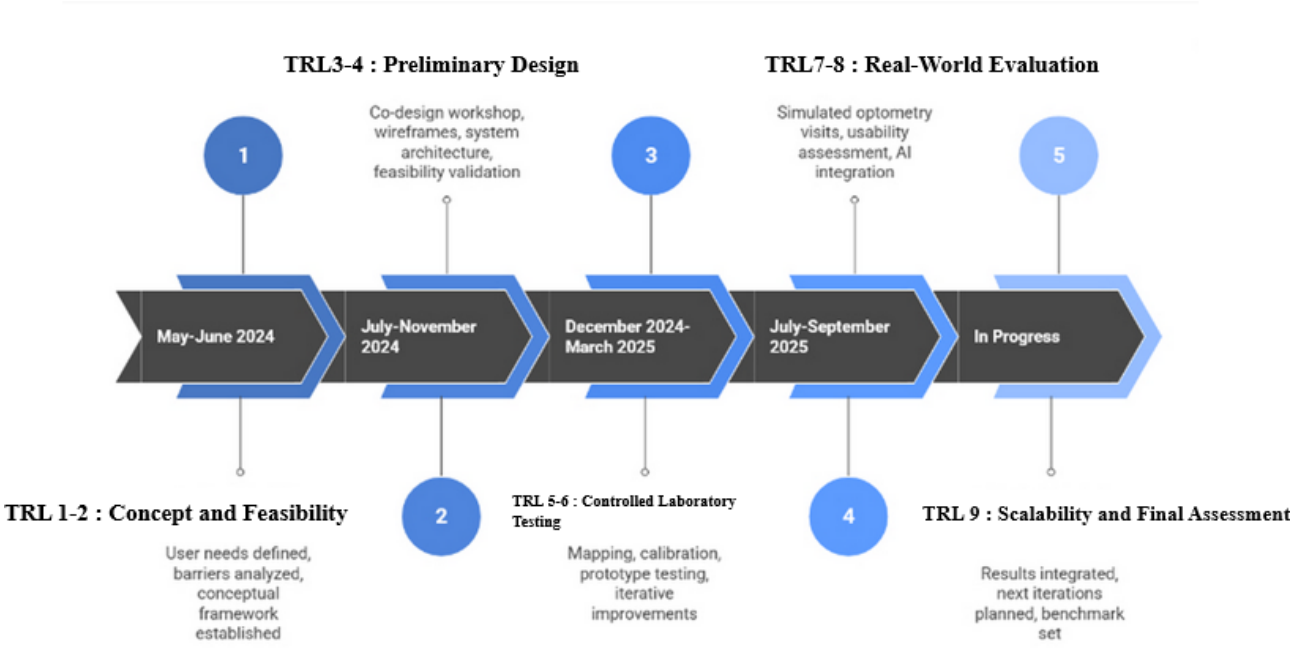
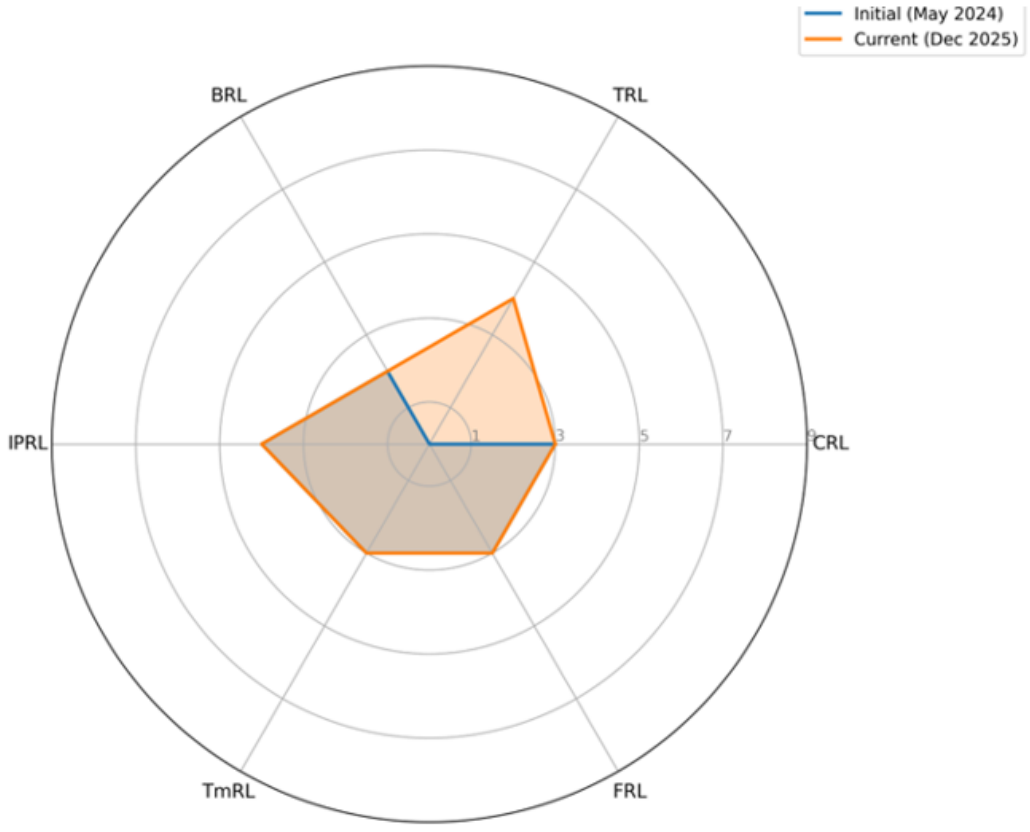


Figure 2. TRL progression across 55 axes, comparing the initial state (May 2024) versus the current state (December 2025). BRL: Business Readiness Level ; CRL: Commercial Readiness Level ; FRL: Financial Readiness Level ; IPRL: Intellectual Property Readiness Level :TMRL: Team/Management Readiness Level ;TRL: *Technology Readiness Level.*



XSL•FO
RenderX

In summary, combining the technology acceptance model for acceptability, Bird for co-design, and TRL for maturity creates an integrated logic chain that strengthens the relevance, robustness, and sustainability of the innovation.

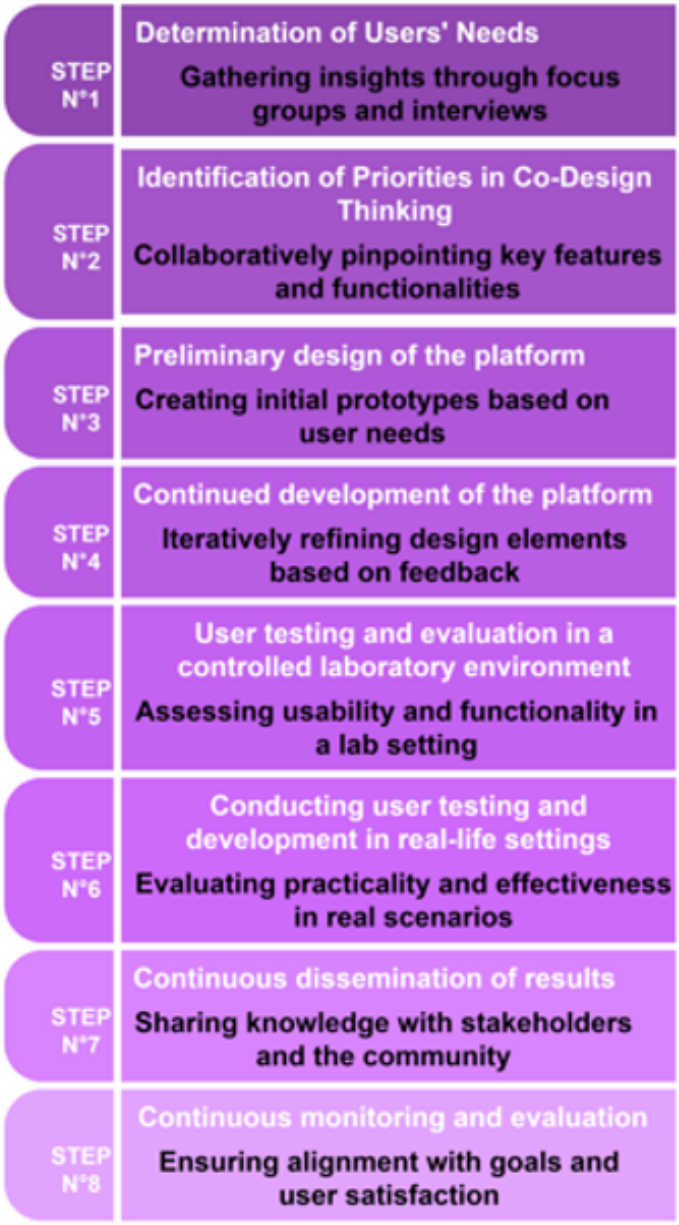
Steps of the Co-Design Process

This study resulted in the development of eight structured co-design steps, as shown in Figure 3. The structured methodology, aligned with scientific rigor and ethical standards, yielded these key steps. First, regarding the determination of users' needs, focus groups and interviews provided essential qualitative insights into the lived experiences, mobility challenges, and expectations of people with visual impairment. This foundational step ensured that the platform was grounded in real user contexts [36]. Second, regarding the identification of co-design priorities, stakeholders collaboratively identified key features, navigation challenges, and accessibility requirements. This prioritization ensured alignment with user preferences and everyday mobility needs [35,37]. Third, regarding the preliminary platform design, initial prototypes were developed based on users' expressed needs. Early conceptualization enabled rapid feedback and ensured that proposed solutions matched user expectations [36]. Fourth, regarding the iterative platform development, ongoing user feedback guided continuous refinement of functionalities, interface elements, and navigation cues. This iterative process reinforced the adaptability and responsiveness of the design [37]. Fifth, regarding the controlled-environment testing, user evaluations conducted in a laboratory environment yielded important insights into usability and interface functionality, supporting improvements before real-world testing [35]. Sixth, regarding the real-world testing, field testing confirmed the practical relevance, accessibility, and feasibility of the platform in authentic indoor environments. Real usage scenarios provided rich data on actual navigation behaviors and enhanced ecological validity [25,35]. Seventh, regarding the continuous dissemination of results, regular sharing of findings with partners, stakeholders, and community organizations strengthened collaboration and supported transparency throughout the development process [41]. Eighth, regarding the ongoing monitoring and evaluation, continuous assessment ensured alignment with user expectations and project objectives. Iterative adjustments maintained platform relevance and enhanced long-term user satisfaction [37].

37

XSL•FO
RenderX

Figure 3. Structured steps of the co-design process.



To make sprint decisions traceable, user-reported feedback was prioritized using a three-factor rubric: (1) frequency (number of participants or scenarios affected across sessions), (2) severity (risk to safety or task completion; high potential collision, misrouting, or loss of situational awareness), and (3) usability impact (cognitive load, clarity of cues, and effort required to operate hands-free). Items with high severity were addressed first, even if low in frequency; ties were resolved by usability impact. Each sprint started with a prioritization board (backlog review) and ended with a verification step (targeted retesting of high-priority fixes in the laboratory, then real-world scenarios). Representative examples include refining turn-by-turn prompts at decision points to lower cognitive load (medium severity or high impact).

Application of ISO 9241-210:2019 Principles

The co-design methodology adopted in this study was explicitly aligned with the six principles outlined in ISO 9241-210:2019, which provides internationally recognized guidance for HCD of interactive systems [22]. These principles include (1) an explicit understanding of users, tasks, and environments; (2) active user involvement throughout design and development; (3) iterative design processes; (4) design driven and refined by user-centered evaluation; (5) consideration of the entire user experience; and (6) a multidisciplinary design team. In practice, these principles were operationalized through continuous engagement of people with visual impairment and visual rehabilitation specialists across all eight co-design stages, iterative prototype testing in both controlled and real-world settings, and integration of expertise from engineering, vision

XSL•FO
RenderX

science, and human-centered computing. This alignment ensured that the Edge A-Eye platform was developed in accordance with global standards for usability, accessibility, and inclusivity.

Participant Recruitment and Eligibility

A purposive sampling strategy was used to recruit both people with visual impairment and vision rehabilitation specialists. This approach aligns with qualitative research standards emphasizing the involvement of key informants whose experiential knowledge is essential for co-design processes [42]. Vision rehabilitation specialists were selected for semistructured interviews conducted in French, with inclusion criteria requiring that participants be practicing rehabilitation specialists residing in Montréal. Recruitment of people with visual impairment followed three complementary components: (1) French- and English-language focus groups, simulating indoor navigation in grocery store and clinic scenarios to identify barriers and facilitators; (2) a co-design workshop at IVADO, where users collaboratively prioritized needs and defined core functionalities of the navigation platform; and (3) real-life scenario simulations at the LLMRC, evaluating feasibility, acceptability, and user experience in authentic environments. Information sheets and electronic consent forms were distributed via email through partner organizations (CRIR, CNIB Foundation, ASAMM, INLB, and RAAMM) and posted on social media. Inclusion criteria for people with visual impairment required being aged ≥18 years, residence in Canada, self-identification as visually impaired (including low vision, blindness, and deafblindness), semi-independent travelers, regular smartphone use for at least one year, ability to communicate in French or English, basic internet and Zoom (Zoom Communications, Inc) proficiency, stable internet access, absence of intellectual or cognitive disabilities, and voluntary written consent to participate. To ensure diversity and representativeness, the sample considered variations in age, gender, ethnic and cultural background, and type of vision loss. Individuals who did not meet the eligibility criteria were excluded. Participants who were determined eligible by the research team were contacted about their interest in participating in additional research by the laboratory. They were selected at each stage to prioritize a diversity of profiles. An estimated sample size of 13 participants was determined for each stage, based on similar qualitative studies conducted in this field. However, data collection continued until saturation was reached. Participant continuity was ensured across the different phases. A core group of 5 participants (4 people with visual impairment and 1 specialist for people with visual impairment or vision-rehabilitation) was maintained throughout the process, which allowed for improved iterative refinement and ensured coherence in design decisions.

Data Collection

The data collection strategy combined multiple instruments to ensure a comprehensive Edge A-Eye platform during the co-design process. Together, all tools provided a structured framework that covered the entire co-design cycle. By integrating qualitative approaches, the methodology enabled a robust triangulation of findings, thereby strengthening the validity and reliability of the study outcomes.

Focus Group Sessions

To provide initial insights into the needs of people with visual impairment in our current Edge A-Eye project (to inform this study), 2 web-based focus groups were conducted between February and June 2023 to explore barriers and facilitators related to indoor navigation for people with visual impairments. One session was held in English and the other in French via Zoom, each lasting approximately 2 hours. Moderators followed a pretested guide to ensure consistency and encouraged participants to share experiences or imagine completing tasks in 5 everyday scenarios: coffee shop, clinical visit, big-box store, bus travel, and party.

Sessions were recorded, transcribed (English automated and French manual), and validated. Themes were coded independently, with discrepancies resolved by JPN. A total of 2 weeks later, participants ranked the identified themes by importance via email, enabling prioritization of barriers and facilitators for each scenario. The results that informed the initial stages of the Edge A-Eye project were recently published by the laboratory [43].

Semistructured Interviews

On July 24, 2024, a total of 2 vision rehabilitation specialists with visual impairments participated in semistructured interviews conducted via the Zoom platform. One of them joined the research team at a later date to collaborate closely throughout all stages of the Edge A-Eye project. Each session lasted approximately 2 and a half hours and allowed participants to express their perspectives in either language while discussing simulations of 2 scenarios: navigating in a grocery store and attending a clinic appointment. Their expertise provided valuable insights into the barriers and facilitating factors associated with the use of ATs for indoor navigation in Canada, thereby contributing to a deeper understanding of the real needs of people with visual impairment. The interviews were facilitated by SSK, who ensured the structured flow of the interview, managed time effectively, and took detailed notes. A pretested semistructured interview guide (Multimedia Appendix 1) was used to support consistency and rigor in data collection.

Co-Design Workshop

The co-design workshop was hosted on November 1, 2024, by IVADO in Montreal, Quebec. Building on insights gathered from interviews and focus groups, the research team designed and pretested a semistructured interview guide (Multimedia Appendix 2) to frame and support the co-design thinking process. The primary objective of this workshop was to prioritize the needs identified by people with visual impairment, ensuring that their experiential knowledge guided the development of the Edge A-Eye platform from the outset. The workshop brought together 22 participants, including 4 people with visual impairment of diverse ages, genders, and conditions, 1 vision rehabilitation specialist, 4 vision science researchers, 11 technology design researchers, 1 representative from the IVADO, and 1 industry stakeholder. All stakeholders facilitated constructive discussions. A structured brainstorming and prioritization process was used to identify the key features and functionalities most relevant to people with visual impairment.

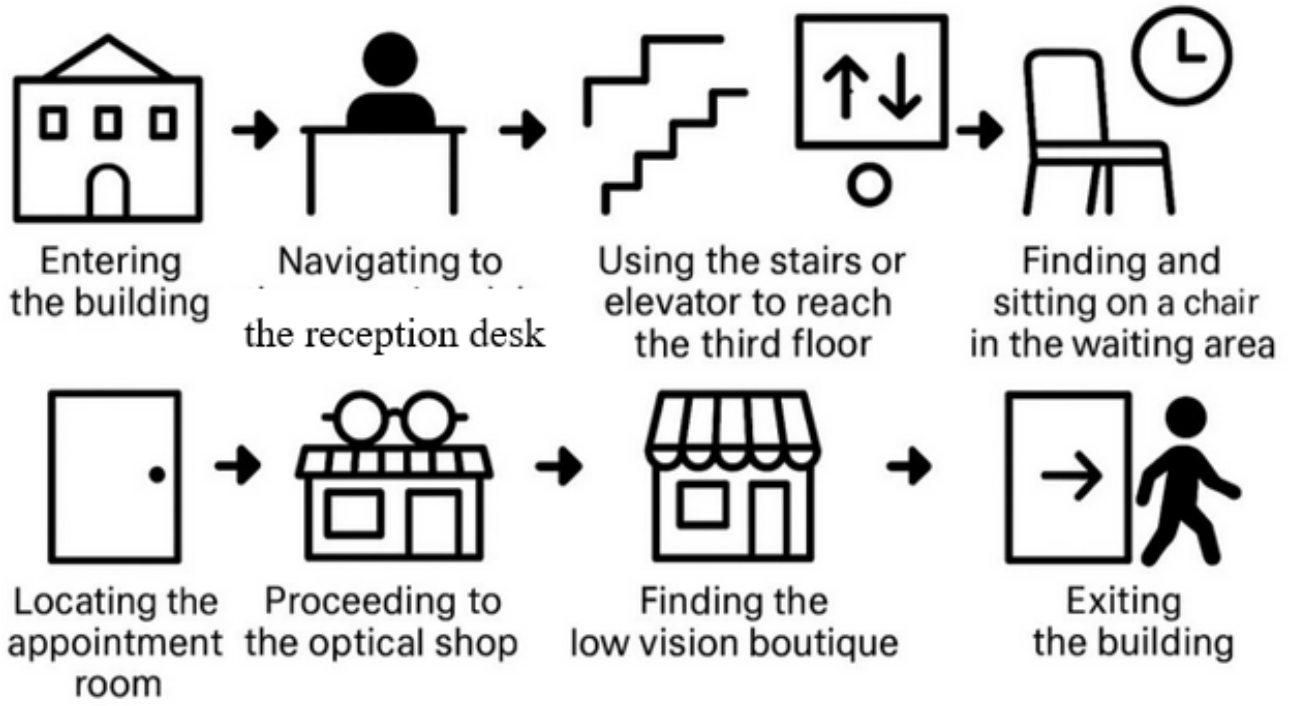
XSL•FO
RenderX

Notetakers were assigned to each table, which included at least 1 visually impaired participant, 1 industry representative, and 1 academic researcher. Each team generated lists of priorities in separate Microsoft Excel (Microsoft Corp) files, and designated presenters reported thematic categories to SSK. These themes were compiled and reviewed for accuracy by team members. The responses were then synthesized and summarized collaboratively by SSK, JPN, DA, PRP, and PAN, and subsequently used to produce presentations and documentation aimed at effective knowledge transfer.

Simulating Scenarios in Real-Life Contexts

Real-life testing of people with visual impairment with vision rehabilitation specialists and vision researchers was conducted at the LLMRC site between July 17 and September 19, 2025. For this evaluation, a simulated optometry visit was designed, consisting of eight sequential tasks: (1) entering the building, (2) navigating to the reception desk, (3) using the stairs or elevator to reach the third floor, (4) finding and sitting on a chair in the waiting area, (5) locating the appointment room, (6) proceeding to the optical shop, (7) finding the low-vision boutique, and (8) exiting the building. Figure 4 presents the sequential tasks of the simulated optometry visit.

Figure 4. Sequential tasks for real-life clinical testing for Edge A-Eye.



As part of this real-life indoor navigation simulation stage, four complementary data collection instruments were used to enable a nuanced and multidimensional understanding of the needs and experiences of people with visual impairment:

1. Eligibility questionnaire (Multimedia Appendix 3): administered during the initial recruitment phase, before any testing. It consisted of closed-ended questions designed to verify inclusion and exclusion criteria, as well as to collect sociodemographic data (age, gender, visual condition, and experience with ATs). Its early administration ensured that only eligible participants were included in this study, while also establishing a baseline profile of the sample.
2. Prescenario questionnaire (Multimedia Appendix 4): administered immediately before the navigation scenario. It included questions aimed at assessing participants' expectations, confidence levels, and comfort regarding indoor navigation with ATs. By being administered just before the scenario, it captured initial perceptions and provided a reference point for comparison with postscenario feedback.
3. Walking interview questions (Multimedia Appendix 5): conducted interactively during the real-life scenario. This semistructured guide comprises open-ended questions designed to capture real-time perceptions, obstacles encountered, and immediate reactions related to indoor navigation. Notetakers documented responses during navigation, ensuring the collection of contextualized qualitative data. The objective was to complement quantitative measures with rich insights into user experience, frustrations, and facilitators.
4. Postscenario questionnaire (Multimedia Appendix 6): administered immediately after the completion of the navigation scenario. It included both closed-ended and open-ended questions, aimed at evaluating satisfaction, perceived effectiveness of the technology, usability, and suggestions for improvement. Administered while impressions were still fresh, it allowed the documentation of real navigation needs as well as recommendations for refinement.

XSL•FO
RenderX

Data Analysis

Thematic analysis, conducted through an inductive-deductive approach [44], identified user needs, preferences, and concerns regarding accessibility, safety, and privacy [43]. The inductive approach allowed themes to emerge directly from the data, while the deductive approach ensured consistency with existing theoretical frameworks on AT and user experience [44]. The combination of these perspectives provided a robust analytical framework, blending structured reasoning with open-ended exploration, and ensuring a comprehensive understanding of qualitative data [44]. The thematic analysis procedure followed the Braun and Clarke [45] model with NVivo (Lumivero). Interviews and focus groups were fully transcribed and validated for accuracy. The detailed procedure of this thematic analysis, applied to focus groups and expert interviews, is described in a recently published study by our laboratory [43]. In parallel, a descriptive qualitative approach was adopted to synthesize feedback and identify recurring ideas related to the needs of people with visual impairment and their co-design experiences. Meeting transcripts, facilitator logs, and session summaries were reviewed by the research team to highlight recurring themes, key insights, and decisions made collaboratively with participants. Further, 2 researchers (JPN and HA) independently analyzed the session notes to ensure consistency of interpretation and strengthen analytic trustworthiness through iterative discussion and consensus. A detailed log of all interactions related to the Edge A-Eye project was maintained, and information was systematically organized into categories to support the decision-making process. Following the evaluation of these sessions, as well as qualitative interviews, the priority areas necessary for the co-design of the application were defined. Table 2 presents the six phases of thematic analysis as described by Braun and Clarke [45], outlining the progression from familiarization with the data to the final reporting of results.

Table . Six phases of thematic analysis.

Steps	Description	Team members involved
Familiarization	Reading and rereading full interview transcripts and written statements to become deeply familiar with the data and identify initial ideas.	SSK, DA, PRP, HA, JPN
Generating initial codes	Organizing all data into meaningful groups through the creation of preliminary codes.	SSK, DA, PRP
Searching for themes	Grouping codes from transcripts and Excel files into subthemes and then into overarching themes.	SSK, DA, PRP, JPN, HA
Reviewing themes	Checking the coherence of themes in relation to the codes and the entire dataset, resolving discrepancies through discussion.	SSK, DA, PRP, JPN, HA
Defining and naming themes	Clearly defining the scope, content, and labels of each theme.	SSK, DA, PRP, HA, JPN
Producing the report	Synthesizing analyses and producing the final report of results.	SSK, DA, PRP, HA, JPN

Preliminary Analyses to the Edge A-Eye Team

To accelerate the transmission of preliminary results among members of the Edge A-Eye research team and promote an agile framework for technology development, descriptive statistics were shared on a weekly basis during testing. This was based on the results from participants in each of the four phases of real-life testing, indicated above. This included sociodemographic information, the visual status of participants, the ability of participants to complete tasks, the use of technological devices, as well as distance and direction errors of participants. Although more detailed analyses are underway for the scenarios in real-life contexts, this is beyond the scope of this manuscript, illustrating the co-design methodology carried out during the Edge A-Eye project.

Results

Overview

This section summarizes the main findings from the co-design process and subsequent evaluations. It begins by outlining the participants' characteristics to highlight the diversity of profiles included in this study, followed by a description of the key stages in the development of the Edge A-Eye platform, the identified needs, prioritized requirements, and performance outcomes observed during real-world testing.

Evolution of the Edge A-Eye System Across Co-Design Phases

The steps of the co-design process are shown in the results below, as well as how each of these steps informed the subsequent steps of design and development of the platform.

Determining User Needs (Focus Groups and Semistructured Interviews)

The needs assessment phase combined qualitative insights from two web-based focus groups and semistructured interviews with adults living with visual impairments across Canada to inform the subsequent phases of this project. Across all datasets, the analysis converges toward a coherent understanding of the factors shaping indoor navigation for people with visual impairment, emphasizing the interaction between environmental accessibility, human support, and the limitations of existing technologies. A total of 20 adults with visual impairments

XSL•FO
RenderX

participated in the two focus groups (one in English and one in French). Participants represented a wide diversity of demographic and clinical profiles, including age groups (18 - 49 and 50+ years), levels of vision (low-vision, legal blindness, or complete blindness), etiologies of vision loss (visual acuity loss, visual field loss, or combined deficits), and mobility aids (white cane, guide dog, support cane, walker, or none). All were regular smartphone users, fluent in English or French, and able to travel independently at least occasionally. This purposeful sampling ensured maximum variability across navigation habits and accessibility needs. Data were collected between February and June 2023, beginning with the two focus groups, followed by an individual email-based ranking exercise to consolidate and prioritize user-reported needs. All participants completed each component of this study. In TA cross-scenarios, participants consistently reported structural barriers that undermine autonomy during indoor navigation. These include inaccessible or inconsistent signage, cluttered or unpredictable spaces, poor lighting or contrast, difficulties in locating specific destinations, and ineffective or uncomfortable social interactions when requesting help. Conversely, users emphasized that supportive staff, advance preparation, reliable multisensory cues, and access to clear and accessible information significantly facilitate safe and confident mobility. These facilitators were echoed across both language groups and multiple contexts, confirming their central role in indoor wayfinding. The need for coherent environmental design, context-aware guidance, and multimodal orientation cues also emerged strongly from the focus groups. Participants described scenarios where auditory cues, tactile markers, or consistent spatial organization substantially improved their navigation performance. At the same time, concerns were raised regarding the usability and reliability of current technologies, particularly in complex indoor spaces where GPS is not effective and where mobile apps may provide insufficient precision or unclear instructions.

To complement user perspectives, 2 vision rehabilitation specialists participated in this study as expert validators. Their contribution was essential for linking experiential needs to clinical practice. They confirmed the centrality of personalized and contextually integrated solutions, highlighting that indoor navigation support must be adaptable to diverse functional profiles and embedded into broader rehabilitation pathways. Their recommendations reinforced the importance of developing systems capable of offering fine-grained directional guidance, dynamic distance cues, reliable landmark recognition, and flexible notification settings adapted to users' mobility strategies. Overall, the combined findings from the focus groups, semistructured interviews, and expert validation identify a clear set of priorities for accessible indoor navigation: (1) ensuring cohesive and perceivable environmental information, (2) integrating multisensory cues that align with real-world mobility practices, (3) designing intuitive and trustworthy technologies, and (4) strengthening inclusive human support systems that promote autonomy without compromising user comfort. These results directly informed the subsequent co-design and prototype development phases, ensuring that Edge A-Eye remains grounded in the lived realities and expectations of its intended users. The barriers and facilitators by scenario are summarized in Table 3.

Table . Barriers and facilitators (focus groups French or English).

Scenario	Facilitators (English)	Facilitators (French)	Barriers (English)	Barriers (French)
Coffee shop	Human assistance; preparation; accessible physical space	Human assistance; accessible payment; sensory cues	Difficulty locating exact spot; poor interactions; inaccessible signage	Lighting; inaccessible signage; mobility difficulties
Hospital or clinic	Human assistance; mobility support; preparation	Human assistance; preparation; accessible signage	Locating exact destination; poor interactions; mobility difficulties	Inaccessible signage; unsuccessful interactions; mobility difficulties
Big box store	Human assistance; preparation; accessible website	Accessible website; preparation; human assistance	Mobility difficulties; locating exact place; locating specific item	Asking staff; inaccessible signage; mobility difficulties
Party	Preparation; accessible space; human assistance	Human assistance; accessible space; alternative strategies	Mobility challenges; poor interactions; loud environment	Fear of spilling; noise; mobility difficulties
Bus	Accessible signage; preparation; accessible space	Accessible signage; accessible space; human assistance	Inaccessible signage; poor interactions; mobility difficulties	Inaccessible signage; locating exact place; unexpected events

Prioritization of Requirements (Co-Design Workshop)

The intersectoral co-design workshop held on November 1, 2024, brought together people with visual impairment, a vision rehabilitation specialist, researchers from the University of Montreal or ÉTS, representatives from IVADO or Mathematics of Information Technology and Complex Systems, and contributors from the INLB. This collaborative session played a pivotal role in translating the needs expressed during focus groups, walking interviews, and real-world evaluations into a structured set of actionable design requirements. The outcomes were organized around three core axes: physical accessibility, virtual accessibility, and privacy or security.

Physical Accessibility

This axis captures elements related to the user's interaction with the built environment. Discussions highlighted the need to support safe, efficient navigation in complex indoor settings,

XSL•FO
RenderX

consistent with the barriers identified during earlier stages (eg, inaccessible signage, low contrast, difficulty identifying doors, and risk of collisions). Priority requirements were (1) door-level guidance, including precise identification and approach support; (2) accessible emergency evacuation plans; (3) automated sign reading and audio announcement using camera-to-speech conversion; (4) AI-based object recognition to support proactive awareness; (5) emergency mode for urgent navigation needs; (6) indoor maps translated into audio instructions; (7) obstacle detection with real-time feedback; and (8) noise adaptive alerts and haptic notifications.

These priorities reflect the need for robust environmental interpretation features that enhance independence and reduce cognitive strain.

Virtual Accessibility

Virtual accessibility encompasses the interaction between the user and the digital interface, including personalization, control mechanisms, and flexibility in information delivery. Participants emphasized simplicity, adaptability, and the ability to tailor navigation to situational needs. Priority requirements were (1) voice or haptic command options enabling hands-free use; (2) fully customizable settings (volume, speech rate, contrast, or language); (3) flexible distance representations (meters, steps, or clock-face orientation); (4) exhaustive but filterable information, preventing overload; (5) two navigation modes: focused mode for essential, task-specific guidance and exploratory mode for broader environmental awareness; (6) estimated time of arrival or floor indicators for vertical navigation; (7) full compatibility with screen readers (eg, VoiceOver or TalkBack [Samsung]); and (8) user-defined alert configurations.

These requirements ensure that the system remains intuitive, adaptable, and accessible across varied contexts.

Privacy and Security

Privacy and data protection emerged as essential conditions for user trust and adoption. Participants stressed the importance of strict control over personal data, especially given the use of camera-based and location-based technologies. Priority requirements were (1) opt-out options for any nonessential data collection; (2) strict data minimization, limiting processing to navigation-critical information; (3) Al-based identity protection; (4) automatic face blurring in camera streams; (5) clear and strict privacy policies; (6) user-controlled data storage; and (7) robust data protection mechanisms, including encryption and secure authentication.

These safeguards uphold ethical standards and ensure compliance with privacy by design principles. The co-design workshop crystallized this project's direction by refining user needs into a coherent set of prioritized requirements. Together, the 3 axes of physical accessibility, virtual accessibility, and privacy or security form a comprehensive framework guiding the Edge A-Eye development toward a solution that is (1) inclusive, (2) technically feasible, (3) context-aware, (4) secure, and (5) deeply aligned with the lived experiences of people with visual impairment. Key outputs are organized in Table 4.

Table . Co-design priorities by axis.

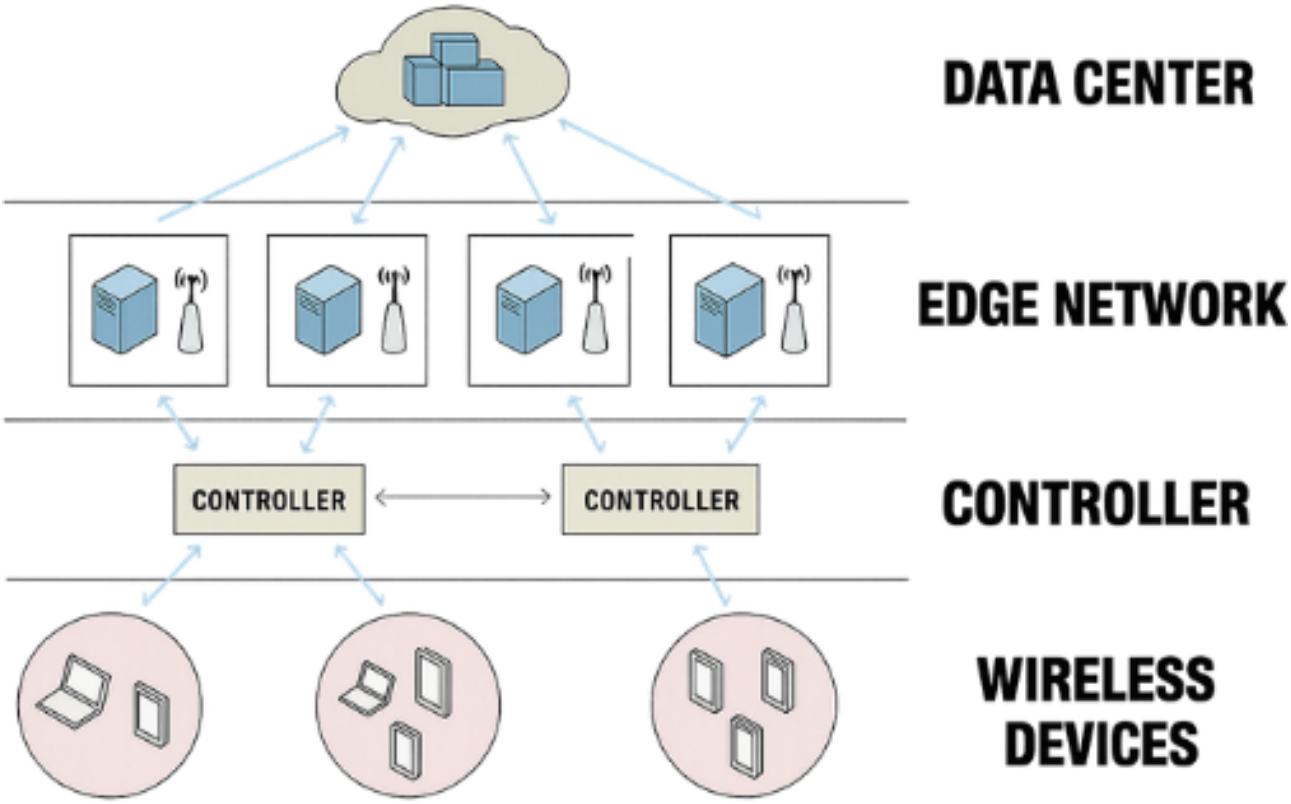
Axis	Key themes	Priority requirements
Physical accessibility	Doors or entrances; signage; object identification; stress; routing; collisions; environmental distractors	Precise door guidance; evacuation plans; camera-based signage reading; product identification; emergency mode; obstacle detection; sound/haptic alerts
Virtual accessibility	Hands-free use; adaptability; instructions; product information; filtering; localization; accessibility; preferences	Voice/haptic control; adjustable volume/speed/contrast; flexible distance units; detailed product info; targeted information modes; estimated time of arrival and floor tracking; screen reader compatibility; customizable alerts
Privacy/security	Geolocation; personal data; camera use; human assistance; guarantees; hyperlinks	Opt-out; data minimization; aliases; facial blurring; strict retention policies; user-controlled storage; safe hyperlink identification

System Architecture Edge A-Eye

The system aims to implement an intelligent framework that dynamically composes microservices running on mobile devices, edge, and cloud environments, and then optimally allocates resources to create personalized services for each individual with respect to privacy and security. We will also develop digital-twin models based on the behaviors of people with visual impairment. Figure 5 presents the system model of the Edge A-Eye Platform.

XSL•FO
RenderX

Figure 5. System model of the Edge A-Eye platform.



A sophisticated assistive system has been proposed, featuring a mobile app and a cloud-based management platform specifically for individuals with visual impairments. To address navigational challenges faced by those with visual impairment, an advanced automatic navigation system has been developed. This system uses edge computing, federated training, and AI computer vision for object detection and anomaly identification. By integrating these advanced technologies, the system provides visual cues, auditory feedback, and communication features, enhancing path-following accuracy and effective obstacle avoidance. The proposed framework is structured into three layers: (1) cloud computing, (2) edge network, and (3) mobile devices.

When mobile devices lack the computational power to complete tasks within the workflow, they send service requests to edge nodes to obtain the required resources. Controllers are tasked with executing control functions on one or more edge nodes and periodically exchanging messages to maintain synchronization. This method ensures efficient resource management on the edge node.

Server side, we noted three key elements, which are the following: (1) have microservices for the edge-environment to automatically update the map and dataset; (2) frequently store and update the map with obstacles. Additionally, train server-side "anomaly detection" AI model with dataset updated from each device; and (3) AI and training microservices.

Regarding the edge environment, edge computing allows devices in remote locations to process data at the "edge" of the network, either by the device or a local server. Further, when data needs to be processed in the central datacenter, only the most important data is transmitted, thereby minimizing latency. The five key elements were (1) a microservice that is called from people with visual impairment devices, with the people with visual impairment device status and controller signal. It will calculate and draw a full path for people with visual impairment to go on; (2) to update locally the map with patient devices sensor status and Wi-Fi fingerprint; (3) to have received data and produce abnormality detection to warn people with visual impairment and calculate safety; (4) an AI assistant, which can receive raw sound from people with visual impairment device and transfer to text, producing with local large language model, receiving the output it positions to navigate the people with visual impairment; and (5) a federated model, which can share the dataset from a different device and produce it on the Edge A-Eye local server.

Edge A-Eye System Architecture and User Interaction in a Real-World Setting

It illustrates, first, system architecture (Figure 6, left): cloud services (handles data processing and updates), edge computing (performs mapping and localization tasks), and a mobile device (collects sensor data and communicates with edge or cloud layers). Second, user interaction (Figure 6, right): a person navigating an indoor hallway using a white cane and a smartphone mounted on the chest for hands-free operation, voice or haptic feedback (directions communicated via audio and vibration), and environment (entryway and typical indoor obstacles).

XSL•FO
RenderX

Figure 6. Edge A-Eye system architecture and user interaction in a real-world setting.

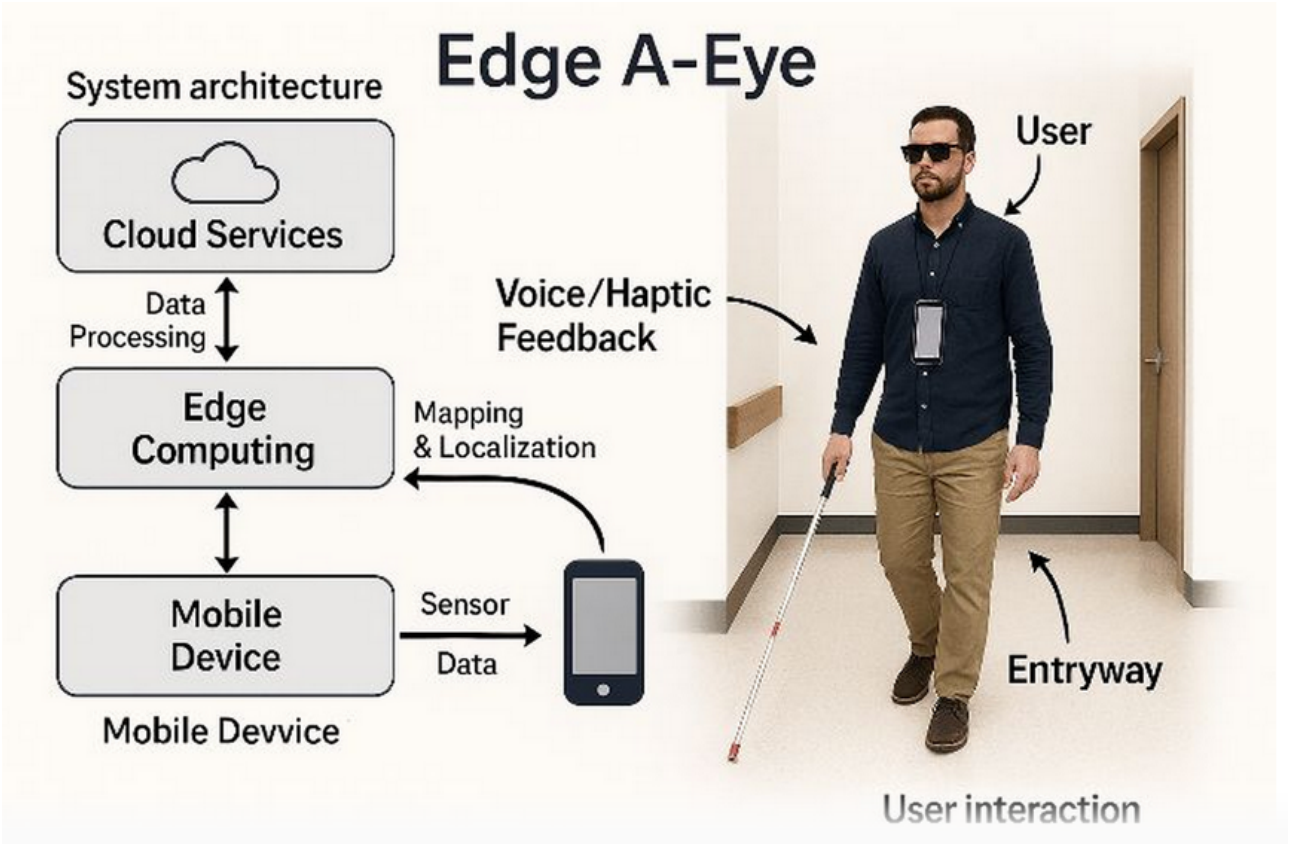


Figure 6 shows the layered architecture (cloud, edge, or mobile device) and illustrates how the user engages with the platform through voice or haptic feedback while navigating an indoor environment.

Iterative Testing of the Edge A-Eye Platform

Based on the insights from focus groups, interviews and the co-design process, we decided a clinical site was appropriate for real-world testing and the ÉTS building for laboratory testing. Therefore, before testing the Edge A-Eye platform, a detailed mapping and calibration process was conducted at both the ÉTS and LLMRC. This step ensured that navigation scenarios were realistic and adapted to the needs of people with visual impairment. Key outcomes include (1) regarding the identification of critical zones: entrances, reception areas, staircases, elevators, and waiting rooms being mapped as priority navigation points; (2) regarding the contrast and lighting analysis: variations in floor and wall contrasts, lighting intensity, and potential glare sources being documented to optimize visual cues for low-vision users; (3) regarding the spatial marker calibration: distances between key landmarks (doors, corridors, or elevators) being measured and integrated into the navigation algorithm for accurate route planning; (4) regarding the environmental complexity assessment: obstacles such as furniture, signage placement, and noise sources being cataloged to anticipate challenges during navigation; and (5) regarding the scenario definition: two main navigation scenarios being established which were academic setting (ÉTS): navigating from building entrance to designated laboratories; and clinical setting (LLMRC): simulating an optometry visit, including reception, waiting area, consultation room, and optical shop. These findings provided the foundation for precise route planning and adaptive guidance features in the Edge A-Eye platform, ensuring that subsequent usability tests reflected real-world conditions.

A total of 1 vision rehabilitation specialist and people with visual impairment were accompanied by 4 academic and industry partners to evaluate the usability and robustness of the Edge A-Eye platform. These tests were performed on a biweekly basis in the controlled laboratory environment at the ÉTS, with the participation of academic and industry partners alongside a clinician specialized in visual impairment. This stage was critical for assessing ergonomics and robustness, while also identifying technical limitations and guiding necessary adjustments. The collaboration between users and technical experts provided a unique opportunity to refine the solution before moving into real-world testing. For developers, this phase offered several strategic advantages: (1) early detection of technical issues, allowing corrective measures before large-scale deployment; (2) direct feedback from vision rehabilitation specialists and experts, ensuring that design choices were validated against real needs; (3) controlled testing conditions, which facilitated precise measurement of performance and usability; (4) strengthened collaboration with interdisciplinary partners, enabling developers to integrate clinical, industrial, and experiential perspectives into technical refinements; (5) iterative improvement cycle, reducing risks of failure in later stages and accelerating the path toward a stable beta version. This laboratory evaluation demonstrated the essential role of developers in bridging technical feasibility with user experience, ensuring that the Edge A-Eye platform evolved into a reliable, inclusive, and adaptable

XSL•FO
RenderX

solution. A video demonstration of this process is available in Multimedia Appendix 7.

Navigation Performance Across Eight Real-World Tasks

This phase aimed to capture the concrete challenges experienced by people with visual impairment during indoor navigation. A total of 13 people with visual impairment completed 8 structured navigation tasks in clinical indoor environments. Performance varied by lighting, contrast, spatial clarity, and noise levels. These results present explicit, measurable findings that directly inform system requirements. Participants were selected to include a diversity of technological competency, age, and visual status. The demographic characteristics can be seen in Multimedia Appendix 8. Participants were asked to complete the tasks independently using their current mobility tools and devices. The results of participants after completing the 8 sequential tasks (Figure 4) varied considerably across tasks in the success of independent completion of the task, key needs and preferences of the users, as well as the factors in the scene. These results underline the need for the Edge A-Eye platform to have task-dependent capabilities. The results are summarized in Table 5 below.

Table . Task-based navigation performance and contextual scene factors.

Task (distance; m)	Success rate	Key needs/preferences	Scene factors
Entrance (10)	7/13 (54)	Notification at 6 - 8 m; door type; landmarks	High luminance and contrast facilitate entry
Reception (2.5)	8/13 (62)	Immediate notification; positioning	Contrast improves desk detection
Stairs/elevator (7 - 13)	11/13, directed (85)	Distance/direction; buttons; handrail	Varies by layout
Waiting area (2.5 - 7)	11/13 (85)	Direction to seat; 54% prefer human help	High contrast helpful
Examination room (54)	4/13 (31)	Detailed cues; distance; position; tactile/braille	Fine-grained cues essential
Optical store (18 guided)	—[a]	Assistance preferred (85%); satisfaction 100%	—
Low-vision store (55)	7/13 (54)	Category clarity; accessible payment; directions	Environmental factors cause confusion
Exit (8)	13/13 (100)	Less command guidance; door type; landmarks	High luminance aids orientation

[a]Not available.

A video demonstration of the Edge A-Eye platform iterative testing with an Android (Google LLC) phone, including route planning, voice command, and camera integration at the LLMRC, is available in Multimedia Appendix 9.

Continuous Dissemination of Results

The dissemination of the results of the Edge A-Eye project was structured around several complementary channels, ensuring scientific, clinical, and intersectoral visibility. Preliminary design and development progress has been shared at scientific conferences (both in vision sciences and engineering), at vision rehabilitation centers, at associations for those with visual impairment, and among interested industry leaders. This dissemination strategy made it possible: (1) to strengthen the scientific legitimacy of this project among peers; (2) to anchor the results in clinical realities and field practices; (3) to foster intersectoral synergies and opportunities for technology transfer; and (4) to support a continuous improvement dynamic by integrating stakeholder feedback at each stage.

Ongoing Monitoring and Evaluation of the Edge A-Eye Project (May 2024 to December 2025)

The comprehensive evaluation phase applied the TRL framework [40] to monitor technological maturity, identify gaps, and guide improvements. Continuous feedback from real-world testing, clinical partners, and technical teams enabled iterative refinements, ensuring robustness, usability, and long-term viability. This adaptive process-maintained alignment with user needs and operational constraints, paving the way for scalable implementation.

TRL 1-2: Concept and Feasibility (May to June 2024)

For TRL 1-2, we configured the following: (1) defined user needs through focus groups and interviews; (2) analyzed barriers and facilitators in indoor navigation scenarios; and (3) established a conceptual framework; no technology developed yet.

TRL 3-4: Preliminary Design (July to November 2024)

For TRL 3-4, we configured the following: (1) conducted co-design workshop (November 1, 2024) to prioritize needs and core features; (2) developed initial wireframes and system architecture; (3) validated feasibility and user alignment; (4) regarding the features and functions included: SITUM (*Smart Indoor Tracking and Indoor Mapping*) indoor mapping for navigation in ÉTS building using Wi-Fi fingerprinting (French only), development of user interface with TalkBack, directional information with gyroscope, obstacle detection, optical character recognition, free of charge, protection of user information (use of test phone); and (5) regarding the features and functions desired: bilingual, accessible with iPhone, customizable directional information, hands-free use, voice activation,

XSL•FO
RenderX

accessible interface (text-to-speech and high contrast), haptic door, and obstacle detection.

TRL 5-6: Controlled Laboratory Testing (December 2024 to March 2025)

For TRL 5-6, we configured the following: (1) performed mapping and calibration at ÉTS and LLMRC; (2) tested prototypes in controlled settings with academic and industry partners; (3) identified technical limitations and applied iterative improvements; (4) regarding the features or functions included: voice activation of navigation, bilingual use (French- and English-language), navigational guidance at 10-second intervals only, activation with physical buttons on phone; (5) regarding the features or functions excluded: gyroscope during route navigation (to permit hands-free use after issues related to manipulation of phone), constant haptic obstacle detection (cognitive overload of testers); and (6) regarding the features or functions desired: prioritize turn-by-turn directions at critical moments, use of camera to read room numbers, rapid door detection near entryways, and use with iOS (Apple Inc).

TRL 7-8: Real-World Evaluation (July to Sept 2025, Ongoing)

For TRL 7-8, we configured the following: (1) simulated optometry visits at LLMRC with 13 participants; (2) assessed usability, accessibility, and adaptability in authentic contexts; (3) integrated AI features (guidance with optical character recognition [OCR]); (4) regarding the features or functions included: optimized turn-by-turn directions (near approach intersection), integration of floor maps, route planning, use of OCR at destination, and voice activation of destinations at multiple locations; (5) regarding the features or functions excluded: automated integration of OCR (to allow for user to call functions with button or voice); and (6) regarding the features or functions desired: mapping with geomagnetic indoor positioning to improve scalability, use with iOS, customizable voice activation of features, specific detection of objects desired for navigation using YOLO (You Only Look Once; ie, bathrooms, chairs, doors, or desks).

TRL 9: Scalability and Final Assessment (In Progress)

The results from the previous phases, integrated with laboratory and real-world testing, inform the team as to the next iterations of the Edge A-Eye platform. This structured progression through TRL demonstrates how iterative, user-centered development can transform an initial concept into a robust, inclusive, and scalable AT solution, setting a benchmark for future innovations in indoor navigation for persons with visual impairments. Figure 1 presents ongoing monitoring and evaluation of the Edge A-Eye project (May 2024 to December 2025).

Figure 2 presents the Edge A-Eye TRL radar, summarizing technological maturity and development progress.

Stakeholder Contributions

In sum, each stage played an irreplaceable role in the co-design of the Edge A-Eye app. People with visual impairment ensured experiential relevance, clinicians guaranteed clinical integration, researchers contributed scientific rigor, and industry partners reinforced technical and economic viability. Together, these stakeholders transformed an idea into an inclusive and sustainable solution.

Discussion

Principal Findings

The results of the Edge A-Eye project demonstrate that users desire an indoor navigation platform that is customizable, accessible, and user-friendly. This meant that the platform required accessibility with text-to-speech as well as various screen settings. Additionally, voice activation was of particular interest to the users, as this allowed those with lower technological competencies to use the platform, as well as allowing ease of use with their mobility aids. As only certain types of information are required during the journey, the Edge A-Eye project allowed us to examine and prioritize those pieces of information to avoid cognitive overload while walking in real-world settings. As technical and implementation challenges remain to provide scalability and provide an eventual final assessment of the platform, we continue the co-design process to ensure that it meets users' needs. This study marks the first time a co-design process for indoor navigation in Canada has been documented. Additionally, our novel approach demonstrates how integrating vision rehabilitation with engineering and computer science can improve user-centered outcomes for indoor navigation.

Human-Centered Co-Design and Collective Intelligence

The Edge A-Eye project illustrates the transformative impact of human-centered co-design in AT development. By engaging people with visual impairment as co-creators rather than passive testers, this project ensured that lived experiences were embedded at the core of the innovation process. This approach aligns with the six principles of UCD outlined by Ortiz-Escobar et al [17] which emphasize early and continuous user involvement as a foundation for usability, ownership, and long-term adoption. Compliance with ISO 9241-210:2019 strengthened the inclusivity and usability of the Edge A-Eye platform by embedding internationally recognized HCD principles into every stage of development, thereby enhancing adoption potential and reducing abandonment risk [22]. Slattery et al [21] similarly argue that embedding user perspectives across all stages of design enhances both relevance and acceptability, creating solutions that reflect real needs rather than abstract assumptions. Our results confirm that HCD (ISO 9241 210) remains essential but is not sufficient on its own to ensure adoption and scalability in AT [22]. Organizational and regulatory constraints can limit iterative development and hinder large-scale deployment, as illustrated by Chavarria et al [19]. In parallel, the sociomaterial perspective proposed by Rong and Hansopaheluwakan-Edward [24] helps explain why certain constraints (eg, rules, artifacts, and material properties) may become enablers depending on how they are configured. This lens also clarifies our own methodological choices, particularly in the alignment of technical, clinical, and experiential priorities throughout the co-design process. This project also mobilized collective intelligence, defined as the ability of diverse actors to collaborate effectively toward shared goals. By integrating experiential knowledge from people with visual impairment

XSL•FO
RenderX

with the expertise of vision researchers, rehabilitation specialists, and developers, Edge A-Eye fostered a shared vision of inclusivity and equity. This convergence of perspectives strengthened the legitimacy of this project and ensured that technical decisions were aligned with social realities. Such collaborative intelligence echoes the developmental evaluation framework by Patton [46], which advocates adaptive learning environments in complex innovation ecosystems. Ultimately, the co-design process transformed diversity into a strategic asset, ensuring that the platform was not only technologically robust but also socially relevant and ethically grounded. This section also highlights the unique contributions of the Edge A-Eye project, including its interdisciplinary co-design approach integrating vision science, engineering, AI, and clinical rehabilitation, and its iterative design process that demonstrably improved usability and user satisfaction compared to conventional solutions.

Iterative Development and Contextual Evaluation

The success of the Edge A-Eye project was also driven by its reliance on iterative development cycles, which allowed the platform to evolve progressively in response to user feedback. Each prototype, from initial wireframes to functional beta versions, was tested, evaluated, and refined through systematic feedback loops. This iterative agility reflects the principles of inclusive design, where adaptation and responsiveness to real-world contexts are prioritized [17]. By embedding feedback at every stage, the current project reduced risks of failure, avoided costly redesigns, and accelerated the stabilization of the beta version. Equally important was the dual strategy of laboratory-based and real-world evaluations. Controlled laboratory testing provided precise measurements of ergonomics, usability, and technical robustness, while real-world deployments captured the complexities of authentic navigation environments, including noise, crowding, and spatial ambiguity. This combination aligns with best practices in human-computer interaction and accessibility research, which stress the importance of contextual evaluation in producing intuitive and reliable tools [21]. By balancing controlled refinement with ecological validity, the current project ensured that Edge A-Eye was not only technically feasible but also practically usable in diverse environments. This iterative and contextualized approach strengthened both the sustainability and the acceptability of the final product, positioning it as a reliable solution for inclusive indoor navigation.

Knowledge Dissemination and Continuous Improvement

Active knowledge dissemination was a cornerstone of the Edge A-Eye project, ensuring transparency, collaboration, and ecosystem building around inclusive navigation technologies. Dissemination sessions at ÉTS and the LLMRC, combined with presentations at international scientific conferences, created opportunities for constructive exchanges and systematic feedback. These dialogues validated preliminary results, identified areas for refinement, and opened perspectives for technology transfer into rehabilitation practices. The iterative integration of feedback fostered a feedback-rich environment, enhancing agility, usability, and long-term impact. This process reflects that of the description of the developmental evaluation approach by Patton [46], which embeds continuous learning and adjustment into the innovation cycle. Furthermore, recognition at scientific conferences through awards and international visibility reinforced the current project's credibility and positioned it as a benchmark in inclusive AT. Ultimately, dissemination was not only about presenting results but about co-constructing the next steps, ensuring that this project remained scientifically rigorous, socially relevant, and sustainable over time.

Implications From Edge A-Eye Project

Overall, the Edge A-Eye project highlights how co-design, collective intelligence, and iterative development shaped the platform, ensuring that technological robustness was consistently aligned with social relevance. By examining methodological choices, knowledge dissemination strategies, and the ethical commitment to equity and inclusion, the development of the Edge A-Eye platform underscores this project's dual contribution: advancing scientific understanding of AT design while offering practical recommendations for sustainable, user-driven innovation. Beyond methodological insights, this project provides concrete evidence of impact, linking design principles to measurable outcomes observed during laboratory and real-world evaluations. Concerning other specifications are the following: first, regarding the scientific, this project contributes to the literature on user-centered and co-design methodologies in ATs, specifically in the context of indoor navigation for people with visual impairment. It underscores the value of integrating experiential knowledge by users and rehabilitation specialists into system design to enhance functionality and relevance [31]. Second, methodologically, the current research provides a replicable framework that combines iterative prototyping, participatory design, mixed-context testing, and ongoing evaluation. This approach aligns with best practices in inclusive design and can inform future projects [47]. Third, regarding the ethical and equity, diversity, and inclusion (EDI)–related, by embedding principles of EDI throughout the process, the current project upholds a rights-based approach to innovation. Including users with diverse visual profiles and backgrounds fosters fairer outcomes and helps reduce systemic barriers in digital environments [8,29,30]. Fourth, regarding empowerment, autonomy, and social participation, the current project advances the autonomy of people with visual impairment by positioning them as co-creators rather than passive beneficiaries. This empowerment contributes to increased self-efficacy and social participation, key determinants of well-being and quality of life [6,16,48]. Fifth, regarding the practical and technological, grounding development in actual user contexts ensures a better fit between the platform's functionalities and daily realities, promoting long-term usability and adoption. Sixth, regarding policy and advocacy, the current study provides evidence supporting the integration of co-design and UCD principles into public innovation strategies and digital accessibility policies. It reinforces the importance of involving end users in shaping the technologies that impact their lives [8,29,30].

XSL•FO
RenderX

Limitations

Despite the relevant findings from this approach, certain limitations must be acknowledged to contextualize these contributions. The relatively limited number of participants may affect the generalizability of the findings. Including a broader and more diverse sample in terms of age, technological literacy, and types of visual impairment would strengthen future iterations of the platform. Additionally, while this project prioritized user involvement, variations in individual experiences may result in different levels of usability or satisfaction across user groups. The technical ecosystem also presented limitations: compatibility issues with different devices and ATs could restrict access for some potential users. Addressing these interoperability challenges is crucial for ensuring broader implementation and adoption. Understanding these strengths and limitations offers valuable guidance for refining the platform, scaling the innovation, and informing similar projects within the field of inclusive technology and assistive navigation.

Conclusions

This study demonstrates that a sustained participatory co-design process can significantly strengthen the development of indoor navigation technologies for people with visual impairments. Through iterative stages, the integration of user, clinical, engineering, and industry perspectives led to meaningful refinements of the Edge A-Eye platform, including improvements in navigation clarity, interpretation of environmental cues, hands-free interaction, and optimization for iPhone compatibility. User feedback directly informed several critical design decisions, such as reducing unnecessary obstacle alerts, removing gyroscope-dependent interactions due to usability and battery constraints, and addressing challenges related to digital accessibility and varying levels of technological literacy. These findings confirm that co-design not only enhances usability but also ensures alignment with real-world mobility needs, safety considerations, and user expectations. Methodologically, this work contributes a replicable, multiphase co-design framework tailored to the rehabilitation and AT context in Canada. These technological advancements translate into tangible benefits for people with visual impairment, including greater independence, improved safety, and smoother navigation in daily life, reinforcing the social impact of the Edge A-Eye platform. Beyond methodological insights, the Edge A-Eye platform demonstrated measurable improvements during iterative testing. Laboratory and real-world evaluations confirmed continuous enhancement of navigation clarity and environmental interpretation. Across two indoor sites, guidance performance improved with optimized turn-by-turn directions and reduced cognitive load. User satisfaction was high, with participants reporting smoother navigation and fewer directional errors compared to initial prototypes. These findings provide concrete evidence of technological progress and reinforce the platform's potential for real-world adoption.

Recommendations grounded in study findings are the following: (1) prioritize scenario-specific route guidance and minimize cognitive load to prevent information overload, (2) maintain accessibility features that accommodate diverse technological literacy levels and common mobile accessibility practices, and (3) continue iterative refinement to ensure usability and safety across varied indoor environments.

Broader future research directions are the following: (1) conduct long-term evaluations in diverse real-world settings with larger and more heterogeneous people with visual impairment populations to validate adoption and performance stability, (2) explore the integration of emerging technologies (eg, augmented reality, advanced environmental mapping, and multimodal feedback) to enhance navigation efficiency and situational awareness; (3) develop strategies for scalability and affordability to ensure equitable access across socioeconomic and linguistic groups; and (4) promote knowledge sharing and interdisciplinary collaboration to sustain innovation and inclusivity in AT development.

By embedding these priorities, future iterations of Edge A-Eye can advance autonomy, confidence, and day-to-day safety for people with visual impairment, while contributing to a broader framework for inclusive, context-aware technological innovation.

Acknowledgments

We sincerely thank all participants, rehabilitation specialists, and community partners for their invaluable contributions throughout the co-design process. We also acknowledge the support of collaborating organizations, including the CNIB (Canadian National Institute for the Blind) Foundation, Regroupement des Aveugles et Amblyopes du Montréal Métropolitain, and Institut Nazareth et Louis-Braille, for their assistance in participant recruitment and knowledge sharing. All figures were generated using artificial intelligence.

Funding

This research was supported by the academic-industrial partnership Mathematics of Information Technology and Complex System Alliance (IT37317) from the Natural Sciences and Engineering Research Council of Canada with industry financial support from VMWare by Broadcom, Inc. Additional institutional support was provided by the Université de Montréal, École de Technologie Supérieure, and CRIR (Centre de Recherche Interdisciplinaire en Réadaptation du Montréal Métropolitain). The funder had no involvement in this study's design, data collection, analysis, interpretation, or the writing of this paper.

XSL•FO
RenderX

Data Availability

Deidentified qualitative data supporting the findings of this study are available upon reasonable request from the corresponding author. Multimedia appendices include supplementary materials such as questionnaires and prototype demonstration videos.

Authors' Contributions

Conceptualization: SSK, HA, JPN
Methodology: SSK, JPN, JG, KKN, PAN, DA, PRP, SMR, NG-R, BP
Software or prototype development: JG, KKN, PAN, HN, SMR, NG-R, BP
Data collection: SSK, JPN, JG, PAN, HN, DA, PRP, SMR, NG-R, BP
Formal analysis or interpretation: SSK, JPN, HA, KKN, DA, PRP, SMR, NG-R
Supervision: HA, KKN, JPN
Project administration: JPN, KKN
Writing original draft: SSK, HA, JPN
Writing review & editing: all authors contributed to the critical revision of the manuscript, approved the final version, and accept responsibility for its content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guide for the Edge-A-Eye project.
[DOCX File, 25 KB - rehab_v13i1e81347_app1.docx]

Multimedia Appendix 2
Co-design workshop.
[DOCX File, 25 KB - rehab_v13i1e81347_app2.docx]

Multimedia Appendix 3
Questionnaire for selection of participants for Edge A-Eye.
[DOCX File, 23 KB - rehab_v13i1e81347_app3.docx]

Multimedia Appendix 4
Prescenario questionnaire.
[DOCX File, 22 KB - rehab_v13i1e81347_app4.docx]

Multimedia Appendix 5
Clinical walkthrough questionnaire (barriers and facilitators).
[DOCX File, 28 KB - rehab_v13i1e81347_app5.docx]

Multimedia Appendix 6
Postscenario questionnaire.
[DOCX File, 21 KB - rehab_v13i1e81347_app6.docx]

Multimedia Appendix 7
Video on Edge A-Eye app.
[MP4 File, 22997 KB - rehab_v13i1e81347_app7.mp4]

Multimedia Appendix 8
Participant demographics for real-world testing (n=13).
[DOCX File, 26 KB - rehab_v13i1e81347_app8.docx]

Multimedia Appendix 9
Video on navigation.
[MP4 File, 192243 KB - rehab_v13i1e81347_app9.mp4]

XSL•FO
RenderX

Checklist 1
COREQ checklist.
[PDF File, 439 KB - rehab_v13i1e81347_app10.pdf]

References

1. Bourne R, Steinmetz JD, Flaxman S. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. Lancet Glob Health 2021 Feb;9(2):e130-e143. [doi: 10.1016/S2214-109X(20)30425-3] [Medline: 33275950]
2. Vision impairment and blindness. World Health Organization. 2026 Feb 10. URL: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment [accessed 2026-05-15]
3. Cloutier E, Grondin C, Lévesque A. Canadian survey on disability, 2017: concepts and methods guide. : Statistics Canada; 2018 Nov 28 URL: https://publications.gc.ca/collections/collection_2018/statcan/89-654-x/89-654-x2018001-eng.pdf [accessed 2026-05-15]
4. Dos Santos ADP, Ferrari ALM, Medola FO, Sandnes FE. Aesthetics and the perceived stigma of assistive technology for visual impairment. Disabil Rehabil Assist Technol 2022 Feb;17(2):152-158. [doi: 10.1080/17483107.2020.1768308] [Medline: 32501732]
5. da Silva PBE, Leal AS, Ferraz NN. Usability of smartphone apps as reading aids for low vision patients. Disabil Rehabil Assist Technol 2022 Oct;17(7):848-852. [doi: 10.1080/17483107.2020.1820086] [Medline: 32930007]
6. El-Taher FEZ, Taha A, Courtney J, Mckeever S. A systematic review of urban navigation systems for visually impaired people. Sensors (Basel) 2021 Apr 29;21(9):3103. [doi: 10.3390/s21093103] [Medline: 33946857]
7. Assistive technologies [Article in French]. World Health Organization. URL: https://www.who.int/fr/activities/preventing-noncommunicable-diseases/assistive-technology [accessed 2026-05-15]
8. WHO & UNICEF. Global Report on Assistive Technology. 2022. URL: https://iris.who.int/server/api/core/bitstreams/17a1cbf6-be26-466c-9fca-ab80eeb2d192/content [accessed 2026-05-29]
9. Giudice NA, Legge GE. Blind navigation and the role of technology. In: The Engineering Handbook of Smart Technology for Aging, Disability, and Independence: John Wiley & Sons, Inc; 2008:479-500. [doi: 10.1002/9780470379424]
10. Burton MJ, Ramke J, Marques AP, et al. The Lancet Global Health Commission on Global Eye Health: vision beyond 2020. Lancet Glob Health 2021 Apr;9(4):e489-e551. [doi: 10.1016/S2214-109X(20)30488-5] [Medline: 33607016]
11. Golubova E, Starke SD, Crossland MD, Wolffsohn JS. Design considerations for the ideal low vision aid: insights from de-brief interviews following a real-world recording study. Ophthalmic Physiol Opt 2021 Mar;41(2):266-280. [doi: 10.1111/opo.12778] [Medline: 33533098]
12. Starke SD, Golubova E, Crossland MD, Wolffsohn JS. Everyday visual demands of people with low vision: a mixed methods real-life recording study. J Vis 2020 Sep 2;20(9):3. [doi: 10.1167/jov.20.9.3] [Medline: 32876678]
13. Pissaloux E, Velazquez R. Mobility of Visually Impaired People: Fundamentals and ICT Assistive Technologies: Springer; 2026. [doi: 10.1007/978-3-031-91550-5]
14. Wiener WR. Foundations of Orientation and Mobility: AFB Press; 2010, Vol. 2. URL: https://www.aph.org/product/foundations-of-low-vision-clinical-and-functional-perspectives-2nd-edition/
15. Corn AL, Erin JN. Foundations of Low Vision: Clinical and Functional Perspectives: AFB Press; 2010. URL: https://www.aph.org/product/foundations-of-low-vision-clinical-and-functional-perspectives-2nd-edition/
16. Goldschmidt M. Orientation and mobility training to people with visual impairments. In: Pissaloux E, Velazquez R, editors. Mobility of Visually Impaired People: Springer; 2026:385-408. [doi: 10.1007/978-3-031-91550-5_14]
17. Ortiz-Escobar LM, Chavarria MA, Schönenberger K, et al. Assessing the implementation of user-centred design standards on assistive technology for persons with visual impairments: a systematic review. Front Rehabil Sci 2023;4:1238158. [doi: 10.3389/fresc.2023.1238158]
18. Keates S, Clarkson J. Countering design exclusion. In: Countering Design Exclusion: Springer; 2004:140-156. [doi: 10.1007/978-1-4471-0013-3_11]
19. Chavarria MA, Ortiz-Escobar LM, Bacca-Cortes B, et al. Challenges and opportunities of the human-centered design approach: case study development of an assistive device for the navigation of persons with visual impairment. JMIR Rehabil Assist Technol 2025 Aug 18;12:e70694. [doi: 10.2196/70694] [Medline: 40825539]
20. Tang J, Zhu Y, Jiang G, et al. Human-centred design and fabrication of a wearable multimodal visual assistance system. Nat Mach Intell 2025;7(4):627-638. [doi: 10.1038/s42256-025-01018-6]
21. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17. [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
22. ISO 9241-210:2019 ergonomics of human-system interaction – part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2026-05-15]
23. Sanders EBN. From user-centered to participatory design approaches. In: Design and the Social Sciences: CRC Press; 2002. [doi: 10.1201/9780203301302]
24. Rong A, Hansopaheluwakan-Edward N. Uncovering facilitators and constraints in co-design with visually impaired children: a sociomateriality perspective. Des J 2025 Jul 4;28(4):742-760. [doi: 10.1080/14606925.2025.2477929]

XSL•FO
RenderX

25. Casanova E, Guffanti D, Hidalgo L. Technological advancements in human navigation for the visually impaired: a systematic review. Sensors (Basel) 2025 Apr 1;25(7):2213. [doi: 10.3390/s25072213] [Medline: 40218728]
26. Benz C, Scott-Jeffs W, McKercher KA, et al. Community-based participatory-research through co-design: supporting collaboration from all sides of disability. Res Involv Engagem 2024 May 10;10(1):47. [doi: 10.1186/s40900-024-00573-3] [Medline: 38730283]
27. Kaboré SS, Ngangue P, Soubeiga D, et al. Barriers and facilitators for the sustainability of digital health interventions in low and middle-income countries: a systematic review. Front Digit Health 2022;4:1014375. [doi: 10.3389/fdgth.2022.1014375] [Medline: 36518563]
28. Wacnik P, Daly SR, Verma A. Participatory design: a systematic review and insights for future practice. Des Sci 2025;11:e21. [doi: 10.1017/dsj.2025.10009]
29. Université de Montréal. Montréal Declaration for a Responsible Development of AI. 2018. URL: https://declarationmontreal-iaresponsable.com/la-declaration/
30. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
31. Pullin G. Design Meets Disability: The MIT Press; 2011.
32. Vial S, Boudhraâ S, Dumont M, Tremblay M, Riendeau S. Developing a mobile app with a human-centered design lens to improve access to mental health care (Mentallys Project): protocol for an initial co-design process. JMIR Res Protoc 2023 Aug 22;12:e47220. [doi: 10.2196/47220] [Medline: 37606978]
33. Wilkinson CR, De Angeli A. Applying user centred and participatory design approaches to commercial product development. Design Studies 2014 Nov;35(6):614-631. [doi: 10.1016/j.destud.2014.06.001]
34. Boger J, Jackson P, Mulvenna M, et al. Principles for fostering the transdisciplinary development of assistive technologies. Disabil Rehabil Assist Technol 2017 Jul;12(5):480-490. [doi: 10.3109/17483107.2016.1151953] [Medline: 27052793]
35. Lindgaard G, Dillon R, Trbovich P, et al. User needs analysis and requirements engineering: theory and practice. Interact Comput 2006 Jan;18(1):47-70. [doi: 10.1016/j.intcom.2005.06.003]
36. Barontini F. User-centered indoor navigation system for the visually impaired. In: Wearable Haptic Devices for Realistic Scenario Applications: Springer; 2025:79-107. [doi: 10.1007/978-3-031-70539-7_5]
37. Bird M, McGillion M, Chambers EM, et al. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021 Mar 1;7(1):12. [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]
38. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
39. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
40. The KTH Innovation Readiness Level. URL: https://kthinnovationreadinesslevel.com [accessed 2026-05-15]
41. Alizada M, Lampinen A, Idestam-Almquist P, et al. Smartphone-based indoor navigation for blind and visually impaired people. In: AAATE 2025 Proceedings: Springer; 2025. [doi: 10.1007/978-3-032-01628-7_14]
42. Savoie-Zajc L. Comment peut-on construire un échantillonnage scientifiquement valide [Article in French]. Rech Qual 2007;Hors-série(5):99-111 [FREE Full text]
43. Puri PR, Coutaller A, Gwade F, Kabore SS, Annan D, Nemargut JP. Perspectives from canadian people with visual impairments in everyday environments outside the home: qualitative insights for assistive technology development. JMIR Rehabil Assist Technol 2025 Jul 29;12:e73380. [doi: 10.2196/73380] [Medline: 40729678]
44. Azungah T. Qualitative research: deductive and inductive approaches to data analysis. QRJ 2018 Nov 15;18(4):383-400. [doi: 10.1108/QRJ-D-18-00035]
45. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
46. Patton MQ. Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use: Guilford Press; 2011.
47. Southwick D, Resch G, Ratto M. Chapter 30: iterative prototyping and co-design. In: Knowledge, Innovation and Impact: Springer; 2021:231-238. [doi: 10.1007/978-3-030-34390-3_30]
48. Guerrero LA, Vasquez F, Ochoa SF. An indoor navigation system for the visually impaired. Sensors (Basel) 2012;12(6):8236-8258. [doi: 10.3390/s120608236] [Medline: 22969398]

Abbreviations

AI: artificial intelligence
ASAMM: Association des Sports Pour Aveugles du Montréal Métropolitain
AT: assistive technology
CERC: Clinical Research Ethics Committee

XSL•FO
RenderX

CNIB: Canadian National Institute for the Blind
COREQ: Consolidated Criteria for Reporting Qualitative Research
CRIR: Centre de Recherche Interdisciplinaire en Réadaptation du Montréal Métropolitain
EDI: equity, diversity, and inclusion
HCD: human-centered design
INLB: Institut Nazareth et Louis-Braille
ISO: International Organization for Standardization
IVADO: Institut de Valorisation des Données
LLMRC: Lethbridge-Layton-Mackay Rehabilitation Centre
NAT: navigation-based assistive technology
O&M: orientation and mobility
OCR: optical character recognition
RAAMM: Regroupement des Aveugles et Amblyopes du Montréal Métropolitain
TA: TA
TRL: technology readiness level
UCD: user-centered design
YOLO: You Only Look Once
ÉTS: École de Technologie Supérieure

Edited by S Munce; submitted 06.Aug.2025; peer-reviewed by A Gupta, JR Cooperstock, M Chavarria, S Mittal, Sunny, CL Au, VR Cano; revised version received 28.Jan.2026; accepted 10.Feb.2026; published 31.Jul.2026.

Please cite as:
Kaboré SS, Annan D, Puri PR, Mejia Romero S, Gingras-Royer N, Gauthier J, Nhung H, Nguyen PA, Pelletier B, Nguyen KK, Alami H, Nemargut JP
Indoor Navigation for People With Visual Impairment in Canada: Participatory Co-Design and Interdisciplinary Study of the Edge A-Eye Platform
JMIR Rehabil Assist Technol 2026;13:e81347
URL: https://rehab.jmir.org/2026/1/e81347
doi:10.2196/81347

© Soutongnoma Safiata Kaboré, Deborah Annan, Prajjol Raj Puri, Sergio Mejia Romero, Nathalie Gingras-Royer, Justin Gauthier, H ng Nhung, Phuong Anh Nguyen, Benoit Pelletier, Kim Khoa Nguyen, Hassane Alami, Joseph Paul Nemargut. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Accuracy in the Estimation of Self-Reported Knee Brace Wear Time in Young Adults With a Symptomatic Knee Following ACL Reconstruction: Secondary Analysis of a Pilot Randomized Controlled Trial

Matthew Savage[1,2], BSc (Hons); Benjamin F Mentiplay[1,2,3], PhD; Harriette Slater[1,2], MPhysio; Fernanda Serighelli[1,2], BPhysio (Hon); David L Carey[1,2,3], PhD; Jamon L Couch[1,2,4], MPhysioPrac; Andrea M Bruder[1,2,5], PhD; Adam G Culvenor[1,2], PhD

[1]La Trobe Sport and Exercise Medicine Research Centre, School of Allied Health, Human Services and Sport, La Trobe University, Plenty Rd, Bundoora, Melbourne, Victoria, Australia

[2]Australian IOC Research Centre, La Trobe University, Melbourne, Australia

[3]Sport and Exercise Science, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Australia

[4]Arthritis Research Canada, Vancouver, BC, Canada

[5]Department of Physiotherapy, Podiatry and Prosthetics and Orthotics, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Australia

Corresponding Author:
Adam G Culvenor, PhD
La Trobe Sport and Exercise Medicine Research Centre, School of Allied Health, Human Services and Sport, La Trobe University, Plenty Rd, Bundoora, Melbourne, Victoria, Australia

Abstract

Background: Knee braces may improve symptoms and physical function following anterior cruciate ligament reconstruction (ACLR). However, their effectiveness depends on adherence, which typically relies on self-reported wear time, prone to recall and response bias. Objective measures (eg, temperature sensors), validated in footwear and orthotics research, offer a potentially more accurate alternative to self-reporting. Despite this, there is no research comparing self-reported and sensor-measured wear times in a knee brace.

Objective: This study aimed to determine how well self-reported wear times reflect sensor-measured data in a slim-fit knee brace.

Methods: Young adults (aged 18-45 years), 1-8 years post-ACLR, with a symptomatic knee (the 4 Knee injury and Osteoarthritis Outcome Score subscales [$KOOS_4$] score <80/100) wore a slim-fit brace during a 6-week feasibility trial. This study reports a secondary analysis of participants allocated to the brace group. Self-reported wear times were recorded in daily logs. An undisclosed, embedded temperature sensor recorded temperature every 10 minutes. A wear detection algorithm identified brace donning and doffing. These data were used to calculate aggregated measures (ie, summary measures across the entire 6-week intervention period, including cumulative wear time, average daily wear time, and total number of days worn) and repeated measures (daily wear duration, 3- and 7-day rolling averages, where wear time is averaged over consecutive days). Agreement between self-reported and sensor-based measures was assessed using concordance correlation coefficients (CCCs) and limits of agreement (LoA).

Results: Of the 14 randomized participants, 10 (30% male [n=3]; mean age: 33, SD 6 years; time post-ACLR: 4, SD 1 years) had both temperature sensor and self-reported wear data. Six participants (60%) underreported average daily wear time (mean 29, SD 24 minutes across all 10 participants), while nine (90%) overreported the number of days worn (mean 9, SD 6 days across all 10 participants). Daily wear time showed moderate agreement between the sensor and self-reporting (CCC 0.70, 95% CI 0.58-0.79), but wide LoA (−223 to 217 minutes). Using 3- or 7-day rolling averages narrowed LoA (−47 to 36 minutes per day and −14 to 10 minutes per day, respectively) and slightly improved CCCs (0.74, 95% CI 0.58-0.85, and 0.73, 95% CI 0.51-0.86, respectively). Greater agreement was observed with more aggregated outcomes; for total 6-week wear time, the CCC was 0.84 (95%CI 0.50-0.95). When expressed as daily average wear time, the CCC was excellent (0.92, 95% CI 0.73-0.98), although daily LoA remained wide (−68 to 32 minutes), indicating substantial individual variability between self-reported and sensor-based measures. For the total number of days worn, the CCC was moderate (0.64, 95% CI 0.15-0.88) and LoA was wide (−10 to 22 days).

XSL•FO
RenderX

Conclusions: Self-reported daily brace wear time is inaccurate compared to wear time measured by the temperature sensor. Aggregated data and rolling averages showed better agreement. Future intervention studies should consider objective adherence measures. Failing this, averaging self-reported wear time across the intervention period could improve accuracy.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12623001027606; https://tinyurl.com/2spr7bnu

(JMIR Rehabil Assist Technol 2026;13:e79725) doi:10.2196/79725

KEYWORDS

rehabilitation; knee brace; knee osteoarthritis; physiotherapy; knee injury; sensor; wearable technology

Introduction

Slim-fit knee braces can improve pain and function after anterior cruciate ligament reconstruction (ACLR) [1,2], but adherence to wearing them is critical for these benefits [3]. Most studies use self-reported measures of adherence [1-5], which are prone to recall, response, and social desirability bias [6] and missing data [3]. Objective measures are needed to track wear time accurately [4]. One promising, low-cost method to objectively measure adherence (ie, wear time) is with an embedded temperature sensor. Such sensors are highly accurate at detecting wear times in footwear and orthotics when compared to reference standard data (ie, camera [7], smartphone application [8]). However, no research compares self-reported and sensor-measured wear times in a knee brace, which has implications for evaluating adherence in intervention studies. Without accurate adherence data, the effectiveness of brace interventions may be misestimated. We aimed to determine how well self-reported wear times reflect temperature sensor-measured data in a slim-fit knee brace.

Methods

Ethical Considerations

Ethics approval was granted by the La Trobe University Ethics Committee (HEC23249) and written informed consent was gained from all participants prior to enrollment in the study. This trial involved limited disclosure, where participants in the brace group were not informed of the presence of a small temperature sensor embedded in the brace. As part of our ethical approval, all participants were fully debriefed about the sensor at the end of the trial, where a debriefing form was signed. All participant data were deidentified prior to analysis. Identifiable information was stored separately from study data on secure, password-protected servers accessible only to the research team. Participants received an AUD $25 (USD $18.05) gift voucher for their time, and the knee brace used in the study was provided to participants at no cost.

Study Design

This study is a secondary analysis of a 6-week feasibility trial (ACTRN12623001027606) [9] investigating a slim-fit knee brace in 21 young adults with a symptomatic knee following ACLR. No formal power calculation was performed for this exploratory analysis, consistent with our previous investigations [10], which includes only the subset of participants allocated to the brace group. Eligibility criteria included: (1) 1-8 years post-ACLR; (2) aged 18-45 years; and (3) a symptomatic knee (mean score <80/100 from the 4 Knee injury and Osteoarthritis Outcome Score subscales [$KOOS_4$] [11]). Exclusion criteria were: (1) knee reinjury or lower-limb surgery in last 3 months; (2) routine use of a knee brace in the last 3 months; (3) received treatment for their index knee in the last month; (4) another condition affecting physical function; (5) pregnant; or (6) unable to understand English. For this analysis, 14 participants randomized to the brace group were included. Participants were advised to wear the brace for at least 1 hour per day, especially during knee-aggravating activities. Self-reported wear time over the intervention period was measured via a daily log on a smartphone application or paper. Several reminders were sent if a day was missed. Objective wear time was recorded via a small, commercially available temperature sensor (Orthotimer; Balingen, Germany; 9 mm×13 mm×4.5 mm) embedded in the proximal region on the side of the knee brace to ensure it was not noticeable to participants (Figure 1). The sensor uses a lithium dry cell battery (3.0 V/5.5 mAh) with a lifespan of >18 months, is waterproof, can withstand high pressures and temperatures, and has up to 400 days of storage capacity. The sensor has a temperature precision of ±0.1 °Celsius, and measured temperature once every 10 minutes. Participants were not informed of the sensor to prevent biasing self-reported wear time data. Awareness of the presence of the sensor was assessed during 6-week follow-up interviews.

XSL•FO
RenderX

Figure 1. The slim-fit knee brace (GenuTrain; Bauerfeind, Germany) and temperature sensor (Orthotimer; Balingen, Germany).



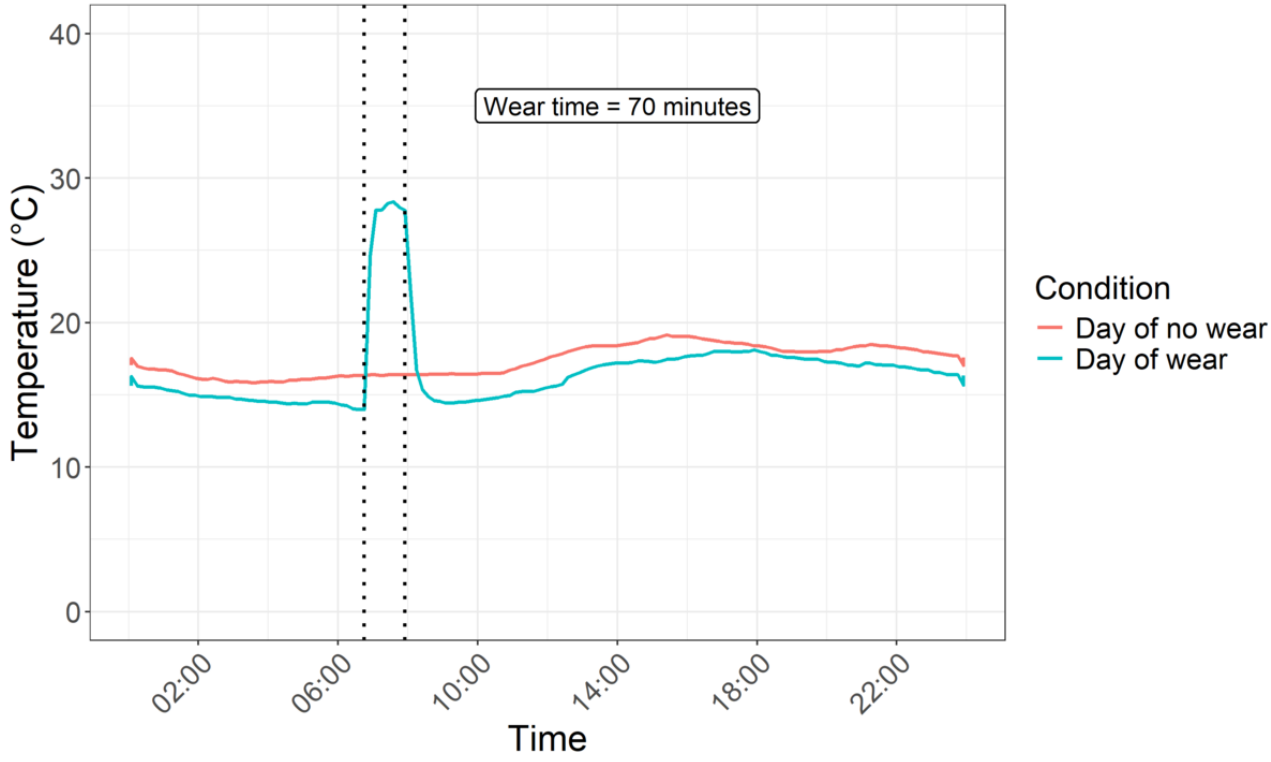
Data Analysis

Temperature Sensor Data Processing

A wear detection algorithm in Microsoft Excel (Multimedia Appendix 1) detected temperature spikes and drops to identify brace donning and doffing, respectively (Figure 2). Compared to algorithms using absolute or relative threshold values, this is the most accurate approach, and it is not influenced by ambient temperatures [8].

XSL•FO
RenderX

Figure 2. Temperature profile over 2 different days for the same participant, comparing wear (total 70 minutes) and no wear conditions.



Statistical Analysis

Our analyses included (1) a description of average daily knee brace wear times (total minutes worn divided by the number of days reported in the daily log or recorded by the temperature sensor) and the number of days worn over the intervention period (days with non-zero wear time); and (2) analyses of agreement for total daily wear times, 3- and 7-day rolling averages, total wear time during the 6-week intervention period (a single cumulative value in minutes per participant, distinct from daily or rolling averages), average daily wear time over the intervention period, and the number of days worn. All data analyses were performed in R-studio (version 4.4.0). To assess agreement between the temperature sensor and self-reported wear times—accounting for repeated measures (ie, daily agreement data for each day across the intervention period for each participant)—we used concordance correlation coefficients (CCCs) and 95% limits of agreement (LoA). We considered CCC values as follows: <0.5 = poor, 0.5 - 0.75=moderate, 0.75 - 0.9=strong, and >0.9=excellent, consistent with common statistical practice [12]. We interpreted LoA width relative to the scale of measurement (eg, minutes per day or total days), with narrower ranges indicating less variability between methods. Sensitivity analyses assessed the impact of handling missing self-reported data, with (1) missing data imputed as "no wear" days and (2) analyses repeated excluding 1 participant who accounted for most (13 of 15) missing values.

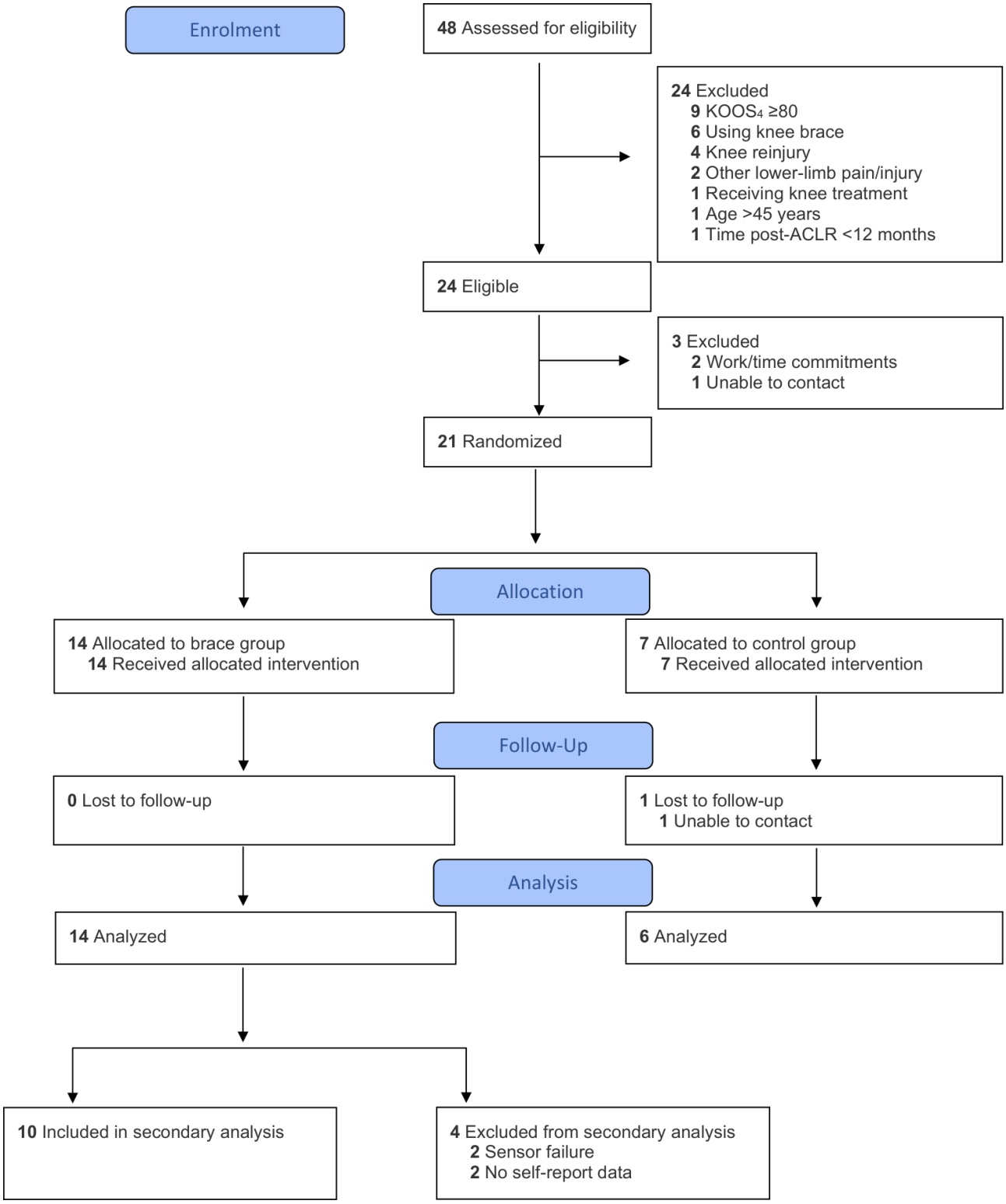
Results

Participant Characteristics

Two participants were missing all temperature sensor data due to software malfunction, and 2 completed no days of self-reported data (Figure 3). The 10 remaining participants contributed to analysis (30% male [n=3]; mean age: 33, SD 6 years; body mass index: 27, SD 4 kg/m^2; time post-ACLR: 4, SD 1 years; time between baseline and follow-up: 7, SD 2 weeks). No participant reported being aware of the brace temperature sensor during the intervention period.

XSL•FO
RenderX

Figure 3. Participant flow through the secondary analysis of the 6-week feasibility trial. ACLR: anterior cruciate ligament reconstruction; $KOOS_4$: 4 Knee injury and Osteoarthritis Outcome Score subscales.



Analysis of Agreement

There was variation between temperature sensor and self-reported data (Figure 4). Over the intervention period, 6 participants (60%) underreported the average daily knee brace wear time, and 4 (40%) overreported (mean 29, SD 24 minutes across all 10 participants) (Figure 5A). In contrast, 9 participants (90%) overreported the number of days they wore the brace, with only 1 (10%) underreporting (mean 9, SD 6 days across all 10 participants) (Figure 5B).

For repeated measures of daily wear time (ie, total minutes each day), we observed moderate agreement between the temperature sensor and self-report (CCC 0.70, 95% CI 0.58-0.79), with wide LoA (−223 to 217 minutes per day). Using 3-day rolling averages narrowed the LoA (−47 to 36 minutes per day), and 7-day rolling averages yielded the narrowest LoA (−14 to 10

XSL•FO
RenderX

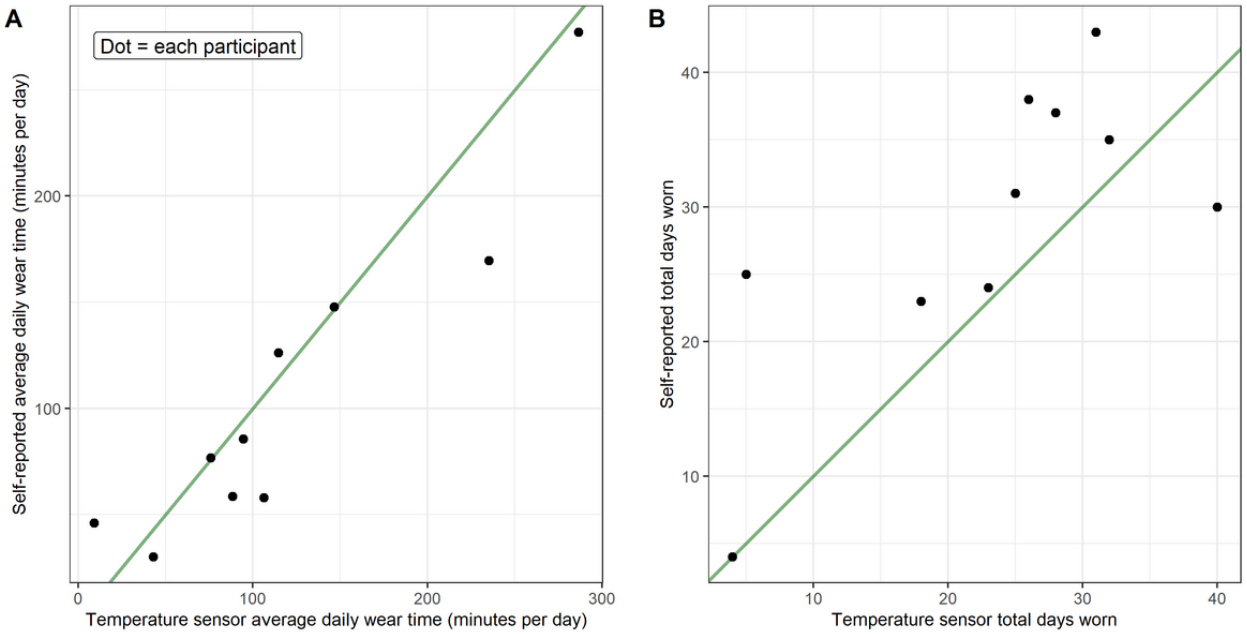
minutes per day), with slightly improved CCCs (0.74, 95% CI 0.58-0.85, and 0.73, 95% CI 0.51-0.86, respectively), indicating improved agreement. For data aggregated over longer periods, agreement was strong when evaluating total wear time across the entire intervention period, with a higher CCC (0.84, 95% CI 0.50-0.95). LoA were wider however (−3065 to 1474 minutes), due to the larger scale of total minutes across the entire intervention period. When expressed as a daily average wear time, the CCC was excellent (0.92, 95% CI 0.73-0.98) and LoA narrowed (−68 to 32 minutes), though remained relatively wide, indicating substantial individual participant variability between self-reported and sensor-based measures. For the total number of days worn, the CCC was moderate (0.64, 95% CI 0.15-0.88) and LoA was wide (−10 to 22 days). Sensitivity analyses with imputed missing data or excluding 1 participant slightly altered LoA but not CCCs (Multimedia Appendix 2).

Figure 4. Individual participant knee brace wear time (minutes per day) comparing temperature sensor and self-report across the intervention period. The x-axis represents days of the intervention period and the y-axis represents daily wear time (minutes). Each point represents an individual participant's wear time measured by the temperature sensor (in blue) or self-report (in red). Post-intervention testing was scheduled for 6 weeks after baseline. However, due to individual differences in participant availability, the exact timing varied slightly between participants.



XSL•FO
RenderX

Figure 5. Agreement between temperature sensor and self-reported knee brace wear. Comparison of (A) average daily wear time (minutes per day, across the entire intervention period) and (B) total days worn, measured by the temperature sensor (x-axis) and self-report (y-axis). Each point represents a single participant. The green line represents the line of equality, indicating perfect agreement between methods.



Discussion

Our exploratory study highlights discrepancies between self-reported and temperature sensor-measured knee brace wear time. While daily knee brace wear time from the temperature sensor and daily log measures demonstrated moderate correlation, the wide LoA reflected high variability and poor accuracy in day-to-day self-reporting. Two participants (14%) failed to complete their daily logs, further highlighting the limitations of relying on self-reported knee brace wear times.

Averaging wear time over the entire intervention period, or using 3- and 7-day rolling averages, improved agreement between the temperature sensor and self-reported measures, likely due to recall bias being washed out over time. Indeed, averaging daily wear time across the entire intervention period showed excellent agreement, emphasizing that self-reports are more accurate when viewed as long-term averages, smoothing discrepancies in day-to-day reporting. This supports the common approach of reporting average wear time per day across a study period [2,13-15], which our results suggest is the optimal method for increasing accuracy of self-reported wear time estimates. However, we found poor agreement between the temperature sensor and self-reporting when knee brace wear time was assessed based on the number of days worn, with participants tending to overreport. Previous studies show that participants tend to overestimate knee brace wear time [16], with the same phenomenon observed with other interventions (eg, exercise, footwear, braces) [3,17,18], perhaps due to response or social desirability bias, where participants may overreport adherence to appear more compliant. Similar discrepancies between self-reported and objectively measured adherence have been reported in broader health interventions, including physical activity monitoring [19] and medication adherence programs, such as in antiretroviral therapy [20]. These findings further emphasize the importance of objective adherence measurement (eg, temperature sensors) to accurately measure knee brace wear time.

While CCCs indicate moderate-excellent agreement, the wide LoA reflect substantial variability. Our results indicate that daily self-reported wear time could be expected to differ from sensor-measured wear time by up to 3.5 hours higher or lower per day (LoA −223 to 217 minutes), potentially misjudging adherence and influencing clinical or research decisions if they are based on single day wear times. However, aggregating wear over 7-day periods reduced this discrepancy (LoA −14 to 10 minutes), with self-reported wear time underestimating sensor-measured wear by up to 14 minutes or overestimating it by up to 10 minutes. Using the average daily wear across the entire intervention period reduced variability (LoA −68 to 32 minutes), with self-reported wear time underestimating sensor-measured wear by up to 68 minutes or overestimating it by up to 32 minutes, suggesting that longer-term averages smooth day-to-day reporting errors.

In clinical settings, where adherence to knee brace wear is important for treatment outcomes [3], reliance on self-reported wear time may misclassify adherence and lead to suboptimal management. From a research perspective, inaccurate adherence measurement may bias assessments of knee brace intervention effectiveness [3]. However, these findings are exploratory and should be interpreted with caution given the small sample size. Future studies should validate temperature-based wear detection for knee braces and evaluate adherence methods in larger samples.

Key limitations should be acknowledged. Firstly, the temperature sensor, used in this study, has not been specifically validated for knee braces, though it has demonstrated high accuracy in detecting wear time in footwear and orthotics compared to reference standards, such as cameras or smartphone applications [7,8]. Knee-specific factors (eg, skin contact, brace

XSL•FO
RenderX

positioning, local sweating) could influence temperature measurements and thus sensor accuracy. While we expect sensor performance to be broadly transferable, direct validation in knee braces would strengthen confidence in these findings. Secondly, while temperature spikes were used to infer brace donning and doffing, environmental factors may have influenced temperature changes (eg, transitioning to a hot car or air-conditioned room). Finally, the small sample size (n=10) limits the generalizability of our findings and likely contributes to the wide limits of agreement observed. As such, our results should be interpreted as preliminary and hypothesis-generating rather than definitive estimates of adherence accuracy.

These exploratory findings highlight potential limitations of relying solely on self-reported knee brace wear. Future studies could consider objective measures of knee brace wear adherence (eg, temperature sensors) to complement self-reported data and mitigate the impact of recall and response bias. In settings where objective monitoring is not available, averaging self-reported wear time across the entire intervention period could improve accuracy compared to day-to-day reporting.

Acknowledgments

Declaration of generative artificial intelligence (AI) and AI-assisted technologies in the writing process: During the preparation of this manuscript, the authors used ChatGPT to assist with improving the clarity and readability of individual sentences. All content was written by author MS, with subsequent input and revisions from co-authors. The authors reviewed and edited the manuscript in full and take complete responsibility for the content of the published article.

Funding

MS is supported by a La Trobe University Graduate Research Scholarship. JLC is supported by an Australian Government Research Training Program scholarship. AGC is a recipient of a National Health and Medical Research Council (NHMRC) of Australia Investigator Grant (GNT2008523). No other authors received funding for the submitted work. For the purposes of open access, the authors have applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The funders had no role in any part of the study or in any decision about publication.

Data Availability

Further data are available from the corresponding author on reasonable request.

Authors' Contributions

MS: data curation, formal analysis, investigation, methodology, project administration, validation, visualization, writing – original draft, writing – review and editing. BFM: formal analysis, methodology, supervision, validation, writing – review and editing. HS: data curation, formal analysis, investigation, project administration, validation, writing – review and editing. FS: data curation, investigation, project administration, writing – review and editing. DLC: formal analysis, methodology, software, validation, visualization, writing – review and editing. JLC: methodology, writing – review and editing. AMB: supervision, validation, writing – review and editing. AGC: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, writing – review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Wear detection algorithm.
[DOCX File, 52 KB - rehab_v13i1e79725_app1.docx]

Multimedia Appendix 2
Sensitivity analyses.
[DOCX File, 20 KB - rehab_v13i1e79725_app2.docx]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 1170 KB - rehab_v13i1e79725_app3.pdf]

References

1. Birmingham TB, Bryant DM, Giffin JR, et al. A randomized controlled trial comparing the effectiveness of functional knee brace and neoprene sleeve use after anterior cruciate ligament reconstruction. Am J Sports Med 2008 Apr;36(4):648-655. [doi: 10.1177/0363546507311601] [Medline: 18192493]

XSL•FO
RenderX

2. Sole G, Lamb P, Pataky T, Klima S, Navarre P, Hammer N. Immediate and 6-week effects of wearing a knee sleeve following anterior cruciate ligament reconstruction: a cross-over laboratory and randomised clinical trial. BMC Musculoskelet Disord 2021 Aug 4;22(1):655. [doi: 10.1186/s12891-021-04540-x] [Medline: 34348704]
3. Davies G, Yeomans D, Tolkien Z, et al. Methods for assessment of patient adherence to removable orthoses used after surgery or trauma to the appendicular skeleton: a systematic review. Trials 2020 Jun 8;21(1):507. [doi: 10.1186/s13063-020-04456-2] [Medline: 32513246]
4. Devanand DB, Kedgley AE. Objective methods of monitoring usage of orthotic devices for the extremities: a systematic review. Sensors (Basel) 2023 Aug 25;23(17):7420. [doi: 10.3390/s23177420] [Medline: 37687876]
5. Hart HF, Crossley KM, Ackland DC, Cowan SM, Collins NJ. Effects of an unloader knee brace on knee-related symptoms and function in people with post-traumatic knee osteoarthritis after anterior cruciate ligament reconstruction. Knee 2016 Jan;23(1):85-90. [doi: 10.1016/j.knee.2015.05.006] [Medline: 26117486]
6. Horne R, Weinman J. Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. J Psychosom Res 1999 Dec;47(6):555-567. [doi: 10.1016/s0022-3999(99)00057-4] [Medline: 10661603]
7. Lutjeboer T, van Netten JJ, Postema K, Hijmans JM. Validity and feasibility of a temperature sensor for measuring use and non-use of orthopaedic footwear. J Rehabil Med 2018 Nov 7;50(10):920-926. [doi: 10.2340/16501977-2494] [Medline: 30299524]
8. Menz HB, Bonanno DR. Objective measurement of adherence to wearing foot orthoses using an embedded temperature sensor. Med Eng Phys 2021 Feb;88(1):19-24. [doi: 10.1016/j.medengphy.2020.12.002] [Medline: 33485509]
9. Savage M, Mentiplay BF, Bruder AM, et al. A Slim-Fit Knee Brace for Young Adults With a Symptomatic Knee Following Anterior Cruciate Ligament Reconstruction: A Feasibility Randomized Controlled Trial. J Sport Rehabil 2026 May 11;2026:1-11. [doi: 10.1123/jsr.2025-0430] [Medline: 42114824]
10. Couch JL, King MG, De Oliveira Silva D, et al. Diagnostic performance of self-reported knee crepitus using a Knee injury and Osteoarthritis Outcome Score item. Osteoarthr Cartil 2025 Jun;33(6):764-769. [doi: 10.1016/j.joca.2024.12.004] [Medline: 39709087]
11. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998 Aug;28(2):88-96. [doi: 10.2519/jospt.1998.28.2.88] [Medline: 9699158]
12. Parker RA, Scott C, Inácio V, Stevens NT. Using multiple agreement methods for continuous repeated measures data: a tutorial for practitioners. BMC Med Res Methodol 2020 Jun 12;20(1):154. [doi: 10.1186/s12874-020-01022-x] [Medline: 32532218]
13. Callaghan MJ, Parkes MJ, Hutchinson CE, et al. A randomised trial of a brace for patellofemoral osteoarthritis targeting knee pain and bone marrow lesions. Ann Rheum Dis 2015 Jun;74(6):1164-1170. [doi: 10.1136/annrheumdis-2014-206376] [Medline: 25596158]
14. Priore LB, Lack S, Garcia C, Azevedo FM, de Oliveira Silva D. Two weeks of wearing a knee brace compared with minimal intervention on kinesiophobia at 2 and 6 weeks in people with patellofemoral pain: a randomized controlled trial. Arch Phys Med Rehabil 2020 Apr;101(4):613-623. [doi: 10.1016/j.apmr.2019.10.190] [Medline: 31830433]
15. Stetter BJ, Fiedler J, Arndt M, Stein T, Sell S. Impact of a semi-rigid knee orthotic intervention on pain, physical activity, and functional capacity in patients with medial knee osteoarthritis. J Clin Med 2024 Mar 7;13(6):38541761. [doi: 10.3390/jcm13061535] [Medline: 38541761]
16. Holden MA, Murphy M, Simkins J, et al. Knee braces for knee osteoarthritis: a scoping review and narrative synthesis of interventions in randomised controlled trials. Osteoarthr Cartil 2024 Nov;32(11):1371-1396. [doi: 10.1016/j.joca.2024.08.010] [Medline: 39218202]
17. Grubhofer F, Gerber C, Meyer DC, et al. Compliance with wearing an abduction brace after arthroscopic rotator cuff repair: a prospective, sensor-controlled study. Prosthet Orthot Int 2019 Aug;43(4):440-446. [doi: 10.1177/0309364619837794] [Medline: 30895872]
18. Nicolson PJA, Hinman RS, Wrigley TV, Stratford PW, Bennell KL. Self-reported home exercise adherence: a validity and reliability study using concealed accelerometers. J Orthop Sports Phys Ther 2018 Dec;48(12):943-950. [doi: 10.2519/jospt.2018.8275] [Medline: 30053792]
19. Larsen RT, Wagner V, Korfitsen CB, et al. Effectiveness of physical activity monitors in adults: systematic review and meta-analysis. BMJ 2022 Jan 26;376:e068047. [doi: 10.1136/bmj-2021-068047] [Medline: 35082116]
20. Kanters S, Park JJH, Chan K, et al. Interventions to improve adherence to antiretroviral therapy: a systematic review and network meta-analysis. Lancet HIV 2017 Jan;4(1):e31-e40. [doi: 10.1016/S2352-3018(16)30206-5] [Medline: 27863996]

Abbreviations

ACLR: anterior cruciate ligament reconstruction
CCC: concordance correlation coefficient
$KOOS_4$: 4 Knee injury and Osteoarthritis Outcome Score subscales
LoA: limits of agreement

XSL•FO
RenderX

Edited by S Munce; submitted 26.Jun.2025; peer-reviewed by KV Sangaraju, S Turnbull; revised version received 30.Mar.2026; accepted 09.Apr.2026; published 27.May.2026.

Please cite as:
Savage M, Mentiplay BF, Slater H, Serighelli F, Carey DL, Couch JL, Bruder AM, Culvenor AG
Accuracy in the Estimation of Self-Reported Knee Brace Wear Time in Young Adults With a Symptomatic Knee Following ACL Reconstruction: Secondary Analysis of a Pilot Randomized Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e79725
URL: https://rehab.jmir.org/2026/1/e79725
doi:10.2196/79725

© Matthew Savage, Benjamin F Mentiplay, Harriette Slater, Fernanda Serighelli, David L Carey, Jamon L Couch, Andrea M Bruder, Adam G Culvenor. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 27.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Functions and Sensors of Smart Walkers From 2015 to 2024: Scoping Review

Nicole Strutz[1,2], Dr rer medic; Hanna Brodowski[3], Dr rer medic; Stephan Schulze[2], Dr phil

[1]MUC.HEALTH, Munich University of Applied Sciences, Landsbergerstraße 187, Munich, Bavaria, Germany

[2]Department of Orthopedic and Trauma Surgery, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

[3]Department of Physiotherapy, Pain and Exercise Research Luebeck (P.E.R.L.), Institute of Health Sciences, Universität zu Lübeck, Lübeck, Germany

Corresponding Author:
Nicole Strutz, Dr rer medic
MUC.HEALTH, Munich University of Applied Sciences, Landsbergerstraße 187, Munich, Bavaria, Germany

Abstract

Background: Early mobilization and mobility are essential components of the recovery process following surgery and trauma-related hospitalization. In addition to personalized support from physiotherapists and health care professionals, assistive devices such as walkers play a crucial role in facilitating safe and effective mobility.

Objective: This scoping review aims to provide a comprehensive overview of the current state of the literature on the design, sensor technologies, and functional applications of smart walkers and to assess the extent to which existing studies reflect clinical use cases.

Methods: Peer-reviewed English articles published between 2015 and 2024 were identified by searching PubMed, CINAHL, SSCI, and IEEE, focusing on the topic of smart walkers. Secondary analyses and walkers with 2 wheels or fewer were excluded in abstract screening. Study screening and selection were performed according to the Joanna Briggs Institute guidelines for scoping research and reported following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. The Rayyan systematic review management software was used for study selection. The articles included were analyzed with respect to the sensor technologies used, their functional capabilities, and their application scenarios.

Results: Of the 800 articles screened, 44 (5.5%) met the inclusion criteria. Most of these articles were research reports (n=36, 81.8%) and were conducted in laboratory-based environments (n=30, 68.2%). Most studies evaluated smart walkers in asymptomatic populations (n=29, 65.9%), with half (n=22, 50%) involving younger adults. Among the sensor modalities reported, camera-based and light detection and ranging–based sensors were most prevalent for half of the implementations. Light detection and ranging–based sensors can be categorized according to their primary functions: gait analysis (n=11, 25%), collision detection (n=9, 36%), and navigation (n=5, 11.4%). Load sensors (n=10, 22.7%) and ultrasonic sensors (n=11, 25%) were among the most frequently cited sensor modalities in the literature. Load sensors, also known as force sensors, are integrated into the handlebars, frame, forearm supports, or chest pads of smart walkers. These sensors measure the user's load, providing essential data for calculating body weight support or inferring the user's intention to move.

Conclusions: The smart walkers described in the literature were predominantly tested in asymptomatic and younger populations. Bridging the gap between current laboratory-based research and real-world clinical environments, as well as the daily lives of end users, remains a critical objective. Addressing the specific needs of older adults through comprehensive requirements analyses and iterative testing continues to be an ongoing challenge, yet these processes can serve as integral components of research and development projects.

Trial Registration: OSF Registries osf.io/ctpf4; https://osf.io/ctpf4

(JMIR Rehabil Assist Technol 2026;13:e78480) doi:10.2196/78480

KEYWORDS

walking aid; navigation; gait analysis; older adults; clinical setting; user feedback; observer feedback

Introduction

In contemporary society, we are confronted with the dual challenges of an aging population and evolving demands in daily routines. These challenges are compounded by age-related impairments in walking ability, often due to lower limb discomfort [1] and/or compromised balance [2], which necessitate the use of mobility aids. The selection of appropriate mobility aids (eg, wheelchair, crutches, and walker) has to be tailored to the individual's remaining functional capacities [3].

XSL•FO
RenderX

Such aids are essential not only for addressing mobility issues but also for mitigating the long-term consequences of reduced compensatory ability (eg, muscle atrophy, altered gait patterns, and further decline in balance) [4]. Without appropriate intervention, this progressive decline may lead to a further reduction in residual capacities, potentially creating a cycle of dependency and physical deterioration [5].

Mobility aids are not restricted to indoor use; they are also vital for maintaining autonomy and mobility outside the home, such as for visiting medical practitioners, shopping, or engaging in social activities. The primary target group for these devices is older adults, for whom mobility aids play a vital role in maintaining participation in daily life and preserving independence in later years [6].

In hospital settings, individuals remain dependent on mobility aids to prevent the adverse effects of immobility, including muscle atrophy, joint contractures, and psychological deterioration. Although clinical settings typically prioritize disease diagnosis and treatment, preserving mobility and promoting early mobilization [7,8] with the aid of assistive devices is an essential objective for health care providers [9]. In this context, walkers and other gait-assisting devices play a crucial role, especially in early mobilization.

Technological advancements in mobility aids have been ongoing for several years, leading to the development of functionally enhanced walkers. These devices now integrate various assistive technologies, such as motorized support for propulsion, navigation systems, balance detection, counter-steering mechanisms in response to balance loss, and compensatory features for sensory deficits [10,11]. Additionally, smart walkers may serve as valuable tools in biomechanical diagnostics within orthopedic clinics when gait and/or postural analysis are implemented or for individualized risk prediction of falls [12]. According to Valadão et al [13], smart walkers are "...walkers that contain, besides the mechanical structure to support the user, electronics, control systems and sensors, in order to allow a better user experience and minimize the risks of falling."

The aim of this scoping review is to identify and map the current applications of functionally enhanced walkers, commonly referred to as smart walkers. The study aims to answer the following primary research question: What are the current functions of smart walkers, and in which settings are they applied?

Secondary research questions include the following: (1) Which types of sensors are integrated into smart walkers? (2) What are the specific application scenarios for these devices, including contextual settings and target user groups?

Methods

Study Design

This scoping review was conducted in accordance with the Joanna Briggs Institute guidelines for scoping research and reported following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [14] guidelines. The literature search was performed in January and February 2025. The study protocol was registered on the Open Science Framework platform (registration number: 10.17605/OSF.IO/CTPF4).

Information Sources and Search Strategy

The search encompassed publications from January 2015 to December 2024, using a predefined search string. The search string for the PubMed database is provided in Textbox 1. The search strings for 4 other databases can be found in Multimedia Appendix 1. An initial search was performed in PubMed, SSCI, CINAHL, IEEE, and Cochrane Library, which informed the development of a comprehensive search strategy subsequently applied across all sources. Final searches were conducted in the 5 databases. After removing duplicates and screening titles and abstracts, full-text reviews were conducted.

XSL•FO
RenderX

Textbox 1. Search string for the PubMed database.

PubMed

1. Walker* [Title/Abstract]
2. rollator* [Title/Abstract]
3. device supported [Title/Abstract]
4. walker* [MeSH Terms]
5. #1 OR #2 OR #3 OR #4
6. smart [Title/Abstract]
7. intelligent [Title/Abstract]
8. robot* [Title/Abstract]
9. artificial [Title/Abstract]
10. ai [Title/Abstract]
11. #6 OR #7 OR #8 OR #9 OR #10
12. DNA [Title/Abstract]
13. device-based [Title/Abstract]
14. molecular [Title/Abstract]
15. Covid* [Title/Abstract]
16. exoskeleton
17. teaching
18. walker* and avant*
19. #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18
20. #5 AND #11 NOT #19

Inclusion and Exclusion Criteria

Articles published in English in peer-reviewed journals that focused on smart walkers and independent walking users were included. Exclusion criteria encompassed secondary analyses (eg, reviews), smart walking aids with 2 wheels or fewer, or studies where the term “walker” referred to nonsmart devices.

Study Selection

The Rayyan (Rayyan Systems Inc) systematic review management software was used for study selection. Four independent reviewers (HB, MW, SS, and NS) initially screened titles and abstracts to exclude articles that did not meet the inclusion criteria. Subsequently, full-text screening was performed independently by 2 of the aforementioned reviewers. Discrepancies were resolved through discussion with a third reviewer until consensus was achieved.

Data Extraction

Data from each included study were extracted independently by HB, MW, NS, and SS using a structured data extraction framework. Any conflicts were resolved through discussion with the principal investigator (NS). The reviewers’ professional backgrounds include doctoral (PhD) degrees in physiotherapy and health sciences, with experience in health technologies related to mobility in older patients, as well as physiotherapy care and therapy for older patients. They also have research experience and have conducted studies on health technologies that address mobility. Furthermore, the researchers are sports scientists and nursing professionals with expertise in geriatric care for older clients and 10 years of experience in clinical gait analysis.

Data extracted from each publication included general publication details (year, author, journal, and country), study population and setting, and characteristics of smart walkers (hardware, software, and technological features such as sensors and feedback mechanisms), as well as technological features (sensors and feedback mechanisms). These data are presented in the *Results* section.

As per the scoping review methodology, the quality of included studies was not appraised.

Results

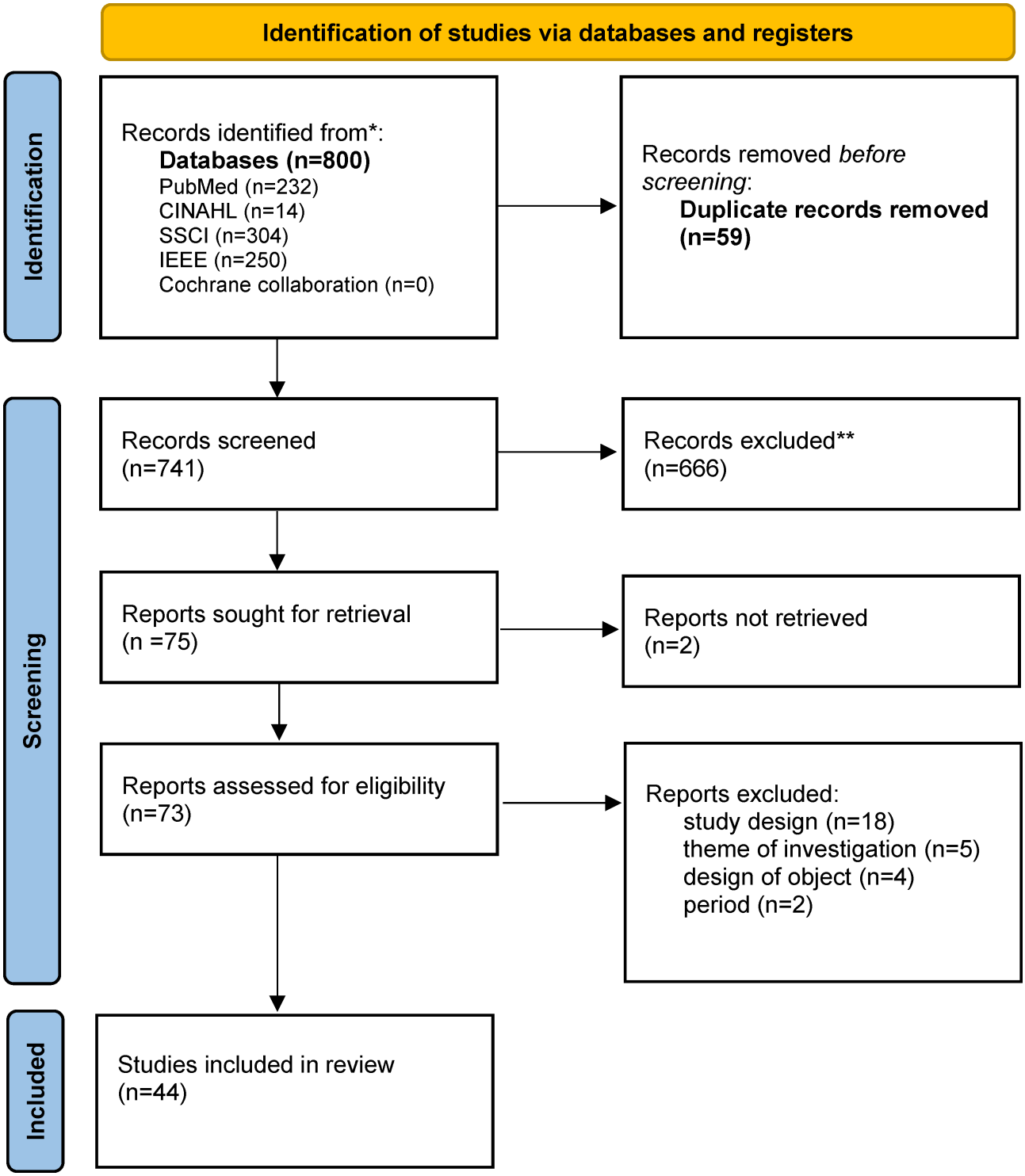
Overview

An initial pool of 800 articles was identified for screening (duplicates removed n=59, 7.4%). After title and abstract screening, 666 (83.3%) articles were excluded based on the predefined criteria related to theme and publication period. A total of 75 (9.4%) records were sought for retrieval. Although the original authors were contacted, 2 (0.3%) studies could not be included due to missing full texts. The remaining 73 (9.1%) articles underwent full-text review, leading to the exclusion of 29 (3.6%) studies due to study design, theme of investigation, and design of object or period. Ultimately, 44 (5.5%) articles

XSL•FO
RenderX

met the inclusion criteria. The entire selection process is summarized in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart [15] (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart for scoping reviews. *Number of records identified from each database. **Number of records excluded.

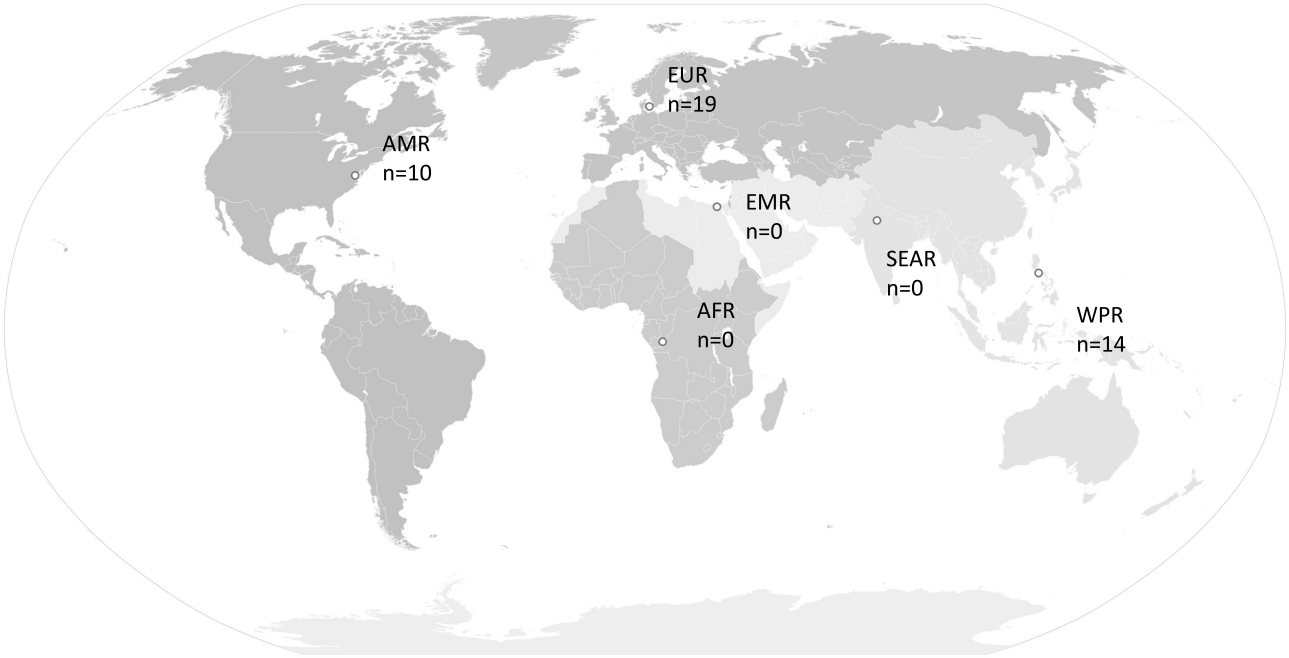


Description of Included Studies

Of the 800 articles screened, 44 (5.5%) met the predefined inclusion criteria. Geographically, the research was predominantly conducted in Europe (n=19, 43.2%; Figure 2).

XSL•FO
RenderX

Figure 2. Geographical distribution of included studies. Countries were grouped according to the World Health Organization regions. AFR: African Region; AMR: Region of the Americas; EMR: Eastern Mediterranean Region; EUR: European Region; SEAR: South-East Asia Region; WPR: Western Pacific Region.



Most of these studies were research reports (n=36, 81.8%). Sample sizes varied considerably across the included studies, ranging from a minimum of 1 participant to a maximum of 43 participants. A total of 5 (11.4%) studies did not report the number of participants. Most evaluations of smart walkers were conducted exclusively or partially in asymptomatic populations (n=29, 65.9%), with half (n=22, 50%) of the studies involving younger adults. A total of 14 (31.8%) studies involved older adults (6 studies without specifying age group and 1 without any person; Table 1).

XSL•FO
RenderX

Table . Overview of study characteristics grouped by year, setting, study type, and population.

Study, country	Year of publication	N	Setting	Study type	Population
Alves et al [16], Portugal	2016	6	Clinical	Research report	Symptomatic and younger adults
Caetano et al [17], Portugal	2016	Not specified	Clinical and rehabilitation	Research report	Not specified
Ballesteros et al [18], Netherlands, Spain, and Italy	2016	43	Laboratory and clinical	Research report	Symptomatic, asymptomatic, younger, and older adults
Hellström et al [19], Sweden	2016	12	Laboratory	Research report	Asymptomatic, symptomatic, and older adults
Valadão et al [13], Brazil	2016	Not specified	Laboratory	Research report	Asymptomatic[a]
Ballesteros et al [20], Spain	2017	19	Rehabilitation	Research report	Symptomatic[a]
Mun et al [21], Singapore	2017	10	Laboratory	Research report	Asymptomatic and younger adults
Ohnuma et al [22], Japan	2017	5	Not specified	Research report	Asymptomatic and older adults
Paulo et al [23], Portugal	2017	30	Laboratory	Research report	Asymptomatic and younger adults
Taghvaei et al [24], Japan	2017	4	Not specified	Research report	Asymptomatic and younger adults
Zhang and Ye [25], not applicable	2017	9	Laboratory	Research report	Not specified
Werner et al [26], Germany	2018	42	Clinical	Randomized controlled trial	Asymptomatic, symptomatic, and older adults
Andreetto et al [27], Italy	2019	29	Laboratory	Research report	Asymptomatic, symptomatic, younger, and older adults
Ferrari et al [28], Italy	2019	29	Laboratory	Research report	Asymptomatic, symptomatic, and older adults
Li et al [29], Japan	2019	3	Laboratory	Research report	Asymptomatic and younger adults
Moreira et al [30], Portugal	2019	3	Rehabilitation	Research report	Symptomatic[a]
Pérez-Rodríguez et al [31], Spain	2019	34	Clinical	Randomized controlled trial	Symptomatic and older adults
Sato et al [32], Japan	2019	11	Community dwelling	Research report	Asymptomatic and older adults
Scheidegger et al [33], Columbia	2019	5	Laboratory	Research report	Symptomatic and younger adults
Sierra et al [10], Columbia	2019	7	Laboratory	Research report	Asymptomatic and younger adults
Werner et al [34], Germany	2019	25	Laboratory	Research report	Symptomatic and older adults
Alazem et al [35], Canada	2020	5	Laboratory	Research report	Symptomatic, children, and younger adults
Jiménez et al [36], Brazil	2020	15	Laboratory	Research report	Not specified
Yeoh et al [37], Japan	2020	18	Not specified	Research	Asymptomatic and younger adults

XSL•FO
RenderX

Study, country	Year of publication	N	Setting	Study type	Population
Werner et al [38], Germany	2020	33	Laboratory	Research report	Asymptomatic and older adults
Zhao et al [39], China	2020	Not specified	Laboratory	Research report	Not specified
Chang et al [40], Taiwan	2021	Not specified	Laboratory	Research	Asymptomatic[a]
Mostofa et al [41], United States	2021	11	Community dwelling	Research report	Asymptomatic and older adults
Moustris et al [42], Greece	2021	32	Rehabilitation	CCT[b]	Asymptomatic, symptomatic, younger, and older adults
Orenius et al [43], Finland	2021	19	Not specified	CCT	Older adults[c]
Sierra et al [44], Columbia	2021	10	Laboratory	Research report	Asymptomatic and younger adults
Sierra et al [45], Columbia	2021	10	Laboratory	Research report	Symptomatic and older adults
Zhang et al [46], China	2021	8	Laboratory	Research	Asymptomatic and younger adults
Fernandez-Carmona et al [47], Spain	2022	11	Laboratory	Observational study	Younger and older adults[c]
Palermo et al [48], Portugal	2022	14	Laboratory	Research report	Asymptomatic and younger adults
Jacobs et al [49], United States	2023	11	Laboratory	Observational study	Asymptomatic and younger adults
Schwarz et al [50], Germany	2023	Not applicable	Laboratory	Research report	Younger adults[c]
Wang et al [51], China	2023	1	Laboratory	Research report	Asymptomatic[a]
Gong et al [52], China	2024	3	Laboratory	Research	Asymptomatic and younger adults
Jiang et al [53], China	2024	3	Laboratory	Observational study	Asymptomatic and younger adults
Machado et al [54], Brazil	2024	20	Laboratory	Observational study	Asymptomatic and younger adults
Mori et al [55], Japan	2024	5	Laboratory	Research report	Asymptomatic and younger adults
Sierra et al [56], United Kingdom	2024	14	Laboratory	Research report	Asymptomatic and younger adults
Zhang et al [57], Japan	2024	20	Community dwelling	Research report	Asymptomatic[a]

[a]No information about age.

[b]CCT: controlled clinical trial.

[c]No information about health status.

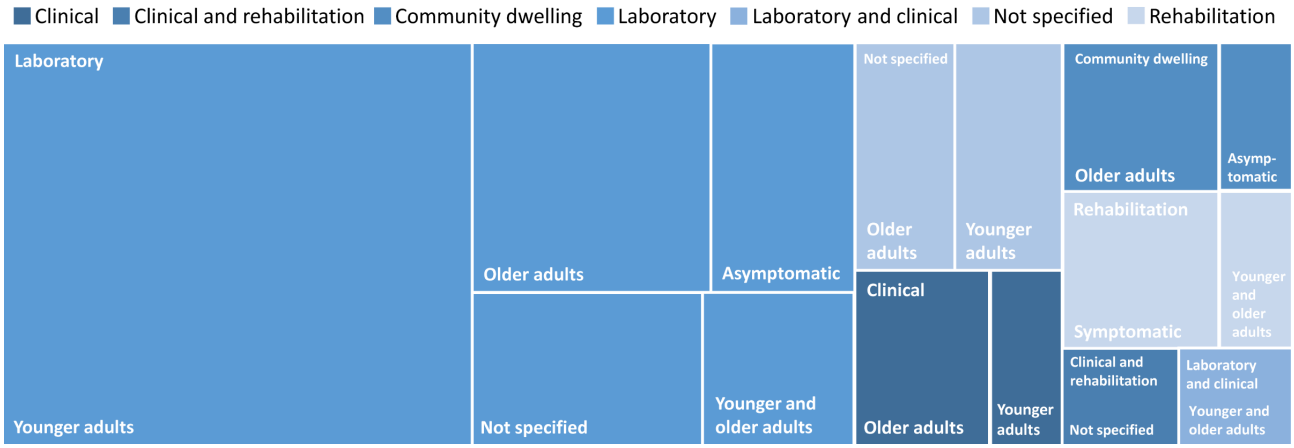
Settings and Participants

The studies were conducted across various settings. Investigations involving smart walkers were most frequently conducted in laboratory-based environments (n=30, 68.2%).

Four groups of participants were evident in the articles included: symptomatic children or younger adults (n=5, 11.4%), asymptomatic younger adults (n=19, 43.2%), symptomatic older adults (n=8, 18.2%), and asymptomatic older adults (n=9, 20.4%). Several articles did not specify the age (n=6, 13.6%) or health status (n=5, 11.4%) of participants. A total of 4 (9.1%) articles provided no information on the age and health status. As there were studies with multiple participant groups, the number of participant groups exceeded the number of articles. The group of participants consisting of asymptomatic younger adults was the most frequently studied group across the articles (n=19, 43.2%; Table 1, Figure 3).

XSL•FO
RenderX

Figure 3. Proportion of different settings and participants.



Walker-Specific Sensor Technologies and Functionalities

The included articles described various types of sensors and functions that constitute the extent of smart technology integrated into modern walkers. The sensors used were fundamental to the operation of these intelligent walkers. They encompassed a range of sensor types, including camera systems (2D cameras, depth cameras, and infrared cameras), optical sensors that used laser lights (2D laser scanners and light detection and ranging [LiDAR]), force and load sensors, ultrasound sensors, and inertial measurement units (IMUs; Table 2).

Among the sensor modalities reported, camera- and LiDAR-based sensors were the most prevalent, accounting for half of the implementations. Camera sensors were described either as standard 2D cameras (n=11, 25%) [10,17,23,24,27,28,31,33,42,44,56] or depth cameras. Depth cameras were used for multiple purposes, including gait analysis (n=3, 6.8%) [16,41,55] and intention recognition (n=3, 6.8%) [23,42,52], by capturing gaze direction. Additionally, depth cameras were used for collision avoidance through object detection (n=4, 9.1%) [17,23,25,42], and 1 (2.3%) study specifically described the use of depth cameras for fall prediction purposes [24]. LiDAR-based sensors can be categorized according to their primary functions: gait analysis (n=11, 25%), collision detection (n=9, 20.5%), and navigation (n=5, 11.4%).

In addition, load sensors (n=10, 22.7%) and ultrasonic sensors (n=11, 25%) were among the most frequently cited sensor modalities in the literature. Furthermore, load and force sensors were typically integrated into the handlebars of smart walkers (n=10, 22.7%) [10,17,18,23,39,44,45,47,48,56]. These sensors were described as “the simplest approach to onboard sensor-based gait analysis” [47] and were used to assess the user’s balance during ambulation [17]. Zhao et al [39] further detailed handgrip pressure sensors that can potentially prevent falls by detecting abnormal pressure or rapid pressure changes, indicating a fall risk. For instance, the smart walker developed by Ballesteros et al [18] incorporated force sensors in the handlebars alongside encoders to measure wheel rotation.

Beyond handlebar-mounted sensors, force sensors were also embedded elsewhere in the walker’s structure, such as in the frame, forearm supports, or chest pads (n=7, 15.9%) [16,32,46,51-54]. Sensors placed within the frame measured the user’s weight load, providing critical data for calculating body weight support (n=3, 6.8%) [46,51,52]. Additionally, Machado et al [54] and Jiménez et al [36] reported on force sensors located in forearm supports, which were used to infer the user’s intention to move, thereby enabling more responsive and adaptive assistance.

Ultrasound sensors were described in a quarter of the articles (n=11, 25%) and were used for collision avoidance [10,13,27,28,30,31,33,41,44,45,48]. IMUs were described with a similar frequency (n=12, 27.3%) [10,26,27,30,33,36,39,44,45,54-56]. An analysis of their functions revealed that they were mainly used to determine position and orientation, measure accelerations, record angular velocities, or support odometry sensors.

XSL•FO
RenderX

Table . Sensors and their function in smart walkers.

Functions	Sensors
Safety and collision detection	LiDAR[a], microphone, ultrasound sensors, camera, and infrared
Navigation	LiDAR, optical encoder, and IMUs[b]
Propulsive support and reduction of speed	Odometry sensors, force and load sensors, and IMUs
Body weight support	Force and load sensors
Intention to move	Force and load sensors and camera (detecting gaze)
Movement analysis (gait, balance, and postural control)	LiDAR, depth camera, and 2D laser range finder
Fall risk detection and falls	3D load sensors and depth camera

[a]LiDAR: light detection and ranging.

[b]IMU: inertial measurement unit.

With regard to sensory-based feedback generated, the following distinctions can be reported: user feedback, which provides relevant information on, for example, navigation details or collision warning during the use of the smart walker (Table 3), or observer feedback, which, for example, provides valuable additional information such as gait-related or system-relevant measurement data (Table 4) to the observer or therapist.

Table . User feedback mechanisms (multiple entries possible; N=44).

Category of observer feedback	Entries, n (%)
Haptic or tactile	10 (22.7)
Visual, user interface or navigation	13 (29.6)
Assistive control-based feedback	7 (15.9)
Virtual, multimodal or social	5 (11.4)
No user feedback	10 (22.7)
Unclear	3 (6.8)

Table . Observer feedback mechanisms (multiple entries possible; N=44).

Category of observer feedback	Entries, n (%)
Objective measurement data, gait analysis and sensor data	27 (61.4)
User experience, interaction and system data	14 (31.8)
No observer feedback	10 (22.7)
Unclear	4 (9.1)

Discussion

Principal Findings

This scoping review was conducted to provide a comprehensive overview of the current application settings, sensors, and functionalities of smart walkers.

Various sensors, such as LiDAR, depth camera, and 2D laser range finder, which could be used for a clinical assessment of the gait pattern or for measuring balance, have been integrated into smart walkers. Analysis of the intended use of sensors revealed a wide variety of applications. For instance, IMUs were mainly used to determine position and orientation, measure system accelerations, record angular velocities, and support odometry. The gait analysis functionality and the validity of the measured parameters depend on the sensor types used and their placement on the smart walker.

Analysis of the included studies indicates that validation data for gait-related sensor technology are often missing or only partially reported. Only 8 (22.9%) of the 35 studies measuring gait-relevant data used reference systems such as the one used in the study by Werner et al [34], IMU-based systems (MTw Awinda, Xsens Technologies, B.V.) [48], or a marker-based camera system operating at 120 Hz (Vicon motion analysis system) [44]. These studies provided the most methodologically robust evidence regarding measurement accuracy and are best suited for comparison with established clinical measurement systems. An additional 17 (38.6%) studies reported some form of validation but without a reference standard, mostly in the form of internal error measures (eg, root mean square error, difference metrics, and trial comparisons). While these data provide valuable technical insights, they allow only limited conclusions regarding external validity. Some articles provided no information on the validation of gait sensors, which is partly due to the studies' focus. In many studies, the primary objective

XSL•FO
RenderX

was different, such as navigation, safety, human-robot interaction, or technology design, meaning that gait analysis was only of marginal or no relevance. Overall, there is a clear gap between technical innovation and methodological transparency: although smart walker systems increasingly integrate complex sensor sets, validation against established gold standards remains uncommon. This lack of consistent validation limits the transferability of findings to clinical practice.

This review demonstrates that smart walkers that have been investigated in laboratory-oriented settings and among user populations are not representative of real-world applications. Most research has taken place in controlled laboratory environments and primarily involves asymptomatic and/or younger users of smart walkers. In contrast, older adults (representing an increasing global aging population) are key drivers of the expanding market for walking aids and walkers [58]. It is undisputed that younger adults have fewer functional limitations, better balance abilities, and overall exhibit fewer fall-related risk factors than older adults. As noted by Maranesi et al [59], research in this field should move beyond focusing solely on the technical functionality of gait devices and walkers. Our results illustrate a gap in effectively incorporating these devices into relevant settings, such as clinical environments or the daily lives of end users. This underscores the need for a comprehensive understanding of the physical, social, and cognitive needs of older adults. In 2025, a funded research and development initiative was launched in Germany to develop a prototype of a smart walker. The SmartRoll project adopts a user-centered study design, systematically identifying the needs of patients, professional caregivers, and physiotherapists [60]. These insights are subsequently evaluated with the target population—patients on an acute geriatric ward—as part of an integrated research and development effort. Additionally, the project involves the integration and testing of sensors capable of performing gait analysis in this vulnerable group of older adults. This scoping review indicates that the search for smart walkers in the clinical setting implies their use for research purposes within clinical environments. The relevance of smart walkers in clinical contexts is driven by the urgent need to address age-related skill decline [26] and mobility impairment [45]. This paper is embedded in a study that examines the current relevance of deploying smart rollators in clinical settings.

Older individuals who rely on assistive devices often do so due to underlying health conditions or following adverse events such as falls. Consequently, it is not surprising that users of walkers exhibit a fourfold increased risk of falls compared to individuals walking without such aids [61]. This population is particularly vulnerable during acute hospital stays. The use of smart walkers equipped with gait analysis and balance measurement functions could potentially be used in hospital settings to assess fall risk among patients. Indicators of fall risk are primarily identified through parameters that suggest an unstable gait balance and gait pattern. These parameters include, but are not limited to, foot angle at contact, step length, maximum foot height, and a high SD in step characteristics [1].

It should be noted that research comparing 2-, 3-, and 4-wheeled walkers is limited. Two-wheeled walkers often need to be partially lifted, requiring strength and coordination and affecting gait parameters. As the poor maneuverability and lower lateral stability of 2-wheeled walkers may pose additional risk factors [62], they were excluded from this review.

Conclusions

The smart walkers described in the literature were predominantly tested in asymptomatic and young populations. Developing smart walkers that explicitly address the specific needs of older adults—through systematic requirements analyses—and iterative testing throughout the development process remains an ongoing challenge. Integrating these steps as core components of research and development projects is essential to enhance clinical relevance.

Acknowledgments

The authors would like to thank Moritz Waage for his support with data extraction.

Funding

This work was supported by Federal Ministry of Education and Research (Germany; grant 16SV9335).

Authors' Contributions

NS conceptualized the scoping review; contributed substantially to drafting the *Methods*, *Results*, *Discussion*, and *Conclusions* sections; and prepared the original manuscript. HB contributed through review and editing of the manuscript. SS contributed to drafting the *Introduction* and *Discussion* sections and supervised the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings for the SSCI, CINAHL, IEEE Xplore, and Cochrane Library databases.
[DOCX File, 18 KB - rehab_v13i1e78480_app1.docx]

XSL•FO
RenderX

Checklist 1
PRISMA-ScR checklist.
[PDF File, 168 KB - rehab_v13i1e78480_app2.pdf]

References

1. Osoba MY, Rao AK, Agrawal SK, Lalwani AK. Balance and gait in the elderly: a contemporary review. Laryngoscope Investig Otolaryngol 2019 Feb;4(1):143-153. [doi: 10.1002/lio2.252] [Medline: 30828632]
2. Young WR, Ellmers TJ, Kinrade NP, Cossar J, Cocks AJ. Re-evaluating the measurement and influence of conscious movement processing on gait performance in older adults: development of the Gait-Specific Attentional Profile. Gait Posture 2020 Sep;81:73-77. [doi: 10.1016/j.gaitpost.2020.07.008] [Medline: 32683216]
3. Werner C, Ullrich P, Geravand M, Peer A, Hauer K. Evaluation studies of robotic rollators by the user perspective: a systematic review. Gerontology 2016;62(6):644-653. [doi: 10.1159/000444878] [Medline: 27008089]
4. Wang J, Li Y, Yang GY, Jin K. Age-related dysfunction in balance: a comprehensive review of causes, consequences, and interventions. Aging Dis 2024;16(2):714-737. [doi: 10.14336/AD.2024.0124-1] [Medline: 38607735]
5. Peterson KF, Adams-Price C. Fear of dependency and life-space mobility as predictors of attitudes toward assistive devices in older adults. Int J Aging Hum Dev 2022 Apr;94(3):273-289. [doi: 10.1177/00914150211027599] [Medline: 34191644]
6. Bertrand K, Raymond MH, Miller WC, Martin Ginis KA, Demers L. Walking aids for enabling activity and participation: a systematic review. Am J Phys Med Rehabil 2017 Dec;96(12):894-903. [doi: 10.1097/PHM.0000000000000836] [Medline: 29176406]
7. Schaller SJ, Anstey M, Blobner M, et al. Early, goal-directed mobilisation in the surgical intensive care unit: a randomised controlled trial. Lancet 2016 Oct 1;388(10052):1377-1388. [doi: 10.1016/S0140-6736(16)31637-3] [Medline: 27707496]
8. Okada Y, Unoki T, Matsuishi Y, Egawa Y, Hayashida K, Inoue S. Early versus delayed mobilization for in-hospital mortality and health-related quality of life among critically ill patients: a systematic review and meta-analysis. J Intensive Care 2019;7:57. [doi: 10.1186/s40560-019-0413-1] [Medline: 31867111]
9. Grund S, Gordon AL, van Balen R, et al. European consensus on core principles and future priorities for geriatric rehabilitation: consensus statement. Eur Geriatr Med 2020 Apr;11(2):233-238. [doi: 10.1007/s41999-019-00274-1] [Medline: 32297191]
10. Sierra M SD, Garzón M, Múnera M, Cifuentes CA. Human-robot-environment interaction interface for smart walker assisted gait: AGoRA Walker. Sensors (Basel) 2019 Jun 30;19(13):2897. [doi: 10.3390/s19132897] [Medline: 31262036]
11. Sun Y, Xiao C, Chen L, et al. A review of intelligent walking support robots: aiding sit-to-stand transition and walking. IEEE Trans Neural Syst Rehabil Eng 2024;32:1355-1369. [doi: 10.1109/TNSRE.2024.3379453] [Medline: 38502616]
12. Parsons R, Blythe RD, Cramb SM, McPhail SM. Inpatient fall prediction models: a scoping review. Gerontology 2023;69(1):14-29. [doi: 10.1159/000525727] [Medline: 35977533]
13. Valadão C, Caldeira E, Bastos-Filho T, Frizera-Neto A, Carelli R. A new controller for a smart walker based on human-robot formation. Sensors (Basel) 2016 Jul 19;16(7):1116. [doi: 10.3390/s16071116] [Medline: 27447634]
14. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
15. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
16. Alves J, Seabra E, Caetano I, et al. Considerations and mechanical modifications on a smart walker. Presented at: 2016 International Conference on Autonomous Robot Systems and Competitions (ICARSC); May 4-6, 2016. [doi: 10.1109/ICARSC.2016.30]
17. Caetano I, Alves J, Goncalves J, Martins M, Santos CP. Development of a biofeedback approach using body tracking with active depth sensor in ASBGo Smart Walker. Presented at: 2016 International Conference on Autonomous Robot Systems and Competitions (ICARSC); May 4-6, 2016. [doi: 10.1109/ICARSC.2016.34]
18. Ballesteros J, Urdiales C, Martinez AB, van Dieën JH. On gait analysis estimation errors using force sensors on a smart rollator. Sensors 2016 Nov 10;16(11):1896. [doi: 10.3390/s16111896]
19. Hellström T, Lindahl O, Bäcklund T, et al. An intelligent rollator for mobility impaired persons, especially stroke patients. J Med Eng Technol 2016 Jul;40(5):270-279. [doi: 10.3109/03091902.2016.1167973] [Medline: 27078084]
20. Ballesteros J, Urdiales C, Martinez AB, Tirado M. Automatic assessment of a rollator-user's condition during rehabilitation using the i-Walker platform. IEEE Trans Neural Syst Rehabil Eng 2017 Nov;25(11):2009-2017. [doi: 10.1109/TNSRE.2017.2698005] [Medline: 28459694]
21. Mun KR, Lim SB, Guo Z, Yu H. Biomechanical effects of body weight support with a novel robotic walker for over-ground gait rehabilitation. Med Biol Eng Comput 2017 Feb;55(2):315-326. [doi: 10.1007/s11517-016-1515-8] [Medline: 27193227]
22. Ohnuma T, Lee G, Chong NY. Development of JARoW-II active robotic walker reflecting pelvic movements while walking. Intel Serv Robot 2017 Apr;10(2):95-107. [doi: 10.1007/s11370-016-0212-7]
23. Paulo J, Peixoto P, Nunes UJ. ISR-AIWALKER: robotic walker for intuitive and safe mobility assistance and gait analysis. IEEE Trans Hum Mach Syst 2017 Dec;47(6):1110-1122. [doi: 10.1109/THMS.2017.2759807]

XSL•FO
RenderX

24. Taghvaei S, Jahanandish MH, Kosuge K. Autoregressive-moving-average hidden Markov model for vision-based fall prediction-an application for walker robot. Assist Technol 2017;29(1):19-27. [doi: 10.1080/10400435.2016.1174178] [Medline: 27450279]
25. Zhang H, Ye C. RGB-D camera based walking pattern recognition by support vector machines for a smart rollator. Int J Intell Robot Appl 2017 Feb;1(1):32-42. [doi: 10.1007/s41315-016-0002-6] [Medline: 28409180]
26. Werner C, Moustris GP, Tzafestas CS, Hauer K. User-oriented evaluation of a robotic rollator that provides navigation assistance in frail older adults with and without cognitive impairment. Gerontology 2018;64(3):278-290. [doi: 10.1159/000484663] [Medline: 29183035]
27. Andreetto M, Divan S, Ferrari F, Fontanelli D, Palopoli L, Prattichizzo D. Combining haptic and bang-bang braking actions for passive robotic walker path following. IEEE Trans Haptics 2019;12(4):542-553. [doi: 10.1109/TOH.2019.2912570] [Medline: 31034420]
28. Ferrari F, Divan S, Guerrero C, et al. Human–robot interaction analysis for a smart walker for elderly: the ACANTO interactive guidance system. Int J Soc Robotics 2020 May;12:479-492. [doi: 10.1007/s12369-019-00572-5]
29. Li P, Yamada Y, Wan X, et al. Gait-phase-dependent control using a smart walker for physical training. IEEE Int Conf Rehabil Robot 2019;2019:843-848. [doi: 10.1109/ICORR.2019.8779563] [Medline: 31374735]
30. Moreira R, Alves J, Matias A, Santos C. Smart and assistive walker – ASBGo: rehabilitation robotics: a smart–walker to assist ataxic patients. In: Sequeira JS, editor. Robotics in Healthcare: Field Examples and Challenges: Springer; 2019:37-68. [doi: 10.1007/978-3-030-24230-5]
31. Pérez-Rodríguez R, Moreno-Sánchez PA, Valdés-Aragonés M, et al. FriWalk robotic walker: usability, acceptance and UX evaluation after a pilot study in a real environment. Disabil Rehabil Assist Technol 2020 Aug;15(6):718-727. [doi: 10.1080/17483107.2019.1617795] [Medline: 31172819]
32. Sato W, Tsuchida Y, Li P, Hasegawa T, Yamada Y, Uchiyama Y. Identifying the effects of assistive and resistive guidance on the gait of elderly people using a smart walker. IEEE Int Conf Rehabil Robot 2019 Jun;2019:198-203. [doi: 10.1109/ICORR.2019.8779556] [Medline: 31374630]
33. Scheidegger WM, de Mello RC, Sierra M SD, et al. A novel multimodal cognitive interaction for walker-assisted rehabilitation therapies. Presented at: 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR); Jun 24-28, 2019. [doi: 10.1109/ICORR.2019.8779469]
34. Werner C, Chalvatzaki G, Papageorgiou XS, Tzafestas CS, Bauer JM, Hauer K. Assessing the concurrent validity of a gait analysis system integrated into a smart walker in older adults with gait impairments. Clin Rehabil 2019 Oct;33(10):1682-1687. [doi: 10.1177/0269215519852143] [Medline: 31131630]
35. Alazem H, McCormick A, Nicholls SG, Vilé E, Adler R, Tibi G. Development of a robotic walker for individuals with cerebral palsy. Disabil Rehabil Assist Technol 2020 Aug;15(6):643-651. [doi: 10.1080/17483107.2019.1604827] [Medline: 31012754]
36. Jiménez MF, Mello RC, Bastos T, Frizera A. Assistive locomotion device with haptic feedback for guiding visually impaired people. Med Eng Phys 2020 Jun;80:18-25. [doi: 10.1016/j.medengphy.2020.04.002] [Medline: 32446757]
37. Yeoh WL, Choi J, Loh PY, Saito S, Muraki S. The effect of horizontal forces from a smart walker on gait and perceived exertion. Assist Technol 2022 Mar 4;34(2):204-212. [doi: 10.1080/10400435.2020.1744771] [Medline: 32216620]
38. Werner C, Geravand M, Korondi PZ, Peer A, Bauer JM, Hauer K. Evaluating the sit-to-stand transfer assistance from a smart walker in older adults with motor impairments. Geriatr Gerontol Int 2020 Apr;20(4):312-316. [doi: 10.1111/ggi.13874] [Medline: 32006458]
39. Zhao X, Zhu Z, Liu M, et al. A smart robotic walker with intelligent close-proximity interaction capabilities for elderly mobility safety. Front Neurorobot 2020;14:575889. [doi: 10.3389/fnbot.2020.575889] [Medline: 33192437]
40. Chang YH, Sahoo N, Chen JY, Chuang SY, Lin HW. ROS-based smart walker with fuzzy posture judgement and power assistance. Sensors (Basel) 2021 Mar 29;21(7):2371. [doi: 10.3390/s21072371] [Medline: 33805520]
41. Mostofa N, Feltner C, Fullin K, et al. A smart walker for people with both visual and mobility impairment. Sensors (Basel) 2021 May 17;21(10):3488. [doi: 10.3390/s21103488] [Medline: 34067717]
42. Moustris G, Kardaris N, Tsiami A, et al. The i-Walk lightweight assistive rollator: first evaluation study. Front Robot AI 2021;8:677542. [doi: 10.3389/frobt.2021.677542] [Medline: 34604315]
43. Ingemar Orenius T, Paloniemi S, Hurri H, et al. Piloting a smart rollator: user experiences with technology-related motivation and physical activity. Gerontechnology 2021 Jan;20(2):1-10. [doi: 10.4017/gt.2021.20.2.420.01]
44. Sierra M SD, Múnera M, Provot T, Bourgain M, Cifuentes CA. Evaluation of physical interaction during walker-assisted gait with the AGoRA Walker: strategies based on virtual mechanical stiffness. Sensors (Basel) 2021 May 7;21(9):3242. [doi: 10.3390/s21093242] [Medline: 34067133]
45. Sierra M SD, Garcia A DE, Otálora S, et al. Assessment of a robotic walker in older adults with Parkinson's disease in daily living activities. Front Neurorobot 2021;15:742281. [doi: 10.3389/fnbot.2021.742281] [Medline: 34970132]
46. Zhang C, Guo S, Xi FJ. Man-machine interaction-based motion control of a robotic walker. Technol Health Care 2021;29(4):749-769. [doi: 10.3233/THC-202503] [Medline: 33074202]

XSL•FO
RenderX

47. Fernandez-Carmona M, Ballesteros J, Díaz-Boladeras M, Parra-Llanas X, Urdiales C, Gómez-de-Gabriel JM. Walk-IT: an open-source modular low-cost smart rollator. Sensors (Basel) 2022 Mar 8;22(6):2086. [doi: 10.3390/s22062086] [Medline: 35336255]
48. Palermo M, Lopes JM, André J, Matias AC, Cerqueira J, Santos CP. A multi-camera and multimodal dataset for posture and gait analysis. Sci Data 2022 Oct 6;9(1):603. [doi: 10.1038/s41597-022-01722-7] [Medline: 36202855]
49. Jacobs M, Stramel DM, Shair M, Agrawal SK. Evaluating gait stability and muscle activation in different hand holding conditions using the robotic walker-mTPAD. Sensors (Basel) 2023 Jun 28;23(13):5996. [doi: 10.3390/s23135996] [Medline: 37447845]
50. Schwarz M, Gauda V, Kaiser M, Kuttruff HKJ, Rentzsch J, Thür M. Networked multi-sensor rollator with AI-assisted fall detection. Presented at: 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS); Jun 22-24, 2023. [doi: 10.1109/CBMS58004.2023.00320]
51. Wang W, Gong T, Song Z, Wang Z, Ji J. Simulation study on assist-as-needed control of a rehabilitation robotic walker. Technol Health Care 2023;31(S1):293-302. [doi: 10.3233/THC-236025] [Medline: 37066930]
52. Gong T, Chen D, Wang G, et al. Multimodal fusion and human-robot interaction control of an intelligent robot. Front Bioeng Biotechnol 2024;11:1310247. [doi: 10.3389/fbioe.2023.1310247] [Medline: 38239918]
53. Jiang X, Zong S, Huang J. Lower limb motion measurements during walking by setting rotation angle and period of the motor for driving the chest pad of a rollator. Presented at: 2024 International Symposium on Intelligent Robotics and Systems (ISoIRS); Jun 14-16, 2024. [doi: 10.1109/ISoIRS63136.2024.00037]
54. Machado F, Loureiro M, Bezerra M, Zimerer C, Mello R, Frizera A. Virtual obstacle avoidance strategy: navigating through a complex environment while interacting with virtual and physical elements. Sensors (Basel) 2024 Sep 25;24(19):6212. [doi: 10.3390/s24196212] [Medline: 39409252]
55. Mori Y, Yokoyama S, Yamashita T, Kawamura H, Mori M. Development of assistance level adjustment function for variable load on a forearm-supported robotic walker. Sensors (Basel) 2024 Oct 6;24(19):6456. [doi: 10.3390/s24196456] [Medline: 39409496]
56. Sierra M SD, Harris N, Múnera M, Cifuentes CA. Socially assistive walker for daily living assistance in older adults. Front Robot AI 2024;11:1401663. [doi: 10.3389/frobt.2024.1401663] [Medline: 39220205]
57. Zhang B, Wang Z, Li Z. Mobility Aid Design for the Elderly (MADE): a design thinking approach using a smart walker as a case study. Humanit Soc Sci Commun 2024 Nov 5;11:1469. [doi: 10.1057/s41599-024-04007-z]
58. Bora D. Walkers and rollators market size, share & trends analysis report. : Straits Research; 2024 URL: https://straitsresearch.com/report/walkers-and-rollators-market [accessed 2025-06-03]
59. Maranesi E, Amabili G, Cucchieri G, Bolognini S, Margaritini A, Bevilacqua R. Understanding the acceptance of IoT and social assistive robotics for the healthcare sector: a review of the current user-centred applications for the older users. In: Scataglini S, Imbesi S, Marques G, editors. Internet of Things for Human-Centered Design: Springer; 2022:331-351. [doi: 10.1007/978-981-16-8488-3_16]
60. Strutz N, Meyer-Feil T, Schwesig R, Schulze S, Gissendanner H. Development of a smart walker for clinical settings: a protocol of an exploratory mixed-methods study. BMJ Open 2025 Dec 31;15(12):e105342. [doi: 10.1136/bmjopen-2025-105342] [Medline: 41475804]
61. Roman de Mettelinge T, Cambier D. Understanding the relationship between walking aids and falls in older adults: a prospective cohort study. J Geriatr Phys Ther 2015;38(3):127-132. [doi: 10.1519/JPT.0000000000000031] [Medline: 25594520]
62. Nickerson KA, Diaz K, Muir BC. Circumstances of falls among older adult walker users in long-term care and the associated walker design deficits. Assist Technol 2024 May 3;36(3):248-253. [doi: 10.1080/10400435.2023.2267623] [Medline: 37796449]

Abbreviations

IMU: inertial measurement unit
LiDAR: light detection and ranging
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

XSL•FO
RenderX

Edited by A Scano; submitted 03.Jun.2025; peer-reviewed by A Opheim, G Amabili; revised version received 03.Feb.2026; accepted 20.Feb.2026; published 26.May.2026.

Please cite as:
Strutz N, Brodowski H, Schulze S
Functions and Sensors of Smart Walkers From 2015 to 2024: Scoping Review
JMIR Rehabil Assist Technol 2026;13:e78480
URL: https://rehab.jmir.org/2026/1/e78480
doi:10.2196/78480

© Nicole Strutz, Hanna Brodowski, Stephan Schulze. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Health Care Providers' Perspectives on a Hybrid Outpatient Stroke Telerehabilitation Program: Qualitative Implementation Study

Marina B Wasilewski[1,2,3], PhD; Shaghayegh Mirbaha[1], MSc; Karl Wong[1], MHSc; Gary Siu[1], PT, MScPT, MHM; Siobhan Donaghy[1], MSc; Justin Huynh[2], MScOT; Sarmiya Uruthiralingam[2], MScOT; Michelle L A Nelson[4,5], PhD; Elizabeth Linkewich[6], PhD; Ron Lacombe[1]; Matthew Godleski[1], MD

[1]St. John's Rehab, Sunnybrook Research Institute, 285 Cummer Ave, North York, ON, Canada

[2]Department of Occupational Science and Occupational Therapy, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[3]Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[4]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[5]Bruyère Health Research Institute, Ottawa, ON, Canada

[6]University Hospital, London Health Science Centre, London, ON, Canada

Corresponding Author:
Marina B Wasilewski, PhD
St. John's Rehab, Sunnybrook Research Institute, 285 Cummer Ave, North York, ON, Canada

Abstract

Background: Although patient outcomes are improved by stroke rehabilitation, the suggested amount of therapy is rarely maintained. The COVID-19 pandemic aggravated this situation further due to disruptions in health care. One solution was the rapid and extensive transition to virtual care. A hybrid outpatient stroke telerehabilitation program (HOSTP) was introduced by St John's Rehab—a tertiary rehabilitation hospital in Toronto, Ontario. The HOSTP integrated in-person and virtual care in an effort to alleviate long-standing obstacles that challenge stroke rehabilitation.

Objective: This study explored health care providers' (HCPs) experiences with the HOSTP and their perspectives on its implementation, quality, and impact to determine the modifications needed to optimize its delivery and sustainability.

Methods: A qualitative implementation study was conducted, with semistructured interviews conducted among HCPs involved in the HOSTP. The interview guide was informed by the CFIR (Consolidated Framework for Implementation Research). In total, 14 HCPs were recruited and interviewed from St John's Rehab outpatient program. Interview transcripts were analyzed using a 2-stage analytic approach involving inductive thematic analysis, followed by deductive categorization using CFIR.

Results: Four main themes were identified across CFIR domains: (1) adaptability and flexibility of the hybrid care model (intervention characteristics), (2) alignment with patient needs and resources (outer setting), (3) the impact of organizational resources and infrastructure (inner setting), and (4) variability in provider confidence and perceptions of virtual care (characteristics of individuals). Key determinants were identified as adaptability, patient-related factors, resource availability, and provider beliefs about virtual care.

Conclusions: Our findings suggest that, from HCPs' viewpoints, optimizing virtual care processes and resources may support access and care quality within hybrid outpatient stroke rehabilitation. HCPs viewed maintaining virtual care as important for supporting ongoing access and patient-centered care. Lastly, optimizing the benefits and mitigating the drawbacks of hybrid care can ensure future integration of virtual care into standard outpatient stroke rehabilitation.

(JMIR Rehabil Assist Technol 2026;13:e83081) doi:10.2196/83081

KEYWORDS

stroke telerehabilitation; virtual care; hybrid care; implementation evaluation; outpatient care

Introduction

Stroke prevalence has increased by 85% over the past 30 years and continues to be a leading cause of adult disability and rehabilitation needs worldwide [1]. In Canada, approximately 878,000 people are living with the effects of stroke, with nearly 90,000 individuals experiencing a new stroke each year [2]. While advances in stroke care have reduced mortality, the individual and societal burden of morbidity remains significant [3]. Half of the people with stroke experience moderate to severe impairments (eg, hemiplegia, speaking difficulties, or cognitive issues) [4] that require long-term support with daily living [3].

Participation in stroke rehabilitation has been widely demonstrated to reduce stroke recurrence, improve functional

XSL•FO
RenderX

outcomes, and optimize quality of life [5]. Despite these benefits, the recommended amount of therapy is often not available or achieved [6], with outpatient stroke rehabilitation being particularly underused [7]. This highlights persistent access barriers and the need for innovative models of care that can mitigate these barriers and enhance access to outpatient rehabilitation [7,8]. The advent of the COVID-19 pandemic has exacerbated the general underuse of stroke rehabilitation due to unprecedented disruptions in health care delivery across settings and jurisdictions [1,9]. Recent studies have demonstrated that fewer rehabilitation-eligible people with stroke were discharged to rehabilitation services throughout the pandemic when compared to prepandemic circumstances [10]. One response to these disruptions was the rapid and widespread pivot to virtual care or "telerehabilitation."

Although stroke telerehabilitation capacity has gradually increased across Canada from 71 hospitals offering the service in 2009 to 307 offering it in 2022, the pandemic was a period of unprecedented expansion in telerehabilitation provision [11]. Telerehabilitation refers to the remote provision of clinical rehabilitation services using information and communication technologies either synchronously (eg, phone calls and videoconferences) or asynchronously (eg, email) [12]. Telerehabilitation delivered through virtual-only care or hybrid program models, with hybrid care models combining in-person and virtual modalities to support more flexible service delivery, whereby rehabilitation services are provided through a combination of in-person and virtual modalities [13]. Telerehabilitation enables the provision of services to manage functional and psychosocial recovery of people with stroke, with evidence demonstrating it to be as effective as usual in-person care at improving motor function, performance of activities of daily living, independence, and self-efficacy [14-16]. Despite these positive impacts and the important role telerehabilitation can play in reducing disability and improving quality of life [17], its uptake prior to the COVID-19 pandemic was slow, with little focus placed on telerehabilitation delivery or its translation from research into stroke practice [1].

Throughout the pandemic, telerehabilitation often served as the only link to care for people with stroke when outpatient services were suspended or reduced. Although virtual care was mobilized out of necessity during the pandemic, it has the potential to mitigate long-standing resource and geographic barriers that limit the provision of and access to stroke treatment. Notably, telerehabilitation precludes the need for travel, which reduces costs and expands care to people with stroke living in rural communities as well as those facing transportation challenges [18]. In a 2021 survey of over 3000 Canadians living with stroke and other vascular conditions, half wanted the option of virtual care, and 60% rated virtual care as good as in-person. This trend in user perspectives is supported by recent recommendations and best practices for community-based stroke rehabilitation that emphasize both virtual and hybrid care models should be leveraged to provide comprehensive care across the continuum [18,19].

The pandemic has driven enhanced uptake of telerehabilitation, recognition of its efficacy, and provider willingness to integrate it into future practice [18]. Thus, the time is now to work towards maintaining this momentum and translating it into opportunities for improvement and sustainability in order to continue meeting the health care needs of Canadian people with stroke [18]. Although hybrid models may improve access flexibility [1,18], there is limited understanding of how they are integrated into routine outpatient practice, including sustainability, workflow, and provider-level factors. Despite growing adoption of hybrid models of care, there remains limited real-world evidence on how these programs are implemented in practice and what factors impact their sustainability in outpatient stroke rehabilitation settings. By evaluating the implementation of current stroke telerehabilitation initiatives, real-world evidence can be used to optimize and broaden this care model. This includes exploring the complex process of integrating virtual care into usual stroke rehabilitation [1] and understanding for whom stroke telerehabilitation works best and under what circumstances [18,19].

The overarching goal of this study was to evaluate the implementation of an existing hybrid outpatient stroke telerehabilitation program (HOSTP) for people with stroke. Specifically, this study advances knowledge by providing CFIR (Consolidated Framework for Implementation Research)-mapped insights into barriers and facilitators impacting the implementation and sustainability of hybrid outpatient stroke rehabilitation in real-world settings beyond the pandemic era. We aimed to:

1. Explore the experiences of health care providers (HCPs) with the HOSTP and their perspectives on its implementation, quality, and impact.
2. Determine organization, human, and resource modifications that can be made to optimize future iterations of the HOSTP for ongoing delivery.

Methods

HOSTP Implementation

The HOSTP is offered through the outpatient program of St John's Rehab—a tertiary rehabilitation hospital in Toronto, Ontario. From 2015 - 2020, St John's Rehab admitted an average of 2500 patients per year to its outpatient program, with over 300 being people with stroke. In response to the COVID-19 pandemic, the St John's Rehab outpatient program launched a virtual care initiative to mitigate the service disruptions experienced by community-residing people with stroke. Services that are provided virtually include one-on-one appointments with HCPs, group sessions, and education classes. As the pandemic progressed and some in-person services were restored, this program was adapted into a "hybrid care model." All patients who participate in the HOSTP undergo, at a minimum, a virtual intake assessment. Modes of care (virtual, in-person, or hybrid) were determined based on clinical judgment and patient needs. Virtual care was primarily delivered using Zoom for Healthcare (Zoom Communications, Inc). After that, patients may receive virtual therapy, in-person therapy, or a combination of both, as needed. The HOSTP serves over 100 people with stroke annually, with patients typically participating in therapy twice a week for 8 - 12 weeks. People with stroke receive approximately 16 - 24 treatment sessions per care discipline.

XSL•FO
RenderX

A total of 30 care providers are involved in the HOSTP, including physiotherapists (PTs), occupational therapists (OTs), PT assistants (PTAs), OT assistants (OTAs), speech language pathologists, social workers, and physicians. HCPs work in various capacities, including clinical, managerial, and administrative roles.

The HOSTP represents an informative implementation case due to its scale (services over 100 people with stroke annually), hybrid design, and duration. It is embedded in a well-established outpatient program at a tertiary rehabilitation hospital, which provides a unique opportunity to evaluate the real-world implementation and impact of this type of hybrid telerehabilitation program.

Ethical Considerations

This study was approved by the Research Ethics Board at Sunnybrook Health Sciences Center (REB# 5144). To protect confidentiality in this single-site study, all data were deidentified, and roles and identifying details were generalized to minimize identification risk within this single-site setting. Researchers followed up with interested participants, went through the informed consent forms with them, and obtained consent prior to interviews.

Design

We used a qualitative implementation science approach. Using a qualitative descriptive approach [20], we gained insight into the experiences of HCPs with the HOSTP and explored their perspectives on its implementation, quality, and impact. Qualitative description produces a report that remains near to participants' own words, making it more meaningful to key stakeholders and easily taken up for the type of actionable change we hope to achieve through this project [20,21]. We also conducted this study through the lens of implementation science by using the CFIR to inform the development of the interview guide and guide the deductive categorization and reporting of findings following an initial inductive thematic analysis [22]. The CFIR outlines 5 domains that have been associated with effective implementation of health care innovations: (1) intervention characteristics (eg, relative advantage and disadvantage of intervention, adaptability and trialability, and intervention complexity), (2) inner setting (eg, networks and communication, implementation climate, and tension for change), (3) outer setting (eg, external policies and incentives, and patient needs), (4) characteristics of individuals (eg, self-efficacy, stakeholders' knowledge and beliefs about intervention), and (5) implementation process (eg, engaging appropriate individuals, reflecting, and evaluating). The CFIR guided the topics and questions included in our semistructured interview guide and the categorization of thematic data (see Multimedia Appendix 1 for the complete interview guide). We adhered to recognized reporting criteria by applying the Standards for Reporting Qualitative Research (SRQR; Checklist 1) [23] and the Consolidated Criteria for Reporting Qualitative Research (COREQ; Checklist 2) checklists [24]. These reporting guidelines were used to ensure comprehensive reporting of the qualitative study design.

Participants and Recruitment

HCPs were recruited using purposeful convenience sampling. This sampling technique helped us identify participants with experience directly relevant to the topic of interest while aiming to capture a range of perspectives across disciplines [25]. We purposively sampled for diversity in HCP disciplines (eg, OTs, physiotherapists, OTAs, PTAs, social workers, patient care managers, and service coordinators). To be included in the study, HCPs must have (1) worked full- or part-time at St John's Rehab outpatient department, and (2) delivered care through the HOSTP at least twice in 6 months prior to recruitment. We recruited participants until data saturation was achieved (ie, no new data emerged in subsequent interviews) [26]. An email requesting participation was sent out to HCPs on the mailing list for the outpatient stroke and neurology unit based on the eligibility criteria. The email described the study's purpose, benefits/risks of participating, time commitment, and confidentiality and freedom to withdraw at any time.

Data Collection

One-on-one semistructured interviews were conducted with HCPs using Zoom (Zoom Communications, Inc). The interviews explored participants' experiences with the HOSTP and were informed by the CFIR. Based on the CFIR, the interview guide included open-ended questions regarding experiences with virtual care (likes and dislikes; characteristics of the individuals), what made virtual care easier/harder, what aspects are going well and can be improved (intervention characteristics), and how virtual care promotes/challenges rehabilitation care in the community (implementation process) [22]. Interviews lasted between 30 and 45 minutes and were audio-recorded, professionally transcribed, and checked for accuracy. Transcripts were used for analysis.

Data Analysis

We used a 2-stage analytic approach. First, we conducted an inductive thematic analysis following the steps outlined by Braun and Clarke [27], whereby data were deconstructed into isolated fragments followed by reconstruction into overarching themes that described the higher-level messaging in the data pertaining to HCPs' experiences with the HOSTP. Coding was conducted iteratively, with an initial codebook established and refined through regular team meetings and discussions. Discrepancies in coding were resolved through consensus among the research team. This was followed by a more deductive analysis whereby the experiences of HCPs with the HOSTP were categorized according to the CFIR framework. Three independent researchers (JH, SU, and SM) completed the coding process, and 4 additional researchers (MW, KW, GS, and SD) contributed to the thematic and deductive analyses. Final themes were developed through a collaborative approach among the research team, guided by recurring patterns and relevance to the research objectives.

Rigor

Analytic rigor was enhanced through multiple strategies. Data saturation was assessed iteratively, with recruitment continuing until no new data emerged. The coding team collaboratively determined that saturation had been reached when repetitive

XSL•FO
RenderX

comments were heard from participants and no new information was gathered. Coding was conducted by 3 researchers (SM, JH, and SU) using qualitative analysis software, NVivo (Lumivero), with discrepancies discussed and resolved in team meetings until consensus was achieved. Triangulation between multiple individuals throughout analysis, having regular team meetings, and exercising reflexivity was maintained through ongoing team discussions about the study team's own biases and experiences that may influence data interpretation. Member checking was not conducted due to budget and time restrictions, which are common challenges for undertaking this practice [28]. Studies reassure, however, that member checking is just one of many valuable techniques (eg, triangulation or reflexivity) that can be used to establish credibility and trustworthiness [29]. An audit trail documenting coding decisions, theme establishment, and analytic processes was kept to ensure that the research process was transparent. Finally, appropriate quotes were provided in the final report to support interpretations of the data, thereby enhancing the reliability of study findings.

Results

Overview

We interviewed 14 HCPs, comprised of OTs (n=4), physiotherapists (n=4), social workers (n=1), OTA (n=1), and PTA (n=1), patient care managers who have a professional clinical background in physiotherapy (n=2), and a service coordinator (n=1). Most participants were female (n=12; see Tables 1 and 2 for characteristics and demographic information of HCPs). All 5 CFIR domains were represented in participants' narratives: intervention characteristics, outer setting, characteristics of the individual, inner setting, and process. Below, we outline the details of each domain, with representative quotes included from HCPs, and the specific CFIR constructs italicized in parentheses (Table 3). We also conducted a complementary thematic analysis to identify cross-cutting themes that extend across these domains, offering a more integrated perspective on the overarching insights from the interviews. These themes represented HCPs' views on the (1) adaptability and flexibility of the hybrid care model, (2) alignment with patient needs and resources, (3) the role of organizational resources and infrastructure, and (4) variability in provider confidence and perceptions of virtual care. Discussions pertaining to each theme were fairly consistent across participants, with some variation observed across disciplines.

Table . Characteristics of the health care providers.

Role	Description	Health care providers, n (%)	Years of practice, mean (SD)
Occupational therapist	Uses daily activities to treat the physical, mental, developmental, and emotional conditions that affect a patient's capacity to carry out day-to-day duties	4 (28.6)	15.75 (7.80)
Physiotherapist	Assess and treat individuals whose movement and function are affected by varied factors, such as aging, conditions, injury, or diseases	4 (28.6)	17.25 (9.21)
Social worker	Works with different stakeholders to identify users' needs, provide practical and emotional support, as well as assist users and their caregivers to improve their quality of life	1 (7.14)	>25
Occupational/physical therapy assistants	Similar scope of practice as OTs[a] and PTs[b], but are only licensed to work under the supervision of registered OTs and PTs	2 (14.29)	10.5 (2.12)
Patient care manager/project manager	Leads a clinical team and is in charge of determining the direction of patient care	2 (14.29)	18 (5.66)
Service coordinator	Complete and oversee a variety of professional assignments to assist and provide services such as casework, treatment, outreach development, and referrals to rehabilitation patients	1 (7.14)	15

[a]OT: occupational therapist.

[b]PT: physiotherapist.

XSL•FO
RenderX

Table . Demographic information of health care providers.

Characteristics	Health care providers (n=14)
Age in years, mean (SD)	43.29 (6.09)
Assigned sex at birth	
Male	3 (21.43)
Female	11 (78.57)
Current gender identity	
Male	3 (21.43)
Female	11 (78.57)

XSL•FO
RenderX

Table . Health care providers' (HCPs) insights about the hybrid outpatient telerehabilitation program for people with stroke, categorized by the CFIR (Consolidated Framework for Implementation Research) domain and construct.

CFIR domain and construct/subconstruct	Example
Intervention characteristics	
Adaptability: the degree to which an innovation can be adapted, tailored, refined, or reinvented to meet local needs.	• "We are given the flexibility to choose whether we want to see our clients virtually, or we want to be able to see them in person. That definitely makes a big difference, and that definitely has helped us to grow as a team and to be able to provide the best care to our clients." (HCP09, PCM[a])
Relative advantage: stakeholders' perception of the advantage of implementing the intervention versus an alternative solution.	• "There is a subset of people who sometimes have trouble getting to [REHAB HOSPITAL], either for geographic reasons or transportation reasons or because they have become ill suddenly and now cannot attend, but maybe they are well enough to do things, just not well enough to come in means me having that option to go virtually means that the care can continue in those situations." (HCP07, PT[b]) • "It is good to see how they do it in their home environment. It's also good to see how they walk in their home environment because sometimes patients tell you that […] Again, seeing them in their home environment also adds another layer of depth to everything." (HCP08, PT) • "It is a nice feature, especially when we were not letting family members come in. It's nice that then they can have a family member there to add information or just help them answer some of those questions." (HCP11, PT) • "I think the virtual intake assessments have been a really big win for us both on our side as administrators but also from feedback from our clinicians. It's given us the opportunity to touch base with a patient a lot sooner, but in a way that the patient doesn't actually have to come here for things that can be done virtually. So, conversations about what the goals of therapy are, past medical history, all that stuff can be done." (HCP04, PM[c]) • "With discharge planning, I think that's where you can probably utilize teach-back to really see whether or not someone has been able to meet their goals. You can observe what they're doing, and you can observe them doing the strategies within a different setting, within their home setting. So, I think from a discharge planning perspective, that may be helpful. There may be some pieces with warm handovers." (HCP03, PT) • "It certainly added to the progress of the patient because even if it's virtual and that patient can't come in, you still have some amount of therapy, so it's better than nothing." (HCP13, OT[d])
Outer setting	
Patient needs and resources: the extent to which patient needs are accurately known and prioritized by the organization.	• "Clients who are fairly independent or have family members that can set them up to be independent, who have a good space in their home, good connectivity, and good equipment on their end, because that can pose problems if they don't. So, in some ways, it's more privileged people that benefit most from it." (HCP02, OT) • "As an administrator, when we talk about transitioning someone from our care to the community care, that if we set them up and make them comfortable […]. So, if they're comfortable accessing virtual care, you've eliminated one barrier for patients accessing their own after-rehab programs, and you set them up to be more confident in that way. So, I think it does have a large role in discharge planning." (HCP04, PM) • "I think clients who have experience using video chat, especially Zoom, clients who are fairly independent or have family members that can set them up to be independent, who have a good space in their home, good connectivity, and good equipment on their end, because that can pose problems if they do not. … more privileged people benefit most from it. So those people who are disadvantaged …, they are the people who suffer the most from using virtual care and would benefit more from coming in person, if they can." (HCP02, OT)

XSL•FO
RenderX

CFIR domain and construct/subconstruct	Example
Cosmopolitanism: the degree to which an organization is networked with other external organizations.	• "Virtual care can help to facilitate a discharge into the community a little bit better, so it's not so drastic in terms of you've been coming twice a week for 12 weeks, and then all of a sudden it goes from that to, you're on your own. Maybe there's a transition period in the final weeks of care where people are more in their home environment." (HCP04, PM) • "So, if we could connect with the community OT or the physio community therapist that's seeing the patient for a home safety assessment and give our input as well, that would be ideal. The community OT or the community rehab would coordinate care with the outpatient rehab team." (HCP13, OT)
Inner setting	
Tension for change: the degree to which stakeholders perceive the current situation as intolerable or needing change.	• "COVID has definitely accelerated it. It has made it necessary…Sometimes, these things happen, and they force organizations and programs to adapt. And we had to because we had no other choice. We couldn't shut down our operations. People still were having strokes and amputations and all that stuff. We needed to find a way to be able to see our patients." (HCP04, PM)
Learning climate: a climate in which leaders express their own fallibility and need for team members' assistance and input, and team members feel that they are essential, valued, and knowledgeable partners in the change process.	• "It was definitely a learning curve just as we were going and just giving advice to each other. We would also have weekly meetings or regular meetings both as an outpatient department [and] as a neuro team." (HCP08, PT)
Networks and communications: the nature and quality of webs of social networks, and the nature and quality of formal and informal communications within an organization.	• "Since we started the whole virtual care intake process. We have working groups here that are looking at all of that constantly. We try to touch base every few months through our neuro team meetings as well as through the working group. So, definitely, our technology has improved." (HCP14, OT)
Available resources: the level of resources organizational dedicated for implementation and ongoing operations including physical space and time.	• "We've talked a lot about maybe having an extra external camera so you could have it farther back so that the person on the other end could see your whole body for an exercise demonstration, or at least having a separate screen sometimes so that if you wanted to tilt the screen down to show your feet, you could still see the person's face at the same time." (HCP07, PT) • "Our Zoom rooms are primarily little, rectangular boxes that were meant as offices and not a Zoom space, and there's a limited amount that you can do in this space. Unfortunately, we're a hospital. We're not going to suddenly find new spaces. So you'll have to be creative about using space in a way that's more conducive to therapy." (HCP05, service coordinator) • "We have started a new program or a new part of the program where we do the intake as virtual. And that's just the first session. That's to get us more of the past medical history and social history to fill out what we call our neuro database. But then afterwards, they're usually booked for in-person treatment and assessment." (HCP08, PT) • "[There is a document] that guides patients and clinicians through what to expect about virtual care at [REHAB HOSPITAL], including things like how to access Zoom, how to connect with Zoom, how to troubleshoot if they have issues, whom to contact if they have issues, and things like that." (HCP04, PM)
Compatibility: the degree of tangible fit between the meaning and values attached to the intervention by involved individuals, and how the intervention fits with existing workflows and systems.	• "I would say from an OT/PT perspective, the vast majority of clients are coming in for face-to-face intervention. Within the role of speech, there is still, I think, more virtual happening from a speech/language pathology perspective, but OT and PT have transitioned really to the vast majority coming in for treatment." (HCP01, OT)
Characteristics of Individuals	

XSL•FO
RenderX

CFIR domain and construct/subconstruct	Example
Knowledge and beliefs about the intervention: individuals' attitudes toward and value placed on the innovation, as well as familiarity with facts, truths, and principles related to the innovation.	• "My hope is that we're not really replacing anything that we do in person. Eventually, when things settle down, everything that we do in person will still be in person, but we can leverage virtual care and the infrastructure we already have in place to enhance our program." (HCP04, PM)
Self-efficacy: individual belief in their own capabilities to execute courses of action to achieve implementation goals.	• "When I first started, I didn't feel very prepared for it. I don't think anyone really did. It was like, let's see how this goes so that we could reopen. But the administration and the team were very good at trying to fill in any of those gaps that we didn't feel comfortable with or troubleshooting a lot of the tech issues surrounding Zoom. It didn't take so long to feel less confident. But in the beginning, it was definitely like, we were all just trying to figure it out." (HCP11, PT)
Individual stage of change: characterization of the phase an individual is in, as s/he progresses toward skilled, enthusiastic, and sustained use of the innovation.	• "I had no clue how it would go, and a long time ago, I said to myself, you have to move with the times. You have to deal with change. You don't have a choice. I just dealt with it. You have to be adaptable." (HCP06, social worker)
Process	
Planning: the degree to which a scheme or method of behavior and tasks for implementing an innovation are developed in advance, and the quality of those schemes or methods.	• "When we put it together was at the beginning of COVID, so we really didn't have a clear idea of how we were going to provide virtual care and really what was going to be the scope of virtual care." (HCP11, PT)
Engaging: attracting and involving appropriate individuals in the implementation and use of the innovation through a combined strategy of social marketing, education, role modeling, training, and other similar activities.	• "I think having our clinicians and patients involved both from a planning perspective as well as from an implementation perspective really helped us to a, make sure that there's engagement in the process and then b, to be able to find solutions that were meaningful both from a patient perspective as well as from a staff perspective." (HCP03, PT)
Reflecting and evaluating: quantitative and qualitative feedback about the progress and quality of implementation, accompanied by regular personal and team debriefing about progress and experience.	• "We launched a virtual care initiative around April and May this year. We've had a lot of feedback from staff and patients on what some of the gaps were and what we could do to improve virtual care. And so, we had lots of ideas that related to documentation, consent, how to prepare patients better for virtual care." (HCP04, PM)

[a]PCM: patient care manager.

[b]PT: physiotherapist.

[c]PM: project manager.

[d]OT: occupational therapist.

Domain 1: Intervention Characteristics (Features of the Intervention That Affect Implementation Success)

Participants affirmed that "the hybrid model of care has become normalized in the overall program of outpatients, or at least the neuro program" (HCP07). HCPs spoke of the benefits of having "the flexibility to choose whether we want to see our clients virtually, or we want to be able to see them in person, … has helped us to grow as a team and to be able to provide the best care to our clients" (HCP09; adaptability). HCPs did note, however, that despite the benefits of virtual care, there were challenges and complexities in offering virtual care to the "stroke population as the majority of assessments (ie, balance, gait, and safety) require hands-on contact" (HCP09). Participants were pragmatic in their perspectives on the HOSTP, though, recognizing that "if It is this [virtual care] or nothing, this is definitely better than nothing" (HCP11), as virtual care "certainly adds to the progress of the patient and you still have some amount of therapy" (HCP13; relative advantage).

HCPs felt that the hybrid model of care provided an "opportunity to touch base with patients a lot earlier and quicker" when it came to simple follow-up appointments that "can be done virtually, such as conversations about what the goals of therapy are and past medical history" (HCP04; relative advantage). HCPs also found virtual intake assessments to be especially beneficial in terms of optimizing and accelerating care provision since patient histories were already completed and did not have to be repeated.

Other advantages of virtual care included the ability to increase geographical reach by reducing or eliminating travel to health care facilities, providing an opportunity to "see what [the] patient is doing in their own home environment versus just what they are doing here [outpatient clinic]" (HCP04; relative advantage). Seeing patients in their home environment was especially helpful and provided the unique advantage of being able to witness patients' organic interactions with "family members that could not come [to the in-person appointment]

XSL•FO
RenderX

because they were at work or they could not travel this far" (HCP04), as well as family members' additional assistance with "providing medical history if there is a memory loss, translating if there is a language barrier, and being prepared to provide hands-on support" (HCP05; relative advantage). This extended to discharge planning, where HCPs explained that doing this virtually allowed them "to see how the patients are functioning at home at discharge" (HCP08).

Domain 2: Outer Setting (External Impacts on the Intervention Implementation)

HCP comments reflected that the "needs and resources of the patient" was the primary "outer setting" construct that drove their valuation of virtual care. HCPs recognized that the use of virtual modalities for stroke rehabilitation was highly "client-dependent" (HCP02) and that several factors may impact patients' participation in and benefit from virtual care. Such factors included level of functioning (higher-functioning or lower-functioning), dependency of patients on their caregivers (highly dependent or relatively independent), comfort levels with using technology, and access to technology (needs and resources of the patients).

One area where HCPs felt that virtual care was especially well-aligned with patient needs was during discharge planning. Virtual care was viewed as a way to facilitate reintegration by "providing feedback to the patient in the home environment, as that can be a little bit more of a smooth transition versus a drastic stop to care" (HCP04; needs and resources of the patients). HCPs felt that virtual care allowed them to better "[connect] online and [give] our input to a community OT or the community therapist that's seeing the patient [who] was referred for a home safety assessment at the same time" (HCP13; cosmopolitanism). These cross-setting communications were viewed as optimizing care integration for patients by creating a "warmer handover to the next episode of care" (HCP03; cosmopolitanism). Thus, the "patient feels comfortable, as well as the person you are sending the patient to now knows that the information, and so they can do the best intervention they can do" (HCP03).

Domain 3: Inner Setting (Features of the Implementing Institution)

At the onset of the COVID-19 pandemic, there was a strong tension for change given that infection prevention and control regulations resulted in "the entire department [shutting] down within a couple of weeks" (HCP04). HCPs explained that "COVID definitely made [us] think outside the box and still be able to provide the services we need" (HCP14; tension for change). HCPs felt that the pandemic accelerated the uptake of virtual care, along with the realization that different care models and approaches are needed to ensure continuity of care.

As the COVID-19 pandemic progressed, HCPs encountered a learning curve with regard to technology, administration, and care provision, which left them and their department "constantly looking into growing" (HCP09; learning climate). Nevertheless, they felt that team dynamics and workflow were well-established and optimally positioned to deliver a hybrid model of care for the population with stroke. Participants explained that "it is a good communication that exists between the therapist and the management and the client" (HCP09) that leads to a continuum of teamwork towards "refining the model and just looking at how we can use it to our advantage" (HCP10; network and communication). One of the key advantages identified was the fact that virtual care streamlined appointment scheduling since "[HCPs] do not do the setup for arranging the meetings. It is the office that does all of the scheduling" (HCP06) which, in turn, freed up HCPs and allowed them to "focus strictly on providing care for the patient, [since] everything is set up for us" (HCP06; network and communication). Another perceived "win" with regard to virtual care was the virtual intake assessment conducted by the outpatient team. HCPs viewed this as an "opportunity to touch base with a patient a lot sooner...virtually. So, conversations about therapy goals, past medical history, all that can be done.... It is a lot of work for many of our stroke patients to come to the hospital, so we have prevented them from having to come... and it has actually helped direct our assessment a little bit better" (HCP04).

Conversely, HCPs felt like the organization could have been more prepared in terms of having appropriate physical space and equipment set up for virtual care (eg, having cameras external to the ones built into laptops; available resources). One HCP explained that "the physical space where we do our Zoom sessions is a little bit small for some of the physio-type, exercises I might want to demonstrate. It is a combination of the size of the room and [the fact] that right now, we are typically just using a laptop with a little camera on the laptop" (HCP07). In some cases, however, these issues were addressed by the "comprehensive patient-readiness and clinician-readiness documents" (HCP05) designed by the outpatient team (available resources). Further, a few participants suggested the idea of "looking at how we can reconfigure the office rooms to allow therapists to have a little bit more of a conductive environment" (HCP04), as well as "having an external camera or a separate screen so that the person on the other end could see your [provider] whole body for an exercise demonstration" (HCP07).

HCPs' profession strongly influenced their perception of the virtual components of the hybrid model, with social workers feeling especially positive about virtual care, believing that "the techniques are the same, the counseling is the counseling. I just had to depend more certainly over the phone on the nonverbal [cues, such as] tone of the voice, expressions, and what they say" (HCP06). On the contrary, occupational therapy and physiotherapy practices entail more hands-on care, which led these HCPs to view virtual care as less practical, particularly for neurorehabilitation, compared to in-person care (compatibility). As one HCP reflected:

> *I would say from an OT/PT perspective, the vast majority of clients are coming in for face-to-face intervention...OT and PT have transitioned really to the vast majority coming in [in-person] for treatment.* [HCP01, OT]

Domain 4: Characteristics of Individuals (Characteristics of the Implementing Individuals)

Although HCPs recognized the importance of hybrid models of care, particularly during the COVID-19 pandemic, for

XSL•FO
RenderX

outpatient rehabilitation of people with stroke, they believed virtual care was "an adjunct" (HCP05) to in-person treatment and did not view it as "the only tool" (HCP05; knowledge and beliefs about the intervention). HCPs varied in how confident they were in their capabilities to use the hybrid model of care (self-efficacy). Some were "very tech savvy" (HCP04) and felt more comfortable using virtual modalities, while others felt less comfortable and "did not have any experiences" with novel technologies (HCP04). During the early stages of hybrid care implementation, HCPs reported that most providers and staff "did not have any [or had insignificant] experiences [with] virtual care" and were "trying to adapt to virtual care" by receiving "step-by-step instructions from [other colleagues] learning how to apply technology" (HCP01; individual stage of change). However, over time, the majority of HCPs became more skilled in using technology to provide care (eg, Zoom for Healthcare software).

Domain 5: Process (Phases of Implementation)

Providers described how the sudden onset of the COVID-19 pandemic affected the outpatient service provisions, resulting in very limited time for detailed planning (planning). As one participant explained, "The entire department shut down within a couple of weeks, and we had to implement an infrastructure for virtual care within those couple of weeks" (HCP04). Although there was a lack of formal evaluation initially when the program began, as the pandemic progressed, both HCPs and patients were involved to a greater extent in the "planning and implementation process" (HCP03; engaging), which ultimately ensured that the outpatient department was "able to find solutions that were meaningful both from a patient perspective and staff perspective" (HCP03). Team debriefing about the progress of the hybrid care model further improved the implementation, as "we have worked on it as a department," "ran focus groups" (HCP07), and "did some self-reflection to see how we could better incorporate it into the way in which we do what we do" (HCP03; reflecting and evaluating).

Discussion

Overview

Our study aimed to explore the implementation of a HOSTP launched at St John's Rehab. According to the CFIR, for an innovation to be implemented successfully, accounts must be taken of the innovation's core components and adaptable periphery, the structural and cultural context through which the implementation occurs, the social context within which the organization resides, and an active change process. Our findings highlighted that all 5 CFIR domains were discussed by HCPs as impacting the implementation of the HOSTP. This study extends the telerehabilitation implementation literature by examining how HCP perceptions of and experiences with the HOSTP evolved from its initial pandemic deployment to integration into routine outpatient care. Although the study reflects HCPs' perspectives within a single HOSTP (rather than objective measures of patient outcomes or system performance), our findings begin to elucidate how perceived flexibility, clinical fit across disciplines, and organizational learning processes influence provider buy-in and sustainability over time.

Several stay-at-home mandates and lockdown orders during the COVID-19 pandemic fast-tracked the adoption of virtual care in Canada [30]. In response to the pandemic, provinces looked enthusiastically to virtual care as a means of safely providing health care services by mitigating the risk of in-person viral transmission. While participants in our study viewed virtual care as being "better than nothing" at the onset of the pandemic, our findings strongly emphasize that it would be difficult to retain buy-in from HCPs for a virtual-only or "virtual first" approach during the postpandemic era. Findings from our study illustrated that HCPs do not view virtual-only models of care positively, especially when the virtual modality does not align with a hands-on approach needed for their clinical role. Other studies have highlighted that this is particularly pronounced among physiotherapists due to the nature of their practice [31], which was echoed by physiotherapists in our study as their therapeutic domains require hands-on physical activities and examinations with their patients. Adding to the literature, our findings suggest that OTs also view certain therapeutic activities within their scope of practice (eg, gait assessment) as requiring in-person care. Conversely, social workers found virtual care to align with their usual clinical tasks.

Across HCP accounts, the hybrid model was viewed as supporting both patient- and family-centered care by increasing flexibility, continuity, and caregiver involvement, while preserving the option for necessary in-person, hands-on rehabilitation. Consistent with Picker principles of person-centered care [32], participants described how virtual components enabled faster access to care, smoother transitions between outpatient and community services, and clearer communication by allowing HCPs to observe patients' home environments and tailor recommendations accordingly [32-34]. HCPs also perceived that hybrid delivery supported greater involvement and support for family caregivers by enabling them to attend appointments remotely, contribute contextual information, and participate more actively in discharge planning when in-person attendance was challenging [18,35]. Together, these perceived benefits were viewed as supporting a more integrated and responsive model of care that balanced patient preferences, caregiver engagement, and continuity of care, while preserving the need for in-person, hands-on rehabilitation when clinically indicated.

Recommendations for Future Optimization of Hybrid Models of Stroke Care

Having the flexibility to choose the care provision approach, either through virtual streams or in-person, showcased the potential of the hybrid model of care to be adaptable to the patients' and settings' needs (domain 1: intervention characteristics). This echoes findings from previous studies indicating that for a novel care program to be adaptable, modifications of an existing program must fit the organization's capacity and needs without compromising its integrity and being flexible enough to meet the patient's needs and settings [36]. The hybrid model of care was positively viewed as a facilitator for community reintegration, aligning well with patient needs during discharge planning. Evidence has indicated that virtual self-management approaches enhance community reintegration, such as socializing and engaging in daily activities that enhance

XSL•FO
RenderX

well-being in community-dwelling adults with stroke [37]. Consequently, it is recommended that organizations capitalize on hybrid models of care that enable providers to maintain connections and support people with stroke virtually and in-person once they return to the community (domain 2: outer setting).

Learning environments that involve HCPs and patients need to be fostered as part of the change process to improve knowledge and build skills among virtual care users. Previous studies have demonstrated that in an effective learning environment, the workplace culture is caring and compassionate, leading to the provision of high-quality care and experience for staff, patients, and caregivers [38]. As a result, cultivating supportive learning environments may be important to ensuring that supportive relationships and transformative policies and processes that support innovation and learning are in place (domain 3: inner setting). HCPs in our study reported that virtual care enabled them to work better as a team and to streamline scheduling and communication, which created more time for patient care. This is a key benefit of the virtual aspect of the hybrid care model in as far as it mobilizes Canadian Stroke Best Practice recommendations that virtual hybrid outpatient stroke rehabilitation services should offer case coordination approaches that include regular team communication to discuss assessments, management, goals, and feedback (domain 3: inner setting) [18]. HCPs discussed how upper management involved them in the implementation process and accepted their feedback to iteratively enhance the HOSTP program, which made them feel heard and validated. As a result, ongoing evaluations may help capture the various stages of virtual care implementation and the particular considerations related to each stage (domain 5: process).

Our findings point to several challenges to implementation that may need to be addressed in the future to ensure that hybrid models of outpatient stroke care continue to provide high-quality, patient-centered care. First, choices and alternatives regarding virtual care should be provided to HCPs, and extreme caution should be exercised to avoid "technological determinism," in which virtual care is considered the best and only mode of care provision [39]. In our study, this was apparent as many HCPs, particularly physiotherapists and OTs, expressed how the relative advantages of virtual care could be diminished for specific professions that require more hands-on assessments and interventions (domain 1: intervention characteristics).

Secondly, HCPs may benefit from receiving training and educational resources that help them both attain and maintain critical competencies needed to provide safe and appropriate virtual care. Despite the rapid adoption of digital tools and technologies in response to the COVID-19 pandemic, most HCPs, patients, families, and caregivers did not receive adequate training on how to use digital tools in health care systems (domain 3: inner setting). In our study, this meant that HCPs had varying levels of self-efficacy in using technology, subsequently leading to very individualized stages of change when it came to adapting to and adopting virtual care (domain 4: individual characteristics).

Thirdly, adequate infrastructure and resources may be important for supporting successful virtual care implementation [40]. Consistent with growing evidence [40,41], our findings suggest that better resourcing and infrastructure in terms of physical space and technological equipment can encourage HCP uptake of this modality and potentially improve the quality of virtual care sessions (domain 3: inner setting). Overall, our findings are consistent with prior research illustrating the acceptance, feasibility, and potential benefits of stroke telerehabilitation [42]. They also align with existing qualitative and implementation studies that identify user perceptions, resource availability, and workflow considerations as key factors influencing implementation [42,43]. In this way, our findings confirm prior evidence regarding the importance of these determinants. However, much of the existing research has emphasized virtual-only implementation during the pandemic, with less attention to the integration and sustainability of hybrid models. Our findings underscore that hybrid models afford flexibility but also ensure that clinical judgment, patient needs, and discipline-specific requirements drive the nature of care delivery.

Strength and Limitations

A notable strength of this study was our use of a robust theoretical framework to comprehensively evaluate a novel HOSTP that was rapidly deployed during the COVID-19 pandemic [44]. Our combined use of a qualitative approach and implementation science evaluation allowed us to provide rich and in-depth descriptions of HCPs' experience with the HOSTP while also identifying key implementation successes and areas for improvement. One limitation of our study is that the HOSTP was provided by a rehabilitation hospital in an urban setting, thus potentially limiting the transferability of our findings to programs offered in rural areas (where, eg, issues may arise around technology access, internet reliability, and geographical distance to in-person care). Additionally, this study was conducted at a single site with a relatively small sample size, which may, in turn, limit the transferability of findings to other settings. An additional limitation related to our recruitment approach where participants were recruited using purposeful convenience sampling from a single outpatient unit mailing list. This approach may have introduced selection bias, as those who chose to participate may have been more involved, more available, or held more favorable perspectives toward the program. Therefore, the findings may not fully represent the perspectives of all HCPs, particularly those who may have been less involved or held more critical perspectives. Another limitation relates to the overrepresentation of physiotherapists and OTs in our study. The inclusion of additional health professionals (eg, social workers, speech-language pathologists) could have enriched topics pertaining to the alignment of virtual care with these practice disciplines. Furthermore, this study focuses on HCPs' perspectives only. Patient and family caregiver perspectives were also collected in a related study and will be reported separately, which will further inform understanding of the implementation and impact of hybrid stroke rehabilitation. Thus, while our study has begun to elucidate issues related to this topic, there may be additional nuances that were not captured.

XSL•FO
RenderX

Conclusion

This study adds insights into the implementation considerations of the HOSTP program, one of the few outpatient hybrid programs established for the population with stroke using the CFIR framework. The adaptability and relative advantages of the HOSTP, as well as its use and value, were predominantly driven by how well it met patient needs, alongside HCPs' beliefs, self-efficacy with virtual care, and their involvement in program planning and implementation. These findings suggest that therapy type and specific session goals must guide the choice of delivery medium. Not all rehabilitation translates safely to a video interface, and clinicians need to exercise judgment in that triage. Caution should be exercised to ensure that future virtual care initiatives work to expand access, rather than create new disparities. This includes leveraging hybrid models that afford the flexibility and patient-centeredness of virtual modalities, while ensuring it does not replace in-person care for those that need it.

Funding

This study was funded by an AFP Innovation Fund Grant. The study was funded for a 2-year period, starting in 2023. The study was funded for CAD $104,999 (US $76,029.25).

Data Availability

The data that support the findings of this study are available on reasonable request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Authors' Contributions

Conceptualization: MW, MN, EL, SD, KW, GS, MG, RL
Funding acquisition: MW, MG
Investigation: MW, MG
Methodology: MW, MN, EL, SD, KW, GS, MG
Project administration: MW, MG
Resources: MW, MG
Software: MW, MG
Supervision: MW, MG
Validation: MW, MG
Data curation: SM, JH, SU
Formal analysis: SM, JH, SU
Interpretation of the data: SM, JH, SU
Writing—original draft: MW, SM, MG
Writing—review & editing: MW, SM, JH, SU, MN, EL, SD, KW, GS, MG, RL
Patient partnership: RL

Conflicts of Interest

None declared.

Multimedia Appendix 1
Health care providers' interview guide.
[PDF File, 371 KB - rehab_v13i1e83081_app1.pdf]

Checklist 1
SRQR checklist.
[PDF File, 559 KB - rehab_v13i1e83081_app2.pdf]

Checklist 2
Completed COREQ 32-item checklist.
[PDF File, 328 KB - rehab_v13i1e83081_app3.pdf]

References

1. Stephenson A, Howes S, Murphy PJ, et al. Factors influencing the delivery of telerehabilitation for stroke: a systematic review. PLoS ONE 2022;17(5):e0265828. [doi: 10.1371/journal.pone.0265828] [Medline: 35544471]
2. Fact sheet: stroke statistics. Government of Canada, Ontario Stroke Network. 2019. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/stroke-canada-fact-sheet.html [accessed 2026-05-29]

XSL•FO
RenderX

3. Garnett A, Ploeg J, Markle-Reid M, Strachan PH. Formal health and social services that directly and indirectly benefit stroke caregivers: a scoping review of access and use. Can J Nurs Res 2022;54(2):211-233. [doi: 10.1177/08445621211019261] [Medline: 35130749]
4. Kuys SS, Bew PG, Lynch MR, Brauer SG. Activity limitations experienced by people with stroke who receive inpatient rehabilitation: differences between 2001, 2005, and 2011. Arch Phys Med Rehabil 2014;95(4):741-746. [doi: 10.1016/j.apmr.2013.12.008] [Medline: 24365092]
5. Prvu Bettger JA, Kaltenbach L, Reeves MJ, et al. Assessing stroke patients for rehabilitation during the acute hospitalization: findings from the get with the guidelines-stroke program. Arch Phys Med Rehabil 2013;94(1):38-45. [doi: 10.1016/j.apmr.2012.06.029] [Medline: 22858797]
6. Sentinel stroke national audit programme clinical results 23/24. data.gov.uk. 2015. URL: https://www.data.gov.uk/dataset/470bb015-8602-4bf9-8749-4a57839c7187/sentinel-stroke-national-audit-programme-clinical-results-23-24 [accessed 2026-05-29]
7. Teasell R, Meyer MJ, McClure A, et al. Stroke rehabilitation: an international perspective. Top Stroke Rehabil 2009;16(1):44-56. [doi: 10.1310/tsr1601-44] [Medline: 19443347]
8. Ayala C, Fang J, Luncheon C, et al. Use of outpatient rehabilitation among adult stroke survivors - 20 states and the district of Columbia, 2013, and four states, 2015. MMWR Morb Mortal Wkly Rep 2018;67(20):575-578. [doi: 10.15585/mmwr.mm6720a2] [Medline: 29795076]
9. Sheng S, Wang X, Gil Tommee C, et al. Continued underutilization of stroke care during the COVID-19 pandemic. Brain Behav Immun Health 2021;15:100274. [doi: 10.1016/j.bbih.2021.100274] [Medline: 34589777]
10. Thau L, Siegal T, Heslin ME, et al. Decline in rehab transfers among rehab-eligible stroke patients during the COVID-19 pandemic. J Stroke Cerebrovasc Dis 2021;30(8):105857. [doi: 10.1016/j.jstrokecerebrovasdis.2021.105857] [Medline: 34022581]
11. Stroke month 2022 key messages and backgrounder for media. Heart and Stroke Foundation of Canada. 2022. URL: https://www.heartandstroke.ca/-/media/pdf-files/what-we-do/news/stroke-month-2022-backgrounder-for-media.pdf?rev=a88365a3f5a94dae9ebc8e8f4d0fc310 [accessed 2026-05-29]
12. Brennan DM, Mawson S, Brownsell S. Telerehabilitation: enabling the remote delivery of healthcare, rehabilitation, and self management. Stud Health Technol Inform 2009;145:231-248. [Medline: 19592797]
13. Ganeshan S, Jackson H, Grandis DJ, et al. Clinical outcomes and qualitative perceptions of in-person, hybrid, and virtual cardiac rehabilitation. J Cardiopulm Rehabil Prev 2022;42(5):338-346. [doi: 10.1097/HCR.0000000000000688] [Medline: 35420563]
14. Appleby E, Gill ST, Hayes LK, Walker TL, Walsh M, Kumar S. Effectiveness of telerehabilitation in the management of adults with stroke: a systematic review. PLoS ONE 2019;14(11):e0225150. [doi: 10.1371/journal.pone.0225150] [Medline: 31714924]
15. Laver KE, Adey-Wakeling Z, Crotty M, Lannin NA, George S, Sherrington C. Telerehabilitation services for stroke. Cochrane Database Syst Rev 2020;1(1):CD010255. [doi: 10.1002/14651858.CD010255.pub3] [Medline: 32002991]
16. Chen J, Jin W, Zhang XX, Xu W, Liu XN, Ren CC. Telerehabilitation approaches for stroke patients: systematic review and meta-analysis of randomized controlled trials. J Stroke Cerebrovasc Dis 2015;24(12):2660-2668. [doi: 10.1016/j.jstrokecerebrovasdis.2015.09.014] [Medline: 26483155]
17. Duncan PW, Bernhardt J. Telerehabilitation: has its time come? Stroke 2021;52(8):2694-2696. [doi: 10.1161/STROKEAHA.121.033289] [Medline: 34192896]
18. Salbach NM, Mountain A, Lindsay MP, et al. Canadian stroke best practice recommendations: virtual stroke rehabilitation interim consensus statement 2022. Am J Phys Med Rehabil 2022;101(11):1076-1082. [doi: 10.1097/PHM.0000000000002062] [Medline: 35767008]
19. Schwamm LH, Audebert HJ, Amarenco P, et al. Recommendations for the implementation of telemedicine within stroke systems of care: a policy statement from the American Heart Association. Stroke 2009;40(7):2635-2660. [doi: 10.1161/STROKEAHA.109.192361] [Medline: 19423851]
20. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
21. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
22. Keith RE, Crosson JC, O'Malley AS, Cromp D, Taylor EF. Using the Consolidated Framework for Implementation Research (CFIR) to produce actionable findings: a rapid-cycle evaluation approach to improving implementation. Implement Sci 2017;12(1):15. [doi: 10.1186/s13012-017-0550-7] [Medline: 28187747]
23. Standard for reporting qualitative research: a synthesis of recommendations. Equator Network. URL: https://www.equator-network.org/reporting-guidelines/srqr/ [accessed 2029-05-29]
24. Consolidated criteria for reporting qualitative research (COREQ): 32-item checklist. Equator Network. URL: https://www.equator-network.org/reporting-guidelines/coreq/ [accessed 2026-05-29]
25. Etikan I. Comparison of convenience sampling and purposive sampling. AJTAS 2016;5(1):1. [doi: 10.11648/j.ajtas.20160501.11]

XSL•FO
RenderX

26. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Kullman SM, Chudyk AM. Participatory member checking: a novel approach for engaging participants in co-creating qualitative findings. Int J Qual Methods 2025;24:16094069251321211. [doi: 10.1177/16094069251321211]
29. Motulsky SL. Is member checking the gold standard of quality in qualitative research? Qualitative Psychology 2021;8(3):389-406. [doi: 10.1037/qup0000215]
30. Patterson PB, Roddick J, Pollack CA, Dutton DJ. Virtual care and the influence of a pandemic: necessary policy shifts to drive digital innovation in healthcare. Healthc Manage Forum 2022;35(5):272-278. [doi: 10.1177/08404704221110084] [Medline: 35775144]
31. Braund H, Dalgarno N, Ritsma B, Appireddy R. Exploring virtual care clinical experience from non-physician healthcare providers (VCAPE). SSM Qual Res Health 2023;3:100289. [doi: 10.1016/j.ssmqr.2023.100289] [Medline: 37283887]
32. The Picker principles of person centred care. Picker. URL: https://picker.org/who-we-are/the-picker-principles-of-person-centred-care/ [accessed 2026-05-29]
33. Zhang X, McJoynt T, Furst JW, Myers JF. Establishing a patient-centered virtual care model across the continuum of care. J Prim Care Community Health 2022;13:21501319221088823. [doi: 10.1177/21501319221088823] [Medline: 35323069]
34. Latulippe K, Provencher V, Boivin K, et al. Using an electronic tablet to assess patients' home environment by videoconferencing prior to hospital discharge: protocol for a mixed-methods feasibility and comparative study. JMIR Res Protoc 2019;8(1):e11674. [doi: 10.2196/11674] [Medline: 31344677]
35. Appireddy R, Khan S, Leaver C, et al. Home virtual visits for outpatient follow-up stroke care: cross-sectional study. J Med Internet Res 2019;21(10):e13734. [doi: 10.2196/13734] [Medline: 31593536]
36. Zullig LL, Bosworth HB. Selecting, adapting, and sustaining programs in health care systems. J Multidiscip Healthc 2015;8:199-203. [doi: 10.2147/JMDH.S80037] [Medline: 25931825]
37. Huijbregts MPJ, McEwen S, Taylor D. Exploring the feasibility and efficacy of a telehealth stroke self-management programme: a pilot study. Physiother Can 2009;61(4):210-220. [doi: 10.3138/physio.61.4.210] [Medline: 20808482]
38. Mash B, Edwards J. Creating a learning environment in your practice or facility. S Afr Fam Pract (2004) 2020;62(1):e1-e5. [doi: 10.4102/safp.v62i1.5166] [Medline: 32787390]
39. Campbell EM, Sittig DF, Guappone KP, Dykstra RH, Ash JS. Overdependence on technology: an unintended adverse consequence of computerized provider order entry. AMIA Annu Symp Proc 2007;2007:94-98. [Medline: 18693805]
40. Romanchych E, Desai R, Bartha C, Carson N, Korenblum M, Monga S. Healthcare providers' perceptions of virtual-care with children's mental health in a pandemic: a hospital and community perspective. Early Interv Psychiatry 2022;16(4):433-443. [doi: 10.1111/eip.13196] [Medline: 34309196]
41. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016;11(1):146. [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
42. Chen Y, Chen Y, Zheng K, et al. A qualitative study on user acceptance of a home-based stroke telerehabilitation system. Top Stroke Rehabil 2020;27(2):81-92. [doi: 10.1080/10749357.2019.1683792] [Medline: 31682789]
43. Wiley E, Moncion K, Noguchi KS, et al. Facilitators and barriers to participation in exercise-based telerehabilitation in individuals with stroke: a scoping review of quantitative and qualitative studies. Telemed J E Health 2025;32(1):28-48. [doi: 10.1177/15305627251378059] [Medline: 41058538]
44. Ruskin D, Borsatto J, Szczech K, et al. "Working together": perspectives of healthcare professionals in providing virtual care to youth with chronic pain during the COVID-19 pandemic. Int J Environ Res Public Health 2023;20(6):4757. [doi: 10.3390/ijerph20064757] [Medline: 36981665]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
COREQ: Consolidated Criteria for Reporting Qualitative Research
HCP: health care provider
HOSTP: hybrid outpatient stroke telerehabilitation program
OT: occupational therapist
OTA: occupational therapist assistant
PT: physiotherapist
PTA: physiotherapist assistant
SRQR: Standards for Reporting Qualitative Research

XSL•FO
RenderX

Edited by J Sarvestan; submitted 27.Aug.2025; peer-reviewed by PC Wang, TA Wani; revised version received 24.Apr.2026; accepted 27.Apr.2026; published 15.Jun.2026.

Please cite as:
Wasilewski MB, Mirbaha S, Wong K, Siu G, Donaghy S, Huynh J, Uruthiralingam S, Nelson MLA, Linkewich E, Lacombe R, Godleski M
Health Care Providers' Perspectives on a Hybrid Outpatient Stroke Telerehabilitation Program: Qualitative Implementation Study
JMIR Rehabil Assist Technol 2026;13:e83081
URL: https://rehab.jmir.org/2026/1/e83081
doi:10.2196/83081

© Marina Wasilewski, Shaghayegh Mirbaha, Karl Wong, Gary Siu, Siobhan Donaghy, Justin Huynh, Sarmiya Uruthiralingam, Michelle L A Nelson, Elizabeth Linkewich, Ron Lacombe, Matthew Godleski. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 15.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical Outcomes of Individualized Electrostimulation Using a Wearable Electro Suit and Qualitative Feedback From a Mixed Cohort of Survivors of Stroke and Spinal Cord Injury With Spasticity: Case Series

Jia Min Yen[1,2], MMed; Nur Shafawati Kamsani[3]; Hwa Sen Lai[1]; Ning Tang[1]; Effie Chew[1,2]

[1]Division of Rehabilitation Medicine, Department of Medicine, National University Hospital, 5 Lower Kent Ridge Road, Singapore, Singapore

[2]Rehabilitation Medicine, Department of Medicine, Alexandra Hospital, Singapore, Singapore

[3]Department of Rehabilitation, Alexandra Hospital, Singapore, Singapore

Corresponding Author:
Jia Min Yen, MMed
Division of Rehabilitation Medicine, Department of Medicine, National University Hospital, 5 Lower Kent Ridge Road, Singapore, Singapore

Abstract

Background: Various forms of electrical stimulation have been integrated into the multimodal management of spasticity. However, high-frequency electrical stimulation can potentially induce muscle fatigue. The Exopulse Mollii Suit (EMS) is a multichannel full-body garment that delivers low-frequency (20 Hz), low-amplitude (20 V), subthreshold sensory stimulation aimed at reducing spasticity.

Objective: Primarily, we examined the effects of a single session of the EMS on spasticity in 7 participants with chronic stroke or cervical spinal cord injury (SCI), specifically those with upper or lower limb spasticity impacting function and gait who were able to walk with minimal or no assistance (Functional Ambulatory Category scores of 2 - 5). We assessed the impact on gait and ambulatory function, as well as user perceptions of usability and acceptability.

Methods: Participants wore the EMS for 60 minutes, consisting of 30 minutes of standardized goal-directed activities performed in two 15-minute blocks, interspersed with 15-minute rest breaks. Measurements included the Modified Tardieu Scale with surface electromyography for spasticity and functional mobility tests (Functional Ambulatory Category, 10-meter walk test, 5 times sit-to-stand test, and step test). Spatiotemporal gait parameters were quantified using a markerless vision-based motion capture system using the OpenPose BODY25 pose estimation model.

Results: On the basis of the Modified Tardieu Scale and surface electromyography signals, improvements in spasticity were only observed in 2 participants. However, 4 participants demonstrated faster walking speeds. Improvements in the 5 times sit-to-stand test and step test were noted in 3 and 4 participants, respectively. Spatiotemporal gait parameters revealed improvements in gait symmetry in 6 participants. Qualitative feedback based on the Assistive Technology Usability Questionnaire for People With Neurological Diseases (NATU Quest) returned positive results in 3 participants. Overall outcomes, defined as meeting the individualized goals of each participant, were positive in 4 participants.

Conclusions: This case series provides preliminary evidence that a single session with the EMS may offer benefits for functional mobility and gait quality for individuals with spasticity resulting from stroke or SCI. To our knowledge, this is the first report examining the effects of the EMS in participants with SCI and the first to include spatiotemporal gait parameters associated with its use. However, the small sample size, variable outcomes, and lack of a control group necessitate caution in interpreting these findings and preclude definitive conclusions regarding the efficacy of the EMS. Larger, controlled trials with repeated sessions of EMS use are required to establish the effectiveness and optimal application of the EMS for spasticity management.

(JMIR Rehabil Assist Technol 2026;13:e81522) doi:10.2196/81522

KEYWORDS

electrostimulation; Mollii Suit; assistive technology; spasticity; gait

Introduction

Spasticity is a common and disabling complication of central neurological diseases, particularly following stroke [1] and spinal cord injury (SCI) [2]. It is characterized by a velocity-dependent increase in tonic stretch reflexes [3] and involuntary muscle activation [4], which disrupt motor control. When synchronized agonist and antagonist muscle activity is

XSL•FO
RenderX

disturbed, fluid movement is no longer possible. Consequently, the ability to perform activities of daily living, including transfers and ambulation, is significantly impaired.

Given the profound impact of spasticity on function and quality of life, effective management is crucial during rehabilitation to restore mobility—a top priority following neurological injury [5]. Spasticity treatments have largely focused on oral pharmacological agents and interventions, such as botulinum toxin injections, alcohol neurolysis, or intrathecal baclofen, as nonpharmacological interventions have limited efficacy in isolation [6-8]. While pharmacological treatments may pose a risk of systemic side effects, such as drowsiness, risks associated with physical modalities, such as electrical stimulation are generally minimal when used appropriately. Hence, various forms of electrical stimulation have been integrated into the multimodal approach to the management of spasticity [9]. Among these, neuromuscular electrical stimulation (NMES) using 0 to 100 Hz biphasic symmetrical waveform electrical stimulation [10] and transcutaneous electrical nerve stimulation (TENS) using high frequencies of more than 100 Hz [11] have been shown to be effective in reducing spasticity. Their efficacy in reducing spasticity is further enhanced when combined with physical therapy interventions, such as range of motion and functional task training [12]. The proposed mechanisms include increased presynaptic inhibition of hyperactive stretch reflexes, a reduction in neural circuitry hyperexcitability, and a decrease in cocontraction of spastic antagonists by the use of TENS [13,14]. In addition, NMES has been shown to improve walking dysfunction in stroke survivors, leading to improved gait speed and balance [15]. However, the potential limitation of high-frequency electrical stimulation is the induction of muscle fatigue [16], which may in turn exacerbate spasticity [17].

The Exopulse Mollii Suit (EMS; Inerventions AB) is a multichannel full-body garment (jacket and pants) consisting of 58 electrodes. It is powered by a 4×AAA battery-powered control unit that transmits low-energy electrical pulses to the garments via the connectors. The suit comes in 37 sizes, starting from 104 cm up to 5XL for women and men. In contrast to NMES and TENS, the EMS uses 2 mA monophasic square waves to deliver low-frequency (20 Hz), low-amplitude (20 V) electrical stimulation through prepositioned electrodes to stimulate up to 40 key muscle groups simultaneously. Its pulse width is adjustable from 30 to 175 microseconds, with a period length of 50 milliseconds between pulses, to allow for varying intensities of electrical stimulation. During stimulation, the pulse width increases by 5 microseconds every 5 seconds from 25 microseconds until the preset pulse width is reached. This targeted suprasensory and submotor electrical stimulation aims to modulate the stretch reflex and mitigate spasticity via reciprocal inhibition of antagonistic muscles mediated by Renshaw cells [18,19]. Apart from being noninvasive and nonpharmacological, the EMS may reduce the risk of muscle fatigue associated with higher-frequency stimulation and can potentially be self-administered at home, making it a compelling option for improving function with minimal adverse effects.

In this study, we aimed primarily to examine the acute effects of individualized electrostimulation using the EMS on spasticity in participants with chronic stroke or SCI. Secondarily, we aimed to assess the impact of a single session on gait and ambulatory function. We also assessed user perceptions of the usability and acceptability of the EMS.

Methods

Study Design and Participants

This was a case series of participants with spasticity from chronic stroke or SCI. Participants were included if they had upper or lower limb spasticity that impacted function and gait and were able to walk with minimal or no assistance (ie, Functional Ambulatory Category [FAC] scores of 2 - 5) [20]. Exclusion criteria included preexisting medical conditions preventing safe use of the EMS, such as the presence of implanted electromagnetic devices (eg, shunts, pacemakers, and intrathecal baclofen pumps), epilepsy, malignancy, uncontrolled cardiac conditions, and poor skin condition. Patients with recent initiation of oral muscle relaxants or botulinum toxin injections in the past 6 months were also excluded. A convenience sample was identified by the study team and contacted by the principal investigator for participation.

Study Protocol

An interview for goal selection was performed for each participant before the study. All outcome measures were performed before and after a single session of stimulation with the suit.

Participants were fitted with appropriately sized suits to ensure good contact between the electrodes and the skin. They donned and doffed the EMS with assistance, ensuring that it was correctly worn—for example, zippers were positioned on the outside of the forearm, forming a straight line between the elbow and the little finger, and the seam was placed above the knee. Stimulation parameters were determined by an experienced physical therapist from Exopulse based on the results of the initial clinical assessment. The stimulation parameters were individualized and set to stimulate the antagonists of spastic muscles. A preset program was used to adjust the intensity of electrical stimulation across levels 6 to 22, depending on both the size of the target muscle group and severity of spasticity. Higher-frequency electrical stimulation is used for larger muscle groups or to manage more severe spasticity. Participants wore the suit for 60 minutes. This intervention period included 30 minutes of standardized goal-directed activities performed in two 15-minute blocks, interspersed with 15-minute rest breaks. These light activities simulated daily activities, such as food preparation or walking, and were goal directed. Figure 1 shows a participant walking with a broad-based quad stick while wearing the EMS. Figure 2 depicts the living room–like setting where the study was conducted.

XSL•FO
RenderX

Figure 1. Participant walking with a broad-based quad stick while wearing the EMS. (A) Individualized stimulation parameters are programmed into the control unit that generates electrical pulses. (B) Jacket and pants containing 58 embedded electrodes in direct contact with the skin delivering electrical stimulation.



XSL•FO
RenderX

Figure 2. Living room–like setting where the study was conducted. (A) Participant could sit on the sofa to rest or read, simulating a home environment. (B) Activities of daily living, such as preparing a drink, could be performed in this area.



Study Outcomes

Outcome measures before and after stimulation included measures of spasticity, using the Modified Tardieu Scale (MTS) [21] paired with surface electromyography (sEMG). Assessments of mobility included FAC [20], 10-meter walk test (10MWT) [22] for walking speed, five times sit-to-stand test [23] for transfer skills, and step test [24] for dynamic balance were scored. All outcomes were assessed by an experienced physiotherapist not involved in determining the stimulation parameters for the participants.

sEMG signals were recorded during high-velocity passive stretch movements (V3) during MTS measurements using Cometa EMG system (version 8.7.6.0; Cometa SRL) via lightweight wireless electromyography electrodes placed on the muscle

XSL•FO
RenderX

belly of a standardized set of target muscles. The sampling rate for sEMG was 2000 Hz. Raw sEMG data were then processed by (1) low-pass filtering at 500 Hz with an eighth-order Butterworth filter and high-pass filtering at 20 Hz with a ninth-order Butterworth filter, (2) full wave rectification, and (3) smoothing with the root mean square (RMS) method. The time domain RMS feature of the sEMG signals was extracted. Three indexes were used for spasticity measurement: (1) The RMS (sEMG) of the agonist signal, (2) the RMS (sEMG) of the antagonist signal, and (3) spasticity index [25]:

$$\text{sEMG index} = \text{RMS}_{\text{agonist}} / (\text{RMS}_{\text{agonist}} + \text{RMS}_{\text{antagonist}})$$

For example, the *sEMG index* of biceps measured during passive elbow extension =RMS of triceps/RMS of triceps +RMS of biceps, while the *sEMG index* of the gastrocnemius measured during passive knee flexion = RMS of gastrocnemius/RMS of rectus femoris+RMS of gastrocnemius.

The spasticity angle derived from the MTS was corroborated with the sEMG index. A larger sEMG index indicates less spasticity.

Full-body sagittal plane videos of participants walking at comfortable walking speed, were recorded using an iPad Pro (Apple) at a height of approximately 1.5 m from the ground before and after use of the EMS. These were used for observational gait analysis, as well as quantification of spatiotemporal gait parameters using a markerless vision-based motion capture system. The system used OpenPose BODY25 pose estimation model [26] for 25-keypoint human pose estimation, gait events detection, and derivation of gait parameters via the coordinate-based treadmill algorithm method [27]. The system had previously been validated to detect gait abnormalities [28,29]. A single most accurately captured gait cycle from each walking trial was used to assess spatiotemporal parameters. Additionally, to quantify gait symmetry of participants with stroke, ratio index (RI) of step length was calculated [30].

RI=X1/X2, where X1 is the step length of left lower limb and X2 is the step length of the right lower limb [31]. In participants with stroke, the step length of the paretic lower limb is used as the numerator.

The Assistive Technology Usability Questionnaire for People With Neurological Diseases (NATU Quest) [32] was used to assess acceptability and perceived effectiveness of the EMS and was performed by the same study team member that performed the initial assessment. Scores range from 0 to 5, with 0 to 2 reflecting negative outcomes and 3 to 5 indicating positive ones. An average score of 3 indicates marginal usability, while scores of 4 or 5 denote high usability.

Ethical Considerations

The study has been reviewed and approved by the Domain Specific Review Board, part of the National Healthcare Group (2023/00832) in Singapore.

Results

Demographics and Baseline Characteristics

Seven participants took part in the study. All participants were male due to recruitment constraints, which may limit generalizability. The characteristics of the participants, patient-defined goals, and goal-directed activities performed during the trial, are outlined in Table 1. Four participants had spastic hemiparesis due to stroke, while 3 had spastic tetraparesis due to cervical SCI. All patients had goals focused on improving function, either related to walking or performing activities of daily living.

XSL•FO
RenderX

Table . Demographics and characteristics of participants[a].

ID	Age (y)	Diagnosis	Time from injury (m)	Motor impairments	Individualized goals	Goal-directed activities performed	FAC[b] score	Walking aid	Clinical gait abnormality
1	68	Left lentiform nucleus bleed	12	Right spastic hemiparesis (MRC[c] grade 4 proximally) with foot drop	Improve hand function, upper limb spasticity, and walking while reducing dependence on the ankle-foot orthosis	Maintaining grip on a cup and walking	3	Broad-based quad stick	Increased pelvic obliquity, decreased knee flexion, and foot drop with high steppage during swing phase
2	65	Left subcortical infarct	20	Right spastic hemiparesis (MRC grade 4 proximally and MRC grade 1 distally)	Increase confidence in walking	Sit-to-stand practice and walking	4	Walking stick	Decreased knee flexion and ankle dorsiflexion during swing phase
3	61	Left thalamic bleed	31	Right spastic hemiparesis (MRC grade 4) with ataxia	Improve walking distance and the ability to perform daily activities	Upper limb activities such as stacking cones and blocks, using a mobile phone, and walking	3	Walking stick	Knee hyperextension during stance phase
4	48	Right subcortical infarct	5	Left spastic hemiparesis (MRC grade 4)	Improve hand function and walking speed	Upper limb activities, including preparing a drink, stacking cones and blocks, and walking	5	Nil	Decreased knee flexion and ankle dorsiflexion during swing phase
5	56	AIS[d] C C5 spinal cord injury	350	Spastic paraparesis	Improve walking and balance	Reading a book, sit-to-stand practice and walking	3	Walking stick	Crouched gait with increased hip adduction and anterior pelvic tilt; decreased hip extension during stance phase and decreased knee extension during swing phase
6	76	AIS C C5 spinal cord injury	50	Spastic tetraparesis	Improve hand dexterity, lower limb spasticity and walking endurance	Upper limb activities including stacking cones and blocks, sit-to-stand practice, and walking	3	Nil	Decreased right hip flexion, knee flexion, and ankle dorsiflexion during swing phase

XSL•FO
RenderX

ID	Age (y)	Diagnosis	Time from injury (m)	Motor impairments	Individualized goals	Goal-directed activities performed	FAC[b] score	Walking aid	Clinical gait abnormality
7	27	AIS D C6 spinal cord injury	30	Spastic tetraparesis	Improve walking speed and endurance	Walking	3	Nil	Increased hip adduction; anterior pelvic tilt; decreased knee extension on swing and stance phase bilaterally, worse on right lower limb; increased ankle plantarflexion on stance phase on right lower limb

[a]All participants were male.

[b]FAC: Functional Ambulatory Category.

[c]MRC: Medical Research Council.

[d]AIS: American Spinal Injury Association Impairment Scale.

Primary Outcome: Spasticity

Only 1 participant with stroke and 1 participant with SCI demonstrated improvement in MTS greater than minimal detectable change of 17.16° [33] (Table 2). The sEMG index corroborated the reduction in spasticity in these muscles (Multimedia Appendix 1).

XSL•FO
RenderX

Table . Pre- and postintervention Modified Tardieu Scale (MTS) measurements.

Participant	Muscle	Before MTS, R2-R1[a] (degrees)	After MTS, R2-R1(degrees)	Change (degrees)	Before MTS, X[b]	After MTS, X	Change in X
Stroke							
1	Right biceps brachii	57	86	29	2	2	0
1	Right gastrocnemius	27	31	4	2	2	0
2	Right biceps brachii	35	38	3	2	2	0
2	Right gastrocnemius	30	22	−8	3	3	0
3	Right biceps brachii	38	35	−3	2	2	0
3	Right gastrocnemius	30	45	15	2	2	0
4	Left biceps brachii	13	0	−13	2	0	−2
4	Left triceps brachii	93	0	−93[c]	2	0	−2
4	Left rectus femoris	35	40	5	2	2	0
4	Left gastrocnemius	30	25	−5	2	2	0
Spinal cord injury							
5	Right rectus femoris	17	40	23	2	2	0
5	Right gastrocnemius	0	36	36	0	2	2
5	Left rectus femoris	13	41	28	2	2	0
5	Left gastrocnemius	0	30	30	0	2	2
6	Right biceps brachii	58	0	−58[c]	2	0	−2
6	Right rectus femoris	20	45	25	2	2	0
6	Left rectus femoris	28	15	−13	2	2	0
7	Right rectus femoris	32	66	34	2	2	0
7	Right tibialis anterior	40	28	−12	2	2	0

[a]R2: the angle of full range of motion, taken at a very slow speed (V1); R1: the angle of muscle reaction in which a catch or clonus is found during a quick stretch (V3).

[b]X: quality of muscle reaction score in the Modified Tardieu Scale.

[c]Greater than minimal detectable change.

Secondary Outcomes: Function

Assessments of Mobility

FAC remains unchanged for all participants after the EMS. However, all participants with SCI and 1 participant with stroke demonstrated improved walking speed greater than SE of the measure of 0.05 m/s [34,35]. Two participants with SCI and 1 participant with stroke showed improvements reaching minimal clinically important difference in five times sit-to-stand test of 2.27 seconds and 1.94 seconds, respectively [36,37].

XSL•FO
RenderX

Furthermore, improvements in the step test were observed in 2 participants with SCI and 2 participants with stroke (Table 3). Data on upper limb function are presented in Multimedia Appendix 1.

Table . Mobility assessments before and after Exopulse Mollii Suit (EMS) use.

	FAC[a] score		10MWT[b] (m/s)			5xSTS[c] (s)			Step test	
	Before EMS	After EMS	Before EMS	After EMS	Change (%)	Before EMS	After EMS	Change (%)	Before EMS	After EMS
Stroke										
Participant 1	3	3	0.19	0.19	0	22	27.68	25.82	4	5
Participant 2	4	4	0.31	0.25	−19.35	22	23.15	5.23	5	4
Participant 3	3	3	0.31	0.29	−6.45	62.85	58.75[d]	−6.52	5	4
Participant 4	5	5	0.77	0.84[d]	9.09	12.94	12.63	−2.40	10	14
Spinal cord injury										
Participant 5	3	3	0.24	0.35[d]	45.83	21.56	21.71	0.70	7	6
Participant 6	3	3	0.35	0.41[d]	17.14	34.88	22.72[d]	−34.86	5	9
Participant 7	3	3	0.36	0.57[d]	58.33	19.90	16.48[d]	−17.20	7	9

[a]FAC: Functional Ambulatory Category.

[b]10MWT: 10-meter walk test.

[c]5xSTS: five times sit-to-stand test.

[d]Greater than SE of the measure and minimal clinically important difference.

Spatiotemporal Gait Parameters

In participants with stroke, gait abnormalities were relatively unchanged after EMS use. However, gait symmetry (RI) improved in 3 of 4 patients with stroke. Single-support time increased for 2 participants and double-support time decreased in 1. Specifically, walking speed, cadence, and step time improved in participants 2 and 4 (Table 4) [38].

XSL•FO
RenderX

Table . Spatial and temporal gait parameters of participants with stroke before (pre) and after (post) using the Exopulse Mollii Suit.

	Participant 1[a]				Participant 2[b]				Participant 3[c]				Participant 4[d]			
	Left		Right (affected)		Left		Right (affected)		Left		Right (affected)		Left (affected)		Right	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Spatial parameters (m)																
Step length	0.37	0.35	0.34	0.30	0.36	0.33	0.45	0.32	0.45	0.42	0.34	0.34	0.87	0.72	0.71	0.76[e]
Stride length	0.67	0.65	0.71	0.69	0.68	0.51	0.81	0.67	0.61	0.65	0.77	0.76	1.39	1.44[e]	1.58	1.48
Temporal parameters (s)																
Double support	1.91	1.97	1.72	2.03	0.94	0.68[e]	0.94	0.85[e]	0.67	0.71	0.37	0.67	0.43	0.45	0.40	0.44
Single support	0.78	0.77	0.56	0.60[e]	0.72	0.83[e]	0.46	0.48[e]	0.48	0.50[e]	0.62	0.37	0.48	0.41	0.58	0.49
Total stance time (% of total gait cycle)	—[f]	—	2.27 (70)	—2.64 (79)	—	—	1.48 (67)	—1.35 (64)	—	—	0.99 (76)	1.05 (70)	0.92 (61)	—0.86 (64)	—	—
Total swing time (% of total gait cycle)	—	—	0.97 (30)	—0.71 (21)	—	—	0.73 (33)	—0.75 (36)	—	—	0.31 (24)	0.45 (30)	0.59 (39)	—0.48 (36)	—	—
Walking speed (m/s)	0.20	0.20	0.22	0.20	0.32	0.25[e]	0.37	0.34[e]	0.39	0.41[e]	0.56	0.49	0.93	1.07[e]	1.10	1.10
Cadence (steps/min)	36.94	36.47	37.03	35.22	56.72	58.33[e]	55.62	59.78[e]	77.45	76.23	83.27	77.87	80.03	88.93[e]	83.27	89.48[e]
Step time	1.20	1.17[e]	2.04	2.12	0.80	0.81	1.36	1.16[e]	0.72	0.62	0.77	0.92	0.80	0.75	0.64	0.59
Stride time	3.25	3.29	3.24	3.41	2.12	2.06	2.16	2.01[e]	1.55	1.57	1.44	1.54	1.50	1.35	1.44	1.34

[a]Ratio index (step length): baseline 0.92; after EMS 0.86.

[b]Ratio index (step length): baseline 1.25; after EMS 0.97.

[c]Ratio index (step length): baseline 0.76; after EMS 0.81.

[d]Ratio index (step length): baseline 1.23; after EMS 0.95.

[e]Greater than minimal detectable change.

[f]Total stance and swing time were only recorded for the stroke-affected side.

In participants with SCI, improved gait patterns were observed after use of the EMS. For instance, there was less hip adduction, increased hip flexion, and knee extension during the swing phase in participant 5. Spatial and temporal parameters also improved. There was a reduction in double-support time, an increase in single-support time, and longer step and stride lengths. Additionally, walking speed, cadence, step time, and stride time improved [39]. The stance-to-swing ratio also improved [40] (Table 5).

XSL•FO
RenderX

Table . Spatial and temporal gait parameters in participants with spinal cord injury before (pre) and after (post) using the Exopulse Mollii Suit.

	Participant 5[a]				Participant 6[b]				Participant 7[c]			
	Left		Right		Left		Right		Left		Right	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Spatial parameters (m)												
Step length	0.37	0.47	0.44	0.47	0.44	0.41	0.09	0.25[d]	0.49	0.58	0.34	0.44
Stride length	0.73	0.87	0.87	0.80	0.49	0.66	0.58	0.69	0.77	1.03	0.83	1.12
Temporal parameters (s)												
Double support	1.89	0.99[d]	1.52	1.13	0.67	0.49	0.65	0.39	0.82	0.58	0.72	0.53
Single support	0.54	0.48	0.71	0.57	0.42	0.55	0.38	0.45	0.53	0.62	0.38	0.37
Total stance time (% of total gait cycle)	2.44 (81)	1.47 (75)	2.23 (80)	1.74 (76)	1.09 (73)	1.04 (66)	1.03 (71)	0.84 (64)	1.35 (77)	1.22 (76)	1.10 (70)	0.90 (63)
Total swing time (% of total gait cycle)	0.54 (19)	0.49 (25)	0.55 (20)	0.55 (24)	0.40 (27)	0.54(34)	0.42 (29)	0.47 (36)	0.40 (23)	0.39 (24)	0.47 (30)	0.53 (37)
Walking speed (m/s)	0.25	0.42[d]	0.30	0.36	0.34	0.45	0.36	0.50	0.44	0.63[d]	0.50	0.73[d]
Cadence (step/min)	40.58	58.56[d]	41.88	53.56	81.45	82.79	81.44	86.26	68.93	73.88	73.13	78.30
Step time (s)	1.37	1.04	1.45	0.99	0.70	0.67	0.77	0.77	0.78	0.74	0.86	0.88
Stride time (s)	2.96	2.05	2.87	2.24	1.47	1.45	1.47	1.39	1.74	1.62	1.64	1.53

[a]Ratio index (step length): baseline 0.90; after EMS 1.09.

[b]Ratio index (step length): baseline 4.89; after EMS 1.64.

[c]Ratio index (step length): baseline 1.44; after EMS 1.32.

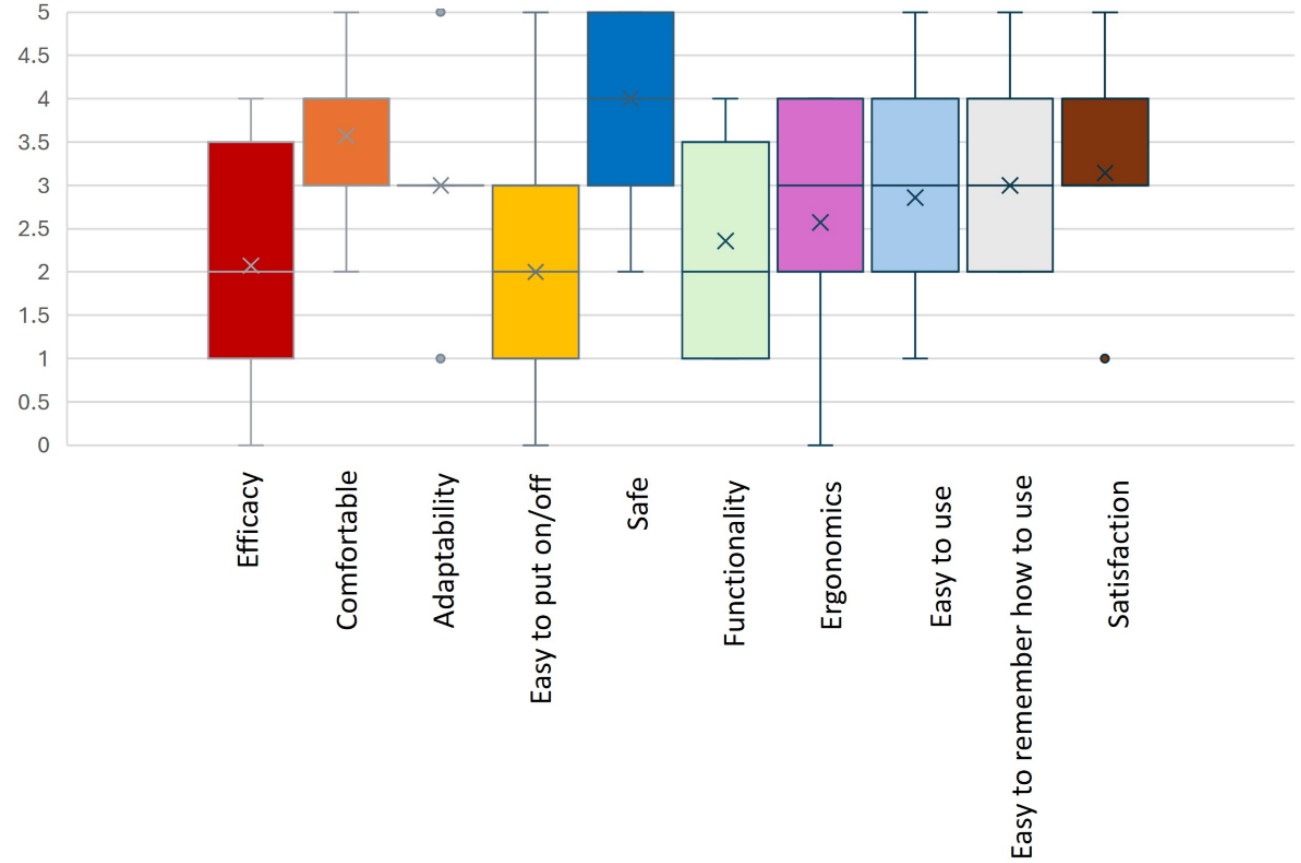
[d]Greater than minimal detectable change.

Qualitative Feedback

The NATU Quest results are shown in Figure 3. On a Likert scale of 1 to 5, the EMS was scored in 10 distinct categories including efficacy, comfort, adaptability, ease of donning and doffing, safety, functionality, ergonomics, satisfaction, ease of use and remembering how to use. Overall, 3 of 7 users returned positive results, as defined by an average overall score of 3 or more. Median scores were positive in comfort, adaptability, safety, ease of remembering how to use, and satisfaction (Figure 3).

XSL•FO
RenderX

Figure 3. Qualitative feedback of Exopulse Mollii Suit.



Summary of Outcomes

An overall outcome was deemed positive if the EMS met the individualized goals specified in Table 1, and negative if these goals were not accomplished. Three participants with SCI and 1 participant with stroke had positive outcomes. Table 6 summarizes the quantitative outcomes and qualitative feedback on the EMS for each patient.

XSL•FO
RenderX

Table . Summary of key outcomes: quantitative and qualitative response to Exopulse Mollii suit (EMS) for each patient.

	Spasticity outcome	Gait outcome	Mobility outcome	Qualitative feedback	Overall outcome
Stroke					
Participant 1	Worsened spasticity angle and sEMG[a] index in right biceps brachii	• Increased single support time • Decreased walking speed, cadence, and step length on stroke affected side	Improved ST[b] and worsened 5xSTS[c]	• Felt the EMS was comfortable and safe • However, the overall usability score indicated notable challenges in usability	Negative
Participant 2	Equivocal	• Improved gait symmetry, increased single support time and decreased double support time • Increased walking speed on stroke affected side	Worsened 10MWT[d] and ST	• Felt the EMS was safe, easy to use and easy to remember how to use it • However, the overall usability score indicated notable challenges in usability	Negative
Participant 3	Equivocal	• Improved gait symmetry • Increased double support time, decreased walking speed on stroke affected side	Improved 5xSTS and worsened ST	• The overall usability score indicated notable challenges in usability	Negative
Participant 4	Improved spasticity angle and sEMG index in left triceps	• Improved gait symmetry • Increased walking speed and increased cadence on stroke affected side	Improved 10MWT and ST	• Felt the EMS was comfortable and safe • However, the overall usability score indicated notable challenges in usability	Positive; improved walking
Spinal cord injury					
Participant 5	Worsened spasticity angle and sEMG index in right and left rectus femoris	• Improved gait symmetry • Decreased double-support time • Increased walking speed and cadence on left lower limb	Improved 10MWT and worsened ST	• Felt the EMS was efficacious, comfortable, adaptable, easy to put on and off, safe, facilitates movement, ergonomic, easy to use, easy to remember how to use it, and is satisfied • The usability score reflected a high level of usability	Positive; improved walking
Participant 6	Improved spasticity angle in right biceps but worsened spasticity angle and sEMG index in right rectus femoris	• Improved gait symmetry	Improved 10MWT, 5xSTS, and ST	• Felt the EMS was comfortable • Overall, the usability score indicates marginal usability	Positive; improved walking and upper limb function

XSL•FO
RenderX

	Spasticity outcome	Gait outcome	Mobility outcome	Qualitative feedback	Overall outcome
Participant 7	Equivocal	• Improved gait symmetry • Increased walking speed on both lower limbs	Improved 10MWT, 5xSTS, and ST	• Felt the EMS was safe, ergonomic and is satisfied • Overall, the usability score indicates marginal usability	Positive; improved walking

[a]sEMG: surface electromyography.

[b]ST: step test.

[c]5xSTS: five times sit-to-stand test.

[d]10MWT: 10-meter walk test.

Discussion

Principal Findings

Our study investigated the acute effects of individualized electrostimulation using a single session with the EMS on spasticity and ambulatory function in participants with stroke and SCI. While improvements in objective measures of spasticity at rest were modest, improvements in ambulatory function and gait parameters were observed.

Four of 7 (57.1%) participants improved in walking speed (10MWT), with only participants with SCI exceeding the minimal clinically important difference of 0.06 m/s [41-43]. The limited response in participants with stroke may be due to the responsiveness of the test, particularly in more severely impaired participants [44]. The limited dose of intervention, as well as other factors impacting on mobility including weakness, impaired muscle synergies, sensory impairment, and impaired endurance may also be contributory. Furthermore, participants who were dependent on physical assistance or had slower walking speed did not show improvements on 10MWT. Functional endurance also influences performance in 10MWT particularly when walking requires considerable effort [45,46]. Outcomes were measured immediately after the EMS session, when participants may have experienced fatigue from the standardized goal-directed activities. Nevertheless, improvements were noted in step test (4 of 7 participants), reflecting better dynamic balance during weight shift and paretic lower-extremity motor control; as well as step-length symmetry (n=6, 85.7%), reflecting increased energy efficiency during walking [47]. Furthermore, step test and step-length symmetry are both predictors of falls risk [48-50]. A trend toward increased single-support time and decreased double-support times in our cohort also suggests improved balance and neuromuscular control [51].

Our findings of modest improvement in spasticity measures (n=2, 28.6%) at rest, including MTS and sEMG indices, corroborated with a previous double-blind crossover study which compared the EMS at 2 different frequencies (20 Hz and 30 Hz) with sham stimulation in patients with chronic stroke and reported no observed reduction in spasticity [52]. In contrast, Palmcrantz et al [53] reported the effectiveness of repeated sessions with the EMS in decreasing wrist flexor spasticity in patients with chronic stroke in a home setting (60 minutes every second day for 6 weeks; total 21 sessions). These disparate findings highlight the need for further investigation into the dose-response relationship and the potential for cumulative effects with repeated EMS sessions. Participants within our cohort who had clinically significant improvements in MTS were also those with greater spasticity, as indicated by larger spasticity angles. Pennati et al [52] also reported high interindividual variability in response to the EMS and suggested that it might be more efficacious in severe spasticity.

The observed improvements in gait symmetry among participants with stroke (3/4, 75%) are encouraging, as they suggest the potential for enhanced walking efficiency. The more pronounced improvements in gait parameters among participants with SCI warrant further investigation into possible differences in the mechanisms of response to the EMS between stroke and SCI populations. These differences may relate to differences in spinal and supraspinal mechanisms of spasticity, as well as varying extent of involvement of the descending pathways responsible for spasticity and its subsequent adaptations [54]. Additionally, presynaptic inhibition of Ia afferent terminals in muscles may remain unchanged in patients with hemiplegic stroke, which could influence responsiveness to the EMS [55]. In this limited series, more participants with SCI reported a favorable response to the EMS compared with participants with stroke. To our knowledge, this is the first report of the effects of the EMS in participants with SCI.

Qualitative feedback was mixed, with a median rating of 2 for ease of donning and doffing, mirroring that of a previous structured interview in patients with stroke and cerebral palsy who had spasticity, who described difficulty in wearing the suit at home for self-administered electrotherapy [56]. This is not unexpected as the rubber electrodes used in the EMS may cause friction against the skin but are necessary to manage impedance and ensure effective electrostimulation [57]. Concerns regarding the ease of donning and doffing, ergonomics, and functionality highlight significant barriers to self-administration and adherence and should be addressed in future iterations of the device to support wider adoption for home use. However, this score should also be interpreted with caution as it reflects the initial impression with a single-session trial. Nevertheless, most participants expressed satisfaction with the session. Others have reported improved patient well-being and muscle-relaxing effect during stimulation [58].

XSL•FO
RenderX

Limitations

We recognize that the absence of a control group and the small sample size in our study limit the reproducibility and generalizability of our findings. Future studies should also aim for greater gender representation. Observed gains in spasticity, gait, and mobility measures could have been confounded by practice effects or neuromuscular warm-up. These factors may directly confound mobility and gait outcomes and hence preclude definitive attribution of the observed changes to EMS use. We designed a structured and clearly defined protocol to standardize the duration of the EMS to 60 minutes for all patients, comprising 30 minutes of standardized light activities performed in two 15-minute blocks, interspersed with 15-minute rest intervals. Furthermore, the use of fixed activity durations and rest periods enhances reproducibility and facilitates replicability in future studies. Study rigor was further strengthened by the involvement of experienced physiotherapists from Exopulse in determining the stimulation parameters, as this process requires specialized expertise. The team from Exopulse had also supported previously published studies [52,53]. In addition, the physiotherapist responsible for outcome assessment was blinded to the stimulation parameters, thereby minimizing the risk of cognitive bias and strengthening the internal validity of the study. To our knowledge, this is the first study to report spatiotemporal gait parameters with EMS use. Future controlled studies with repeated sessions should aim to confirm the efficacy of the EMS in defined patient populations. In addition, further qualitative studies of patient experience may provide insights into functional benefits in real-life activities and participation that are not fully captured by standardized quantitative measures.

Conclusions

With an individualized approach, a single 60-minute session of EMS use may enhance functional mobility and gait quality in some participants with spasticity resulting from stroke or SCI. User qualitative feedback on EMS usability was mixed, with reported challenges in donning and doffing likely reflecting initial experiences with a novel intervention. This finding underscores the need for design enhancements and comprehensive patient onboarding and training to ensure competencies to improve user experience, support compliance, and facilitate independent use at home. Further studies are needed to validate its effectiveness. Overall, the EMS demonstrates potential as a therapeutic, nonpharmacological option for spasticity management, particularly for individuals who require accessible care beyond traditional clinical settings.

Acknowledgments

The authors would like to thank Ms Janelle Phua for assisting with logistical arrangements and Dr Jörgen Sandell for assisting with the study.

Funding

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Data Availability

Data can be made available to interested researchers upon reasonable request by email to the corresponding author.

Authors' Contributions

Conceptualization: JMY, EC
Data curation: NSK, HSL, NT
Formal analysis: JMY, EC
Methodology: JMY
Project administration: NSK, HSL, NT
Writing—original draft: YJM
Writing—review and editing: JMY, NSK, EC

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data on upper limb function.
[DOCX File, 21 KB - rehab_v13i1e81522_app1.docx]

References

1. Wissel J, Manack A, Brainin M. Toward an epidemiology of poststroke spasticity. Neurology 2013 Jan 15;80(3 Suppl 2):S13-S19. [doi: 10.1212/WNL.0b013e3182762448] [Medline: 23319481]
2. Sangari S, Perez MA. Prevalence of spasticity in humans with spinal cord injury with different injury severity. J Neurophysiol 2022 Sep 1;128(3):470-479. [doi: 10.1152/jn.00126.2022] [Medline: 35507475]

XSL•FO
RenderX

3. Lance JW. Symposium synopsis. In: Feldman RG, Young RR, Koella WP, editors. Spasticity: Disordered Motor Control: Year Book Publishers; 1980:485-494.
4. Pandyan AD, Gregoric M, Barnes MP, et al. Spasticity: clinical perceptions, neurological realities and meaningful measurement. Disabil Rehabil 2005;27(1-2):2-6. [doi: 10.1080/09638280400014576] [Medline: 15799140]
5. Olaleye OA, Zaki DA, Hamzat TK. Expectations of individuals with neurological conditions from rehabilitation: a mixed-method study of needs. S Afr J Physiother 2021;77(1):1498. [doi: 10.4102/sajp.v77i1.1498] [Medline: 33604478]
6. Chang E, Ghosh N, Yanni D, Lee S, Alexandru D, Mozaffar T. A review of spasticity treatments: pharmacological and interventional approaches. Crit Rev Phys Rehabil Med 2013;25(1-2):11-22. [doi: 10.1615/CritRevPhysRehabilMed.2013007945] [Medline: 25750484]
7. Khan F, Amatya B, Bensmail D, Yelnik A. Non-pharmacological interventions for spasticity in adults: an overview of systematic reviews. Ann Phys Rehabil Med 2019 Jul;62(4):265-273. [doi: 10.1016/j.rehab.2017.10.001] [Medline: 29042299]
8. Amatya B, Khan F, La Mantia L, Demetrios M, Wade DT. Non pharmacological interventions for spasticity in multiple sclerosis. Cochrane Database Syst Rev 2013 Feb 28;2013(2):CD009974. [doi: 10.1002/14651858.CD009974.pub2] [Medline: 23450612]
9. Reebye R, Jacinto LJ, Balbert A, et al. Multimodal therapy and use of adjunctive therapies to BoNT-A in spasticity management: defining terminology to help enhance spasticity treatment. Front Neurol 2024;15:1432330. [doi: 10.3389/fneur.2024.1432330] [Medline: 39281409]
10. Chasiotis A, Giannopapas V, Papadopoulou M, et al. The effect of neuromuscular electrical nerve stimulation in the management of post-stroke spasticity: a scoping review. Cureus 2022 Nov;14(11):e32001. [doi: 10.7759/cureus.32001] [Medline: 36600817]
11. Fernández-Tenorio E, Serrano-Muñoz D, Avendaño-Coy J, Gómez-Soriano J. Transcutaneous electrical nerve stimulation for spasticity: a systematic review. Neurologia (Engl Ed) 2019 Sep;34(7):451-460. [doi: 10.1016/j.nrl.2016.06.009] [Medline: 27474366]
12. Mahmood A, Veluswamy SK, Hombali A, Mullick A, N M, Solomon JM. Effect of transcutaneous electrical nerve stimulation on spasticity in adults with stroke: a systematic review and meta-analysis. Arch Phys Med Rehabil 2019 Apr;100(4):751-768. [doi: 10.1016/j.apmr.2018.10.016] [Medline: 30452892]
13. Field-Fote EC. Electrical stimulation modifies spinal and cortical neural circuitry. Exerc Sport Sci Rev 2004 Oct;32(4):155-160. [doi: 10.1097/00003677-200410000-00006] [Medline: 15604934]
14. Levin MF, Hui-Chan CW. Relief of hemiparetic spasticity by TENS is associated with improvement in reflex and voluntary motor functions. Electroencephalogr Clin Neurophysiol 1992 Apr;85(2):131-142. [doi: 10.1016/0168-5597(92)90079-q] [Medline: 1373366]
15. Chen S, Gao J, Zhou Y, et al. Implications of neuromuscular electrical stimulation on gait ability, balance and kinematic parameters after stroke: a systematic review and meta-analysis. J Neuroeng Rehabil 2024 Sep 18;21(1):164. [doi: 10.1186/s12984-024-01462-2] [Medline: 39294678]
16. Dreibati B, Lavet C, Pinti A, Poumarat G. Influence of electrical stimulation frequency on skeletal muscle force and fatigue [Article in English, French]. Ann Phys Rehabil Med 2010 May;53(4):266-271. [doi: 10.1016/j.rehab.2010.03.004] [Medline: 20430713]
17. Cheung J, Rancourt A, Di Poce S, et al. Patient-identified factors that influence spasticity in people with stroke and multiple sclerosis receiving botulinum toxin injection treatments. Physiother Can 2015;67(2):157-166. [doi: 10.3138/ptc.2014-07] [Medline: 25931667]
18. Trompetto C, Marinelli L, Mori L, et al. Pathophysiology of spasticity: implications for neurorehabilitation. Biomed Res Int 2014;2014:354906. [doi: 10.1155/2014/354906] [Medline: 25530960]
19. Alvarez FJ, Fyffe RE. The continuing case for the Renshaw cell. J Physiol 2007 Oct 1;584(Pt 1):31-45. [doi: 10.1113/jphysiol.2007.136200] [Medline: 17640932]
20. Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L. Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. Phys Ther 1984 Jan;64(1):35-40. [doi: 10.1093/ptj/64.1.35] [Medline: 6691052]
21. Haugh AB, Pandyan AD, Johnson GR. A systematic review of the Tardieu Scale for the measurement of spasticity. Disabil Rehabil 2006 Aug 15;28(15):899-907. [doi: 10.1080/09638280500404305] [Medline: 16861197]
22. Wade DT, Wood VA, Heller A, Maggs J, Langton Hewer R. Walking after stroke. Measurement and recovery over the first 3 months. Scand J Rehabil Med 1987;19(1):25-30. [Medline: 3576138]
23. Muñoz-Bermejo L, Adsuar JC, Mendoza-Muñoz M, et al. Test-retest reliability of five times sit to stand test (FTSST) in adults: a systematic review and meta-analysis. Biology (Basel) 2021 Jun 9;10(6):510. [doi: 10.3390/biology10060510] [Medline: 34207604]
24. Hill KD. A new test of dynamic standing balance for stroke patients: reliability, validity and comparison with healthy elderly. Physiother Can 1996;48(4):257-262. [doi: 10.3138/ptc.48.4.257]
25. Wang L, Guo X, Fang P, et al. A new EMG-based index towards the assessment of elbow spasticity for post-stroke patients. Annu Int Conf IEEE Eng Med Biol Soc 2017 Jul;2017:3640-3643. [doi: 10.1109/EMBC.2017.8037646] [Medline: 29060687]

XSL•FO
RenderX

26. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell 2021 Jan;43(1):172-186. [doi: 10.1109/TPAMI.2019.2929257] [Medline: 31331883]
27. Zahradka N, Verma K, Behboodi A, Bodt B, Wright H, Lee SC. An evaluation of three kinematic methods for gait event detection compared to the kinetic-based 'gold standard'. Sensors (Basel) 2020 Sep 15;20(18):5272. [doi: 10.3390/s20185272] [Medline: 32942645]
28. Shi B, Yen SC, Tay A, Tan DM, Chia NS, Au WL. Convolutional neural network for freezing of gait detection leveraging the continuous wavelet transform on lower extremities wearable sensors data. Annu Int Conf IEEE Eng Med Biol Soc 2020 Jul;2020:5410-5415. [doi: 10.1109/EMBC44109.2020.9175687] [Medline: 33019204]
29. Shi B, Tay A, Au WL, Tan DM, Chia NS, Yen SC. Detection of freezing of gait using convolutional neural networks and data from lower limb motion sensors. IEEE Trans Biomed Eng 2022 Jul;69(7):2256-2267. [doi: 10.1109/TBME.2022.3140258] [Medline: 34986092]
30. Patterson KK, Gage WH, Brooks D, Black SE, McIlroy WE. Evaluation of gait symmetry after stroke: a comparison of current methods and recommendations for standardization. Gait Posture 2010 Feb;31(2):241-246. [doi: 10.1016/j.gaitpost.2009.10.014] [Medline: 19932621]
31. Viteckova S, Kutilek P, Svoboda Z, Krupicka R, Kauler J, Szabo Z. Gait symmetry measures: a review of current and prospective methods. Biomed Signal Process Control 2018 Apr;42:89-100. [doi: 10.1016/j.bspc.2018.01.013]
32. Masbernat-Almenara M, Rubi-Carnacea F, Opisso E, Duarte-Oller E, Medina-Casanovas J, Valenzuela-Pascual F. Developing an assistive technology usability questionnaire for people with neurological diseases. PLoS One 2023;18(1):e0281197. [doi: 10.1371/journal.pone.0281197] [Medline: 36719908]
33. Paulis WD, Horemans HL, Brouwer BS, Stam HJ. Excellent test-retest and inter-rater reliability for Tardieu Scale measurements with inertial sensors in elbow flexors of stroke patients. Gait Posture 2011 Feb;33(2):185-189. [doi: 10.1016/j.gaitpost.2010.10.094] [Medline: 21131203]
34. Lam T, Noonan VK, Eng JJ, SCIRE Research Team. A systematic review of functional ambulation outcome measures in spinal cord injury. Spinal Cord 2008 Apr;46(4):246-254. [doi: 10.1038/sj.sc.3102134] [Medline: 17923844]
35. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 2006 May;54(5):743-749. [doi: 10.1111/j.1532-5415.2006.00701.x] [Medline: 16696738]
36. Agustín RM, Crisostomo MJ, Sánchez-Martínez MP, Medina-Mirapeix F. Responsiveness and minimal clinically important difference of the five times sit-to-stand test in patients with stroke. Int J Environ Res Public Health 2021 Feb 26;18(5):2314. [doi: 10.3390/ijerph18052314] [Medline: 33652864]
37. Amatachaya S, Khuna L, Thaweewannakij T. Responsiveness and minimal clinically important difference of the five times sit-to-stand test in ambulatory individuals with spinal cord injury: a six-month prospective cohort study. Clin Rehabil 2023 Jan;37(1):109-118. [doi: 10.1177/02692155221122672] [Medline: 36046885]
38. Kesar TM, Binder-Macleod SA, Hicks GE, Reisman DS. Minimal detectable change for gait variables collected during treadmill walking in individuals post-stroke. Gait Posture 2011 Feb;33(2):314-317. [doi: 10.1016/j.gaitpost.2010.11.024] [Medline: 21183350]
39. Nair PM, Hornby TG, Behrman AL. Minimal detectable change for spatial and temporal measurements of gait after incomplete spinal cord injury. Top Spinal Cord Inj Rehabil 2012;18(3):273-281. [doi: 10.1310/sci1803-273] [Medline: 23459394]
40. Iosa M, Fusco A, Marchetti F, et al. The golden ratio of gait harmony: repetitive proportions of repetitive gait phases. Biomed Res Int 2013;2013:918642. [doi: 10.1155/2013/918642] [Medline: 23862161]
41. Musselman KE. Clinical significance testing in rehabilitation research: what, why, and how? Phys Ther Rev 2007;12(4):287-296. [doi: 10.1179/108331907X223128]
42. Tilson JK, Sullivan KJ, Cen SY, et al. Meaningful gait speed improvement during the first 60 days poststroke: minimal clinically important difference. Phys Ther 2010 Feb;90(2):196-208. [doi: 10.2522/ptj.20090079] [Medline: 20022995]
43. Wedege P, Steffen K, Strøm V, Opheim AI. Reliability of three-dimensional kinematic gait data in adults with spinal cord injury. J Rehabil Assist Technol Eng 2017;4:2055668317729992. [doi: 10.1177/2055668317729992] [Medline: 31186937]
44. Scrivener K, Schurr K, Sherrington C. Responsiveness of the ten-metre walk test, Step Test and Motor Assessment Scale in inpatient care after stroke. BMC Neurol 2014 Jun 16;14:129. [doi: 10.1186/1471-2377-14-129] [Medline: 24934859]
45. Pradon D, Roche N, Enette L, Zory R. Relationship between lower limb muscle strength and 6-minute walk test performance in stroke patients. J Rehabil Med 2013 Jan;45(1):105-108. [doi: 10.2340/16501977-1059] [Medline: 23095981]
46. Severinsen K, Jakobsen JK, Overgaard K, Andersen H. Normalized muscle strength, aerobic capacity, and walking performance in chronic stroke: a population-based study on the potential for endurance and resistance training. Arch Phys Med Rehabil 2011 Oct;92(10):1663-1668. [doi: 10.1016/j.apmr.2011.04.022] [Medline: 21872846]
47. Awad LN, Palmer JA, Pohlig RT, Binder-Macleod SA, Reisman DS. Walking speed and step length asymmetry modify the energy cost of walking after stroke. Neurorehabil Neural Repair 2015 Jun;29(5):416-423. [doi: 10.1177/1545968314552528] [Medline: 25288581]

XSL•FO
RenderX

48. Mackintosh SF, Hill KD, Dodd KJ, Goldie PA, Culham EG. Balance score and a history of falls in hospital predict recurrent falls in the 6 months following stroke rehabilitation. Arch Phys Med Rehabil 2006 Dec;87(12):1583-1589. [doi: 10.1016/j.apmr.2006.09.004] [Medline: 17141637]
49. Bower K, Thilarajah S, Pua YH, et al. Dynamic balance and instrumented gait variables are independent predictors of falls following stroke. J Neuroeng Rehabil 2019 Jan 7;16(1):3. [doi: 10.1186/s12984-018-0478-4] [Medline: 30612584]
50. Kumprou M, Amatachaya P, Sooknuan T, Thaweewannakij T, Amatachaya S. Is walking symmetry important for ambulatory patients with spinal cord injury? Disabil Rehabil 2018 Apr;40(7):836-841. [doi: 10.1080/09638288.2016.1277398] [Medline: 28094580]
51. England SA, Granata KP. The influence of gait speed on local dynamic stability of walking. Gait Posture 2007 Feb;25(2):172-178. [doi: 10.1016/j.gaitpost.2006.03.003] [Medline: 16621565]
52. Pennati GV, Bergling H, Carment L, Borg J, Lindberg PG, Palmcrantz S. Effects of 60 min electrostimulation with the EXOPULSE Mollii suit on objective signs of spasticity. Front Neurol 2021;12:706610. [doi: 10.3389/fneur.2021.706610] [Medline: 34721255]
53. Palmcrantz S, Pennati GV, Bergling H, Borg J. Feasibility and potential effects of using the electro-dress Mollii on spasticity and functioning in chronic stroke. J Neuroeng Rehabil 2020 Aug 10;17(1):109. [doi: 10.1186/s12984-020-00740-z] [Medline: 32778118]
54. Mukherjee A, Chakravarty A. Spasticity mechanisms - for the clinician. Front Neurol 2010;1:149. [doi: 10.3389/fneur.2010.00149] [Medline: 21206767]
55. Faist M, Mazevet D, Dietz V, Pierrot-Deseilligny E. A quantitative assessment of presynaptic inhibition of Ia afferents in spastics. Differences in hemiplegics and paraplegics. Brain 1994 Dec;117(Pt 6):1449-1455. [doi: 10.1093/brain/117.6.1449] [Medline: 7820579]
56. Jonasson LL, Sörbo A, Ertzgaard P, Sandsjö L. Patients' experiences of self-administered electrotherapy for spasticity in stroke and cerebral palsy: a qualitative study. J Rehabil Med 2022 Feb 14;54:jrm00263. [doi: 10.2340/jrm.v53.1131] [Medline: 34935050]
57. Gunnarsson E, Rödby K, Seoane F. Seamlessly integrated textile electrodes and conductive routing in a garment for electrostimulation: design, manufacturing and evaluation. Sci Rep 2023 Oct 13;13(1):17408. [doi: 10.1038/s41598-023-44622-5] [Medline: 37833440]
58. Perpetuini D, Russo EF, Cardone D, et al. Use and effectiveness of electrosuit in neurological disorders: a systematic review with clinical implications. Bioengineering (Basel) 2023 Jun 2;10(6):680. [doi: 10.3390/bioengineering10060680] [Medline: 37370612]

Abbreviations

10MWT: 10-meter walk test
EMS: Exopulse Mollii Suit
FAC: Functional Ambulatory Category
MTS: Modified Tardieu Scale
NATU Quest: Assistive Technology Usability Questionnaire for People With Neurological Diseases
NMES: neuromuscular electrical stimulation
RI: ratio index
RMS: root mean square
SCI: spinal cord injury
sEMG: surface electromyography
TENS: transcutaneous electrical nerve stimulation

Edited by S Munce; submitted 30.Jul.2025; peer-reviewed by S Huang, V Block; revised version received 21.May.2026; accepted 08.Jun.2026; published 26.Jun.2026.

Please cite as:
Yen JM, Kamsani NS, Lai HS, Tang N, Chew E
Clinical Outcomes of Individualized Electrostimulation Using a Wearable Electro Suit and Qualitative Feedback From a Mixed Cohort of Survivors of Stroke and Spinal Cord Injury With Spasticity: Case Series
JMIR Rehabil Assist Technol 2026;13:e81522
URL: https://rehab.jmir.org/2026/1/e81522
doi:10.2196/81522

XSL•FO
RenderX

© Jia Min Yen, Nur Shafawati Kamsani, Hwa Sen Lai, Ning Tang, Effie Chew. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Design Implications of Comfort and Usability of Manual Stairclimbing Wheelchair: Ergonomic Assessment and Pilot Study Using Surface Electromyography Inputs

Abhishek Verma, PhD; Rohit Kumar, M Tech; J Ramkumar, PhD

Department of Design, Indian Institute of Technology Kanpur, DJAC Building, Kanpur, India

Corresponding Author:
J Ramkumar, PhD
Department of Design, Indian Institute of Technology Kanpur, DJAC Building, Kanpur, India

Abstract

Background: Stairclimbing wheelchairs offer enhanced mobility for users navigating multilevel environments, yet limited research addresses the ergonomics of lever propulsion-based stair climbing mechanisms. Comprehensive ergonomic assessment integrating both subjective user feedback and objective biomechanical analysis is essential for optimizing assistive device design for comfort and usability.

Objective: This pilot study aims to assess the ergonomic design of a transformable stair-climbing wheelchair through a dual-methodology approach, evaluating plane surface movement accessibility and quantifying muscle activation patterns during lever-propelled stair-climbing operations using surface electromyography (sEMG).

Methods: This 2-part study involved anthropometric measurements from 20 male participants to establish design parameters using 5th and 95th percentile values. Part A assessed plane surface movement with 9 participants (7 healthy, 2 with paraplegia) navigating a simulated urban course featuring a 5° ramp, a 90° turn, and narrow passages across 3 trials. Task completion times and subjective ride easiness ratings were recorded. Part B used a Taguchi-based fractional factorial design to evaluate 3 ergonomic factors, including torso angle (λ), lever distance (L), and lever orientation (ψ), across 7 healthy participants. Maximum voluntary contraction (MVC) was measured for 4 muscles, including biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT).

Results: In Part A, the ramp and 90° turn proved most challenging due to the wheelchair's 65 kg weight and large turning radius (~1450 mm). Driving control scored highest (6/10), while comfort scored lowest due to the tilted seat design. In Part B, a straight torso (λ=0°) consistently reduced muscle strain, particularly for brachioradialis. A lever distance of approximately 50 mm and a neutral to slightly supinated orientation (ψ=0°-30°) optimized muscle effort. Interaction effects revealed high strain configurations (λ=45°; L=100 mm; ψ=−30°) exceeding 75 MVC, while optimal settings reduced strain to approximately 50 MVC.

Conclusions: Optimal ergonomic parameters of λ=0°, L=37.5 mm, and ψ=15° are recommended to minimize fatigue and enhance user comfort. Design improvements should prioritize weight reduction, compact form factor for maneuverability, and adjustable seat tilt. The modular wheelchair design permits customization for diverse user populations. Future research should include larger, gender-diverse participant groups and real-world validation studies.

(JMIR Rehabil Assist Technol 2026;13:e78965) doi:10.2196/78965

KEYWORDS

stairclimbing wheelchair; ergonomic design; surface electromyography; taguchi design; maximum voluntary contraction; likert scale

Introduction

Ergonomic analysis of any newly designed wheelchair or assistive technology intervention is crucial for assessing user comfort and efficiency using already established methodologies. Subjective evaluations, such as Likert scales and visual analog scale (VAS), are used to evaluate user perception about usability and comfort of an assistive technology. Cowan et al [1], assessed paraplegic users, using questionnaires on maneuverability, stability, and comfort, alongside time and kinetic data, providing a comprehensive view of user experience. Van der Woude et al [2], reviewed research and innovation in manual wheelchairs, focusing on rehabilitation, sports, daily life, and health, emphasizing physiological responses during propulsion. However, most studies focus on flat surfaces, with limited research specifically addressing the ergonomics of lever-propulsion–based stair-climbing, indicating a gap that this study aims to address. Dubowsky et al [3] compared kinematics, kinetics, and electromyography in able-bodied and paraplegic

XSL•FO
RenderX

users, analyzing biceps brachii and triceps brachii during propulsion, highlighting axle placement effects.

Studies on specific design variables that may affect muscle activity are crucial for design optimization. Marrow et al [4] examined the influence of grip position on upper-extremity posture and muscle activity, linking pronation, supination, and neutral positions along with torso inclination to muscle demands. Kurup et al [5] looked at push handle height, affecting the distance from the shoulder to the lever, and its impact on muscle activity. Verma et al [6] reviewed how different ergonomic and functional considerations inform manual wheelchair design, emphasizing wheelchair dimensions based on the stair-climbing mechanisms. Among all these studies based on the wheelchair functional or comfort assessments, there is an implicit absence of pilot studies on newly developed stair-climbing technologies that use electromyography to enhance comfort and usability, highlighting the role of electromyography in designing for reduced muscle strain. Such a type of approach has been carried out in this research to discuss implications of integrating electromyography data into design processes for minimizing muscle strain and improving overall user experience in the case of stair-climbing operations.

The ergonomic evaluation of stair-climbing wheelchair functions is necessary for further iterations into the wheelchair design developed by Verma et al [7]. For an assistive technology such as this stair-climbing wheelchair, the goal is to offer assistance in activities of daily living such as mobility, nutrition, personal care, and employment and also act as an enabler of social interactions such as family or relationship responsibilities and interpersonal or community relationships, etc. Both subjective and objective information about the effectiveness of the wheelchair is required for improvement in this design of the stair-climbing wheelchair for both plane surface movement and stair-climbing function. This research focuses on understanding the wheelchair use case for a short duration. Any long-term consequences of the wheelchair design require a more detailed study spanning over years and are out of the scope of this research. This research can be broadly divided into 2 parts, the first one to collect and assess wheelchair mobility on a plane surface when the wheelchair is transformed into a plane-surface wheelchair, and the second part addresses the ergonomics of stair-climbing when the participant performs the pull-push action through a set of levers to rotate stair-climbing wheel.

This pilot study uses an integrated dual-methodology approach to comprehensively evaluate the stair-climbing wheelchair design. Part A focuses on ergonomic assessment of plane surface movement through anthropometric measurements, functional mobility testing, and subjective user feedback. Part B complements this assessment through objective biomechanical analysis using surface electromyography (sEMG) to quantify muscle activation patterns during stair-climbing operations. These 2 methodological components are mutually reinforcing; the subjective assessments from Part A identify user-perceived comfort and usability issues, while the objective sEMG measurements from Part B provide physiological validation of these subjective experiences and reveal underlying biomechanical factors that may not be immediately apparent to users. This integrated approach enables both immediate user feedback and deeper understanding of the physiological demands imposed by specific design parameters, facilitating evidence-based design optimization that addresses both perceived comfort and actual physical strain. The combination of these methodologies provides a comprehensive framework for assistive technology assessment that can inform future design iterations and clinical recommendations.

Methods

Study Design Overview

This 2-part pilot study used a dual-methodology approach to evaluate the ergonomic design of a transformable stair-climbing wheelchair. Part A assessed plane surface movement accessibility through anthropometric measurements, functional mobility testing on a simulated urban course, and subjective user feedback. Part B evaluated lever propulsion ergonomics during stair-climbing using sEMG to quantify muscle activation patterns under varying ergonomic configurations determined by a Taguchi-based fractional factorial design of experiments.

Part A: Ergonomic Assessment of Plane Surface Movement

Participants and Anthropometric Measurements

Anthropometric measurements were collected from 20 male participants (average age 27.5, SD 2.76 years) to establish design parameters for the modular wheelchair. This demographic was selected to determine conservative upper-bound dimensions, recognizing that male participants generally have larger anthropometric dimensions. Measurements were taken with the user seated in an optimal position as shown in Figure 1. The dimensions collected included hip width, buttock-popliteal length, popliteal height, subscapular height, shoulder height, elbow height, knee height, shoulder width, elbow fingertip length, upper limb length, and shoulder grip length. These measurements were used to develop a fit and reach matrix according to user comfort, with 5th and 95th percentile values calculated for design purposes (Table 1).

XSL•FO
RenderX

Figure 1. Anthropometric measurements are required to design wheelchair seating and reach dimensions.

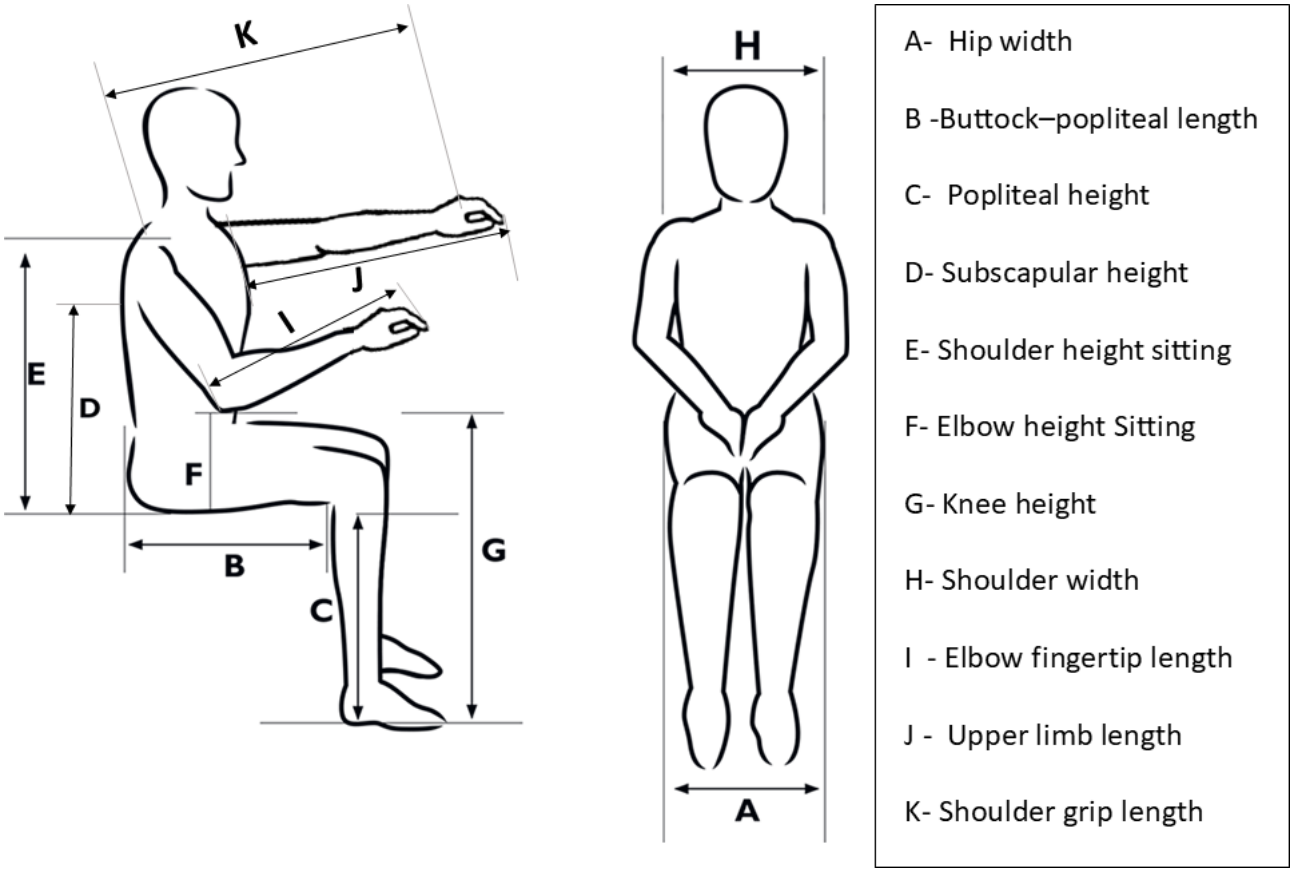


Table . Anthropometric data recorded for wheelchair design (dimensions in centimeters).

Annotation	Anthropometric feature	Mean (SD)	5th -95th percentile
—[a]	Age (years)	27.5 (2.76)	23.95-31.1
—	Height (cm)	173.25 (3.46)	168-178
—	Weight (kg)	71.75 (4.98)	63.95-78.15
A	Hip width (cm)	38.45 (3.26)	34-42.15
B	Buttock-Popliteal length	51 (4.41)	42.95-56.05
C	Popliteal height (cm)	42.1 (2.67)	38.95-47.05
D	Subscapular height (cm)	45.75 (2.22)	43-50.05
E	Shoulder height (cm)	60.85 (3.28)	54.95-60.85
F	Elbow height (cm)	24.85 (2.13)	22-28.05
G	Knee height (cm)	53.95 (4.50)	48.95-60.15
H	Shoulder width (cm)	55.45 (4.5)	50.85-63.15
I	Elbow fingertip length (cm)	36.35 (2.75)	33-41.05
J	Upper limb length (cm)	85.4 (4.28)	78.95-92.1
K	Shoulder grip length (cm)	75.8 (5.01)	67.95-82.15

[a]Not applicable.

XSL•FO
RenderX

Functional Anthropometry for Reach and Clearances

The wheelchair design dimensions were based on standards and guidelines from the Central Public Works Department (CPWD), the Office of Chief Commissioner for Persons with Disabilities, IS 4963 (building code standard), and IS 7454 (wheelchair specification standard). The clearance dimensions are summarized in Table 2.

Table . Clearance dimensions for the wheelchair based on standards and guidelines (in mm).

Reach ranges	Standards and guidelines			
	IS-7454[a] [8]	CPWD[b] [9]	CCD[c] [10]	ISO[d] [11]
Forward high	1350 - 1600	1200	1200	1200
Forward low	Not specified	400	400	400
Forward obstructed high	715 - 830	1100	1100	Not specified
Lateral high	1350 - 1770	1300	1300	Not specified
Lateral low	Not specified	250	250	Not specified
Lateral obstructed high	Not specified	1200	1200	Not specified
Maneuvering clearances				
Circular diameter-360° turn	1500	1500 - 2000	1800	1500

[a]IS: Indian Standard in architecture.

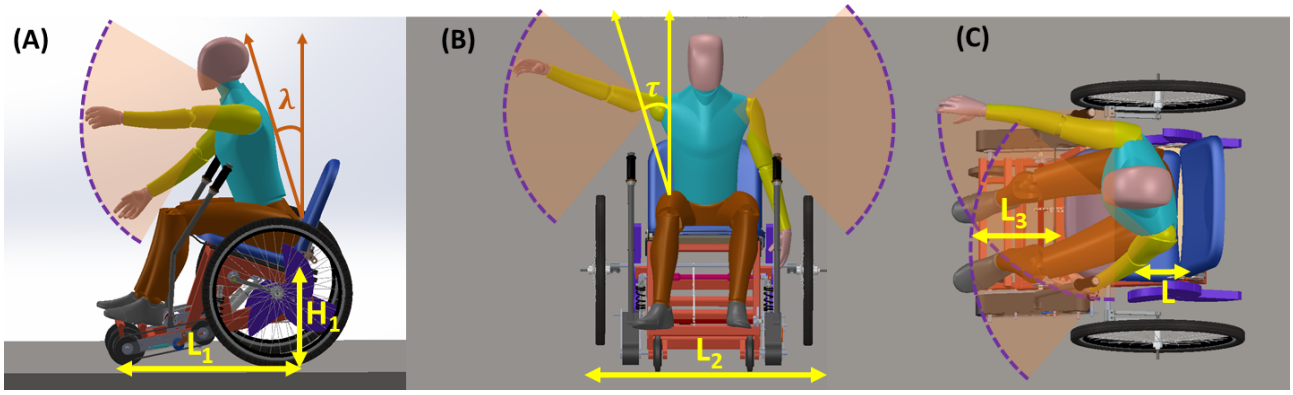
[b]CPWD: Central Public Works Department.

[c]CCD: Chief Commissioner-Disability.

[d]ISO: International Standard Organization.

The wheelchair dimensions for clearances and reach were based on the maximum of maximum (95th percentile) and minimum of minimum (5th percentile) human anthropometric dimensions, respectively. Wheelchair dimensions such as seat length and backrest angle were designed to be adjustable over the 5th and 95th percentile range using an angle-adjustable backrest and a slide-lock mechanism. The reach space design was based on wheelchair height (H1) for vertical reaches and length L1 as shown in Figure 2A-2B, width L2 (Figure 2B), and distance L3 (Figure 2C) for horizontal distances, calculated using a digital human model (DHM).

Figure 2. Maximum reach illustrated by a digital human model in (A) sagittal plane, (B) coronal plane, and (C) transverse plane.



The primary reach relationships derived from the DHM are given by equations (1–4):

Forward high: (1)

$$H1+E\times\cos\lambda+K\times\sin(\pi/4)\geq1200mm$$

Forward low: (2)

$$H1+E\times\cos\lambda-K\times\sin(\pi/4)\geq400mm$$

Lateral high: (3)

$$H1+E+K\times\sin(\pi/4)\geq1300mm$$

Lateral low: (4)

$$H1+E\times\sin(\tau)-K\times\sin(\pi/4)\geq250mm$$

Where H1 is the height of the wheelchair from the ground to the lowest point on the seat (500 mm), E is the shoulder height from the seat, λ is the torso angle the participants make with the coronal plane, K is the shoulder grip length, and τ is the angle the torso makes with the sagittal plane.

The minimum turning distance required for a 360° turn was calculated using Equation (5):

$$Tcircular=2\times\pi\times L12+L22 \quad (5)$$

Where L12+L22 is the turning radius of the wheelchair

For this prototype, L1=600 mm and L2=950 mm

XSL•FO
RenderX

Therefore, the distance required for a 360° turn calculated using DHM will be,

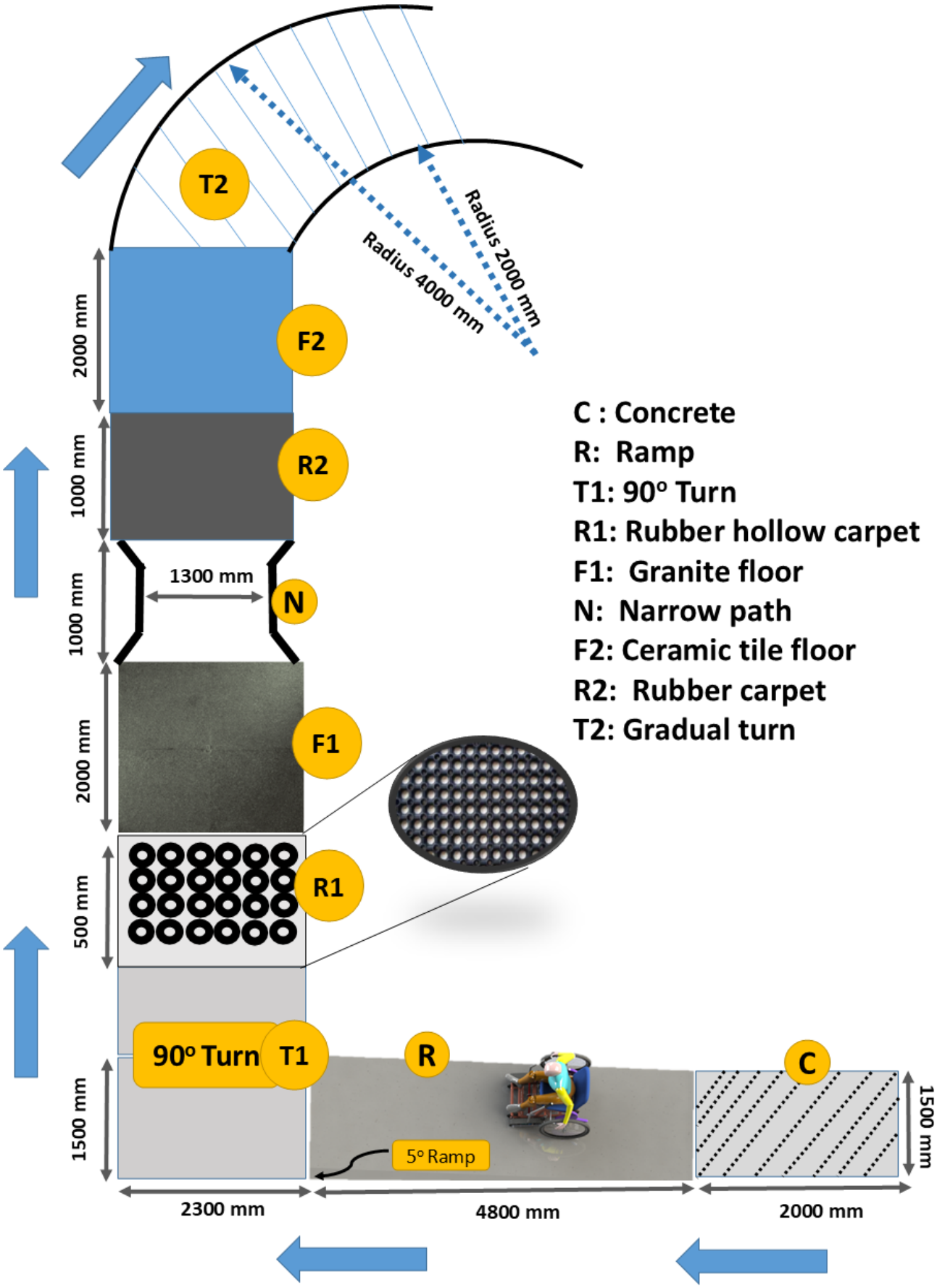
$T_{circular}$=4870 mm.

Simulated Course Design and Protocol

The manual stair-climbing wheelchair is a dual-purpose wheelchair designed to provide multidimensional accessibility on both stairs and plane surfaces. To test the efficacy of wheelchair driving on plane surfaces, a simulated course was designed to test maneuverability and control. The simulated course contained architectural hurdles that a wheelchair user faces in an urban setting, including a 5° ramp followed by a sharp 90° turn and then converging and diverging spaces. The course is illustrated in Figure 3.

XSL•FO
RenderX

Figure 3. Wheelchair course designed for ergonomic design study of plane surface movement configuration of wheelchair.



This pilot study used n=9 participants for plane surface assessment (7 healthy, 2 with paraplegia), following established guidelines for feasibility studies in assistive technology research [12,13]. Two participants with paraplegia provided critical

XSL•FO
RenderX

real-world user validation. The course was designed as a controlled experimental environment to ensure participant safety and enable systematic data collection, including multiple stair configurations with varying heights (15 - 20 cm) and depths (25 - 30 cm), curved pathways, simulated obstacles, and various surface transitions representative of typical urban environments.

The subjective descriptors used for the response for each riding task using Likert scales were (1) strongly favorable (score=4), (2) moderately favorable (score =3), (3) moderately unfavorable (score=2), and (4) strongly unfavorable (score=1). Task completion times were recorded for each of the 9 tasks across 3 trials. After the completion of all tasks, a subjective survey of the whole task was taken on a visual satisfaction scale ranging from 0 (Very negative) to 10 (Very positive) for overall ride comfort, stability, driving control, perception, and maneuverability.

Part B: Lever Propulsion sEMG Study

Muscles Under Study

The biomechanical evaluation of arm and shoulder muscles was performed to identify and mitigate shoulder pain and fatigue generated during lever propulsion. Use of sEMG is a prominent method for recording muscle activation artifacts to understand muscle effort, muscle activation timing, and muscle fatigue. The activation timing study of a muscle group was performed to understand the loading profile of muscle at a specific interval in the work cycle [14]. Surface electromyography also helps in identifying muscles under fatigue by observing an increase in amplitude and decrease in median frequency of the bioelectrical signal [15].

The upper limb and shoulder are the most common sites for any possible stress injury or musculoskeletal disorder due to long-term cyclic loading during lever-based propulsion [16,17]. The muscle groups under study for this specific research go under cyclic dynamic interactions and muscle fiber contractions during the lever-based propulsion cycle and are anisometric (concentric and eccentric) in nature. Therefore, the shoulder and arm joint angles are also required to be recorded while recording sEMG activity. The recorded electromyography signal has artifacts from the velocity- and angle-dependent muscle excitation of the shoulder and elbow extensors and flexors. The muscle activation of the extensor and flexor groups will be significant during both actions of pulling and pushing. But, since only pulling stroke is responsible for stair climbing, only those muscles will show significant dynamic electromyography activity that are associated with pulling action (elbow flexors and shoulder extensors muscle groups).

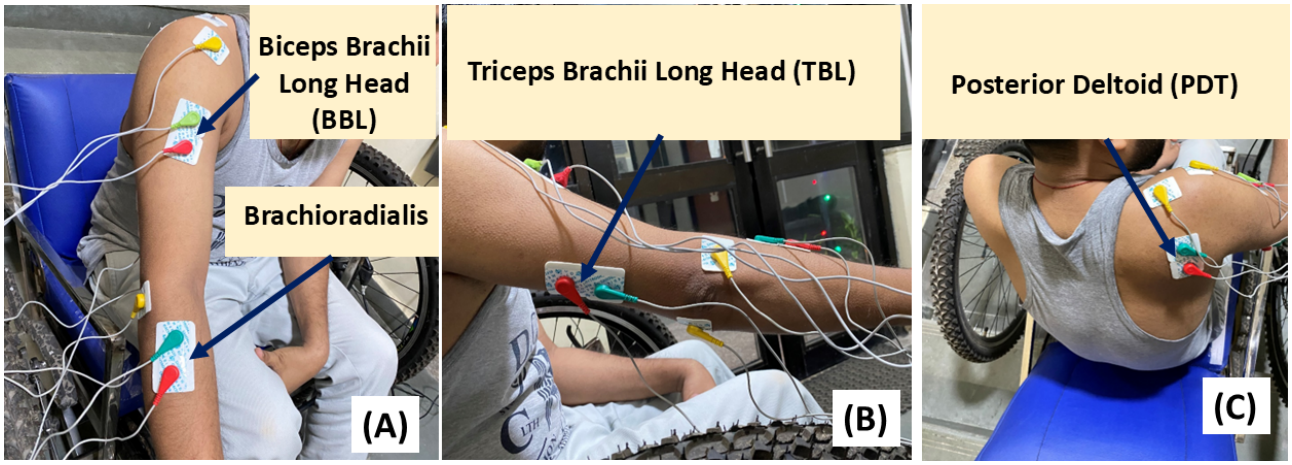
The action of lever propulsion in wheelchairs and its physiological effects are reported by only a few researchers, as the wheelchairs developed with a lever mechanism are very much limited, and in consequence, electromyography-based ergonomic evaluations based on them. There are few studies about lever-based propulsion effort, but they are limited to physiological measurements such as oxygen uptake, respiration ratio, minute ventilation, heart rate, etc [18-20]. With the advent of surface electromyography recording and processing techniques, the idea of using electromyography as a dominant physiological parameter that can influence design is increasing. Push-rim–based propulsion of a plane-surface wheelchair has been highly explored using surface electromyography studies on dominant-side muscle groups such as the deltoid group, triceps group, and bicep groups. As reported in the literature, there are some overlapping muscle groups engaged during both lever propulsion and hand-rim propulsion, such as biceps brachii, posterior deltoid, triceps, pectoralis major, etc [21-23]

The action of lever propulsion is a push-pull action that requires elbow (flexion and extension) and shoulder (flexion and extension) movement. The action of pulling requires flexion of the elbow and extension of the shoulder muscles and vice versa for pushing. The pulling stroke is responsible for rotating the stair-climbing wheel on the staircase. The pushing stroke (moving the lever away from the chest) is idle and requires less effort compared to pulling. Tomiak et al [24], in their research related to bimanual rowing movement, which mimics a lever push-pull activity, group together the pulling and pushing (returning) muscle groups as elbow and shoulder muscles. The primary goal of the ergonomic evaluation in this section is to identify extreme joint loading forces by trying different posture configurations defined by ergonomic factors. Pushing and pulling tasks can cause extreme lower back and upper joint loading, leading to musculoskeletal disorders (MSDs) [25,26].

Four muscle groups were selected for sEMG recording: (1) biceps brachii long head (BBL), a primary muscle involved in flexing the elbow, selected due to availability of higher density of motor neurons (Figure 4A); (2) triceps brachii long head (TBL), located on the back of the upper arm, responsible for extending the elbow during the pushing stroke (Figure 4B); (3) brachioradialis, responsible for forearm and elbow flexion necessary for pulling, with secondary function of supination. and pronation for holding the lever (Figure 4B); and (4) posterior deltoid (PDT), located at the shoulder, extends during shoulder extension required for pulling the lever toward the chest (Figure 4C).

XSL•FO
RenderX

Figure 4. Nickel-chromium (Ni-Cr) alloy electrodes stuck on muscle surface using silver-silver chloride electrode (Ag-Cl) hydrogel (A) location of brachioradialis and BBL muscle group recording, (B) triceps brachii long head (TBL), and (C) posterior deltoid (PDT) muscle group recording.

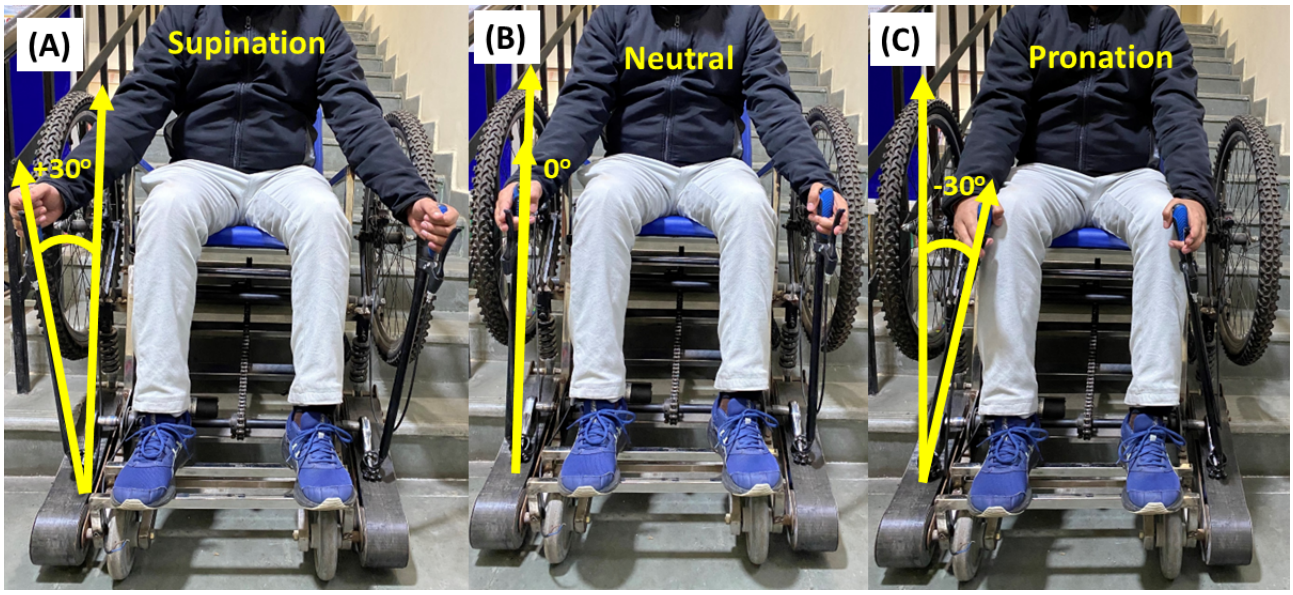


Ergonomic Factors

Three ergonomic factors were evaluated in this study:

(1) Lever orientation (ψ): the range of motion for supination and pronation was limited to ±30°, where supination is +30° (Figure 5A), the neutral position is 0° (Figure 5B), and pronation is −30° (Figure 5C). Although Sullivan et al [27] reported that up to a certain degree of pronation, subjective discomfort is lesser and higher effort can be achieved compared to supination and neutral hand orientation [28], objective evaluation using electromyography was required.

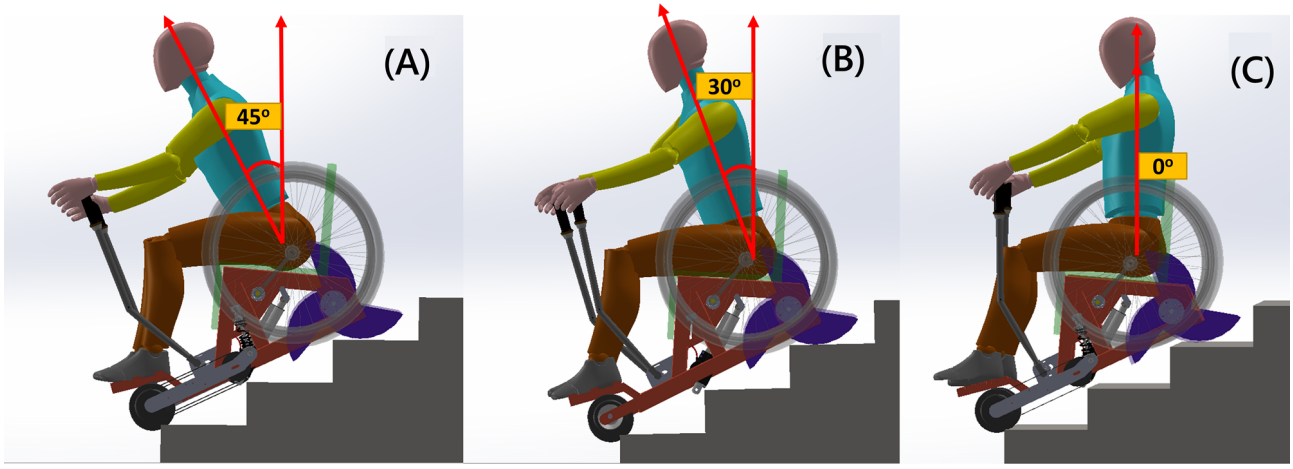
Figure 5. Lever orientation (ψ) as an ergonomic factor to study the effect of (A) supination, (B) neutral, and (C) pronation on muscle activation.



(2) Torso angle (λ): 3 trunk orientations in the coronal plane were studied [29,30], including 45° (Figure 6A), 30° (Figure 6B), and 0° upright (Figure 6C). The maximum low reach can be achieved at 45°, while maximum stability and minimum spinal loading can be attained at 0°.

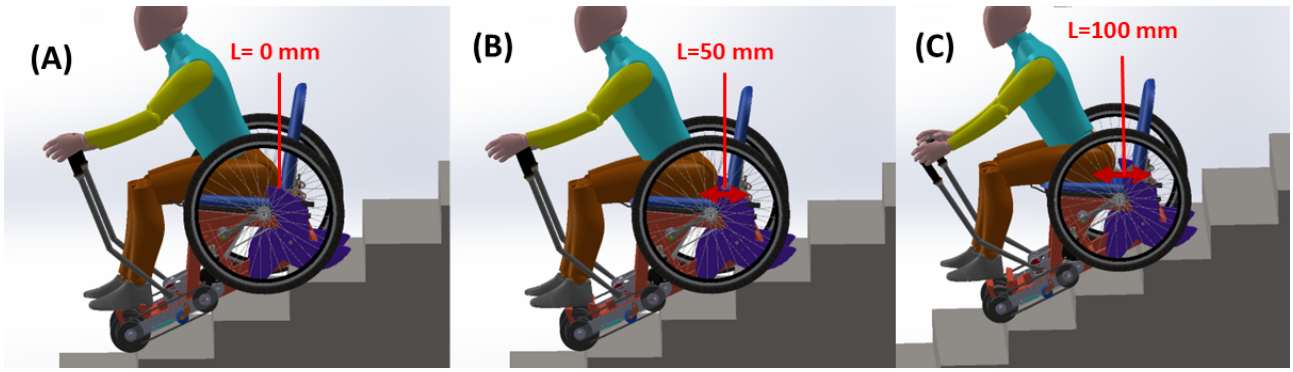
XSL•FO
RenderX

Figure 6. Torso angle made with vertical coronal plane (A) 45°, (B) 30°, and (C) 0°.



(3) Lever distance (L): the distance from the shoulder to the lever (shoulder grip length+L) was varied using a novel seat adjustment mechanism. The distance L was calculated from the lower back of the seated participant to the backrest of the wheelchair. Three levels were tested, including L=0 mm (Figure 7A), L=50 mm (Figure 7B), and L=100 mm (Figure 7C).

Figure 7. Ergonomic factor lever distance L for (A) L = 0, (B) L = 50 mm, and (C) L = 100 mm.



Participants and Experimental Design

Seven healthy participants (S1–S7) were selected for the surface electromyographic study. The anthropometric data of the participants is shown in Table 3.

Table . Anthropometric characteristics of participants in the surface electromyography (sEMG) study.

Participants	Age (years)	Upper arm length (cm)	Lower arm length (cm)	BMI
S1	28	32	24	22.6
S2	33	34	23	26
S3	30	34	26	21
S4	31	36	28	20.7
S5	31	36	27	22.5
S6	33	32	24	24
S7	31	37	29	25

A 3-factor, 3-level experimental design was used. The factors and their levels are shown in Table 4. A full factorial design would require 27 runs; this was reduced to a 1/3 fractional factorial design using the Taguchi method, requiring 9 experimental runs (Table 5). Each participant performed the lever push-pull action while seated in the configuration specified by each experiment.

XSL•FO
RenderX

Table . Factors and their respective levels for identifying ergonomic seating.

Factors	Level 1	Level 2	Level 3
Torso angle (λ)	45°	30°	0°
Lever distance (L)	0 mm	50 mm	100 mm
Lever orientation (ψ)	$–30^o$	0°	$+30^o$

Table . 1/3 fractional factorial design of experiment (Taguchi) for 3 factors, 3 levels.

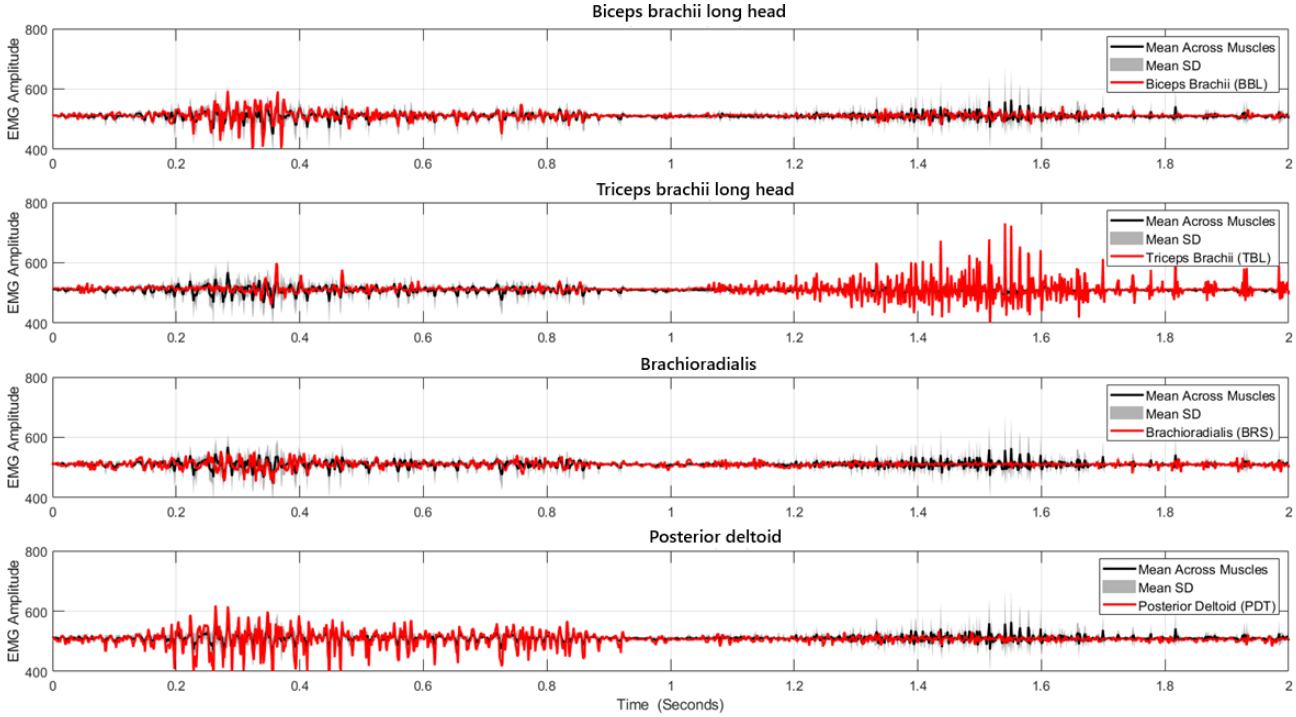
Experiment number	Torso angle level (λ)	Lever distance level (L; in mm)	Lever orientation level (ψ)
E1	45°	0	$–30^o$
E2	45°	50	0°
E3	45°	100	$+30^o$
E4	30°	0	0°
E5	30°	50	$+30^o$
E6	30°	100	$–30^o$
E7	0°	0	$+30^o$
E8	0°	50	$–30^o$
E9	0°	100	0°

Instruments and Electromyography Data Recording

The test wheelchair used in this study is a manually propelled stair-climbing wheelchair with a ratchet-based lever mechanism that provides a to-and-fro motion where pulling toward the chest engages the lever with the propulsion shaft, providing rotational motion to the stair-climbing wheel. For recording sEMG data, an ADS1299-based evaluation board (Texas Instruments) was used, providing 8 analog channels with a common reference, as reported in earlier sEMG studies [31-33]. These 8 channels were configured to create 4 channels with a common-mode rejection ratio for high-quality sEMG data for all 4 muscle groups as shown in Figure 8. An instrumentation gain of 11X and an overall system gain of 2420X were applied. Ultra-low impedance (<100 ohms) nickel-chromium (Ni-Cr) alloy electrodes were used with silver-silver chloride electrode (Ag-Cl) adhesive solid hydrogel at the electrode-skin interface to reduce skin impedance. The sampling frequency was 500 Hz.

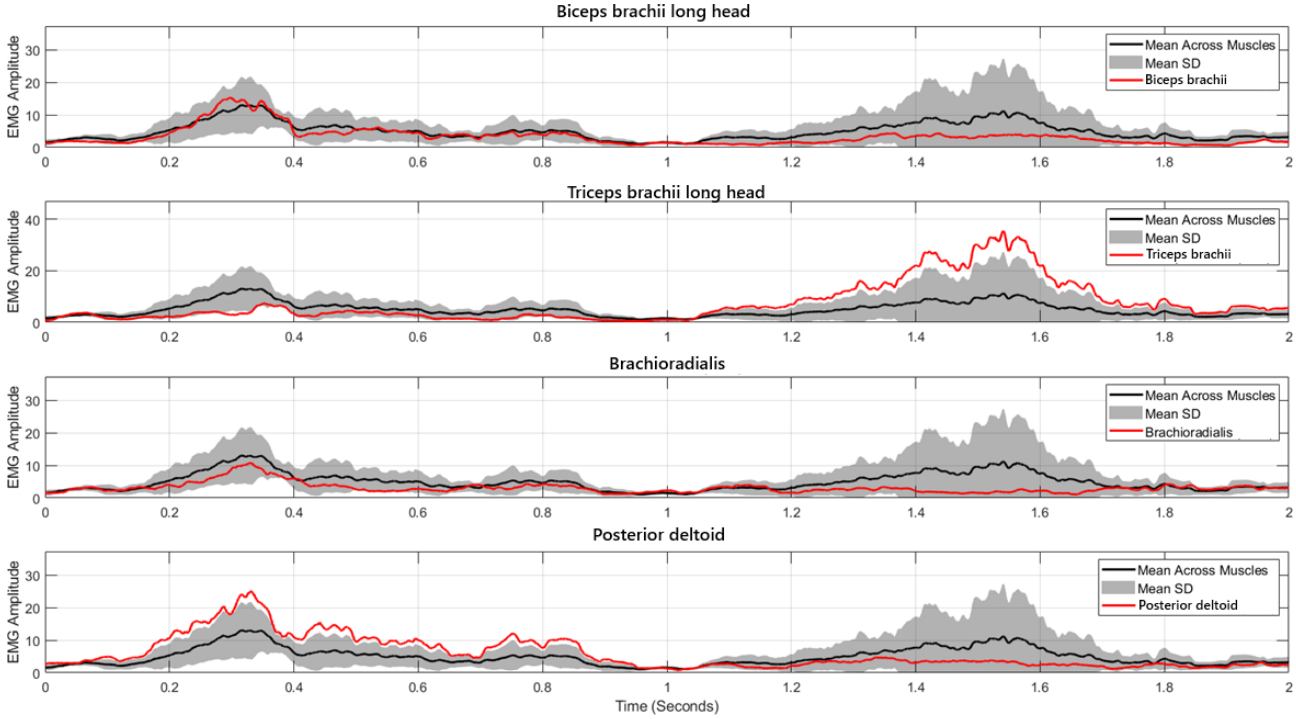
XSL•FO
RenderX

Figure 8. Surface electromyography (sEMG) data recorded on an ADS1299-based board for muscle groups biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT) for participant S1 during experiment E1.



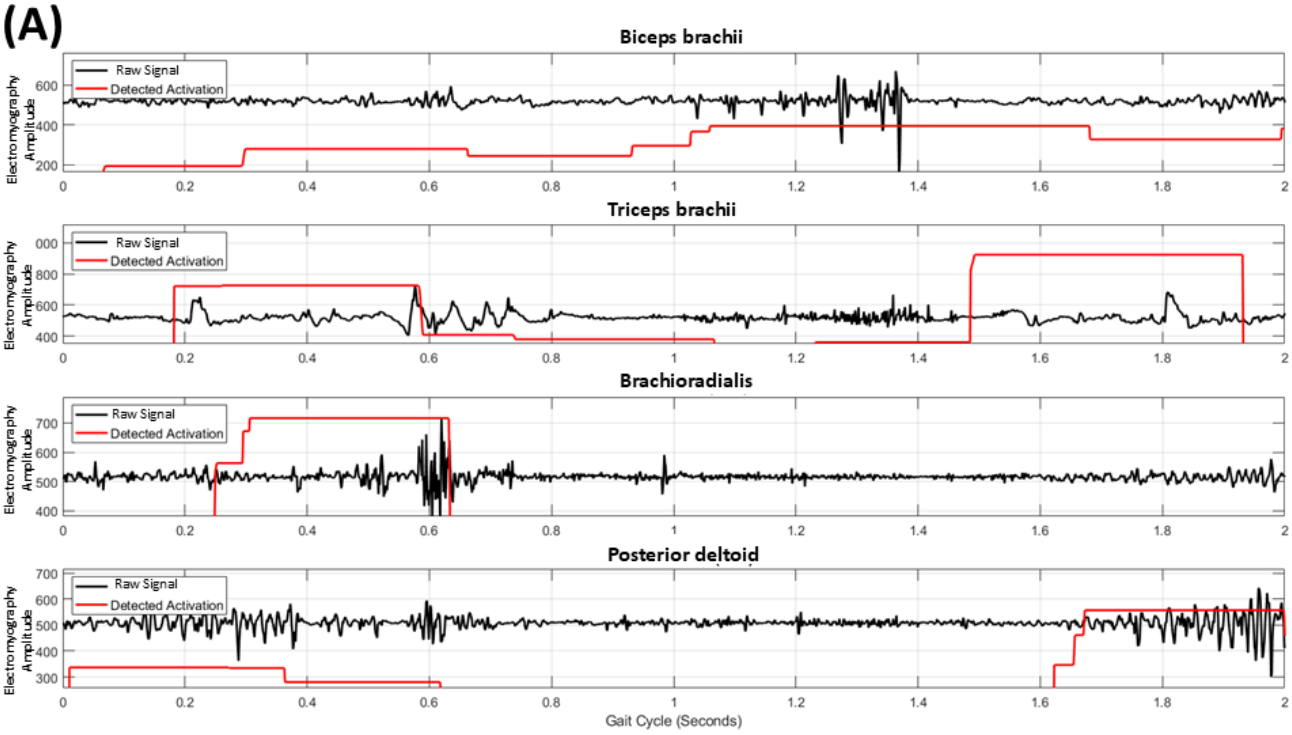
The raw electromyography signal was filtered using a band-pass Butterworth Infinite Impulse Response filter. The filtered electromyography signal was further processed by taking the absolute value and applying a moving average filter (window of 50 samples) as shown in Figure 9. For identification of muscle activation timing based on adaptive threshold, the Hilbert transform was used with a threshold duration of 25 samples to avoid influence of noise or other involuntary artifacts visible in Figure 10A. From the activation timing study, an activation window for the pull-push cycle is illustrated in Figure 10B.

Figure 9. Digitally filtered data for muscle groups BBL, TBL, brachioradialis, and PDT for participant S1 during experiment E1, showing individual activation of muscles following activation timing patterns of pushing and pulling.



XSL•FO
RenderX

Figure 10. Muscle activation detection and its representation (A) Detection of muscle activation timing and amplitude threshold on a digitally processed electromyography signal (participant S1, experiment E8) using the Hilbert transform. (B) Muscle activation window in pull-push cycle of lever propulsion. BBL: biceps brachii long head; PDT: posterior deltoid; TBL: triceps brachii long head.



Electromyography Data Processing: Maximum Voluntary Contraction

Maximum voluntary contraction (MVC) was calculated from the digitally filtered electromyography envelope with absolute values as a quantifiable parameter for normalizing the electromyography signal. The weighted average of envelope amplitude and envelope root-mean square (RMS) was used to calculate MVC [34], expressed as:

$$MVC = w1 \times A + w2 \times RMS \quad (6)$$

Where,A is the maximum amplitude of the electromyography envelope given by

$$A = max(X) - min(X) \quad (7)$$

X={x1,x2,x3……xn}; X is the dataset of filtered electromyography signal recorded over time t.

x_i represents the amplitude of the electromyography signal at i-th time point.

RMS means root mean square of dataset X.

Following Merletti and Parker (2004) [34], the weights were set as w1=0.7 and w2=0.3, prioritizing sensitivity to maximum activation for detecting full recruitment peaks while complementing it with RMS stability to mitigate outlier effects.

Muscle coordination patterns were assessed using the Pearson correlation coefficient “r” to measure the linear relationship between MVC values of muscle pairs [35]. The coefficient is given by:

$$r = \sum i=1n(xi-x-)(yi-y-) \sum i=1n(xi-x-)2 \sum i=1n(yi-y-)2$$

where x_i and y_i are the MVC values for 2 muscles (eg, anterior deltoid and biceps brachii) across all observations, and x- and y- are their means.

XSL•FO
RenderX

Main Effect and Interaction Effect Analysis

Quadratic regression models were fitted for each ergonomic factor–muscle response relationship to characterize the design space. Main effects and interaction effects (AB: λ×L; AC: λ×ψ; BC: L×ψ) were computed by averaging MVC values across paired factor combinations. Interaction effects were visualized using line plots and heatmaps. 3D surface plots were generated to delineate the ergonomic design space. Statistical significance of correlations between anthropometric dimensions and task performance was assessed at $P<.05$.

Ethical Considerations

This study was approved by the Institutional Ethics Committee (IEC) of the Indian Institute of Technology Kanpur (IEC Communication Number: IITK/IEC/2024-25/I/29, Date of Approval: June 20, 2024). The study followed the protocol titled "Design and Development of a multi-functional assistive device for restricted Mobility and Rehabilitation." All procedures involving human participants adhered to the institutional ethical standards and the 1964 Helsinki Declaration and its later amendments or comparable standards. Written informed consent was obtained from all participants, who were informed of objectives, procedures, risks, and their right to withdraw without consequences. No identifying information is included here; data were anonymized and securely stored per institutional and national privacy guidelines. Participants received no financial compensation. For those with paraplegia, extra precautions ensured comfort and safety during testing.

Results

Part A: Plane Surface Movement Results

Task Completion Times

The time taken for each of the 9 participants to complete each task on the simulated course is presented in Table 6.

Table . Time taken for each participant to complete the task (in seconds).

Task	Participants								
	S1	S2	S3	S4	S5	S6	S7	S8	S9
C[a]	5	4	3.5	3	3.5	4.2	4	3.5	5.3
R[b]	12.7	14.5	15	12.5	14	18	13.8	16.4	16.2
T1[c]	21.4	23	25.4	19.5	20.5	24	20	19	22
F1[d]	6	7.2	8.1	6.3	5	4.5	4.3	4	5.5
R1[e]	1.9	2.5	2	2	2	2	2	2.3	2.8
N[f]	10.4	9.8	6	8	7	9.5	6.5	8.3	9.9
F2[g]	2	2	2	2.8	2	2.5	2	2.2	2.3
R2[h]	5.5	3	3.4	4	4	4	5	3.2	4.4
T2[i]	9.2	7	6	7.3	8.2	9	8.5	7	7

[a]C: concrete.
[b]R: ramp.
[c]T1: 90° turn
[d]F1: granite floor.
[e]R1: rubber hollow carpet.
[f]N: narrow path.
[g]F2: ceramic tile floor.
[h]R2: rubber carpet.
[i]T2: gradual turn.

The ramp (R) and sharp 90° turn (T1) emerged as the most time-intensive tasks. The ramp required an average of approximately 14.8 seconds, while the 90° turn required approximately 21.6 seconds. Tasks on smooth floor surfaces (F1 and F2) and rubber carpets (R1 and R2) required relatively less time, averaging 2 - 5 seconds each.

The time required to complete the 9 standardized wheelchair driving tasks, along with participant favorability ratings, is summarized below. All times are reported as mean (SE) in seconds.

Task C (straight corridor): Participants completed the task in 4.1 ± 0.9 s and rated it as moderately favorable. Task R (right turn): 14.8 (1.8)s, rated as moderately unfavorable. Task T1 (tight turn 1): 22.3 (2.1) s, rated as strongly unfavorable. Task F1 (forward maneuver 1): 5.7 (1.2) s, rated as moderately unfavorable. Task R1 (reverse 1): 2.4 (0.4)s, rated as moderately favorable. Task N (narrow passage): 8.5 (1.5) s, rated as strongly favorable. Task F2 (forward maneuver 2): 2.5 (0.5) s, rated as moderately favorable. Task R2 (reverse 2): 4.1 (0.8) s, rated as moderately favorable. Task T2 (tight turn 2): 7.8 (1.1) s, rated as moderately favorable.

XSL•FO
RenderX

These results indicate that tasks involving tight maneuvers (particularly T1) required substantially more time and received the least favorable ratings, whereas simpler forward and reverse tasks were completed quickly and viewed more positively. The pattern highlights the importance of adequate space and clear pathways in wheelchair-accessible environments.

Subjective Ratings and Overall Assessment

In the overall subjective response study for all 9 participants evaluating wheelchair attributes on a visual analog scale from 1 to 10, overall ride comfort received the lowest mean rating of 3.33 (SD 1.58), indicating areas for potential improvement in user experience during mobility. Wheelchair stability was rated moderately higher at a mean of 5.44 (SD 1.74), while driving control emerged as the strongest attribute with the highest average score of 6.00 (SD 1.22), suggesting effective handling and responsiveness in the device's design. Perception of the wheelchair scored a mean of 3.89 (SD 1.17), reflecting possibly mixed sensory feedback, whereas maneuverability—potentially encompassing response aspects—achieved a solid mean of 5.11 (SD 1.17), highlighting reliable navigation capabilities as depicted in the bar graph with error bars representing SDs. The correlation between task difficulty (higher time and low level of easiness) and anthropometric dimensions of the participants was not statistically significant (P=.04), indicating that task difficulty could be directly associated with wheelchair design variables rather than participant characteristics.

Part B: sEMG and MVC Results

Electromyography Amplitude and MVC Values

The average electromyography amplitude and RMS of all 7 partiticipants for each muscle group, along with calculated MVC values, are presented in Table 7. There was no significant statistical correlation between the anthropometric features and electromyography activity in terms of MVC (P=.04). The maximum electromyography amplitude for each muscle averaged from all 9 experiments is illustrated in Figure 12A. Muscle coordination patterns derived from MVC data using the Pearson correlation coefficient are shown in Figure 12B.

Table . Average electromyography amplitude and root mean square (RMS) of all 7 participants for each muscle group along with their calculated maximum voluntary contraction (MVC) values.

Experiment	Electromyography average amplitude				Electromyography root mean square				MVC			
	BBL[a]	TBL[b]	Brachio-radialis	PDT[c]	BBL	TBL	Brachio-radialis	PDT	BBL	TBL	Brachio-radialis	PDT
E1	69.3	94.4	149.6	132.3	7.8	12.2	22.3	17.4	50.85	69.74	111.41	97.83
E2	55.5	74.8	179.0	95.0	7.3	10.5	13.24	12.4	41.04	55.51	129.272	70.22
E3	56.8	105.2	136.1	88.3	9	11.9	14.3	11.5	42.46	77.21	99.56	65.26
E4	52.9	75.1	128.1	82.6	8.3	12.0	13.0	10.0	39.52	56.17	93.57	60.82
E5	59.0	76.5	129.8	83.9	8.6	10.3	15.0	12.0	43.88	56.64	95.36	62.33
E6	56.7	102.5	130.9	121.3	9.0	13.9	16.8	19.0	42.39	75.92	96.67	90.61
E7	47.8	72.9	85.6	77.6	8.6	11.5	11.6	10.1	36.04	54.48	63.4	57.35
E8	48.7	55.8	75.1	62.1	8.0	9.0	11.3	10.0	36.49	41.76	55.96	46.47
E9	56.6	64.0	132.4	64.0	10.8	9.5	14.8	9.7	42.86	47.65	97.12	47.71

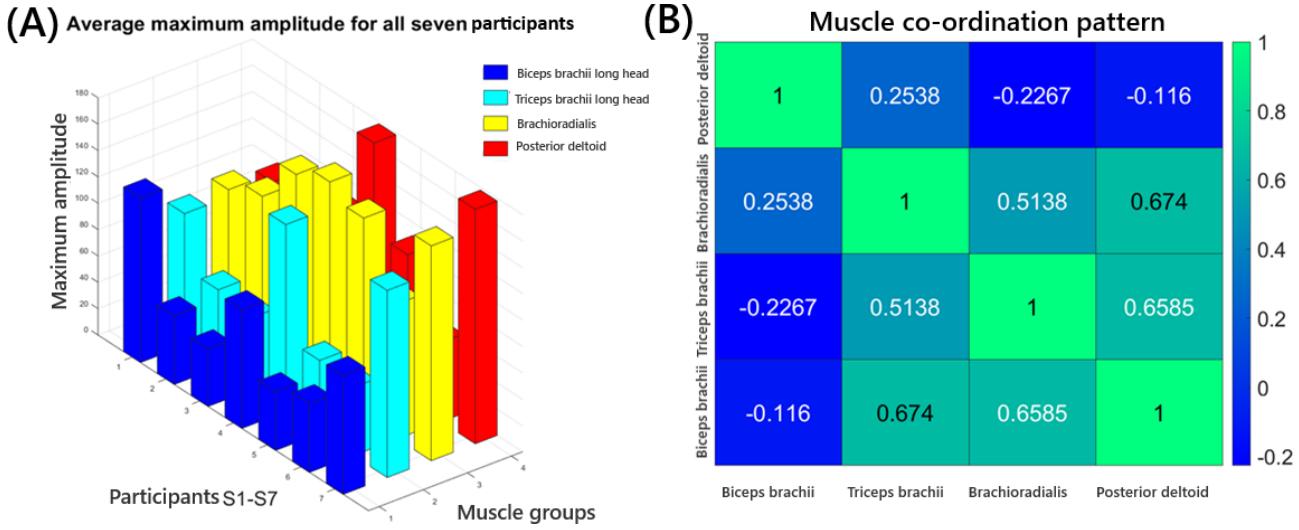
[a]BBL: biceps brachii long head.

[b]TBL: triceps brachii long head.

[c]PDT: posterior deltoid.

XSL•FO
RenderX

Figure 11. Electromyography voltage output for all 4 muscle groups of 7 participants. (A) Average maximum amplitude of electromyography signal at different muscle groups for all 7 participants (S1-S7) calculated from all 9 experiments (E1-E9). (B) Muscle coordination pattern calculated using MVC values. BBL: biceps brachii long head; PDT: posterior deltoid; TBL: triceps brachii long head.



Quadratic Regression Analysis

Quadratic regression models were fitted for each factor-muscle relationship. The regression equations, residual sum of squares, and coefficients of determination are presented in Table 8.

Table . Quadratic regression analysis and its parameters for muscle group response and ergonomic factors.

Relation	Equation	Residual sum of squares	Coefficient of determination (%)
BBL[a]-lever distance	BBL=42.14 - 0.0710×L+.000753×L^2	4.96	4.8
BBL-torso angle	BBL=38.46+.0658× +0.001659× ^{2}	3.98	38.7
BBL-lever orientation	BBL=41.14 - 0.04083×ψ+0.000976×$ψ^2$	4.91	6.8
TBL[b]-lever distance	TBL=60.13 - 0.4210×L+.004890×L^2	11.79	30.6
TBL-torso angle	TBL=47.96+.6270× −0.00429× ^{2}	9.81	52.0
TBL- lever orientation	TBL=53.11+.0051×ψ+0.01057×ψ2	13.05	15.1
Brachioradialis- lever distance	Brachioradi-alis=89.46+.0796×L+.000036×L^2	25.41	2.6
Brachioradialis-torso angle	Brachioradi-alis=72.16+.470× +0.00992× ^{2}	15.35	64.5
Brachioradialis-lever orientation	Brachioradialis=106.7 - 0.0318×ψ −0.02177×ψ2	23.11	19.4
PDT[c]-lever distance	PDT=72.00 - 0.4517×L+.004103×L^2	19.27	9.6
PDT-torso angle	PDT=50.51+.863× −0.00571× ^{2}	14.43	49.3
PDT-lever orientation	PDT=59.58 - 0.2776×ψ −0.01155×ψ2	17.48	25.6

[a]BBL: biceps brachii long head.

[b]TBL: triceps brachii long head.

[c]PDT: posterior deltoid.

XSL•FO
RenderX

Main Effects on MVC

The main effects of the 3 ergonomic factors on MVC for each muscle group are summarized in Table 9 and visualized in the main effect plots (Figure 12).

Table . Main effects on maximum voluntary contraction values for factor A, B, and C.

Muscle	Level 0° (straight)	Level 30° (inclined)	Level 45° (highly inclined)
Factor A: torso angle (λ)			
BBL[a]	38.46	41.93	44.78
TBL[b]	47.96	62.91	67.49
Brachioradialis	72.16	95.20	113.41
PDT[c]	50.51	71.25	77.77
Muscle	Level 0 mm (farthest)	Level 50 mm (neutral)	Level 100 mm (closest)
Factor B: lever distance (L)			
BBL	42.14	40.47	42.57
TBL	60.13	51.30	66.93
Brachioradialis	89.46	93.53	97.78
PDT	72.00	59.67	67.86
Muscle	Level –30° (pronation)	Level 0° (neutral)	Level 30° (supination)
Factor C: lever orientation (ψ)			
BBL	43.24	41.14	40.79
TBL	62.47	53.11	62.78
Brachioradialis	88.01	106.65	86.11
PDT	78.30	59.58	61.65

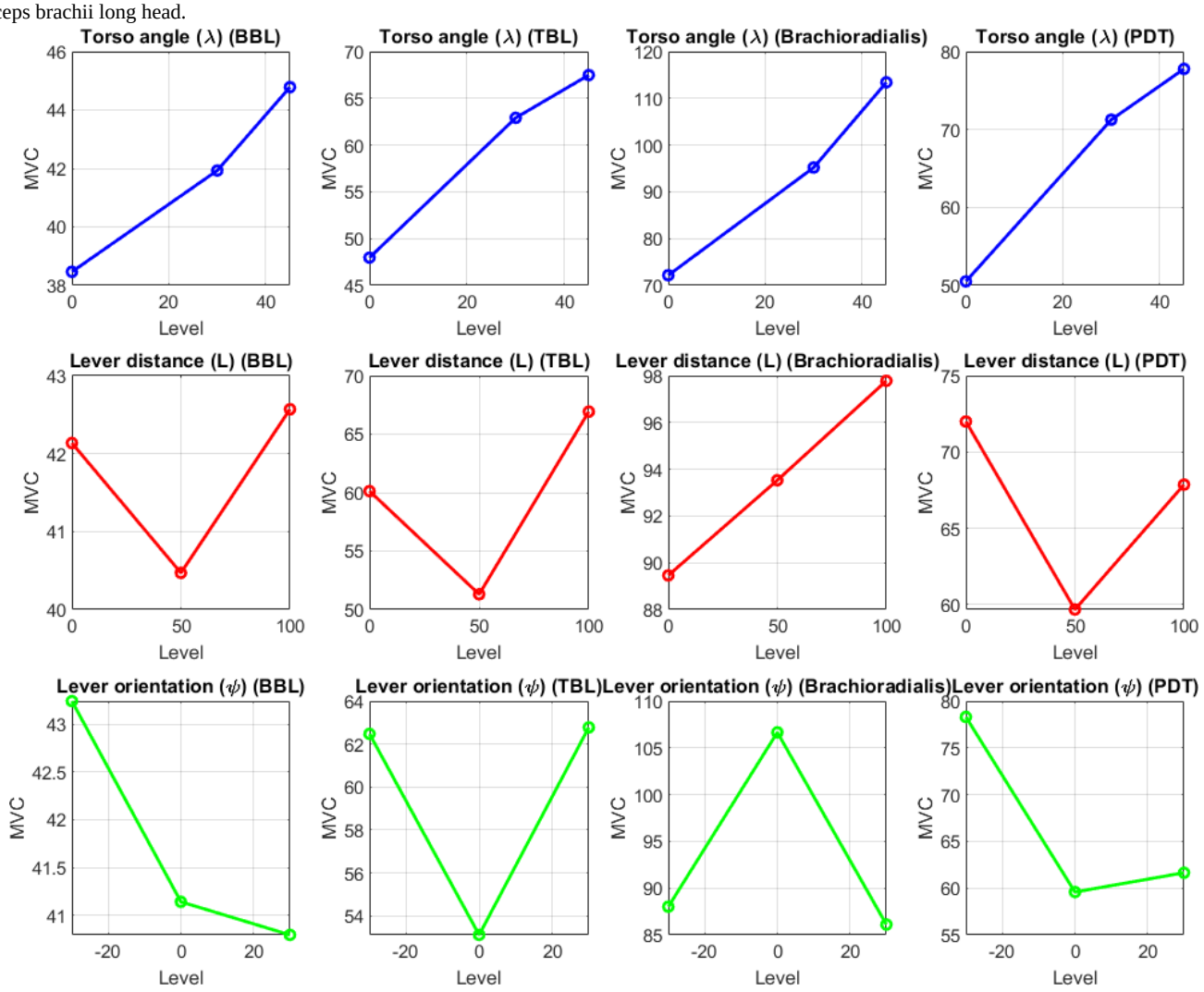
[a]BBL: biceps brachii long head.

[b]TBL: triceps brachii long head.

[c]PDT: posterior deltoid.

XSL•FO
RenderX

Figure 12. Main effect plot derived from the mean of maximum voluntary contraction (MVC) values with respect to ergonomic factors for MVC of biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT) changes with change in torso angle (λ), lever distance (L), and lever orientation angle (ψ). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



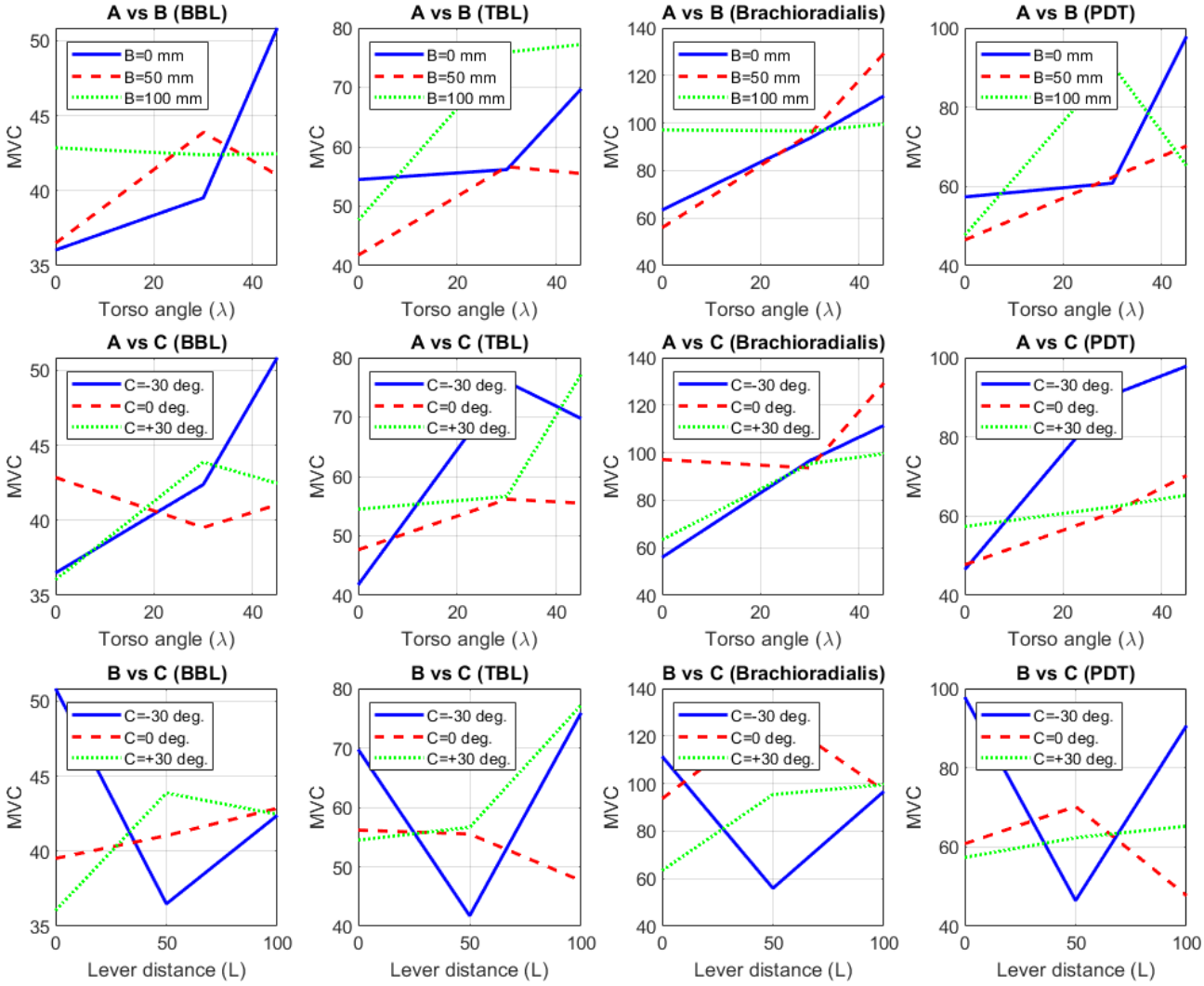
For torso angle (λ), MVC increased with inclination across all muscles, including BBL from 38.46 to 44.78, TBL from 47.96 to 67.49, brachioradialis from 72.16 to 113.41, and PDT from 50.51 to 77.77, with a straight torso (0°) yielding the lowest strain across all muscles, particularly brachioradialis. For lever distance (L), BBL remained relatively stable (40.47 - 42.57, lowest at 50 mm), TBL was optimized at 51.30 (50 mm), brachioradialis rose from 89.46 to 97.78 with decreasing L, and PDT was lowest at 59.67 (50 mm). For lever orientation (ψ), BBL slightly favored 30° (40.79); TBL and PDT were minimized at 0° (53.11 and 59.58, respectively); and brachioradialis peaked at 0° (106.65) but dropped to 86.11 at 30°.

Interaction Effects

Interaction effects (AB: λ×L; AC: λ×ψ; BC: L×ψ) were computed by averaging MVC values across paired factor combinations. These interactions are visualized in line plots (Figure 13) and heatmaps (Figure 14).

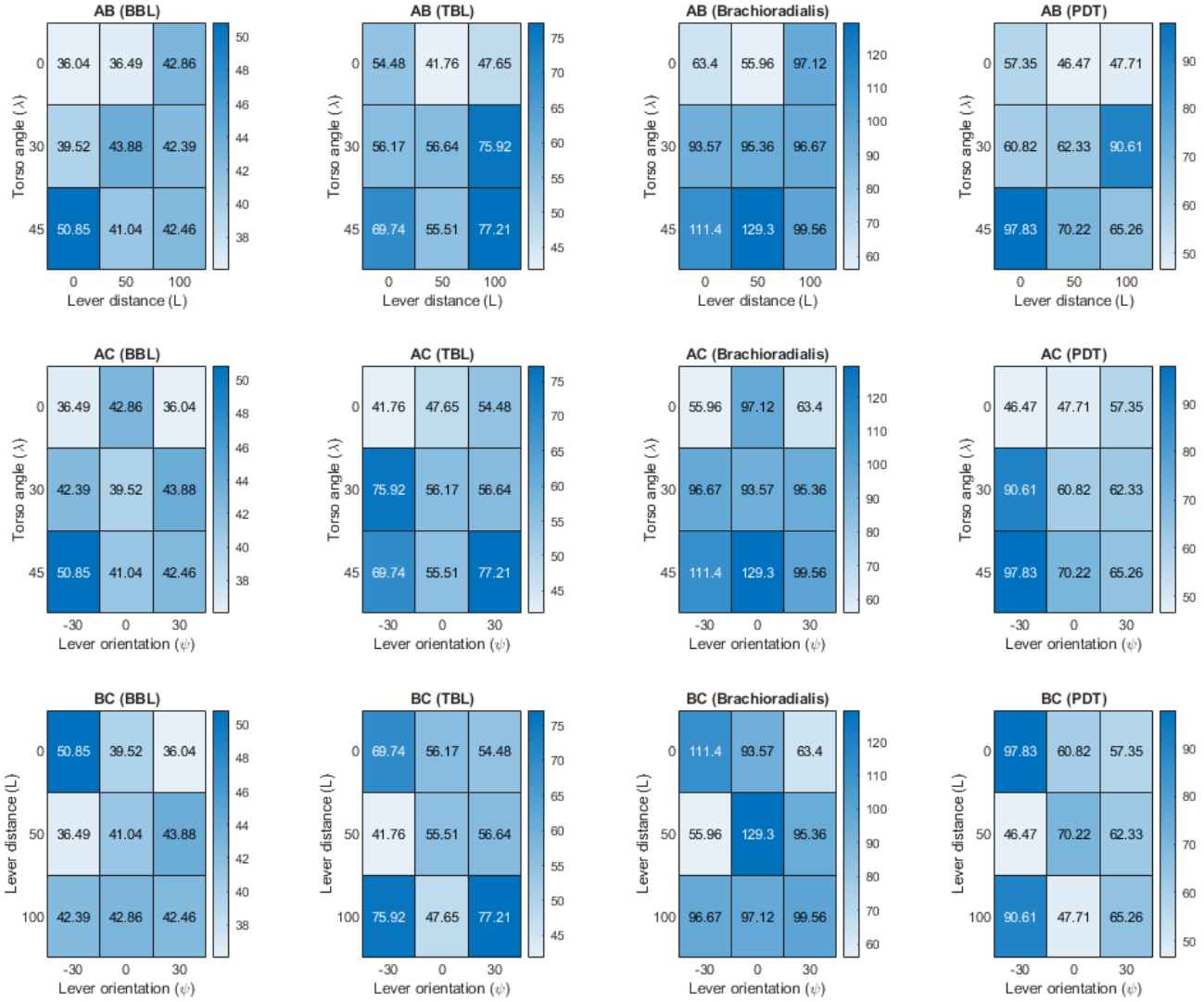
XSL•FO
RenderX

Figure 13. Interaction plot between ergonomics factors lever distance, torso angle, and lever orientation for biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



XSL•FO
RenderX

Figure 14. Interaction heat map between ergonomics factors lever distance, torso angle, and lever orientation for biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



For the AB interaction (Torso angle ×Lever distance), forward lean (λ=45°) combined with long lever distance (L>50 mm) produced elevated effort in BBL and TBL, with lines crossing in Figure 13 and heatmaps showing high-strain zones. Upright posture (λ=0°) at L≈50 mm maintained low effort levels.

For the AC interaction (Torso angle ×Lever orientation), supination (ψ =+30°) combined with λ=0° reduced effort for brachioradialis and PDT, while pronation (ψ = –30°) at λ=45° increased effort, shown by crossing lines and heatmap contrasts. TBL was less affected by this interaction.

For the BC interaction (Lever distance ×Lever orientation), L ≈ 50 mm with ψ =+30° minimized brachioradialis and PDT strain, while a long lever distance with ψ = –30° maximized strain. TBL effort increased with pronation (see Figure 14).

Design Space Optimization

The design space heatmaps (Figure 15) and their associated 3D surface plots (Figure 16) delineated the ergonomic design space. For the AB interaction (λ×L), heatmaps revealed BBL (MVC 38 - 50) and PDT (50-95) with high strain at λ=45°, L=100 mm, and low strain at λ=0°, L=0 mm. The 3D surface plot peaked at 80 MVC where λ=45° and L=100 mm and dipped to 50 MVC at λ=0° and L=0 mm as shown in Figure 16A.

XSL•FO
RenderX

Figure 15. Design space heatmaps for ergonomic factors torso angle (λ) or A, lever distance (L) or B, and lever orientation (Ψ) interacting in AB, AC, and BC configurations. BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.

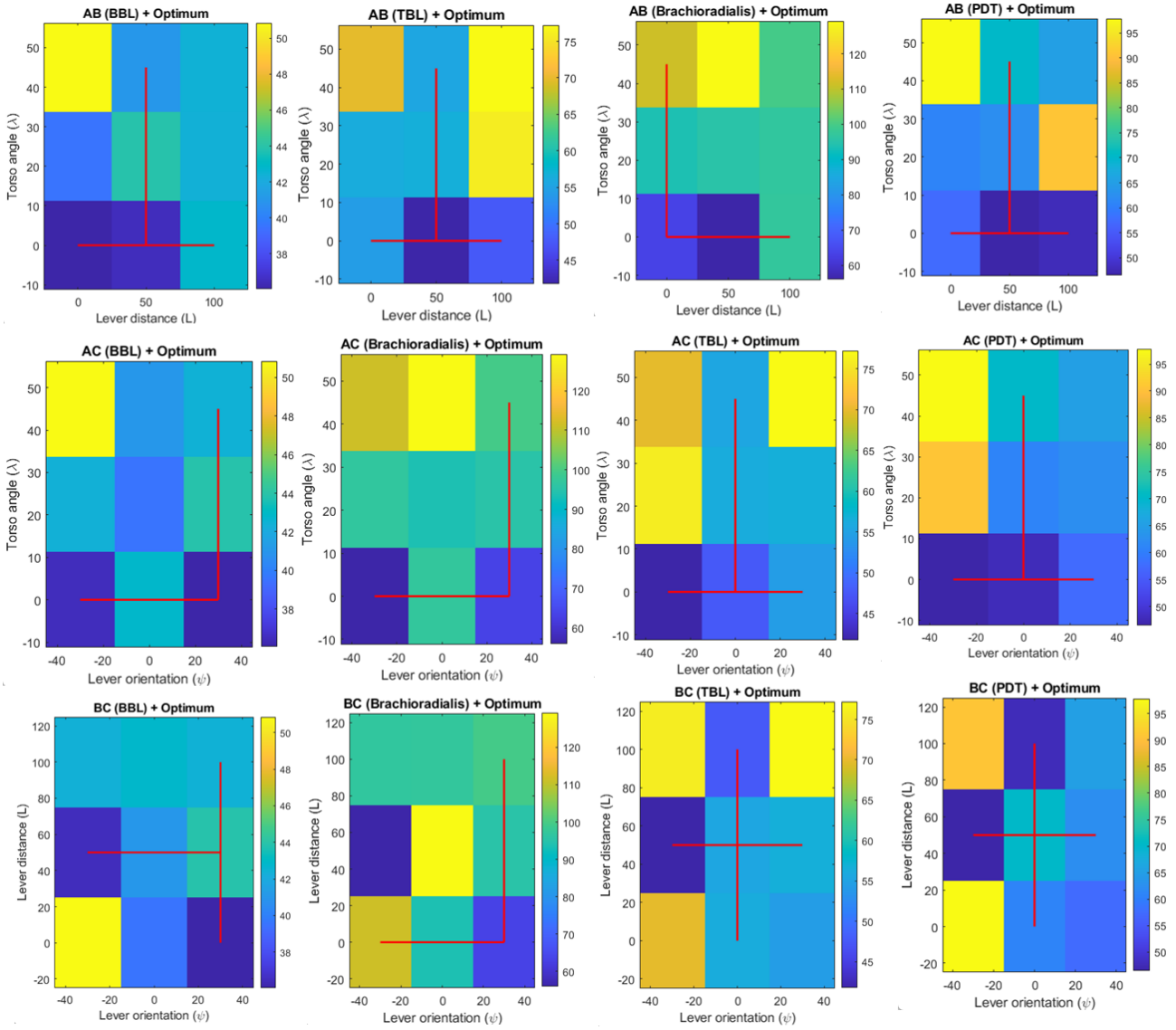
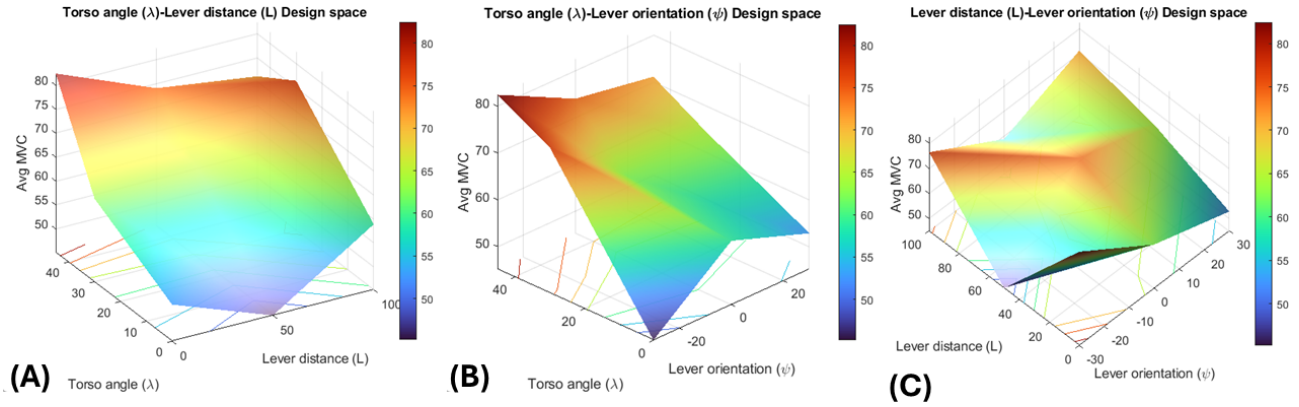


Figure 16. Design space for different ergonomic factors (A) torso angle (λ) and lever distance (L), (B) torso angle (λ) and lever orientation (Ψ), and (C) lever distance (L) and lever orientation (Ψ). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



For the AC interaction (λ × ψ), BBL minimized strain at λ=0°, ψ=30° (supination), and PDT at λ=0°, ψ=0°, with high strain at λ=45° and ψ=−30°. The 3D plot rose to 75 MVC at λ=45°, ψ=−30° and fell to 50 MVC across λ=0° and ψ=0°–30° as shown in Figure 16B. For the BC interaction (L × ψ), BBL was optimized at L=50 mm, ψ=30°, and PDT at L=50 mm, ψ=0°. The 3D plot peaked at 80 MVC for L=100 mm, ψ=−30° and reached approximately 50 MVC near L=50 mm, ψ=0°–30°, as shown in Figure 16C.

XSL•FO
RenderX

Collectively, strain exceeded 75 MVC at extreme configurations (eg, λ=45°, L=100 mm, ψ=-30°) and dropped to approximately 50 MVC at optimal settings (eg, λ=0°, L=0 - 50 mm, ψ=0°-30°).

Optimal Design Parameters

Optimal settings for minimizing individual muscle activity are shown in Table 10. These preferences vary by muscle but consistently favor a straight torso.

Table . Optimum levels for minimum muscle activity for all 4 muscle groups.

Muscle	Torso angle (λ)	Lever distance (L)	Lever orientation (ψ)
BBL[a]	0° (Straight)	50 mm (Neutral)	30° (Supination)
TBL[b]	0° (Straight)	50 mm (Neutral)	0° (Neutral)
Brachioradialis	0° (Straight)	0 mm (Farthest)	30° (Supination)
PDT[c]	0° (Straight)	50 mm (Neutral)	0° (Neutral)

[a]BBL: biceps brachii long head.

[b]TBL: triceps brachii long head.

[c]PDT: posterior deltoid.

Averaging across all muscles for a generalized design, the overall ergonomic recommendation is presented in Table 11 along with optimal wheelchair design parameters in Table 12.

Table . Overall ergonomic recommendation (averaged across muscles).

Parameter	Optimal setting
Torso angle (λ)	0.0° (Straight)
Lever distance (L)	37.5 mm (Between farthest and neutral)
Lever orientation (ψ)	15.0° (Between neutral and supination)

Table . Optimal wheelchair design parameters.

Muscle	Torso angle (in degrees)	Lever distance (mm)	Lever orientation (in degrees)	Torso angle	Lever distance	Lever orientation
BBL[a]	0	50	30	Straight	Neutral	Supination
TBL[b]	0	50	0	Straight	Neutral	Neutral
Brachioradialis	0	0	30	Straight	Farthest from lever	Supination
PDT[c]	0	50	0	Straight	Neutral	Neutral
Overall	0	37.5	15	Straight	Between farthest and neutral	Between Neutral and Supination

[a]BBL: biceps brachii long head.

[b]TBL: triceps brachii long head.

[c]PDT: posterior deltoid.

Discussion

Summary of Key Findings

This study systematically evaluated the ergonomic factors influencing muscle activity and user comfort in a stair-climbing wheelchair design through a dual-methodology approach. In Part A, the plane surface movement study revealed that the ramp and sharp 90° turn were the most challenging tasks, directly attributable to the wheelchair's weight and large dimensions. In Part B, the sEMG analysis demonstrated that a straight torso (λ=0°) consistently minimized muscle strain across all four muscle groups, with an overall ergonomic recommendation of λ=0°, L=37.5 mm, and ψ=15° to balance strain reduction across all muscles.

Design Implications of the Plane Surface Movement Study

The lack of statistical correlation between task difficulty and anthropometric dimensions (*P*=.04) confirms that any difficulty in performing a task can be directly associated with wheelchair design variables, which is the primary goal of this ergonomic assessment. The increase in time and effort at the ramp (R) can be directly correlated with the wheelchair weight increase due to the stair-climbing apparatus installed along with the plane-surface wheelchair. A significant amount of deadweight (net weight of 65 kg) is carried along when the plane surface

XSL•FO
RenderX

mode is active. The next iteration in design will require a detailed study of strength of materials to remove as much weight as possible without compromising strength and durability.

Similarly, the sharp 90° turn (T1) required extreme maneuverability skills at the user's side. The calculated turning distance for a 90° turn of approximately 1450 mm is a direct consequence of the overall length and width of the wheelchair, which plays a critical role in maneuverability and steering experience. The form factor must be reduced not only for addressing the sharp turn problem but also for movement through narrow spaces (Task N), which also took significant time although it had a contradictorily favorable subjective response.

The low comfort scores, particularly during plane surface movement, were attributed to the slightly tilted seat design that, while helping maintain the wheelchair user's center of gravity away from the descent direction during stair-climbing, causes considerable discomfort during plane surface movement. These findings highlight the inherent design tension between stair-climbing functionality and plane surface comfort, suggesting that future designs should incorporate a more dynamic seat tilt adjustment mechanism. Importantly, any design changes inferred from the plane surface movement study must not compromise the primary stair-climbing function.

Interpretation of sEMG Findings and Rationale for Optimal Parameters

The optimal design parameters (λ=0°, L=37.5 mm, and ψ=15°) were systematically determined through Taguchi design of experiments analysis, representing a carefully considered balance of biomechanical efficiency, user comfort, and practical manufacturability.

Torso angle (λ=0°): the neutral torso position minimizes shoulder abduction and maintains a natural arm movement trajectory, reducing rotator cuff stress. The straight torso position showed significant reduction in brachioradialis activation compared to forward-leaning postures (λ=15°, 30°), indicating lower muscular effort. This finding aligns with previous research on wheelchair propulsion biomechanics showing that neutral spine positioning reduces cumulative shoulder joint loading and decreases the risk of overuse injuries. As torso angle increases, muscle effort rises across the forearm (brachioradialis), arm (BBL and TBL), and shoulder (PDT) groups due to anterior torso displacement. The prototype's backrest design mitigates fatigue, with the 0° posture proving optimal during stair-climbing compared to forward-leaning positions.

Lever distance (L=37.5 mm): this distance provides optimal mechanical advantage while maintaining comfortable reach, derived from anthropometric data analysis and biomechanical modeling. Shorter distances resulted in cramped postures and increased BBL activation, while longer distances showed increased PDT activation, suggesting compensation through shoulder muscles. The 37.5 mm distance represents the optimal compromise minimizing total muscular activation across all monitored muscle groups. Prior literature suggests that extreme upper limb reaches elevate MSD risk, a hypothesis confirmed by these results: muscle effort increased with greater L for most muscles.

Lever orientation (ψ = 15°): this slight supination optimizes weight distribution, provides postural support, and maintains shoulder joint alignment. The brachioradialis, BBL, and PDT exhibited the lowest effort at +30° (supination), while TBL effort peaked during the pronation-to-supination transition due to flexion demands. An outward lever orientation consistently reduces muscle activity, lowering fatigue and MSD risk. The 15° angle balances stair-climbing stability requirements with plane surface comfort; neutral seating (0°) showed marginally lower muscle activation but compromised stability during stair-climbing, while greater supination angles (30°) increased discomfort during prolonged sitting and reduced forward reach capability.

The interaction effects, visualized through heatmaps and 3D surface plots, underscored the interdependence of these factors. The interaction plots helped identify critical regions where the combined effects of ergonomic factors diverge, aiding in making informed decisions about the physical dimensions of the wheelchair. For instance, the AB interaction plot showed that forward lean (λ=45°) combined with long lever distance (>50 mm) should be avoided as it spikes effort, while upright posture at L≈50 mm keeps effort low. The consistency between heatmaps and 3D surfaces highlights an ergonomic sweet spot that reduces muscle effort and fatigue, offering clear guidance for refining wheelchair design to enhance user comfort and efficiency during stair-climbing tasks.

The modular design architecture permits adjustment of all 3 parameters to accommodate users with different anthropometric characteristics, strength capabilities, and functional requirements. Individual optimization may be necessary for users at anthropometric extremes or with specific upper extremity limitations.

Weight Reduction Feasibility and Future Design Strategies

This prototype weight of 65 kg represents a significant barrier to usability, particularly for flat surface maneuverability and transportation. This weight is substantially higher than conventional manual wheelchairs (15 - 25 kg) and constitutes a primary design limitation requiring immediate attention. Weight reduction strategies under investigation include the following:

1. Material substitution: replacing steel frame components with aerospace-grade aluminum alloys (estimated 8 - 12 kg reduction) and carbon fiber composites for non–load-bearing elements (estimated 3 - 5 kg reduction).
2. Mechanism optimization: redesigning the transformation mechanism to eliminate redundant components and implementing topology optimization for load-bearing structures (estimated 4 - 6 kg reduction).
3. Component integration: combining multiple discrete parts into unified assemblies to reduce fastener count and material overlap (estimated 2 - 3 kg reduction).

Preliminary engineering analysis suggests a realistic target weight of 45 - 50 kg is achievable through these combined

XSL•FO
RenderX

strategies, representing a 23% - 31% reduction. This improved weight would significantly enhance maneuverability while maintaining structural integrity for stair-climbing operations. Cost-benefit analysis and finite element modeling are ongoing to validate these approaches prior to next-generation prototype development.

Safety Systems, Descending Mechanism, and Wheelchair Adjustability

While this manuscript focuses on ergonomic assessment during upward stair-climbing and plane surface movement, the wheelchair incorporates comprehensive safety and functional systems essential for complete operational capability:

1. Descending mechanism: stair descent uses a controlled ratchet-and-pawl braking system integrated with the stair-climbing wheel assembly. The mechanism permits incremental wheel rotation (step-by-step descent) while preventing uncontrolled backward movement. User-activated brake levers provide variable descent speed control (range: 1 - 4 steps per minute). Antitip sensors provide auditory feedback when the center of gravity approaches the instability threshold.
2. Braking systems: dual independent braking systems ensure redundancy and safety. Primary system lever-operated disc brakes on plane-surface wheels providing immediate stopping capability. Secondary system electromagnetic parking brakes preventing rollback during stationary periods. Emergency brake activation requires <1.5-second response time from full speed, meeting safety standards for powered mobility devices.
3. Adjustability features: the modular design permits extensive customization, including seat height adjustment (400 - 550 mm from ground, 50 mm increments), backrest angle adjustment (0 - 30° recline, 5° increments), footrest height and angle adjustment, lever position adjustment (radial distance 300 - 500 mm, angular position ±30°), and seat width adjustment (400 - 500 mm) to accommodate different pelvic widths. All adjustments use tool-free quick-release mechanisms for rapid configuration changes.

Comprehensive documentation of these systems, including mechanical specifications, safety testing protocols, and user operating procedures, will be presented in subsequent publications focused specifically on mechanical design and safety validation.

Strengths and Limitations

This study has several strengths. It represents one of the first pilot studies to integrate surface electromyography with subjective usability assessment for a stair-climbing wheelchair, providing both user-perceived comfort data and objective biomechanical evidence. The Taguchi-based experimental design efficiently identified optimal ergonomic parameters while minimizing the number of experimental runs. The inclusion of participants with paraplegia, though limited in number, provided critical real-world user validation.

However, several limitations must be acknowledged. The relatively small sample size (n=9 for plane surface movement, including 2 participants with paraplegia; n=7 for sEMG study) is appropriate for this pilot investigation but limits generalizability. Future studies should include larger, more diverse participant groups, including various disability types, severity levels, and demographic characteristics, to establish comprehensive design guidelines.

A significant limitation is the inclusion of only male participants in the anthropometric survey. Female and gender-diverse individuals have different anthropometric characteristics, muscle strength distributions, and ergonomic requirements [36]. Studies have shown that female wheelchair users often experience different pressure distribution patterns, grip strength capabilities, and shoulder biomechanics during propulsion [37]. Future iterations should conduct separate anthropometric studies for different gender groups to ensure optimal comfort, safety, and usability for all users. Additionally, consideration must be given to pregnancy-related ergonomic requirements, hormone-related strength variations, and gender-specific injury risk factors that may influence long-term wheelchair use patterns.

The controlled experimental environment, while necessary for this initial validation study, limits ecological validity. Real-world wheelchair use involves additional challenges including uneven terrain, weather conditions, unpredictable obstacles, crowded spaces, and cognitive demands from environmental distractions [38]. Future research should include (1) field testing in actual urban environments under naturalistic conditions; (2) long-term usability studies capturing daily use patterns and cumulative physical demands; (3) assessment under various weather and lighting conditions to evaluate all-season functionality; and (4) evaluation of cognitive load during real-world navigation tasks, including wayfinding, obstacle avoidance, and multitasking demands typical of community mobility. Long-term usability, durability, and cumulative physical effects require extended longitudinal studies beyond the scope of this pilot investigation, such as the impact of materials for backrest, seat, and lever handle [39-41].

Conclusion

This research provides critical design feedback for the stair-climbing wheelchair. Following a generic product design and development philosophy, the goal was to identify design variables influenced by usability testing. The most basic ergonomic assessment of a wheelchair that is, its reach and clearance, has been established through a detailed anthropometric survey of 20 participants with 5th and 95th percentile data guiding design dimensions.

The plane surface movement study demonstrated that the ramp and sharp 90° turn were the most challenging tasks, attributed to the wheelchair's 65 kg weight and large dimensions requiring approximately 1450 mm for a 90° turn. Subjective feedback indicated low satisfaction, particularly for comfort, due to the tilted seat design. These findings highlight the need for design improvements, including weight reduction, a more compact form factor, and seat tilt adjustments, all while preserving stair-climbing functionality.

The lever-propulsion sEMG study systematically evaluated three ergonomic factors. Key findings indicate that a straight torso (0°) consistently minimizes strain, particularly for

XSL•FO
RenderX

brachioradialis. A neutral lever distance (approximately 50 mm) reduced effort for most muscles, while neutral to slightly supinated orientations (0°-30°) generally optimized performance. Interaction effects guided an overall ergonomic recommendation of λ=0°, L=37.5 mm, and ψ=15°. These settings offer a balanced, practical solution for minimizing fatigue and enhancing user comfort. The modular wheelchair design permits customization for diverse user populations, and future research should include larger, gender-diverse participant groups and real-world validation studies, strengthening the foundation for ergonomic wheelchair optimization.

Funding

This research was funded by the Department of Science and Technology, Government of India.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions to protect participant privacy but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Cowan RE, Nash MS, Collinger JL, Koontz AM, Boninger ML. Impact of surface type, wheelchair weight, and axle position on wheelchair propulsion by novice older adults. Arch Phys Med Rehabil 2009 Jul;90(7):1076-1083. [doi: 10.1016/j.apmr.2008.10.034] [Medline: 19577019]
2. van der Woude LHV, de Groot S, Janssen TWJ. Manual wheelchairs: research and innovation in rehabilitation, sports, daily life and health. Med Eng Phys 2006 Nov;28(9):905-915. [doi: 10.1016/j.medengphy.2005.12.001] [Medline: 16504565]
3. Dubowsky SR, Rasmussen J, Sisto SA, Langrana NA. Validation of a musculoskeletal model of wheelchair propulsion and its application to minimizing shoulder joint forces. J Biomech 2008 Oct 20;41(14):2981-2988. [doi: 10.1016/j.jbiomech.2008.07.032] [Medline: 18804763]
4. Morrow MM, Rankin JW, Neptune RR, Kaufman KR. A comparison of static and dynamic optimization muscle force predictions during wheelchair propulsion. J Biomech 2014 Nov 7;47(14):3459-3465. [doi: 10.1016/j.jbiomech.2014.09.013] [Medline: 25282075]
5. Babu Rajendra Kurup N, Puchinger M, Gfoehler M. A preliminary muscle activity analysis: Handle based and push-rim wheelchair propulsion. J Biomech 2019 May 24;89:119-122. [doi: 10.1016/j.jbiomech.2019.04.011] [Medline: 31053474]
6. Verma A, Shrivastava S, Ramkumar J. Mapping wheelchair functions and their associated functional elements for stair climbing accessibility: a systematic review. Disabil Rehabil Assist Technol 2024 Jan;19(1):200-221. [doi: 10.1080/17483107.2022.2075476] [Medline: 35613308]
7. Verma A, Shrivastava S, Ramkumar J. Development of an optimized stair-climbing mechanism using parametric curved wheels inspired by human foot strides on flat stairs. Eng Res Express 2025 Sep 30;7(3):035436. [doi: 10.1088/2631-8695/ae086f]
8. Product manual for rehabilitation equipment—wheelchairs, folding, adult size according to IS 7454:1991. Bureau of Indian Standards. 2021. URL: https://www.bis.gov.in/wp-content/uploads/2021/08/IS-7454-Product-Manual.pdf [accessed 2026-03-24]
9. Guidelines and space standards for barrier-free built environment for disabled and elderly persons. Central Public Works Department Ministry of Urban Affairs and Employment, Government of India. 1998. URL: https://cpwd.gov.in/Publication/aged&disabled.pdf [accessed 2026-03-24]
10. Planning a barrier free environment. Office of the Chief Commissioner for Persons with Disabilities Ministry of Social Justice and Empowerment, Government of India. 2001. URL: https://www.sparsh.mp.gov.in/Public/SparshPublicPages/BarreirFreeEnvironment.aspx [accessed 2026-04-10]
11. International organization for standardization. wheelchairs — maximum overall dimensions. ISO. 1985. URL: https://www.iso.org/standard/13810.html [accessed 2026-03-25]
12. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]
13. Turner TL, Balmer DF, Coverdale JH. Methodologies and study designs relevant to medical education research. Int Rev Psychiatry 2013 Jun;25(3):301-310. [doi: 10.3109/09540261.2013.790310] [Medline: 23859093]
14. Saleeby PW. An introduction to the International Classification of Functioning, Disability and Health (ICF). Int J Disabil Hum Dev 2016;15(1):1-3. [doi: 10.1515/ijdhd-2015-0027]
15. P. w. d. act, “the persons with disabilities (equal opportunities, protection of rights and full participation) act, 1995, published in part II, section 1 of the extraordinary gazette of india. ministry of law, justice and company affairs (legislative department),”

XSL•FO
RenderX

ministry of law, justice and company affairs (legislative department). 1995 URL: https://www.indiacode.nic.in/bitstream/123456789/12890/1/the_persons_with_disabilities_act%2C_1995_no._1_of_1996_date_01.01.1996.pdf [accessed 2026-03-13]
16. Guidelines and space standards for barrier free built environment for disabled and elderly. : Central Public Works Department; 1998 URL: https://cpwd.gov.in/publication/aged&disabled.PDF [accessed 2026-03-11]
17. Chief commissioner for persons with disabilities, “planning a barrier free environment,”. 2001 URL: https://www.sparsh.mp.gov.in/Common/SJ/PDFFiles/PlanningForBarrierFreeEnvironment.PDF [accessed 2026-03-13]
18. De Luca CJ. The use of surface electromyography in biomechanics. J Appl Biomech 1997;13(2):135-163. [doi: 10.1123/jab.13.2.135]
19. Louis N, Gorce P. Surface electromyography activity of upper limb muscle during wheelchair propulsion: Influence of wheelchair configuration. Clin Biomech (Bristol) 2010 Nov;25(9):879-885. [doi: 10.1016/j.clinbiomech.2010.07.002] [Medline: 20846767]
20. Potvin JR, Bent LR. A validation of techniques using surface EMG signals from dynamic contractions to quantify muscle fatigue during repetitive tasks. J Electromyogr Kinesiol 1997 Jun;7(2):131-139. [doi: 10.1016/s1050-6411(96)00025-9] [Medline: 20719698]
21. Chow JW, Levy CE. Wheelchair propulsion biomechanics and wheelers’ quality of life: an exploratory review. Disabil Rehabil Assist Technol 2011;6(5):365-377. [doi: 10.3109/17483107.2010.525290] [Medline: 20932232]
22. Gellman H, Sie I, Waters RL. Late complications of the weight-bearing upper extremity in the paraplegic patient. Clin Orthop Relat Res 1988 Aug;Available(233):132-135. [Medline: 3402118]
23. Gangelhoff J, Cordain L, Tucker A, Sockler J. Metabolic and heart rate responses to submaximal arm lever and arm crank ergometry. Arch Phys Med Rehabil 1988 Feb;69(2):101-105. [Medline: 3341886]
24. Tomiak T, Gorkovenko AV, Tal’nov AN, et al. The averaged EMGs recorded from the arm muscles during bimanual “Rowing” movements. Front Physiol 2015;6:349. [doi: 10.3389/fphys.2015.00349] [Medline: 26640440]
25. Lui J, MacGillivray MK, Sheel AW, Jeyasurya J, Sadeghi M, Sawatzky BJ. Mechanical efficiency of two commercial lever-propulsion mechanisms for manual wheelchair locomotion. J Rehabil Res Dev 2013;50(10):1363-1372. [doi: 10.1682/JRRD.2013.02.0034] [Medline: 24699972]
26. Hoozemans MJM, Kuijer P, Kingma I, et al. Mechanical loading of the low back and shoulders during pushing and pulling activities. Ergonomics 2004 Jan 15;47(1):1-18. [doi: 10.1080/00140130310001593577] [Medline: 14660215]
27. O’Sullivan LW, Gallwey TJ. Forearm torque strengths and discomfort profiles in pronation and supination. Ergonomics 2005 May 15;48(6):703-721. [doi: 10.1080/00140130500070954] [Medline: 16087504]
28. Roman-Liu D, Tokarski T. Upper limb strength in relation to upper limb posture. Int J Ind Ergon 2005;35(1):19-31. [doi: 10.1016/j.ergon.2004.07.002]
29. Santos Requejo P, Lee S, Mulroy S, et al. Shoulder muscular demand during lever-activated vs pushrim wheelchair propulsion in persons with spinal cord injury. J Spinal Cord Med 2008 Jan;31(5):568-577. [doi: 10.1080/10790268.2008.11754604]
30. Kotajarvi BR, Sabick MB, An KN, Zhao KD, Kaufman KR, Basford JR. The effect of seat position on wheelchair propulsion biomechanics. J Rehabil Res Dev 2004 May;41(3B):403-414. [doi: 10.1682/jrrd.2003.01.0008] [Medline: 15543458]
31. Cavallone P, Vieira T, Quaglia G, Gazzoni M. Electomyographic activities of shoulder muscles during Handwheelchair.Q vs pushrim wheelchair propulsion. Med Eng Phys 2022 Aug;106(1):103833. [doi: 10.1016/j.medengphy.2022.103833] [Medline: 35926952]
32. Mastinu E, Ortiz-Catalan M, Hakansson B. Analog front-ends comparison in the way of a portable, low-power and low-cost EMG controller based on pattern recognition. 2015 Presented at: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Milan p. 2111-2114. [doi: 10.1109/EMBC.2015.7318805]
33. Jiang S, Gao Q, Liu H, Shull PB. A novel, co-located EMG-FMG-sensing wearable armband for hand gesture recognition. Sens Actuators A Phys 2020 Jan;301:111738. [doi: 10.1016/j.sna.2019.111738]
34. Merletti R, Parker PJ. Electromyography: Physiology, Engineering, and Non-Invasive Applications: John Wiley and Sons; 2004. [doi: 10.1002/0471678384]
35. Pearson K. Determination of the coefficient of correlation. Science 1909 Jul 2;30(757):23-25. [doi: 10.1126/science.30.757.23] [Medline: 17838275]
36. Steinfeld E, Maisel J, Feathers D, D’Souza C. Anthropometry and standards for wheeled mobility: an international comparison. Assist Technol 2010;22(1):51-67. [doi: 10.1080/10400430903520280] [Medline: 20402047]
37. Boninger ML, Koontz AM, Sisto SA, et al. Pushrim biomechanics and injury prevention in spinal cord injury: Recommendations based on CULP-SCI investigations. JRRD 2004;42(3sup1):9. [doi: 10.1682/JRRD.2004.08.0103]
38. Tolerico ML, Ding D, Cooper RA, et al. Assessing mobility characteristics and activity levels of manual wheelchair users. J Rehabil Res Dev 2007;44(4):561-571. [doi: 10.1682/jrrd.2006.02.0017] [Medline: 18247253]
39. Muhammad Salman Ajmal H, Muneer R, Saeed A, Tanveer M, Ahsan Saeed M. Synergistic role of green‐synthesized zinc oxide nanomaterials in biomedicine applications. ChemistrySelect 2024 Sep;9(36) [FREE Full text] [doi: 10.1002/slct.202402517]
40. Khan MUA, Ajmal HMS, Hassan HA, Azam A, Malik E. A soft robotic sleeve for physiotherapy: improving elbow rehabilitation in baseball pitchers. Physiother Res Int 2025 Jan;30(1):e70025. [doi: 10.1002/pri.70025] [Medline: 39764679]

XSL•FO
RenderX

41. Verma A, Singh A, Ramkumar J. Topology optimization of mechanical structures in stair-climbing assistive technology. Nanomaterials and Energy 2019 Dec 17;8(2):167-177. [doi: 10.1680/jnaen.19.00033]

Abbreviations

BBL: biceps brachii long head
CPWD: Central Public Works Department
DHM: digital human model
MSD: musculoskeletal disorder
MVC: maximum voluntary contraction
PDT: posterior deltoid
RMS: root-mean square
sEMG: surface electromyography
TBL: triceps brachii long head

Edited by S Munce; submitted 12.Jun.2025; peer-reviewed by D Menychtas, HMS Ajmal; revised version received 04.Jan.2026; accepted 03.Feb.2026; published 13.Apr.2026.

Please cite as:
Verma A, Kumar R, Ramkumar J
Design Implications of Comfort and Usability of Manual Stairclimbing Wheelchair: Ergonomic Assessment and Pilot Study Using Surface Electromyography Inputs
JMIR Rehabil Assist Technol 2026;13:e78965
URL: https://rehab.jmir.org/2026/1/e78965
doi:10.2196/78965

© Abhishek Verma, Rohit Kumar, J Ramkumar. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Social Justice Approach to Assistive Technology and Well-Being of People With Visual Disabilities in Low- and Middle-Income Countries: Qualitative Narrative Study

Luisa Maria Ortiz-Escobar[1,2*], MSc; Mario Andres Chavarria[3,4,5*], PhD; Samia Hurst-Majno[2*], MD, PhD, Prof Dr Med; Oscar Ivan Campo Salazar[3*], PhD; Celia Escobar-Hurtado[6], MSc; Michael Ashley Stein[7,8*], PhD; Minerva Rivas Velarde[2,5*], PhD, Prof Dr

[1]Grupo de Investigación en Ingeniería Biomédica, GBIO, Universidad Autónoma de Occidente, Santiago de Cali, Colombia

[2]Institute of Ethics, History, and Humanities, University of Geneva, Geneva, Switzerland

[3]Grupo de Investigaciones en Biomédica, Universidad Autónoma de Occidente, Santiago de Cali, Colombia

[4]EssentialTech Centre, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[5]School of Health Science, HES-SO Genève, Avenue de Champel 47 1206 Genève, Geneva, Switzerland

[6]Escuela de Rehabilitación Humana, Facultad de Salud, Universidad del Valle, Santiago de Cali, Colombia

[7]Harvard Law School Project on Disability, Harvard University, Cambridge, MA, United States

[8]Faculty of Law Centre for Human Rights, University of Pretoria, Pretoria, South Africa

[*]these authors contributed equally

Corresponding Author:
Minerva Rivas Velarde, PhD, Prof Dr
Institute of Ethics, History, and Humanities, University of Geneva, Geneva, Switzerland

Abstract

Background: The United Nations' third Sustainable Development Goal emphasizes ensuring healthy lives and promoting well-being (WB) for all, which requires effective assistive technology (AT) for persons with disabilities. In low- and middle-income countries (LMICs), however, AT remains largely inaccessible, and high abandonment rates indicate that many existing solutions fail to meet users' needs. To improve AT design and effectiveness, a deeper understanding of users' lived experiences and the ways AT influences WB is essential.

Objective: This study aimed to explore how technology creates opportunities or barriers in the daily lives of persons with visual disabilities in LMICs and how it affects their WB.

Methods: We conducted a qualitative narrative study guided by deductive qualitative analysis, using the capability approach (CA) and disadvantage theory (DT) as theoretical frameworks. Nineteen adults with visual disabilities from Cali, Colombia, participated in in-depth, semistructured interviews. A focus group (n=5) deepened the exploration of shared experiences. Data analysis followed three stages: (1) deductive coding using Nussbaum list of central capabilities and key CA constructs (functionings, conversion factors, and agency); (2) recoding through DT concepts (insecure functioning, corrosive disadvantages, and fertile functionings); and (3) inductive analysis to capture emergent sociocultural themes.

Results: AT shaped both opportunities and constraints in participants' lives. While functionings such as employment, mobility, and affiliation were highly valued, they often remained insecure due to systemic barriers. Corrosive disadvantages—such as unemployment, exclusion, and limited spatial autonomy—undermined multiple capabilities simultaneously. Conversely, fertile functionings such as equitable employment, adaptive sports, and access to well-designed AT supported agency and resilience. The inductive analysis revealed 3 interconnected themes: the aspiration to explore and expand movement, the desire to appear attractive, and the adoption of nonconfrontational strategies to maintain social harmony. These findings highlight how emotional, aesthetic, and cultural dimensions shape the experience and meaning of AT.

Conclusions: While AT research in LMICs often emphasizes availability, it rarely addresses how social norms, structural violence, and fear affect meaningful use. The combined CA and DT lens reveals that AT can either enable or constrain WB depending on how it aligns with users' lived contexts. Designing for fertile functionings—those that support agency, safety, and resilience—is essential. Participatory, context-sensitive design must prioritize not only functionality, but also aesthetic dignity, cultural relevance, and emotional security. Including the voices—and silences—of persons with disabilities in the Global South is crucial for transforming AT from a mere tool into a catalyst for real freedom and WB.

(JMIR Rehabil Assist Technol 2026;13:e72306) doi:10.2196/72306

XSL•FO
RenderX

KEYWORDS

assistive technology; persons with disabilities; visual disabilities; user-centered design; capabilities approach; disadvantage theory; LMIC; low- and middle-income countries

Introduction

Background

The incidence and prevalence of visual impairment worldwide have exceeded World Health Organization (WHO) projections for the second decade of the current century [1,2]. In low- and middle-income countries (LMICs), the rate of persons with severe visual impairments is approximately 2 to 3 times higher than in high-income countries [3].

Visual impairment significantly constrains individual autonomy and social participation, with far-reaching consequences for quality of life [4]. Assistive technology (AT) has emerged as a critical means of mitigating these effects by promoting independence and enabling greater inclusion.

Beyond its practical benefits, AT is increasingly recognized as instrumental in advancing the rights enshrined in the United Nations Convention on the Rights of Persons with Disabilities [5], as well as in supporting broader global commitments such as Sustainable Development Goal 3: Good Health and Well-Being [6].

Thus, AT contributes to overall well-being (WB) by enhancing autonomy, supporting participation in education, employment, and community life, and reinforcing psychosocial WB through increased agency and participation [7,8].

Nevertheless, despite its benefits, access to AT remains severely constrained—particularly in LMICs [9]. This lack of access is further exacerbated by high abandonment rates, which often result from inadequate alignment between technologies and the lived realities of users [10-12]. The literature increasingly points to the need for participatory, person-centered approaches that actively engage persons with disabilities in the design and delivery of AT solutions [13], ensuring that such interventions are both contextually relevant and sustainable.

Although AT development increasingly claims to adopt user-centered design [14], studies have identified a persistent gap between rhetoric and practice. Users are often consulted only to refine existing designs, rather than to inform fundamental decisions based on their needs and experiences [15,16].

To respond to this gap and advance a genuinely person-centered approach, we initiated a study to inform the development of AT for the navigation of persons with visual disabilities in LMICs. This paper presents the results of the first phase, a qualitative inquiry into users' lived experiences, aspirations, and interactions with technology.

While AT research in LMICs often focuses on availability, it neglects how structural and social contexts shape meaningful use. The capability approach (CA) addresses this by emphasizing individuals' real freedoms to achieve valued functionings, not just access to resources [17-19]. However, CA has been critiqued for undertheorizing how power, fear, and structural violence constrain agency [20]. Disadvantage theory (DT) complements CA by introducing concepts such as corrosive disadvantage, enabling deeper analysis of how AT may worsen exclusion under adverse conditions [21]. This combined CA and DT lens highlights how AT can both enable and limit WB in LMICs.

Approaches to Disability, WB, AT, and Justice

The Capabilities Approach

The CA evaluates human development through the lens of equity, emphasizing human flourishing as the ultimate goal. It focuses on the life individuals are able to lead, based on their "capabilities" (practical opportunities) and "functionings" (achievements) [22].

Building on this foundation, this approach recognizes the concept of agency, which is the ability to pursue valued goals. To achieve these goals, individuals need access to necessary resources and "conversion factors," which affect how resources are transformed into functionings [17].

This theoretical lens has gained traction in disability studies over the past 2 decades, where it emphasizes individual agency and enables a nuanced examination of resources and conversion factors. In this context, disability is framed in terms of the deprivation or restriction of capabilities and functionings [23-26].

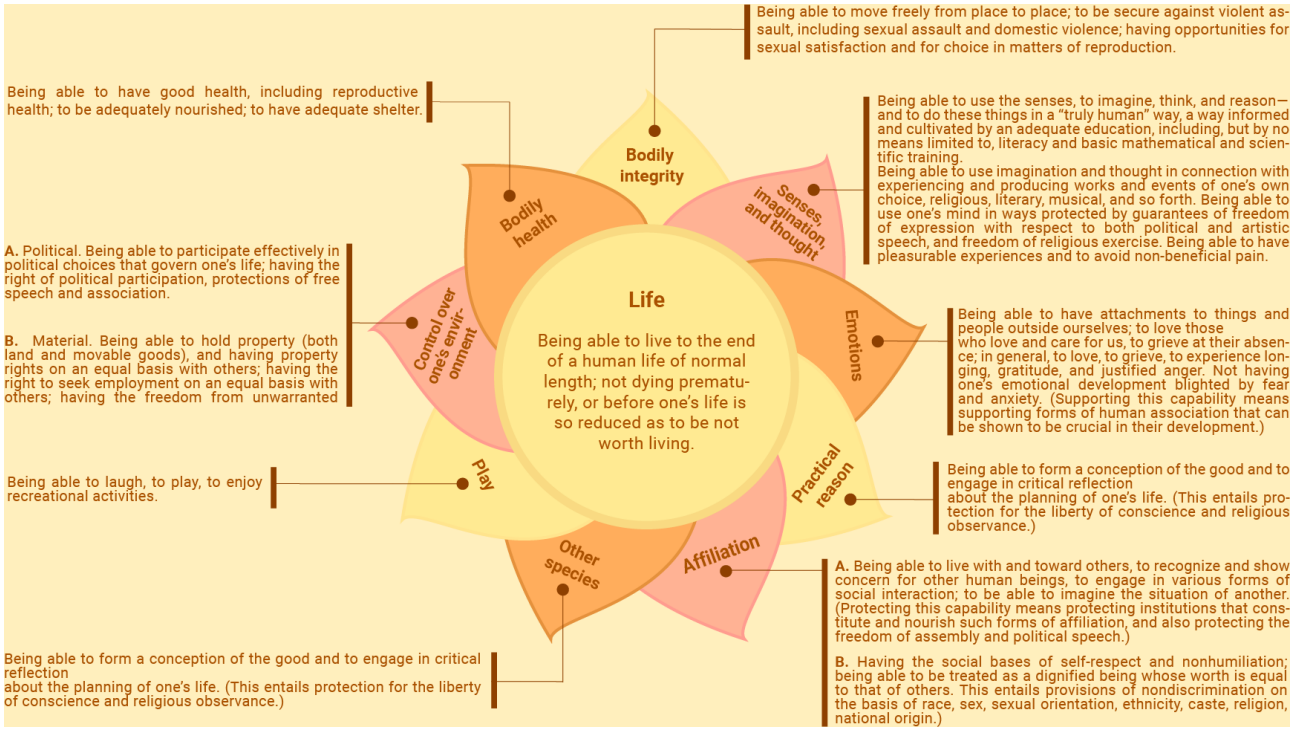
For this study, we chose the CA to analyze the experiences of persons with visual disabilities in a middle-income country, namely Colombia. This framework sheds light on what people value in their lives, guides researchers in identifying daily opportunities and barriers, and aids in evaluating how AT influences the pursuit of a "good life."

Recent literature situates the role of technology within this framework [27,28].

Nussbaum, who has made valuable contributions to the approach, proposed an open list of 10 capabilities, which she argues are central to the realization of self-determined, meaningful, and fulfilling lives [18,29]. This list (Figure 1) has been effectively applied in various contexts, including assessing dignity, monitoring human rights, and guiding social policy to reduce disadvantage [30].

XSL•FO
RenderX

Figure 1. Diagram of the 10 central capabilities proposed by Nussbaum and used as a framework for deductive analysis of participant narratives about technology and well-being.



Disability research focused on understanding quality of life, living conditions, and WB in LMICs is scarce and predominantly quantitative, limiting insights into the lived experiences of disability and personal ideas of flourishing [31,32]. Research on AT within the CA for persons with disabilities is even rarer. This field could benefit from studies like this one, which aim to provide valuable information for developing AT to narrow the capabilities gap. Understanding the lived experience of disability and personal definitions of a good life can guide AT designers in making meaningful and desirable assistive devices (ADs).

Disadvantage Theory

Wolff and De-Shalit [21] developed a theory aimed at achieving equity and social justice by addressing disadvantage. Drawing from Sen CA and Nussbaum contributions, they define disadvantage as a "lack of genuine opportunities for secure functionings" [21].

These authors underscore the plural origins of disadvantages and the necessity of indexing them. They diverge from Sen and Nussbaum by using the terms genuine opportunities and secure functionings instead of capabilities and functionings. While Sen emphasizes freedom to achieve capabilities, Nussbaum defends a universal list of central capabilities, Wolff and De-Shalit focus on how disadvantage is experienced and addressed in real-world contexts. Their terminology reflects a concern with the stability of achieved functionings and the substantive nature of opportunity, aiming to make the framework more actionable for policy design.

They define a disadvantaged person as one whose functionings are involuntarily insecure. Another key concept is "corrosive disadvantage," where one disadvantage negatively impacts another functioning, making it insecure. To counteract this, the authors suggest identifying and addressing corrosive disadvantages while also seeking "fertile functionings," which positively impact and secure other functionings [21].

Persons with disabilities face multiple disadvantages, such as "reduced health, life span, opportunities for fulfilling work, social networks, control over their lives, leisure activities, and living standards" [21,33]. Given the likelihood of facing corrosive disadvantages, it is crucial to prioritize the capabilities and functionings of persons with disabilities to decluster their disadvantages.

WB, Visual Disability, and AT

Sen [34] addresses WB from an individual perspective, asserting that it should be evaluated based on the achievement of functionings that each person values for a good life. He states, "functionings are seen as central to the nature of WB, even though the sources of WB could easily be external to the person." This aligns with the current notion of subjective WB, "a person feeling and thinking his or her life is desirable regardless of how others see it" [35], which is used by organizations and countries to monitor compliance with Sustainable Development Goal 3 [36,37].

Most data on this subject come from high-income countries, showing that persons with disabilities generally have lower WB than those without disabilities, often linked to poverty, educational status, unemployment, and social exclusion [38-41]. A recent study including data from LMICs confirmed this trend, attributing lower WB in persons with disabilities to "differences in living conditions and life experiences in relation to household wealth, level of education, partnership status, and exposure to violence and discrimination in the previous year" [42].

XSL•FO
RenderX

In the context of visual disabilities, common challenges include social isolation, limited mobility, low education, unemployment, stigmatization, and injury risks [43].

ATs are seen as tools to overcome these limitations and improve WB. Despite their potential benefits, not all persons with visual disabilities adopt AT, particularly mobility aids, due to concerns about stigmatization and inadequate obstacle detection [44]. This reluctance indicates that AT does not always enhance WB, underscoring the need to understand user interactions with AT to develop solutions that are more desirable and supportive of users' WB.

Aims of This Study

We explored the role of technology in the lives of persons with visual disabilities in LMICs and aimed to answer the following questions: (1) How can AT create opportunities or barriers in daily life? (2) How does technology affect WB?

Context of Fieldwork

Lived experience, disability, capabilities, and functionings should be understood within their sociocultural context [32,45]. The cultural perception of disability significantly influences both coping strategies and the willingness to use AT. These factors cannot be separated from the historical and political contexts that have shaped the experiences of a particular group [46].

This study took place in Colombia, a country affected by internal armed conflict for nearly 6 decades. The "Comisión de la Verdad (Truth Commission)" reported 8,775,884 victims of human rights violations in the conflict [47].

In 2016, a peace agreement was signed between the Colombian government and the FARC-EP (Revolutionary Armed Forces of Colombia). The historical memory construction process, a process of reconstructing and interpreting the past to dignify victims, clarify the truth about human rights violations, and promote reconciliation in the context of Colombia's armed conflict, highlighted patriarchy, racism, and discrimination based on social class, religion, or ideology as conflict axes [48,49].

Although Colombia adopted the Convention on the Rights of Persons with Disabilities (CRPD) in 2011, disability has been scarcely addressed in the peace agreement and subsequent documents [50]. While this study does not focus on recognized conflict victims, it is important to situate the participants' reality within this sociopolitical context. Living in a region with complex social issues, partly due to the conflict and other factors, means that numerous disadvantaged population groups exist. Recognizing persons with disabilities as one of the "worst-off" groups for prioritizing attention is challenging, especially priorities after the peace agreement priorities include attending to direct victims of the conflict, which continues with other actors present.

The interviews were conducted in Cali, Colombia's third most populous city and economic center. Cali is noted for its multiethnic population, sports culture, and as a cultural hub, but also as one of the most violent cities in Colombia [51].

While ATs can potentially expand the capabilities of people with disabilities, analyzing the geographic, social, political, and economic context of potential users during the design phase is crucial. This consideration is particularly important since using ATs could increase security risks for persons with disabilities.

Methods

Study Design and Theoretical Framework

This study adopted a qualitative narrative approach framed within a deductive qualitative analysis strategy [52]. The research was grounded in the CA [18] and the DT [21] as its core theoretical frameworks, with a critical theory lens [53] guiding the inductive reflection. This hybrid strategy enabled the exploration of how AT supports or constrains human capabilities and WB among people with visual disabilities in low- and middle-income settings.

Participants and Data Collection

This study involved 19 voluntary participants. Eligibility criteria included being 18 years or older; residing in Cali, Colombia, or nearby municipalities; and being blind or visually impaired, without additional hearing, motor, or cognitive disabilities. Eligibility was verified using the Washington Group Questions on Disability Statistics [54]. Purposive and snowball sampling strategies were used. Semistructured interviews (60 - 105 minutes) captured life narratives and good life aspirations. A follow-up focus group (n=5) explored their relationship with technology in greater depth, addressing (1) commonly used technologies, (2) perceived utility, (3) opportunities created, (4) desired opportunities, (5) risks, and (6) features enhancing WB.

To ensure rigor, the study used triangulation through a focus group, member checking with the participants, peer debriefings, and a detailed audit trail. Trustworthiness was further supported by verbatim quotes in the findings, which preserved participant voice and supported transparency.

Data Analysis

Overview

Data were analyzed using deductive qualitative analysis guided by the CA. A preliminary deductive codebook was developed based on Nussbaum 10 central capabilities, supplemented with key CA constructs, including resources, conversion factors (personal, social, and structural), agency, and achieved functionings.

Once these were applied, a second layer of deductive analysis was conducted using DT, focusing on identifying insecure functionings, corrosive disadvantages, and fertile functionings.

In addition to the 2-phase deductive stage, an inductive thematic analysis was performed to identify emergent themes not captured by the theoretical framework. This inductive analysis was critically informed by critical theory, allowing interpretation of power asymmetries in technology design, structural exclusion, and user agency.

XSL•FO
RenderX

Step-by-Step Coding Procedure

Familiarization

The research team read all transcripts multiple times, documenting initial impressions and emotional, social, and technological cues.

Deductive Coding: Phase 1 (CA)

Using Nussbaum capabilities, data were categorized into capability domains and analyzed with subcodes for functionings, resources, and conversion factors.

Deductive Coding: Phase 2 (DT)

Narratives were recoded to assess the security and fertility of functionings and to identify corrosive disadvantages (eg, stigma and economic precarity).

Inductive Coding

Emergent codes such as desire to appear attractive, resistance to confrontation, and wish to explore movement were documented, grouped, and interpreted using critical theory.

Integration

Inductive and deductive codes were reconciled through iterative discussions, using analytic memos. For instance, fear of confrontation was related to insecure functionings under "Control over one's environment," while vanity versus disability was interpreted within "Affiliation."

Synthesis

Final interpretation focused on how AT shaped each participant's capability set, aspirations, and obstacles to achieving the good life.

Disagreements in coding were addressed through iterative team discussions until consensus was reached. A reflexivity journal was maintained by the lead researchers to document positionality, epistemic tensions, and interpretive decisions. All coding was conducted manually without the use of software for qualitative analysis. Transcripts were originally in Spanish, and coding was performed in Spanish by native speakers. English translations were collaboratively developed to preserve semantic and cultural accuracy.

Codebook Summary

Deductive and inductive codes are presented for illustrative purposes (Tables 1-3).

Table . Deductive codebook (capability approach).

Capability	Subcapability	Code	Definition	Conversion factor type	Example quote
Bodily integrity	Bodily sovereignty	Privacy respect	Situations of physical or decisional invasion	Social	"...they touch you… that's not necessary. So, it's like they don't respect your bodily integrity..."
Bodily integrity	Mobility	Inaccessible transport	Lack of access to safe, navigable transport	Structural	"...the audio announcements don't work... so how are we supposed to know if the bus is the one we're waiting for?"
Affiliation	Dignity in service delivery	Discrimination	Being rejected or underestimated during services provision	Social	"...people assume I'm there to beg or ask for something."
Control over one's environment	Material control	Institutional barriers	Bureaucratic or policy-related access issues	Structural	"ATMs don't talk... it's both accessibility and safety."

XSL•FO
RenderX

Table . Deductive codebook (disadvantage theory).

DT[a] code	Definition	Inclusion criteria	Example quote
Insecure functioning	A functioning being achieved but threatened	Describes fragile access to opportunities	"I only go to the bank with someone I trust. ATMs are not accessible and it's unsafe."
Corrosive disadvantage	A single disadvantage that creates multiple deprivations	Barriers that multiply exclusions (eg, inaccessible transport affects work and autonomy)	"If I can't use public transport, I can't go to training, or to work. Everything falls apart."
Fertile functioning	A capability that enables many others	Mentions of tech or contexts that expand autonomy	"My cell phone is essential in my life. It helps me navigate, communicate with others, check when the bus is coming, know where to get off, and get any place."

[a]DT: disadvantage theory.

Table . Integration of inductive themes with deductive frameworks (CA[a] and DT[b]).

Inductive theme	Definition	Example (quote)	Related capability (CA)	Associated concept (DT)	Interpretive notes
Wish to appear attractive	The desire to look good and express personal style, challenging stereotypes of dependency.	"I consider myself vain despite my disability... so, I would buy more clothes... I would dye my hair...."	Affiliation and bodily integrity	Insecure functioning or symbolic disadvantage	Challenges cultural narratives of disability as asexual or unattractive; signals agency in self-image.

[a]CA: capability approach.

[b]DT: disadvantage theory.

Reflexivity and Strategies for Rigor

Reflexivity was maintained via journals kept by both main researchers. These captured assumptions, affective responses, and positionality reflections. Disagreements in coding were addressed through iterative team discussions and memoing. Credibility was supported through member checking and triangulation using interview and focus group data. Saturation was considered achieved after 19 participants, as no new themes emerged during the final interviews and confirmatory focus group.

Researcher-Participant Relationships and Empowerment Strategies

Although this study did not follow a formal participatory design framework, several intentional strategies were used to promote ethical engagement and participant empowerment. Informed consent was provided in accessible formats, including Braille and audio, with participants choosing the version that best suited them. Potential interview locations were identified by the research team and evaluated in consultation with key participants, allowing for informed, context-sensitive decisions.

Participants were clearly informed of the study's aims and the distinct roles of the research teams. Notably, the team responsible for data collection was not involved in AT design, helping to manage expectations and avoid instrumental relationships. Interviews were guided by an ethic of care, privileging participant-led narratives and emphasizing trust, autonomy, and open dialogue.

We also approached participation critically. While participatory methods are often promoted in inclusive design, they may reinforce power asymmetries if not carefully implemented. As Cooke and Kothari [55,56] warn, such approaches can result in a "tyranny of participation," serving dominant interests rather than those of marginalized groups. To avoid this, our study prioritized epistemic justice by centering participants' lived experiences as ends in themselves, not as inputs to a predefined design process. Member checking and a focus group enabled deeper reflection and participatory analysis, contributing to more equitable and meaningful knowledge production in line with the ethical commitments of the CA [17].

Ethical Considerations

The research protocol was approved by the Swiss Federal Institute of Technology Lausanne on research involving humans (056 - 2021 and 068 - 2022), the Human Ethics Committee of Universidad del Valle (008 - 022), Hospital Universitario del Valle (029 - 2022), and Instituto para Niños Ciegos y Sordos del Valle (CEI-2022 - 02), ensuring fully informed consent and participant anonymity. In accordance with ethical guidelines for research involving human participants, compensation was provided in a manner that did not constitute undue inducement. Participants received reimbursement for transportation costs and were offered food and beverages during data collection sessions. Additionally, they received a modest stipend of approximately 40,000 COP (equivalent to less than US $10) as compensation for their time and participation.

XSL•FO
RenderX

Results

Overview

This section presents the main findings of the study, starting with a general description of participant demographics (Table 4) to provide contextual background. The results are then organized into 2 main areas: the first presents the outcomes of the deductive analysis guided by the CA, structured under the subheadings "Notions of Well-Being and Autonomy," "WB and the Good Life: Common Functionings Identified," and "Lived Experiences through the Lens of Capabilities and Technology." The second area includes the inductive themes that emerged beyond the predefined theoretical framework. Although the results of the secondary deductive analysis informed by DT—building upon the initial CA coding—are not included in this section, they are presented in the discussion, where they more effectively enrich the interpretation of structural constraints, genuine opportunities, and the security of valued functionings within participants' lived realities.

Table . Participant demographics.

Participant characteristics	Overall
Age (years), mean (SD)	31.1 (10.9)
Sex, n (%)	
Female	7 (36.8)
Male	12 (63.2)
Marital status, n (%)	
Single	13 (68.4)
Common law	3 (15.8)
Married	3 (15.8)
Education, n (%)	
No formal education	1 (5.3)
Primary	1 (5.3)
Completed high school	6 (31.6)
College or university in process	5 (26.3)
Completed college or university	6 (31.6)
Employment status	
Employed	6 (31.6)
Unemployed	6 (31.6)
Self-employed	2 (10.6)
Student	5 (26.3)
Origin	
Urban	17 (89.4)
Rural	2 (10.6)
Self-reported ethnicity	
Afro-Colombian	2 (10.6)
None	17 (89.4)

Notions of WB

In exploring what constitutes WB for the participants, we inquired about their ideal life and contrasted this with their current experiences. It quickly became clear that disability significantly influences their conception of an ideal life. A couple of participants explicitly noted that their vision of an ideal life had to be "edited" due to their disability, such as stating "my ideal life now that I am blind" or recognizing that their disability constrains their aspirations.

Autonomy and WB

Participants identified autonomy as central to their WB, emphasizing the importance of freedom in decision-making and the ability to make meaningful choices. One participant articulated this need for autonomy as follows:

> *A more assertive life would be one in which people first see me as an equal, rather than focusing on my condition. It would be a life where I can make all my own decisions, where having a disability does not limit my opportunities.* [Male, 45 years]

XSL•FO
RenderX

The emphasis on autonomy extends to living without the constraints of disability-related prejudice and ensuring social acceptance. However, the frustration with societal limitations often leads to a pragmatic view of WB, focusing on achievable goals within their control and avoiding, in the process, being assisted by others. For example, a 29-year-old participant described her ideal life as

> *To have as much autonomy as possible—being confident to go out independently, having a stable job that meets my needs, maintaining good health, and practicing my sport. Total autonomy in my life is what I strive for.* [Participant, 29 years]

WB and the Good Life: Common Functionings Identified

Work is frequently cited as a critical element of WB. It is valued both for providing financial security and as a source of personal fulfillment. Achieving financial stability and having a fulfilling job are integral to their idea of a desirable life. As one participant who has several part-time jobs with limited hours a week, put it,

> *I have a job and I am grateful, but I would like to improve my working conditions to better provide for my family and ensure our own home.* [Female, 28 years]

Ownership of a home is another significant aspiration, often linked to job stability. Other desired functionings include pursuing education, maintaining good health, and excelling in sports. These aspirations reflect a broader understanding of WB that includes personal and professional fulfillment.

Moreover, family stability, emotional support, and social connections are also important components of WB for participants.

When one participant (male, 40 years) compared his current life with his ideals, he became aware that he had achieved much of what he considers a flourishing life. He reflected on his accomplishments with satisfaction while acknowledging areas for improvement. Despite this, he emphasized that his experience is not universal among persons with disabilities and that the limitations imposed by disability affect what can be achieved.

> *At this point, as a person with a visual disability, I believe I have achieved a great deal, despite the challenges. I see others who, unfortunately, lack employment, family, or social connections, or who haven't been able to fully enjoy life. For me, job stability would make my life 100% ideal. I feel fulfilled because I've experienced love, marriage and the journey of being a parent. I have also traveled, engaged socially, and even gone bungee jumping.* [Male, 40 years]

Though the desire for family formation or partnership is less commonly mentioned by those who were single at the time of the interview.

Lived Experiences Through the Lens of Capabilities and Technology

In examining Nussbaum's central capabilities in the context of individuals with visual disabilities, the analysis reveals both opportunities and barriers that impact their daily lives.

Life

This capability is initially affected by the sociopolitical context of the country and specifically by the social and infrastructural conditions of the city. As one participant recounted,

> *We left Cauca, displaced by the violence. We also lived in Putumayo for a few years, but in 2009 they killed my brother, and we had to leave again—displaced by the violence.* [Female 21 years]

(Cauca and Putumayo are departments in southern Colombia that have been heavily affected by armed conflict.)

In some cases, participants explicitly linked their visual impairment to episodes of violence. One participant explained:

> *I've been living in Cali for four years, which is also how long I've had this condition (low vision). I was living in my village at the time, and I came here to study. It was around then that the incident happened — someone attacked me to steal my phone and other belongings.* [Male, 23 years]

Similarly, another participant shared:

> *I lost my sight when I was 20 years old, as a result of a firearm injury. After that—well, thank God—I recovered well after about a year. It was like a sabbatical year, where I didn't really do anything, because at the time, I didn't know how to get my life back on track.* [Male, 29 years]

Beyond the direct consequences of violence, participants also pointed to everyday urban dynamics that fail to prioritize the safety and dignity of people with visual disabilities. In particular, unsafe social behavior—such as the frequent disregard for traffic rules by drivers—emerged as a recurring concern. Falls, collisions, and traffic accidents were identified as major threats to personal safety and mobility. Several participants described incidents in which they had sustained serious injuries.

> *I fell into the gap and ended up on the road, between the station platform and the bus.* [Female, 31 years]

One participant told during the interview that her brother, who was also blind, was the fatal victim of a traffic accident 2 months earlier, when he was run over by a car while crossing the street.

Bodily Health

Participants have access to the Colombian health system and benefit from rehabilitation services that focus on orientation and mobility. Participants from rural areas only accessed these services once they arrived in Cali. Regarding access to basic AT, although there is a route to obtain it through the health system, a couple of participants reported that it is cumbersome. However, they say that canes are more readily available through the mayor's office.

XSL•FO
RenderX

Financial instability has led to periods where participants' nutrition and living conditions have been compromised. As the following quotes illustrate:

> *When I was little and still living with my mom and siblings, we got evicted because we couldn't pay the rent anymore. We ended up moving into the first place we found, but it was full of rats—and I'm really scared of them.* [Female, 54 years]

> *There were hard times, times of real scarcity. Sometimes you had to really stretch things—like cutting back a lot on food—just to pay for basic services, because skipping those payments wasn't an option.* [Female, 21 years]

Bodily Integrity

Participants have received training that enables some degree of autonomy in navigating the city using a white cane. Nonetheless, the freedom to move around safely is frequently compromised by risks of physical injury and insecurity.

Technology plays an essential role in facilitating mobility for participants, with apps aiding navigation and transportation. However, while most participants own smartphones, many do not have consistent access to mobile data to ensure these apps function effectively.

Some participants limit navigation to familiar areas or rely on human assistance due to safety concerns and technological limitations. Experiences of mugging and other safety risks further hinder their mobility, independence, and willingness to use technology in public.

> *My phone was stolen at the station while I used a transportation app to check my bus arrival time for work. It was early, and I thought I was alone since I heard no one nearby.* [Male, 40 years]

Others are not unaware of the risks in public spaces—they just have a lower perception of the danger, or they choose to manage the risk rather than give up the technology altogether. Instead of avoiding it, they take precautions to reduce the chance of getting robbed.

> *I'm one of those people who says that blind folks don't get robbed — and yeah, people laugh when I say that, hahaha. Sure, I know plenty of blind people who've been robbed — but that's because they gave papaya (in Colombia, that means making yourself an easy target). You just can't give papaya! I've been in dangerous situations before, but thank God, I've always gotten out safely. I've never been robbed.* [Male, 39 years]

During the focus group, when participants discussed the disadvantages or risks of using technology, the first issue that came up was personal safety. As illustrated in the following quote:

> *The first thing that comes to mind is safety. For example, around here everyone talks about Lazarillo and all those apps—but I don't use any of them, because I don't take my phone out. We become very visible when we pull out our phones, hold them to our ear, or start using them. That makes me an easy target for someone to steal it. Given the context we live in, it just doesn't work for me—it actually makes me feel even more unsafe.* [Male, 37 years]

After sharing stories about people they knew who had been victims of phone theft, participants went on to describe the strategies they use to avoid becoming targets themselves. For example, the 2 female participants said they use headphones so they do not have to take their phones out of their pockets, and they hide the headphones with their hair.

Safe mobility is essential for achieving various functionings, but its impact extends beyond simply moving freely and avoiding aggression. One participant noted that communication is a major challenge for those unable to move independently. He added that autonomous mobility is crucial for meaningful interactions, including forming personal relationships, such as finding a partner.

> *Not walking alone makes others, even friends, perceive a person with a disability differently. They already have questions about your disability, which you must address when interacting with them. However, if they see you always relying on someone, it becomes much harder to connect with them.* [Male, 24 years]

Several interviewees noted that being a woman heightens their fear during independent travel, whether on public transport or walking in the streets. One participant shared,

> *Regardless of disability, ethnicity, or economic status, being a woman increases the risk of aggression, especially sexual violence. Being a woman in Cali, and in the areas I frequent, is terrifying. People often mock me with idiotic comments. Just the other day, someone said, 'Oh, you're so pretty, but what a pity about that stick (the cane).* [Female, 21 years]

Senses, Imagination, and Thought

With the exception of 2, all participants completed high school, and about half (n=10) pursued higher education. AT, such as screen readers and text-to-speech apps, are widely used. However, those who studied before advanced technologies faced significant barriers, such as manually converting documents to audio or advocating for accommodations. Despite these challenges, most described their education positively but noted limited career options due to inaccessible educational and professional pathways. As a 19-year-old participant explained,

> *I chose law by default. I picked it because I'm good at public speaking, but do I truly enjoy it? No. Our options are limited, even if they seem abundant. While barriers can be overcome, I'm not willing to exhaust myself doing so.* [Participant, 19 years]

Practical Reason

Participants' decision-making reflects strong priorities and resource management, aided by technology for accessing information and connecting with advocacy groups. However, uncertainty and anxiety about the future, especially among those

XSL•FO
RenderX

newly adjusting to disability, hinder long-term planning. As one 19-year-old male shared:

> *I don't think about the future. It doesn't exist—it's just conjecture, you know? I've learned not to stress over it. For me, the future would mean a stable job, continued training, and… that's it.* [Male, 19 years]

While technology plays a vital role in accessing information and social networks, accessibility issues remain a significant challenge.

Affiliation

People mostly connect with family members and others with visual impairments, but they find it hard to build relationships beyond those close groups. Strong support networks often form in educational or social contexts, such as public libraries, and through cultural or adapted sports activities. However, barriers persist, including discrimination, ableist remarks, and inaccessible public infrastructure, which often prioritizes property over safety.

> *Do you think people respect the few podotactile guides that exist? No, they park their cars on them. If you get too close, they say, "You're going to scratch my car," forcing you onto the road and into danger.* [Female, 54 years]

Participants also face employment challenges, frequently working in roles they are overqualified for or in part-time positions with insufficient pay. Visual impairment complicates social interactions, as one participant noted:

> *Sometimes relationships begin with a look, but without that, it's already one point less, isn't it? And then there's the hesitation people feel toward disabilities.* [Male, 45 years]

Accessing services adds to these difficulties. City buses equipped with loudspeakers frequently have these systems turned off, while shopping malls present physical accessibility barriers; additionally, salespeople often fail to recognize persons with disabilities as customers.

> *I've never gone shopping alone. I have no way to orient myself in a mall, and I don't like arriving somewhere unfamiliar and having to wait for someone to help me. Salespeople usually don't realize that a blind person with a cane can be a customer.* [Female, 31 years]

While technology, such as social networks and apps, has improved accessibility, significant issues remain, such as inaccessible ATMs, banking apps, and dating platforms such as Tinder (Match Group, Inc), where user photos lack proper descriptions.

Other Species

Participants express satisfaction with their ability to engage with nature through their other senses. Despite some limitations, they continue to find joy and fulfillment in their sensory experiences and interactions with the natural world.

Play

Available opportunities include participation in adapted sports and recreational activities supported by local initiatives. Participants enjoy activities such as listening to music, radio programs, and audio-described films. However, financial limitations restrict their access to a broader array of recreational options.

Control Over One's Environment

Political involvement is marked by participation in advocacy activities and voting, though challenges such as discriminatory treatment and the need to navigate physical barriers persist.

Material control is constrained by the scarcity of full-time employment and inadequate financial stability. Issues with banking and consumer services, compounded by inaccessibility and discrimination, affect their ability to manage material resources effectively.

Themes Identified Through the Inductive Analysis

The inductive analysis of interviews and focus group discussions revealed 3 interrelated themes that reflect participants' subjective experiences and priorities in everyday life. These themes include the aspiration to express and expand their movement potential, the importance placed on maintaining a positive and attractive personal appearance, and the strategies used to avoid confrontation and maintain social harmony.

Desire to Exploit Movement Potential

A prominent theme in the interviews was the frustration many participants felt regarding their limited mobility. They expressed this frustration over their inability to fully experience freedom of movement, navigate the city and its spaces because of infrastructural barriers, and socialize due to a lack of bodily experience. Additionally, they cited deficiencies in ADs as a contributing factor.

Participants frequently noted that the cane, while essential, places limitations on their freedom of movement and enforces a rigid posture, making their gait different from others.

> *The cane forces you to maintain a fixed posture, which limits your ability to move naturally. Unlike others who can easily turn and walk freely, we can't move the same way.* [Male, 47 years]

Many expressed a strong desire to engage in motor activities that are currently restricted for them, such as riding a bicycle, running, swimming, dancing, and skating. There was a clear enthusiasm for speed.

> *I long to run, but to run alone. I yearn to walk freely, without a cane, moving as I please. I would love to ride a bicycle.* [Male, 47 years]

In the realm of dance, beyond the desire for self-expression through movement and the joy of dancing itself, this activity serves as a crucial means of socialization—especially in this city, renowned as the salsa capital.

> *Well, I like to dance but I don't know how to do it well. I love dancing. ...I do it my own way, in a clumsy*

XSL•FO
RenderX

> *way, but I do it. ... Yet, I would love to master dancing.* [Female, 56 years]

In terms of technology, while some participants acknowledged the crucial role of the cane, they expressed a need for an AD that would enhance their mobility with greater security and fluidity, without compromising their sense of independence or human qualities. Additionally, they emphasized the importance of keeping at least one hand free, as the other one is currently occupied by the cane.

Although many express hope and enthusiasm for the potential of AT in enhancing mobility, others caution that the technology may not address the factors that limit their ability to navigate the city, particularly those related to human interaction. As one participant shared in the following quote:

> *Many variables are involved. The recklessness of drivers is one of them. I've heard that in other countries, people really respect traffic signs and regulations. If that were the case here, one could cross the street when the light is red, because you could trust that drivers would obey the rules. But that's not the case here. Traffic lights are often ignored. Motorcycles also present a significant issue. Even when there's a pedestrian crossing, not all vehicles line up properly. If they did, I would know the boundaries of the space I need to cross. Unfortunately, that doesn't happen. When I'm crossing at a traffic light, I may plan to pass on one side, but I find one car ahead of another, and then a motorcycle suddenly appears.* [Male, 40 years]

Desire to Look Attractive

The theme of looking attractive surfaced notably, particularly among female participants. One participant (female, 29 years) expressed frustration with online shopping due to insufficient garment descriptions. She also desired AD that would enable her to perform personal grooming tasks independently, such as painting her nails and plucking her eyebrows, activities she used to do on her own before acquiring a visual disability.

Another participant (female, 27 years) mentioned that if she had the financial means, she would purchase more clothes and enjoy activities such as hair dyeing and makeovers. Regarding mobility AT, she expressed a desire for a discreet or stylish device that could detect high obstacles and match her clothing. She confessed, somewhat apologetically, "I´m somewhat vain, despite my disability."

A 21-year-old female participant noted her reluctance to use the cane when she had the option of human assistance, explaining that the cane did not complement her style.

> *I think it's more like this: it's not as if I say, 'Oh, that cane matches my style, how exciting, man!' No, it's not like people come up and say, 'What a cool cane, it goes great with your shirt.* [Female, 21 years]

The desire to appear attractive also extends to social and romantic interactions. Participants revealed that concerns about attractiveness influenced their willingness to use certain AT. For example, one participant (male, 40 years) described how societal norms and personal shame affected his acceptance of using a cane.

> *Even though I knew how to get around with my cane, I always relied on my grandmother, who met me at the bus stop every day. One morning, I couldn't let her know I'd be early, so I ended up waiting three hours instead of using my cane. That experience made me realize I needed to start using it. Honestly, the shame I felt was mostly about what the opposite sex might think of me. I guess those feelings are pretty common for young people, shaped by societal pressures and insecurities.* [Male, 40 years]

A female participant (57 years old) noted that her daughter's initial embarrassment about using a cane was linked to teenage concerns about appearance:

> *She felt embarrassed to use the cane, especially because, as girls start to become aware of their attractiveness to boys, they may feel self-conscious. However, she eventually came to accept it.* [Female, 57 years]

Another female participant noted that her embarrassment about using the cane is heightened when she is around peers of her age and the opposite sex who do not have disabilities. She mentioned that she has rarely discussed this issue, as she recently recognized her feelings of shame and struggled to openly acknowledge them.

One male participant (45 years old) described his experience after trying a prototype of an AD that included glasses equipped with an obstacle-detection camera, expressing strong reservations.

> *I wouldn't even go out in public wearing those. First, they pose a danger—I could get robbed. Second, they look terrible. If you wore them to a party, do you really think anyone would find you attractive?* [Male, 45 years]

The Bid to Avoid Confrontation

The theme of avoiding confrontation appears in various contexts throughout the study. For instance, some participants avoid expressing their political views to prevent arguments or physical confrontation.

One participant noted that while he is open about his political convictions, he often faces derogatory remarks about his disability when others disagree with his views.

The desire to avoid confrontation also emerges when participants describe their experiences with ableism. Many report feeling frustrated by ableist treatment but choose to remain silent to avoid conflict. For example, a participant shared an experience:

> *They see you with a cane and automatically lock you in the world of the blind, of the <what a pity>, of <how do you manage to walk by yourself?>. Also, they think they have the right to ask questions and make silly expressions. One day I was with my cane, I sat in the Bus and a woman came up behind my ear*

XSL•FO
RenderX

> *and said to me: <You have to pray a lot so that the spirit of blindness leaves your life>* [Female, 29 years]

Such incidents often deter individuals from using ADs, fearing further stigma.

Another participant recounted an incident where he felt threatened after unintentionally pushing another man while stumbling in a public bathroom. Since he was in a familiar location, he was not using his cane. To prevent a confrontation, he apologized and explained his condition.

This theme underscores the dual role of the cane. While it can signal vulnerability and deter aggression, it can also trigger stigmatization. Participants sometimes avoid using the cane to prevent negative reactions, which may impact their overall functionality and social participation.

Discussion

Principal Findings

The discussion begins by highlighting the centrality of autonomy in participants' conceptions of WB and its implications for AT design ("The Good Life and WB"). It then reorganizes functionings into 3 analytical categories: "Most Valued Functionings," "Insecure Functionings," and—based on their enabling or limiting effects—"Corrosive Disadvantages" and "Declustering Disadvantage." This is followed by a review of participants' experiences with specific technologies in comparison to existing literature ("AT and the Creation of Opportunities and Barriers in Participants' Daily Lives"). The final sections present emergent sociocultural themes ("Themes Identified Through the Inductive Analysis and AT") and integrate all theoretical strands into a situated analysis of agency, risk, and design ("AT Between Opportunity and Constraint").

Main Results

The Good Life and WB

This study aimed to examine how AT creates opportunities and barriers in participants' daily lives and influences their WB. We first explored the concept of an ideal life, which informed our understanding of participants' perceptions of a good life and WB.

This valued way of life includes, but is not limited to, WB [57]. WB, on the other hand, is assessed in terms of the achievement of functionings that people value and incorporate into their conception of a good life [34].

While interviewees widely valued agency, opinions on autonomy varied. A few participants preferred autonomy over independence, noting that dependence is a natural part of life. However, others viewed autonomy as minimizing reliance on others for daily activities.

This distinction is especially important when considering how AT is developed for persons with disabilities. While the desire for independence is a valid and often central driver in AT design, it can sometimes overshadow other dimensions of daily life. Bennett et al [58] argue for reframing AT through a lens of interdependence, emphasizing that technologies are embedded in relationships, routines, and environments—not isolated solutions. Similarly, Joskow et al [59], using the CA, show that the value of AT lies not only in enabling functional tasks but also in supporting participation, emotional WB, and meaningful choice. These perspectives highlight the need to design AT with a more holistic understanding of how persons with disabilities navigate and shape their worlds.

This becomes particularly evident when comparing the experiences of different users. One male participant, reflecting on the challenge of crossing at traffic lights, shared, "An AT device wouldn't work as well for that kind of situation, because it's still dangerous." Even with a guide dog, he noted, it's hard to manage the risks posed by reckless drivers. He also described the value of informal social interactions that occur in public space, "Sometimes I walk a bit on my own and someone will say, 'Hey, where are you headed?'...I've even met people that way—people who later became friends."

In contrast, a female participant expressed a clear preference for technological assistance over human help, "I would much rather have a device than a person helping me, because I'd feel like I'm working together with the device—and that makes me feel more autonomous...it's perfect!" She added, "Regardless of its size or how it looks, having a tool that helps us move around independently is incredibly important." Even if the device is "big, ugly, or heavy," she saw it as empowering—what mattered most was the sense of self-reliance it enabled.

Both perspectives offer valuable insights for AT design. Listening to this diversity of lived experience helps avoid defaulting to a one-size-fits-all model centered solely on individual autonomy. While preferences like those of the second participant—who expressed enthusiasm for any device that enhances independence regardless of size or appearance—may seem appealing to designers because they reduce complexity, relying solely on such responses can be misleading. If designers prioritize these simplified preferences without accounting for broader user expectations—such as comfort, discretion, aesthetics, or social dynamics—they risk developing yet another device that ultimately fails to meet the needs of a wider group of potential users. Truly inclusive design must engage with this heterogeneity, acknowledging not only aspirations for autonomy but also the embodied, emotional, and relational dimensions of navigating everyday life.

Most Valued Functionings

Among the functionings most valued by participants for a good life, material security—primarily through work as a source of income—stood out. Other key functions included affiliation, education, sports, health, and mobility. This aligns with Wolff and De-Shalit findings, which identify core categories of functioning, including life, bodily health, bodily integrity, affiliation, environmental control, and senses, imagination, and thought [21].

Insecure Valued Functionings

Although these functionings are highly valued by participants, they remain insecure for a significant proportion of them. Employment challenges are prevalent, reflecting findings from other studies that report low employability, poor working

XSL•FO
RenderX

conditions, lower wages, and job instability for people who are blind or have low vision [60-63].

Similarly, while participants emphasize the importance of family and social networks, they also highlight barriers to full inclusion and social participation, making affiliation another insecure functioning. These challenges are consistent with existing literature, which associates limited social engagement with reduced WB and identifies significant social stigma and restricted access to goods and services for persons with visual disabilities, in both public and private spaces [62,64-67].

Given the violence, physical barriers in urban areas, and societal attitudes, life and bodily integrity also emerge as insecure functionings. Empirical studies support these findings, citing falls, life-threatening injuries, and higher mortality risks among persons with visual disabilities [68,69].

Education, including higher education and vocational training, is generally seen as a secure functioning. Despite Colombia's high illiteracy rates among persons with disabilities, participants in this study appear to have attained higher levels of education than the national average for persons with disabilities [70,71].

Finally, although all participants have health insurance, some have faced precarious situations that question health's security as a functioning, particularly given the material insecurity caused by unemployment.

Corrosive Disadvantages

Overview

Efforts for social justice should prioritize groups experiencing multiple disadvantages. Participants highlighted unemployment, lack of affiliation, and mobility barriers as significant corrosive disadvantages.

Unemployment or underemployment threatens basic functionings such as housing, food, health, and life, while also limiting leisure, education, and recreation due to financial constraints. Studies show that income loss exacerbates social isolation, diminishes status, and is linked to reduced WB, limited access to healthcare and education, declines in health, and increased mortality risk [72-75].

Insecurity in affiliation similarly undermines other functionings. Participants often face disrespect and unequal treatment, which compromise job security, environmental control, and freedom of movement without risking safety and WB. Research links social exclusion with limited participation in economic, social, and cultural life, resulting in poverty, unemployment, inadequate health care, and poor housing for persons with disabilities [67,76-78].

Restricted mobility constitutes another corrosive disadvantage, as it directly undermines functionings such as access to essential services, physical activity, communication, WB, and social inclusion. Research highlights the importance of sensorimotor abilities in social interaction [79,80], with decreased mobility linked to isolation and lower satisfaction [81].

Declustering Disadvantage

Identifying corrosive disadvantages highlights the importance of fostering fertile functionings—those that support resilience and WB. Participation in adaptive sports is a clear instance, as research shows that each additional year in adaptive sport is associated with a 4% increase in employment likelihood among people with physical disabilities [82]. In our study, involvement in sports provided income, travel subsidies, and social connections. It supported capabilities such as affiliation, bodily integrity, and control over one's environment.

Employment also served as a fertile functioning when conditions were fair, remuneration equitable, and respect for dignity was present. Studies indicate that AT access in the workplace is strongly linked to employment retention and productivity among individuals with visual disabilities [83].

As pivotal conversion factors, AT tools such as screen readers and accessible mobile apps enabled participants to use digital platforms daily—facilitating email, social media, and work tasks [84]. As a result, these technologies not only support function, but also reinforce confidence and autonomy.

Although participants used diverse ATs, this paper focuses on representative examples. Further research should explore a broader array of technologies and how they foster capability and reduce disadvantage across different daily contexts.

Mobility technologies warrant particular attention. Studies show that interactive navigation tools can significantly enhance users' spatial confidence and independence [85]. Participants emphasized that AT, which enables seamless environmental interaction—not just orientation or obstacle avoidance—enhances their movement potential, autonomy, self-efficacy, and engagement with public space.

AT and the Creation of Opportunities and Barriers in Participants' Daily Lives

AT has opened multiple avenues for expanding capabilities among persons with visual disabilities, supporting valued functionings across educational, occupational, and social domains. Participants in this study described how information and communication technologies have enhanced their performance in academic tasks, enabled participation in the labor market, and expanded opportunities for leisure and connection. These tools facilitated informed decision-making, strengthened social networks, and increased personal visibility in both professional and interpersonal contexts [86].

Mobility technologies, including white canes and mobile navigation apps, were also perceived as crucial for promoting spatial autonomy. They allowed individuals to traverse their environments more independently, access public and private spaces, and engage in civic and social life. AT also enabled participants to perform essential tasks such as banking and online shopping, thereby enhancing their control over daily routines and improving their sense of independence.

However, participants also identified significant barriers to AT use—barriers that highlight the ambivalence of technology in their daily lives. As Moser [87] argues, technology does not simply empower—it also has the capacity to withdraw agency

XSL•FO
RenderX

by rendering difference hypervisible or by reinforcing stigma. This tension was evident in the accounts of participants who, while recognizing the utility of AT, expressed concern about its potential to attract unwanted attention, particularly in contexts marked by high levels of urban violence. In these settings, visibility becomes a liability, increasing the risk of theft or social targeting.

This ambivalence was especially pronounced in public spaces. While some participants preferred discreet technologies that do not mark them as disabled, others acknowledged that visible devices, such as white canes, could help signal their needs and elicit assistance. These findings resonate with Dos Santos et al, who found that aesthetic and symbolic dimensions strongly influence the acceptability and use of AT [88]. Similar insights have emerged in studies from the Global South, such as Barbareschi et al [89] work in Kenya, where ADs can either mitigate or exacerbate stigma depending on the sociocultural context.

Thus, the potential of AT to support WB is conditioned not only by its functional properties but by its alignment with users' social environments and identities. When well-designed and contextually appropriate, AT can foster fertile functionings—such as autonomy, affiliation, and environmental control. However, when AT is experienced as stigmatizing, risky, or misaligned with local realities, it can become a source of discomfort or even exclusion. This dynamic reinforces the importance of attending to not just technical efficiency, but also social safety, emotional meaning, and contextual sensitivity in design and policy.

Themes Identified Through the Inductive Analysis and AT

In Latin America, physical appearance holds significant value, rooted in colonial legacies that entrenched hierarchies by emphasizing differences between colonizers and the colonized. These hierarchies associated power with traits such as being European, patriarchal, White, wealthy, heterosexual, and able-bodied. Disability, conversely, was linked to negative characteristics such as unproductivity and rebellion, devaluing disabled bodies and dissociating them from beauty [90,91]. This detachment was reflected in interviews, where participants either avoided discussing physical appearance or addressed it obliquely. One participant remarked, "I am somewhat vain despite my disability."

The desire for attractiveness was particularly pronounced among women, aligning with research that connects beauty ideals to women's roles in Latin American cultures [92,93]. In Colombia and Brazil, beauty is often tied to socioeconomic status and professional opportunities [94,95]. While beauty remains culturally valued, ADs such as canes—symbolizing disability—are perceived as aesthetically unappealing and are sometimes rejected.

Aesthetics influence social interactions beyond economic implications. Participants expressed a desire to appear attractive to the opposite sex but seldom mentioned romantic aspirations, potentially reflecting aspirational frustrations. This tension between appearance and relational expectations emerged more clearly in the focus group discussions, which revealed the need for greater accessibility on dating platforms such as Tinder, highlighting participants' openness to using AT to facilitate social interactions and romantic relationships.

Limited mobility emerged as a recurring challenge, restricting participants' navigation skills and engagement with urban environments. Kukla's concept of spatial agency—the ability to autonomously use space—captures this dynamic. City navigation requires interaction with people and environments, relying on cues such as body language, eye contact, and street signs [96]. Existing mobility AT primarily addresses orientation and obstacle detection but lacks tools for effective environmental engagement. Participants' adaptive strategies highlight the need for improved AT design. Kukla also emphasizes that movement pace and aesthetic norms influence spatial navigation [96]. Participants underscored that AT should enhance autonomy while enabling natural interaction with tasks and people, without disrupting rhythms.

AT design must prioritize the user's holistic experience, accounting for the diverse environments they navigate [97]. This context extends beyond controlled settings, incorporating cultural meanings of disability, prior interactions, and broader social, historical, and political factors. The social realities of designers in the Global North often differ significantly from those of persons with disabilities or users in the Global South, as do their perceptions of aesthetics and its functional role in daily life.

AT Between Opportunity and Constraint: A Situated Analysis of Agency, Risk, and Design

In Colombia, defying dominant social norms often entails serious consequences, frequently enforced through violence. This violence is not limited to state institutions; nonstate actors—including armed groups and criminal organizations—have also enforced normative boundaries. Individuals who challenge socially constructed expectations—such as LGBTIQ+ (lesbian, gay, bisexual, transgender, intersex, and queer) persons, Indigenous peoples, Afro-Colombians, and social leaders—are disproportionately subjected to threats, forced displacement, and even assassination [98,99]. In cities such as Cali, such patterns have at times been reinforced by criminal governance structures that monitored daily life and punished perceived transgressions, exerting coercive control over entire neighborhoods [100]. In these contexts, expressing a nonnormative identity or dissent becomes a high-risk act, revealing how violence operates as a mechanism of civic regulation.

This coercive environment encourages conformity not through democratic consensus but through fear. Individuals internalize the threat of retaliation, adapting their appearance, behavior, and speech to avoid becoming targets [98]. Over time, this normalizes silence and compliance, particularly in communities with limited institutional protection. Violence, therefore, becomes a structuring force that shapes notions of "acceptable" citizenship and sustains systems of exclusion.

Within this broader context, ableism functions as a parallel yet distinct form of disciplinary control. In Colombia, persons with disabilities who defy normative expectations—by asserting

XSL•FO
RenderX

rights, demanding visibility, or rejecting passive roles—often face institutional neglect, attitudinal bias, and spatial exclusion. The Constitutional Court has described this as the product of "a segregating and exclusionary culture" that restricts life trajectories for persons with disabilities [101]. In conflict-affected areas, civic engagement can become a risk in itself, as some persons with disabilities avoid visibility out of fear of losing access to essential services [102]. Ableism reinforces conformity not through overt coercion, but through anticipated punishment—governing behavior, silencing dissent, and normalizing invisibility [103].

From the perspective of the CA, these exclusionary and coercive dynamics represent severe constraints on individuals' real freedoms to pursue lives they have reason to value. Structural violence, institutional neglect, and everyday surveillance undermine central capabilities such as bodily integrity, affiliation, and control over one's environment [18]. These forces not only restrict access to resources but also impair agency itself, fostering anticipatory compliance and silencing non-normative expression.

While the CA emphasizes agency and meaningful choice, it has been critiqued for undertheorizing the sociopolitical and ideological conditions that suppress agency—particularly in contexts marked by systemic domination and fear [20,104]. Although Sen acknowledges power structures through the notion of conversion factors and has addressed issues such as identity-based violence and gender inequality, he provides only a partial theorization of how social and institutional forces constrain choice [20]. In response, critical theorists emphasize that agency is always situated—shaped by hegemonic ideologies, historical inequalities, and structural constraints [20]. These insights enrich the CA by highlighting how oppression is not only material but also discursive and normative.

AT becomes a key analytical entry point in this discussion. Within the CA, AT functions as both a resource and a conversion factor—potentially expanding real freedoms by enabling valued functionings such as mobility, communication, and social participation [9,18]. Yet in contexts shaped by criminal governance, stigma, or weak institutional presence, AT's capacity to enhance capabilities may be compromised. Fear of surveillance, ridicule, or exclusion can deter individuals from using AT, effectively transforming it into a corrosive disadvantage—a concept from DT that refers to resources that worsen users' conditions in hostile environments [21].

In cities such as Cali, where high levels of urban violence intersect with limited protection for vulnerable groups, participants reported avoiding certain technologies out of fear for their personal safety. While navigation apps and mobile-based tools might enhance autonomy in principle, in practice, they often increase visibility and perceived vulnerability in public space. As shown in the findings, such risks lead many to forgo potentially beneficial technologies, not because of a lack of interest or understanding, but because the act of using them carries a social and physical cost that undermines the very freedoms they are meant to support.

This reflects a broader pattern in which structural violence becomes embedded in everyday decision-making. Technology adoption is shaped not only by access and design, but also by how individuals negotiate exposure, safety, and autonomy in hostile or unstable environments. In this sense, the very tools intended to promote inclusion may reinforce exclusion when they are not aligned with the social conditions and lived experiences of their users.

Moreover, AT design and policy often reflect biomedical or market-driven logics, failing to account for cultural preferences, affective experiences, and social risks. Research shows that when AT is perceived as overly medicalized or aesthetically stigmatizing, it is more likely to be abandoned—despite its functional benefits [105,106]. In Latin American contexts, aesthetic appeal and discretion are not merely preferences but protective strategies in public spaces marked by insecurity. Orellano-Colón et al [105] found that Hispanic users preferred AT devices that did not draw attention, valuing discretion and aesthetics alongside functionality. In Colombia, where individuals must navigate public visibility with caution, design decisions can significantly impact the acceptability and sustainability of AT use.

Emotional responses—such as shame, pride, or resignation—often remain unspoken but can be inferred from narrative silences, reflecting adaptive preferences. Rather than ignore these silences, designers should interpret them as meaningful indicators of unmet needs—shaped by structural disadvantage—and take them into account in participatory design.

Given that the aim of this study was to understand the role of technology in the lives of people with visual disabilities in LMICs, and how AT shapes opportunities, barriers, and WB, the findings underscore both the promise and limits of AT. While AT cannot resolve the structural inequalities that underpin marginalization, it can be designed to mitigate their effects. This requires that designers pay careful attention to users' cultural values, local forms of risk management, and context-specific aspirations. Culturally sensitive design that values emotional WB and aesthetic dignity is more likely to foster meaningful adoption and sustained use.

Applying the DT alongside the CA allows for a more nuanced understanding of how AT impacts people's lives. It invites researchers and designers to consider not only what a technology enables, but what it may jeopardize, particularly when social or environmental conditions are hostile. As the findings show, AT that is context-blind risks becoming ineffective or even harmful. Design processes must therefore attend to fertile functionings—those that support broader forms of agency and resilience—while actively avoiding the reproduction of corrosive disadvantages.

Limitations

Although the study used rigorous qualitative methods and achieved thematic saturation, the sample size (N=19) and composition—skewed toward younger and more educated urban residents—may limit the diversity of perspectives captured. Future research should include larger and more demographically

XSL•FO
RenderX

diverse samples, especially individuals from rural areas, older adults, and persons with visual disabilities with intersecting marginalized identities, to further explore how AT and structural disadvantage interact in varied sociocultural contexts.

While the study aimed to include participants from both urban and rural areas, the final sample was predominantly urban. This may have limited the representation of geographically diverse experiences.

Additionally, the snowball sampling strategy, while effective in reaching a specific population, may have introduced bias by concentrating perspectives within existing social networks.

Certain themes, such as romantic relationships, were underexplored due to study design limitations. However, the focus group offered deeper insights, showing that group discussions encourage openness on sensitive topics.

Even if this study does not cover all 10 of Sen's capabilities, it supports the framework's relevance in analyzing AT users' needs. Future research should further explore the specific characteristics of AT and how they align with users' daily realities.

Conclusions

This study sheds light on how AT is woven into the everyday lives of people with visual disabilities in Colombia—not just as tools, but as companions, barriers, enablers, and sometimes sources of tension or silence. In a context marked by inequality, insecurity, and deep social stigma, AT does not operate in a vacuum. Whether it supports or hinders WB depends on how it fits into people's lived realities, relationships, environments, and desires.

By combining the CA with DT and a critical lens, we were able to see not only the functionings participants valued—such as education, work, and independence—but also the emotional and social costs of trying to pursue them. Technologies that might promise autonomy in one context could increase vulnerability in another. A navigation app or white cane, for instance, might enhance freedom, but also attract unwanted attention in unsafe public spaces. For many, using AT meant balancing aspirations with risk.

Participants told stories of resilience, but also frustration. They spoke of wanting to be seen as attractive, of yearning for movement without fear, of choosing silence over confrontation to stay safe. Others left things unsaid—particularly around relationships and intimacy—pointing to desires that are often overlooked, even by well-intentioned designers and researchers.

These findings remind us that the voices of people with disabilities in the Global South need to be at the center of technology design—not just in consultations or interviews, but as full collaborators in research and development. The choices they make, the things they value, and even what they choose not to say are powerful forms of knowledge. Listening deeply to these experiences helps us understand what truly makes a technology meaningful.

To create technologies that genuinely support WB, we must look beyond technical fixes. We must ask, does this technology help someone live the life they have reason to value? Does it support their dignity, their safety, their hopes? These questions cannot be answered by engineers or policymakers alone. They require interdisciplinary collaboration—between designers, social scientists, humanists, and most importantly, people with disabilities themselves.

Ultimately, this study shows that AT holds enormous promise—but only if it is rooted in empathy, shaped by real-world complexity, and open to the diverse ways in which people navigate, resist, and imagine their worlds.

Funding

MR was supported by the Ambizione grant (PZOOP1_186035), funded by the Swiss National Science Foundation and by the Latin American Leading House of the University of St Gallen. LMO-E and MAC were supported by the Fondation Gelbert (0751-2020).

Conflicts of Interest

None declared.

References

1. Blindness and vision impairment. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment [accessed 2026-03-31]
2. World report on vision. : World Health Organization; 2019 URL: https://www.who.int/publications/i/item/world-report-on-vision [accessed 2026-03-31]
3. Rizzo JR, Beheshti M, Hudson TE, et al. The global crisis of visual impairment: an emerging global health priority requiring urgent action. Disabil Rehabil Assist Technol 2023 Apr;18(3):240-245. [doi: 10.1080/17483107.2020.1854876] [Medline: 33332166]
4. Brunes A, B Hansen M, Heir T. Loneliness among adults with visual impairment: prevalence, associated factors, and relationship to life satisfaction. Health Qual Life Outcomes 2019 Feb 1;17(1):24. [doi: 10.1186/s12955-019-1096-y] [Medline: 30709406]

XSL•FO
RenderX

5. Smith EM, Huff S, Wescott H, et al. Assistive technologies are central to the realization of the Convention on the Rights of Persons with Disabilities. Disabil Rehabil Assist Technol 2024 Feb;19(2):486-491. [doi: 10.1080/17483107.2022.2099987] [Medline: 35900971]
6. Disability and development report: realizing the sustainable development goals by, for and with persons with disabilities. : Department of Economic and Social Affairs, United Nations; 2019 URL: https://www.un.org/development/desa/disabilities/wp-content/uploads/sites/15/2019/10/UN-flagship-report-on-disability-and-development.pdf [accessed 2026-03-31]
7. Tebbutt E, Brodmann R, Borg J, MacLachlan M, Khasnabis C, Horvath R. Assistive products and the sustainable development goals (SDGs). Global Health 2016 Nov 29;12(1):79. [doi: 10.1186/s12992-016-0220-6] [Medline: 27899117]
8. Jang HY, Chung D, Oh E, Son Hong GR. Experiences of individuals with severe disabilities using assistive devices: a qualitative study. Disabil Health J 2025 Jul;18(3):101833. [doi: 10.1016/j.dhjo.2025.101833] [Medline: 40194902]
9. Global report on assistive technology. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240049451 [accessed 2026-03-31]
10. Barbet I, Hartmann L, Deville D, Ferreira MS. Design of an assessment tool for implementing assistive technology (AT) reuse programs in France. Resour Conserv Recycl Adv 2022 Nov;15:200094. [doi: 10.1016/j.rcradv.2022.200094]
11. Grott R. On technology abandonment or discontinuance. In: Assistive Technology Service Delivery: Academic Press; 2019:211-216. [doi: 10.1016/B978-0-12-812979-1.00015-1]
12. Sugawara AT, Ramos VD, Alfieri FM, Battistella LR. Abandonment of assistive products: assessing abandonment levels and factors that impact on it. Disabil Rehabil Assist Technol 2018 Oct;13(7):716-723. [doi: 10.1080/17483107.2018.1425748] [Medline: 29334475]
13. Menich N. Each person as an end? The users' choices in the service delivery process for assistive technology in Hungary. Societies (Basel) 2022;12(5):130. [doi: 10.3390/soc12050130]
14. Real S, Araujo A. Navigation systems for the blind and visually impaired: past work, challenges, and open problems. Sensors (Basel) 2019 Aug 2;19(15):3404. [doi: 10.3390/s19153404] [Medline: 31382536]
15. Blackburn SJ, Cudd PA. A discussion of systematic user requirements gathering from a population who require assistive technology. Technol Disabil 2012;24(3):193-204. [doi: 10.3233/TAD-2012-0352]
16. Ortiz-Escobar LM, Chavarria MA, Schönenberger K, et al. Assessing the implementation of user-centred design standards on assistive technology for persons with visual impairments: a systematic review. Front Rehabil Sci 2023;4:1238158. [doi: 10.3389/fresc.2023.1238158] [Medline: 37744430]
17. Sen A. Development as Freedom: Alfred A. Knopf; 1999. URL: https://raggeduniversity.co.uk/wp-content/uploads/2025/01/1_x_senDevelopmentasFreedom-_compressed.pdf [accessed 2026-03-31]
18. Nussbaum MC. Creating Capabilities: The Human Development Approach: Harvard; 2011. [doi: 10.4159/harvard.9780674061200]
19. Drèze J, Sen A. India: Development and Participation: Oxford University Press; 2002. URL: https://archive.org/details/indiadevelopment0000drez_u8n8/page/n7/mode/2up [accessed 2024-03-24]
20. Robeyns I. The capability approach: a theoretical survey. J Hum Dev 2005 Mar;6(1):93-117. [doi: 10.1080/146498805200034266]
21. Wolff J, De-Shalit A. Disadvantage: Oxford University Press; 2007.
22. Sen A. Inequality Reexamined: Oxford University Press; 1995. URL: https://www.jstor.org/stable/2940919 [accessed 2022-12-10]
23. Terzi L. Beyond the dilemma of difference: the capability approach to disability and special educational needs. J Philos Educ 2005 Aug;39(3):443-459 [FREE Full text] [doi: 10.1111/j.1467-9752.2005.00447.x]
24. Mitra S. The capability approach and disability. J Disabil Policy Stud 2006 Mar;16(4):236-247. [doi: 10.1177/10442073060160040501]
25. Trani JF, Bakhshi P, Bellanca N, Biggeri M, Marchetta F. Disabilities through the capability approach lens: implications for public policies. ALTER 2011;5(3):143-157. [doi: 10.1016/j.alter.2011.04.001]
26. Mitra S. Disability, Health and Human Development: Palgrave Pivot New York; 2018. URL: http://link.springer.com/10.1057/978-1-137-53638-9 [accessed 2026-03-05] [doi: 10.1057/978-1-137-53638-9]
27. Haenssgen MJ, Ariana P. The place of technology in the capability approach. Oxf Dev Stud 2018 Jan 2;46(1):98-112. [doi: 10.1080/13600818.2017.1325456]
28. Ryan M. The ethics of dietary apps: technology, health, and the capability approach. Technol Soc 2022 Feb;68:101873. [doi: 10.1016/j.techsoc.2022.101873]
29. Nussbaum MC. Women and Human Development: The Capabilities Approach: Cambridge University Press; 2000. [doi: 10.1017/CBO9780511841286]
30. Reed R. Dignity in transgender lives: a capabilities approach. J Human Dev Capabil 2020 Jan 2;21(1):36-48. [doi: 10.1080/19452829.2019.1661982]
31. Kuper H, Davey C, Banks LM, Shakespeare T. Trials and Tribulations of Collecting Evidence on Effectiveness in Disability-Inclusive Development: A Narrative Review. Sustainability 2018;12(18):7823 [FREE Full text] [doi: 10.3390/su12187823]

XSL•FO
RenderX

32. Shakespeare T. Nasty brutish and short? on the predicament of disability and embodiment. In: Bickenbach J, Felder F, Schmitz B, editors. Disability and the Good Human Life: Cambridge University Press; 2013:93-112. [doi: 10.1017/CBO9781139225632.005]
33. Riddle CA. Disability and disadvantage in the capabilities approach. In: Oxford Handbook of Philosophy and Disability: Oxford University Press:2020. [doi: 10.1093/oxfordhb/9780190622879.013.19]
34. Sen A. Capability and well-being. In: The Quality of Life: Oxford University Press; 1993, Vol. 30:270-293. [doi: 10.1093/0198287976.003.0003]
35. Das KV, Jones-Harrell C, Fan Y, Ramaswami A, Orlove B, Botchwey N. Understanding subjective well-being: perspectives from psychology and public health. Public Health Rev 2020 Nov 19;41(1):25. [doi: 10.1186/s40985-020-00142-5] [Medline: 33292677]
36. OECD. Health at a Glance 2013: OECD Indicators: OECD Publishing; 2013. [doi: 10.1787/health_glance-2013-en]
37. Diener E, Heintzelman SJ, Kushlev K, et al. Findings all psychologists should know from the new science on subjective well-being. Can Psychol 2017;58(2):87-104. [doi: 10.1037/cap0000063]
38. Lucas RE. Long-term disability is associated with lasting changes in subjective well-being: evidence from two nationally representative longitudinal studies. J Pers Soc Psychol 2007 Apr;92(4):717-730. [doi: 10.1037/0022-3514.92.4.717] [Medline: 17469954]
39. Emerson E, Llewellyn G, Honey A, Kariuki M. Lower well-being of young Australian adults with self-reported disability reflects their poorer living conditions rather than health issues. Aust N Z J Public Health 2012 Apr;36(2):176-182. [doi: 10.1111/j.1753-6405.2011.00810.x] [Medline: 22487354]
40. Office for National Statistics. Understanding well-being inequalities: who has the poorest personal well-being. 2018 URL: https://www.ons.gov.uk/releases/thecharacteristicsandcircumstancesofthosewhoreportlowestpersonalwellbeingratings [accessed 2026-03-31]
41. Emerson E, Fortune N, Aitken Z, Hatton C, Stancliffe R, Llewellyn G. The wellbeing of working-age adults with and without disability in the UK: associations with age, gender, ethnicity, partnership status, educational attainment and employment status. Disabil Health J 2020 Jul;13(3):100889. [doi: 10.1016/j.dhjo.2020.100889] [Medline: 32046927]
42. Emerson E, Llewellyn G. The wellbeing of women and men with and without disabilities: evidence from cross-sectional national surveys in 27 low- and middle-income countries. Qual Life Res 2023 Feb;32(2):357-371. [doi: 10.1007/s11136-022-03268-y] [Medline: 36273048]
43. Gothwal VK, Kanchustambam J, Kodavati K, Subramanian A. Young adults with vision impairment in India: loneliness and social networks. Ophthalmic Physiol Opt 2024 Jul;44(5):808-818. [doi: 10.1111/opo.13317] [Medline: 38619126]
44. Zapata MA. Disability affirmation predicts employment among adults with visual impairment and blindness. Rehabil Couns Bull 2022 Jan;65(2):120-128. [doi: 10.1177/0034355220957107]
45. Halvorsen R, Hvinden B, Brown JB, Biggeri M, Tøssebro J. In: Waldschmidt A, editor. Understanding the Lived Experiences of Persons with Disabilities in Nine Countries: Active Citizenship and Disability in Europe: Routledge; 2017. [doi: 10.4324/9781315623924]
46. Ripat J, Woodgate R. The intersection of culture, disability and assistive technology. Disabil Rehabil Assist Technol 2011;6(2):87-96. [doi: 10.3109/17483107.2010.507859] [Medline: 20698763]
47. Ganem Maloof K. Comisión para el esclarecimiento de la verdad. hay futuro si hay verdad: informe final, convocatoria a la paz grande [there is a future if there is truth: final report. a call for a great peace]. : Comisión para el Esclarecimiento de la Verdad; 2022 URL: https://www.comisiondelaverdad.co/ [accessed 2022-03-22]
48. Building memory. National Center for Historical Memory. 2025 May 20. URL: https://centrodememoriahistorica.gov.co/construccion-de-la-memoria-historica/ [accessed 2026-03-05]
49. Yaffe L. Conflicto armado en Colombia: análisis de las causas económicas, sociales e institucionales de la oposición violenta: [Armed conflict in Colombia: analysis of the economic, social, and institutional causes of violent opposition]. RevCS 2011(8):187-208. [doi: 10.18046/recs.i8.1133]
50. Velarde MR, Lord JE, Stein MA, Shakespeare T. Disarmament, demobilization, and reintegration in Colombia: lost human rights opportunities for ex-combatants with disabilities. J Hum Rights 2022 Jan 1;21(1):18-35. [doi: 10.1080/14754835.2021.1969648]
51. Martínez L, Prada S, Estrada D. Homicides, public goods, and population health in the context of high urban violence rates in Cali, Colombia. J Urban Health 2018 Jun;95(3):391-400. [doi: 10.1007/s11524-017-0215-5] [Medline: 29204844]
52. Fife ST, Gossner JD. Deductive qualitative analysis: evaluating, expanding, and refining theory. Int J Qual Methods 2024 Jan;23:16094069241244856. [doi: 10.1177/16094069241244856]
53. Kincheloe JL, McLaren P. Rethinking critical theory and qualitative research. In: Denzin NK, Lincoln YS, editors. The Landscape of Qualitative Research, 2nd edition: Sage; 2005:303-342. [doi: 10.1007/978-94-6091-397-6_23]
54. The Washington Group Short Set on Functioning (WG-SS). Washington Group on Disability Statistics. 2006. URL: https://www.washingtongroup-disability.com/question-sets/wg-short-set-on-functioning-wg-ss/ [accessed 2022-11-20]
55. Cooke B, Kothari U. Participation: The New Tyranny: Zed Books; 2001. [doi: 10.1016/S0738-0593(02)00022-6]
56. Kothari U. In: Cooke B, Kothari U, editors. Power, Knowledge and Social Control in Participatory Development: Zed Books; 2001:139-152.

XSL•FO
RenderX

57. Restrepo-Ochoa DA. La salud y la vida buena: aportes del enfoque de las capacidades de Amartya Sen para el razonamiento ético en salud pública [Health and the good life: contributions of Amartya Sen's capability approach to ethical reasoning in public health]. Cad Saúde Pública 2013 Dec;29(12):2371-2382. [doi: 10.1590/0102-311X00069913]
58. Bennett CL, Brady E, Branham SM. Interdependence as a frame for assistive technology research and design. Presented at: Proceedings of the 20th ACM SIGACCESS Conference on Computers and Accessibility; Oct 8, 2018. [doi: 10.1145/3234695.3236348]
59. Joskow R, Patel D, Landre A, et al. Understanding the impact of assistive technology on users' lives in England: a capability approach. Bioengineering (Basel) 2025 Jul 9;12(7):750. [doi: 10.3390/bioengineering12070750] [Medline: 40722442]
60. McDonnall MC, Sui Z. Employment and unemployment rates of people who are blind or visually impaired: estimates from multiple sources. J Vis Impair Blind 2019 Nov;113(6):481-492. [doi: 10.1177/0145482X19887620]
61. Llorente-Barroso C, Mañas-Viniegra L, Sierra-Sánchez J, García-García F. Disability and employability in the audio-visual sector: the (dis)connection between corporate social sustainability goals and the employment experiences of people with disabilities. EPI 2023;32(6). [doi: 10.3145/epi.2023.nov.04]
62. Tsatsou P. Is digital inclusion fighting disability stigma? Opportunities, barriers, and recommendations. Disabil Soc 2021 May 28;36(5):702-729. [doi: 10.1080/09687599.2020.1749563]
63. Viñarás-Abad M, Sánchez-Valle M, Vázquez-Barrio T. The situation of people with disabilities in the Communication sector in Spain: labor, professional and academic aspects. EPI 2021;30(2):e300202. [doi: 10.3145/epi.2021.mar.02]
64. Tshuma C, Ntombela N, Mabvurira V. Challenges and coping strategies of visually impaired adults in Zeerust, South Africa. J Soc Sci Humanit 2021;18(5):53-65 [FREE Full text]
65. Lamoureux EL, Hassell JB, Keeffe JE. The determinants of participation in activities of daily living in people with impaired vision. Am J Ophthalmol 2004 Feb;137(2):265-270. [doi: 10.1016/j.ajo.2003.08.003] [Medline: 14962415]
66. Jackson SE, Hackett RA, Pardhan S, Smith L, Steptoe A. Association of perceived discrimination with emotional well-being in older adults with visual impairment. JAMA Ophthalmol 2019 Jul 1;137(7):825-832. [doi: 10.1001/jamaophthalmol.2019.1230] [Medline: 31145413]
67. Tobias EI, Mukhopadhyay S. Disability and social exclusion: experiences of individuals with visual impairments in the Oshikoto and Oshana regions of Namibia. Psychol Dev Soc 2017;29(1):22-43. [doi: 10.1177/0971333616689203]
68. Saur R, Hansen MB, Jansen A, Heir T. Visually impaired individuals, safety perceptions and traumatic events: a qualitative study of hazards, reactions and coping. Disabil Rehabil 2017 Apr;39(7):691-696. [doi: 10.3109/09638288.2016.1161836] [Medline: 27027571]
69. Ehrlich JR, Ramke J, Macleod D, et al. Association between vision impairment and mortality: a systematic review and meta-analysis. Lancet Glob Health 2021 Apr;9(4):e418-e430. [doi: 10.1016/S2214-109X(20)30549-0] [Medline: 33607015]
70. International Disability Alliance. Our right to education: a compilation of evidence gathered by opds on progress towards SDG 4 and CRPD article 24. 2021 URL: https://www.internationaldisabilityalliance.org/sites/default/files/sdg4_crpd24_compilation_april_2021_final.pdf [accessed 2024-10-29]
71. UNESCO Institute for Statistics (UIS). Education and disability: analysis of data from 49 countries. 2018 URL: https://unesdoc.unesco.org/ark:/48223/pf0000262805 [accessed 2025-04-07]
72. Jahoda M. Employment and Unemployment: Cambridge University Press; 1982. URL: https://www.cambridge.org/us/universitypress/subjects/psychology/social-psychology/employment-and-unemployment-social-psychological-analysis?format=PB&isbn=9780521285865 [accessed 2024-02-03]
73. Kuklek NM, Cséplő M, Pozsonyi E, Pusztafalvi H. Raising employment and quality of life among people with disadvantages - results of a Hungarian project. BMC Public Health 2021 Sep 23;21(1):1729. [doi: 10.1186/s12889-021-11763-z] [Medline: 34556054]
74. Korpi T. Accumulating disadvantage. Longitudinal analyses of unemployment and physical health in representative samples of the Swedish population. Eur Sociol Rev 2001 Sep 1;17(3):255-273. [doi: 10.1093/esr/17.3.255]
75. Dănăcică DE. Unemployment of highly educated disabled individuals in Romania. Ann Constantin Brancusi Univ Targu Jiu Econ Ser 2023;1:45-53 [FREE Full text]
76. Bates * P, Davis FA. Social capital, social inclusion and services for people with learning disabilities. Disabil Soc 2004 May;19(3):195-207. [doi: 10.1080/0968759042000204202]
77. Pal GC. Disability, intersectionality and deprivation. Psychol Dev Soc J 2011 Sep;23(2):159-176. [doi: 10.1177/097133361102300202]
78. Grech S, Soldatic K. Disability and colonialism: (dis)encounters and anxious intersectionalities. Soc Ident 2015 Jan 2;21(1):1-5. [doi: 10.1080/13504630.2014.995394]
79. Spaulding S. Introduction to debates on embodied social cognition. Phenom Cogn Sci 2012 Dec;11(4):431-448. [doi: 10.1007/s11097-012-9275-x]
80. Wagner-Ferrari A, Kuschel Ruiz C. Understanding the embodied nature of social cognition through visuospatial perspective-taking. LíMITE (Arica) 2022;17:61-76. [doi: 10.4067/S0718-50652022000100204]
81. Schölvinck AFM, Pittens C, Broerse JEW. The research priorities of people with visual impairments in the Netherlands. J Vis Impair Blind 2017 May;111(3):201-217. [doi: 10.1177/0145482X1711100302]

XSL•FO
RenderX

82. Lastuka A, Cottingham M. The effect of adaptive sports on employment among people with disabilities. Disabil Rehabil 2016 Apr;38(8):742-748. [doi: 10.3109/09638288.2015.1059497] [Medline: 26114627]
83. Akkiraju K, Rao P, Srinivas KV, et al. Assistive technology use and employment outcomes among users in a low-income urban setting in Bangalore, India. ACM Trans Access Comput 2012;5(3):1-19. [doi: 10.1145/2160673.2160711]
84. McDonnall MC, Boydstun J, Steverson A. Is one enough? Screen reader use among employed people who are blind or have low vision in the U.S. Disabil Rehabil Assist Technol 2025 Nov;20(8):2725-2735. [doi: 10.1080/17483107.2025.2528858] [Medline: 40614067]
85. Yao F, Zhou W, Hu H. A review of vision-based assistive systems for visually impaired people: technologies, applications, and future directions. arXiv. Preprint posted online on May 20, 2025 URL: https://arxiv.org/abs/2505.14298 [accessed 2026-03-05]
86. Montenegro-Rueda M, Fernández-Batanero JM, Fernández-Cerero J. Impact of ICT on university students with visual impairment. Br J Spec Educ 2023;50(1):28-48. [doi: 10.1111/1467-8578.12433]
87. Moser I. Disability and the promises of technology: technology, subjectivity and embodiment within an order of the normal. Inf Commun Soc 2006 Jun;9(3):373-395. [doi: 10.1080/13691180600751348]
88. Dos Santos ADP, Ferrari ALM, Medola FO, Sandnes FE. Aesthetics and the perceived stigma of assistive technology for visual impairment. Disabil Rehabil Assist Technol 2022 Feb;17(2):152-158. [doi: 10.1080/17483107.2020.176830810.1080/17483107.2020.1768308] [Medline: 32501732]
89. Barbareschi G, Carew MT, Johnson EA, Kopi N, Holloway C. "When They See a Wheelchair, They've Not Even Seen Me"-factors shaping the experience of disability stigma and discrimination in Kenya. Int J Environ Res Public Health 2021 Apr 17;18(8):4272. [doi: 10.3390/ijerph18084272] [Medline: 33920601]
90. Grech S. Decolonising Eurocentric disability studies: why colonialism matters in the disability and global South debate. Soc Ident 2015 Jan 2;21(1):6-21. [doi: 10.1080/13504630.2014.995347]
91. Fanon F. The Wretched of the Earth: Penguin twentieth-century classics; 1963. URL: https://www.penguin.co.uk/books/57385/the-wretched-of-the-earth-by-frantz-fanon-trans-constance-farrington-preface-by-jean-paul-sartre/9780141186542 [accessed 2024-10-02]
92. Gruber E, Kalkbrenner MT, Hitter TL. A complex conceptualization of beauty in Latinx women: a mixed methods study. Body Image 2022 Jun;41:432-442. [doi: 10.1016/j.bodyim.2022.04.008] [Medline: 35533521]
93. Adames HY, Chavez-Dueñas NY. Cultural Foundations and Interventions in Latino/a Mental Health: History, Theory and within Group Differences: Routledge; 2016. [doi: 10.4324/9781315724058]
94. Jarrin A. Towards a biopolitics of beauty: eugenics, aesthetic hierarchies and plastic surgery in Brazil. J Lat Am Cult Stud 2015 Oct 2;24(4):535-552. [doi: 10.1080/13569325.2015.1091296]
95. Figueroa MGM, Moore MR. Beauty, race and feminist theory in Latin America and the Caribbean. Fem Theor 2013 Aug;14(2):131-136. [doi: 10.1177/1464700113483233]
96. Kukla QR. City Living: How Urban Dwellers and Urban Spaces Make One Another: Oxford University Press; 2021. [doi: 10.1093/oso/9780190855369.001.0001]
97. Chavarria MA, Rivas Velarde M. When innovation is perpetrating inequities: the pivotal role of contextualisation in the development of assistive technology for low- and middle-income countries (LMICs). Bioethica Forum 2025;17(1). [doi: 10.24894/BF.2024.17003]
98. Cuerpos excluidos, rostros de impunidad. informe de violencia hacia personas LGBT en Colombia, 2015. Colombia Diversa, Caribe Afirmativo, Santamaría Fundación. 2015. URL: https://colombiadiversa.org/ddhh-lgbt/Informe-Violencia-LGBT-Colombia-DDHH-2015.pdf [accessed 2026-03-31]
99. Mosquera Vera C. Gender and sexuality in the Colombian armed conflict: from patriarchal oppression to feminist resistance. Rocznik Filozoficzno-Społeczny Civitas Hominibus 2023(18):73-84. [doi: 10.25312/2391-5145.18/2023_06cmv]
100. Ordóñez-Valverde J. From gangs to bands: transformations of gang violence in slums in Cali. Soc Econ 2017(32):107-126 [FREE Full text]
101. Corte Constitucional de Colombia. Judgment C-671 of 2014. Bogotá, Colombia: Constitutional Court of Colombia.: Corte Constitucional; 2014. URL: https://www.alcaldiabogota.gov.co/sisjur//normas/Norma1.jsp?i=63728 [accessed 2024-03-05]
102. Rodríguez Caicedo N, León-Giraldo S, González-Uribe C, Bernal O, Gallego-Pantoja L, Carrasquilla G. Access to health services during the Colombian armed conflict: a challenge for the population with disabilities in the department of Meta. BMC Health Serv Res 2023 Jun 13;23(1):628. [doi: 10.1186/s12913-023-09472-x] [Medline: 37312099]
103. Díaz Velázquez E. Citizenship, identity and social exclusion of persons with disability. Polit Sci 2010(1):115-135 [FREE Full text]
104. Zheng Y, Stahl BC. Technology, capabilities and critical perspectives: what can critical theory contribute to Sen's capability approach? Ethics Inf Technol 2011 Jun;13(2):69-80. [doi: 10.1007/s10676-011-9264-8]
105. Orellano-Colón EM, Mann WC, Rivero M, et al. Hispanic older adults' perceptions of personal, contextual and technology-related barriers for using assistive technology devices. Occup Ther Int 2016;23(2):136-145. [doi: 10.1002/oti.1415]
106. MacLachlan M, Banes D, Bell D, et al. Assistive technology policy: a position paper from the first global research, innovation, and education on assistive technology (GREAT) summit. Disabil Rehabil Assist Technol 2018 Jul;13(5):454-466. [doi: 10.1080/17483107.2018.1468496] [Medline: 29790393]

XSL•FO
RenderX

Abbreviations

AD: assistive device
AT: assistive technology
CA: capability approach
CRPD: Convention on the Rights of Persons with Disabilities
DT: disadvantage theory
FARC-EP: Revolutionary Armed Forces of Colombia
LGBTIQ+: lesbian, gay, bisexual, transgender, intersex, and queer
LMIC: low- and middle-income country
WB: well-being
WHO: World Health Organization

Edited by S Munce, U Gopaul; submitted 10.Feb.2025; peer-reviewed by R Wu, T Hodson; revised version received 20.Jul.2025; accepted 08.Oct.2025; published 08.Apr.2026.

Please cite as:
Ortiz-Escobar LM, Chavarria MA, Hurst-Majno S, Campo Salazar OI, Escobar-Hurtado C, Stein MA, Rivas Velarde M
A Social Justice Approach to Assistive Technology and Well-Being of People With Visual Disabilities in Low- and Middle-Income Countries: Qualitative Narrative Study
JMIR Rehabil Assist Technol 2026;13:e72306
URL: https://rehab.jmir.org/2026/1/e72306
doi:10.2196/72306

© Luisa Maria Ortiz-Escobar, Mario Andres Chavarria, Samia Hurst-Majno, Oscar Ivan Campo Salazar, Celia Escobar-Hurtado, Michael Ashley Stein, Minerva Rivas Velarde. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of Upper Body Joint and Hand Motions in Eating Solid Foods With Chopsticks and Semisolid Foods With a Spoon in Healthy Males and Females: Observational Study

Jun Nakatake[1], MMSc, PhD; Shogo Maeda[1], BHS; Shigeaki Miyazaki[1], MMSc, PhD; Hideki Arakawa[1], MD, PhD; Etsuo Chosa[2], MD, PhD

[1]Rehabilitation Unit, University of Miyazaki Hospital, 5200 Kihara Kiyotake-cho, Miyazaki-shi, Miyazaki, Japan

[2]Department of Rehabilitation Medicine, Junwakai Memorial Hospital, Miyazaki-shi, Miyazaki, Japan

Corresponding Author:
Jun Nakatake, MMSc, PhD
Rehabilitation Unit, University of Miyazaki Hospital, 5200 Kihara Kiyotake-cho, Miyazaki-shi, Miyazaki, Japan

Abstract

Background: Foods are not only masticated and swallowed but they also influence the choice of utensils and their use. Comparing the contexts in which different utensils are used with each food form could help in the assessment of individuals experiencing eating difficulties in the food culture unique to East Asian countries.

Objective: Considering East Asian rehabilitation practices, in this study, we evaluated upper body movements involved in eating pickles (solid food) and yogurt (semisolid food) using chopsticks and a spoon, respectively.

Methods: Upper body kinematics, including joint and hand spatiotemporal parameters, were quantified using a 3D inertial motion-capturing system and analyzed in healthy males (n=22; mean age 27.9, SD 5.5 years) and females (n=21; mean age 26.9, SD 4.7 years) across 4 feeding phases (reaching, picking it up, transporting, and inserting food into the mouth) by comparing utensils for eating respective food forms using the Wilcoxon signed rank test.

Results: Both sexes showed smaller maximum shoulder flexion angles with chopsticks for eating solid food across all phases (Males: reaching phase, $P<.001$; picking foods up phase, $P=.04$; transport phase, $P<.001$; and mouth phase, $P<.001$. Females: reaching phase, $P<.001$; picking foods up phase, $P=.007$; transport phase, $P<.001$; and mouth phase, $P<.001$.). Elbow, forearm, and wrist ulnar deviation angle changes were smaller using chopsticks during the “picking up” phase (in both sexes, $P<.001$) compared with using a spoon. However, greater elbow joint angle changes were found during the “reaching” phase (males, $P=.002$; females, $P=.03$) and greater forearm angle changes were found during the “transporting” phase (males, $P=.01$; females, $P=.001$) with chopsticks than with spoons. Regarding hand spatiotemporal parameters, the chopstick condition involved shorter actual distances (males, $P<.001$; females, $P<.001$), lower distance efficiency (males, $P=.04$; females, $P=.001$), and slower speed (males, $P=.001$; females, $P<.001$) during food transport.

Conclusions: The joint angle and hand spatiotemporal parameter characteristics observed during chopstick use for eating solid foods and spoon use for eating semisolid foods in healthy individuals could serve as reference movements in individuals with sensorimotor dysfunctions and inform the selection of adaptive utensils in rehabilitation practice.

(JMIR Rehabil Assist Technol 2026;13:e76239) doi:10.2196/76239

KEYWORDS

activities of daily living; biomechanical phenomena; neck; torso; upper extremity

Introduction

In East Asian countries, such as China, Korea, and Japan, chopsticks are primarily used to take food into the mouth during meals. This utensil serves multiple functions, including picking up solid foods, cutting them into manageable pieces, and securely holding them while bringing them to the mouth. Due to their versatile functions, chopsticks have become an integral part of the food customs of these regions. To compensate for the difficulty in handling semisolid foods, a spoon is typically used. Thus, chopsticks and a spoon are commonly paired with solid and semisolid foods, respectively. Furthermore, since people in Japan have the habit of consuming soup by bringing the bowl to their mouths instead of using spoons, Japanese people more often use chopsticks than do people in other countries.

When patients undergo eating activity rehabilitation, selection of appropriate feeding utensils is essential. Patients, including those recovering from strokes, often use a spoon for its adaptability to foods that are transferred to the mouth,

XSL•FO
RenderX

masticated, and then swallowed. Semisolid foods, regulated for viscosity, are commonly recommended for individuals with dysphagia [1], making the spoon a preferred adaptive device. Chopsticks are not routinely selected due to the difficulty in using them for semisolid foods. As the oral and swallowing abilities improve, patients transition from using spoons for semisolid foods to using chopsticks for solid foods. The food custom in East Asian countries is also focused on the use of chopsticks for eating solid foods, which facilitates the prospect of dining together and improving well-being. However, sensorimotor dysfunctions can impair the ability to perform feeding movements involving chopsticks and spoons. For instance, patients with right hemiparesis can use chopsticks or a spoon only with their dominant right hand if their impairment is mild [2]. Therefore, the selection of chopsticks or a spoon depends on patients' ability to transfer, masticate, and swallow foods, as well as their limb function for manipulating utensils.

Daily movements and postures can be assessed to determine whether patients encounter difficulties performing tasks. Moreover, in the context of rehabilitation practice, assessments are needed to facilitate normal or compensatory movements or to change utensils used by patients while eating [3]. Understanding typical movements and postures when using feeding utensils aids in these assessments. Previous studies have separately reported differences in upper body joint motions across feeding phases during the normal use of chopsticks [4] and spoons [5]. A comparative study on normal upper limb kinematics has shown greater joint motion during spoon use than that during chopstick use when eating pickles, a common form of solid food [6]. However, a study investigating performance while using eating utensils reported that spoon use for noodle strips and tofu cubes prolonged eating time or led to failure to properly manipulate foods [7], suggesting that the choice of utensils should match food type.

In another aspect, proximal body movements, including the neck and trunk joint angles, tilting and rotation angles, or transfer distances, occur during eating [4,5,8-10] and tend to increase when foods are prone to be dropped [8,11]. The movements and postures of these proximal segments are important as normal functions in relation to preventing food dropping and form the basis for fine upper limb function. A recent study that simulated eating movements for fictitious foods using utensils demonstrated larger neck joint motion with a spoon than with chopsticks; however, trunk flexion angles were comparable between utensils [12]. These findings perhaps imply that using a spoon requires compensatory neck joint motion for preventing food drop and trunk stability for maintaining upper extremity functions. Using chopsticks, conversely, may not need neck compensation and may need trunk stability as when using a spoon. However, the actual movement and positioning of the upper body, including the neck and trunk, during solid food intake with chopsticks, compared with those during semisolid food intake with a spoon, remain unclear.

In addition to joint parameters, spatiotemporal parameters are valuable for assessing upper limb motions [13,14]. Considering the importance of rehabilitation practice in East Asia, understanding the normal kinematics of chopstick use while eating solid foods and spoon use while eating semisolid foods not only expands knowledge about the utensils or food forms in isolation but is also crucial, yet currently lacking. Therefore, a comparative study of these 2 food type and utensil type conditions is warranted in clinical settings, as previous research found significant head motion differences in the conditions of eating food forms with these respective utensils [8]. This previous study design is thought to be validated for assessing patients with eating difficulties such as swallowing function deficit with upper limb paresis in the recovery process.

In this study, we aimed to identify the characteristics of joint motions and spatiotemporal parameters during eating solid foods with chopsticks and semisolid foods with a spoon across feeding phases. We hypothesized that upper body joint motions for chopstick use for solid food intake could be reduced compared with those for spoon use for semisolid food intake across all feeding phases; however, the time required may be similar. We also hypothesized that the distance traveled by the hand, particularly during the reaching movements, could be greater when using chopsticks than when using a spoon. Additional parameters, such as speed, could provide further insights into kinematics during clinical observations. Consequently, in this study, we investigated the characteristics of upper body joint angles and hand spatiotemporal parameters across all feeding phases by comparing the use of chopsticks for eating solid foods and of a spoon for eating semisolid foods in healthy adults.

Methods

Participants

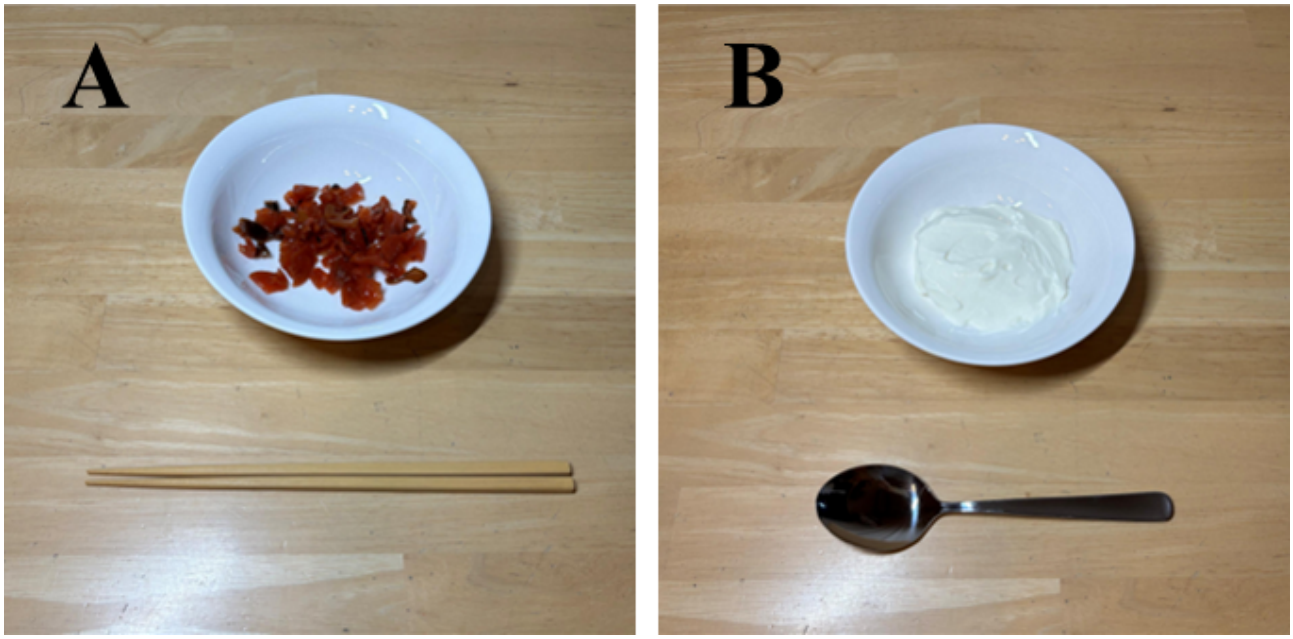
Between April 2013 and October 2017, 50 members of our institution participated per the inclusion criteria. The criteria were age 20‐39 years, right-hand dominance, and no neurological or musculoskeletal disorders. Individuals with left-hand dominance for using chopsticks or a spoon were excluded.

Eating Items and Tasks

Eating solid foods with chopsticks was performed using wooden chopsticks (22.5 centimeters in length, 8 grams in weight) and small Japanese pickled vegetables (Figure 1A). Eating semisolid foods with a spoon was performed using a stainless steel spoon (17.5 centimeters in length, 41 grams in weight) and yogurt (Figure 1B).

XSL•FO
RenderX

Figure 1. Eating items for 2 conditions. (A) Chopstick condition: small Japanese pickled vegetables in a bowl and wooden chopsticks. (B) Spoon condition: yogurt in the bowl and a stainless steel spoon.



The participants sat on a stool 40 centimeters in height, facing a table adjusted to their elbow height. The food bowl was placed in front of them at the same distance of the tip of the extended finger when their shoulder joint was in a neutral position and their elbow flexed at 90°. Neck and trunk movements were not constrained. The participants were asked to eat the foods using utensils held in their right hand while their left hand rested on their left thigh. They were instructed to eat at a comfortable pace and perform the movements sequentially 3 times. This frequency was to prevent fatigue and a full stomach. Based on the assumption that natural daily feeding movements comprise the cycle of chopsticks or a spoon traveling to and from the bowl and the mouth, starting and ending postures were not predefined to capture these natural cyclic daily movement data. The initial measurement position is illustrated in Figure 2.

XSL•FO
RenderX

Figure 2. Initial position for measurements during eating tasks using utensils with the dominant right hand while the left hand rests on the left thigh. Illustrated is an example of the task of eating pickles with chopsticks or yogurt with a spoon. The participant is shown wearing a suit, grooves, and a headband, all containing inertial sensors.



Measurement Instruments and Data Processing

Kinematic data at a frequency of 120 Hz were obtained using a 3D motion analysis system based on 17 inertial sensor units (Xsens MVN system, Xsens Technologies BV). Validation studies for this system have been previously conducted [15-19]. The inertial sensors were attached to participant body parts using a Lycra suit, gloves, and a headband and were confirmed

XSL•FO
RenderX

not to shift on the body surface or interfere with utensils and movements during feeding. The present tasks and measurements were valid; however, they might not be confirmed as sensor slipping or skin artifacts and interference with objects or bodies are recognized sources of error that could compromise the accuracy of joint angle and position data. The biomechanical model of the system consisted of 23 body parts, including the head, neck, 6 vertebrae, pelvis, scapulae, upper arms, forearms, hands, thighs, tibiae, feet, and toes. Joint angles were calculated based on the positions and orientations of the body parts. The neutral joint angle position of the model corresponded to a standing posture with the upper limbs alongside the trunk, the face directed forward, feet parallel at a distance of 1-foot width, and palms facing forward.

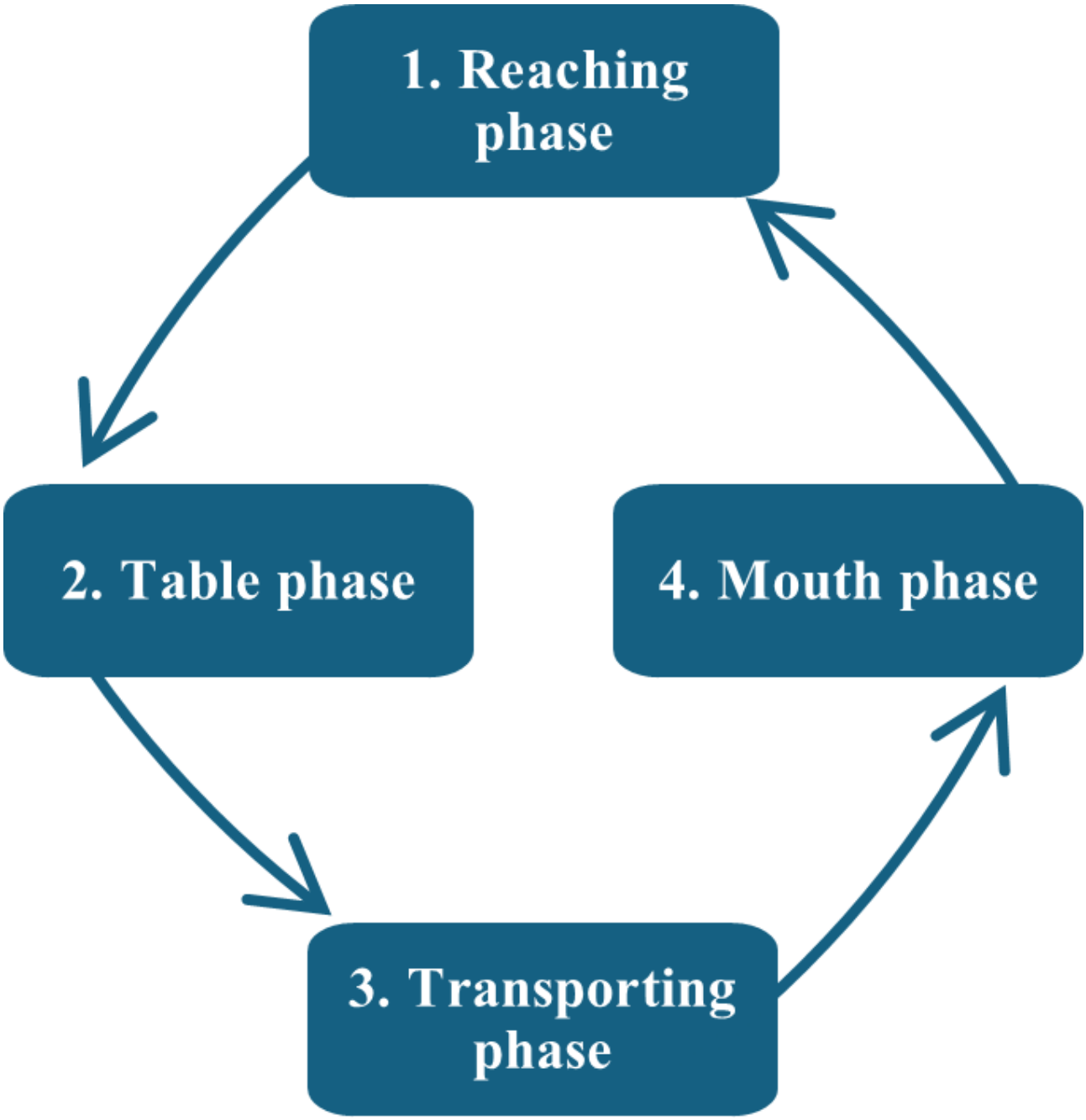
When the participants ate the food 3 times sequentially, there were 2 eating cycles. One success cycle of these 2 eating cycles (or the first cycle if both were successful) in each condition was selected for each participant for analysis. The eating cycle was defined as the interval between the moment after the utensil left the mouth and the moment immediately before it next left the mouth. This was confirmed using synchronized video recordings. To simplify the analysis, failure cycles involving the following were excluded: excessive upper limb elevation during food transport, looking away from the bowl or food, extraneous head movements, separating food, or taking more than 1 scoop of yogurt. Considering performing analysis and maintaining sample size, we decided to focus on analyzing 1 successful eating cycle, which was defined as a natural attempt of daily movements, of 2 cycles regarding each condition in each participant, excluding failed cycles that sometimes were observed.

Pauses during the reaching phase (observed in 84% of chopstick tasks and 65% of spoon tasks in the final analysis) and attempts to pick up the pickles more than once (28%) were frequently observed but not excluded from the analysis. Only participants with complete task data were included in the final analysis.

One eating cycle was divided into the following four phases: (1) reaching phase—from the moment the utensil left the mouth to the point of reaching the food, (2) table phase—picking up or scooping the food, (3) transporting phase—moving the food to the mouth, and (4) mouth phase—inserting the food into the mouth (Figure 3). These phases were defined because upper body joint motions differ during each feeding phase when using a spoon [5], which was expected to apply to chopstick use as well.

XSL•FO
RenderX

Figure 3. One eating cycle contains 4 phases. The arrows indicate the order of eating phase. The reaching phase is when the utensils move from the mouth to the food. The table phase is when the food is picked up or spooned up. The transporting phase is when the food moves from the bowl on the table to the mouth. The mouth phase is when the utensils are in the mouth.



Joint kinematic data from the biomechanical model were analyzed, including right shoulder flexion, abduction, and internal rotation; right elbow flexion and pronation; right wrist dorsiflexion and ulnar deviation (palmar flexion and radial deviation in the model); neck flexion, right lateral flexion, and right rotation (left rotation in the model); and right hip flexion. These joint kinematics were summarized as the maximum and minimum joint angles and ranges of angle change in each feeding phase. Movement time(s) was calculated in all phases.

The right-hand spatiotemporal parameters during the transporting phase included the actual distance traveled (meters), relative distance traveled, and maximum velocity (meters per second). The timing of maximum velocity (%) was calculated as motor control strategy metrics. The number of movement units was used to measure hand movement smoothness. The origin point for the right hand was defined as the midpoint between the styloid processes. A movement unit was defined as the difference between a local minimum and the next maximum velocity value that exceeded an amplitude limit of 20 millimeters per second, with a minimum interval of 150 milliseconds between peaks. This parameter was adapted from a task originally defined for drinking motions [20]. Most spatiotemporal parameters were limited to the transporting phase, as the start and end points of the reaching movement were clearly defined and undisturbed. Movements during the reaching phase showed frequent disruptions, and those during the table and mouth phases were not reaching movements. These hand spatiotemporal parameters could not be interpreted

XSL•FO
RenderX

according to the common concept of reaching movement parameters [13,14] and, consequently, were excluded from analysis.

Data Analysis

All metrics were analyzed separately for each sex because of known differences in daily movements [8,9,21-24] and fit to clinical assessments practically. The data are displayed as medians with IQRs. Since the study design compared movements between eating solid foods with chopsticks and semisolid foods with a spoon, these 2 task conditions were compared using the Wilcoxon signed rank test using IBM SPSS for Windows (version 27.0; IBM Corp), with a significance level of $P<.05$. Effect sizes (r) were calculated and categorized as small (|0.3|), medium (|0.5|), or large (|0.6|) [25].

Ethical Considerations

The protocol for this cross-sectional study, including informed consent using an opt-out method, was approved by the research ethics committee of the Faculty of Medicine of the University of Miyazaki (Miyazaki-shi, Japan; approval number O-1501). Since the sample data were derived from past studies [4,5] and participants could not be contacted directly, information related to this study was published on the local institutional website, giving past participants the opportunity to opt out of the study at any time. All participants provided written informed consent prior to participation in the previous studies [4,5]. The data were anonymized in both the previous and the present studies.

Results

Overview

Forty-three participants with 1 success cycle regarding each condition have been selected, excluding 7 participants with 2 failed cycles regarding 1 of 2 conditions. The analyzed study sample included 22 male and 21 female participants. The mean (SD) age and height of the male participants were 27.9 (5.5) years and 172.0 (5.1) cm, respectively, while those of the female participants were 26.9 (4.7) years and 158.0 (4.4) cm, respectively.

Joint Angles During Feeding Phases

In male participants (Table 1), various maximum and minimum joint angles, including shoulder flexion, abduction, and internal rotation; wrist ulnar deviation; and neck right lateral flexion, were lower when using chopsticks compared with when using spoons across many feeding phases. Additionally, during the reaching phase, the minimum wrist dorsiflexion angle with chopsticks was smaller than that with the spoon. In contrast, most elbow flexion and some neck right rotation angles were greater with chopsticks than with the spoon.

XSL•FO
RenderX

Table . Comparison of maximum and minimum joint angles between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in male participants[a].

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion										
Re	**23 (21 to 30)**	**36 (31 to 42)** [b]	−3.945	<.001	−0.595	**16 (13 to 20)**	**19 (16 to 24)**[b]	−3.1	.002	−0.467
Ta	**21 (16 to 27)**	**23 (19 to 31)**[b]	−2.029	.04	−0.306	19 (14 to 26)	22 (19 to 27)	−1.737	.08	−0.262
Tr	**25 (19 to 31)**	**36 (31 to 45)** [b]	−4.01	<.001	−0.605	**18 (15 to 24)**	**23 (19 to 30)**[b]	−3.1	.002	−0.467
Mo	**25 (19 to 32)**	**41 (33 to 49)** [b]	−4.042	<.001	−0.609	**23 (17 to 29)**	**36 (29 to 44)**[b]	−4.107	<.001	−0.619
Shoulder abduction										
Re	**32 (27 to 38)**	**38 (33 to 47)** [b]	−3.685	<.001	−0.556	**24 (20 to 29)**	**28 (22 to 33)**[b]	−2.808	.005	−0.423
Ta	30 (24 to 35)	31 (24 to 39)	−1.737	.08	−0.262	30 (24 to 34)	30 (23 to 36)	−1.315	.19	−0.198
Tr	**32 (27 to 40)**	**37 (32 to 48)**[b]	−3.847	<.001	−0.58	28 (24 to 34)	31 (24 to 39)	−1.672	.10	−0.252
Mo	**32 (26 to 40)**	**38 (33 to 48)**[b]	−3.847	<.001	−0.58	**31 (26 to 39)**	**36 (31 to 47)**[b]	−3.912	<.001	−0.59
Shoulder internal rotation										
Re	**0 (−6 to 8)**	**2 (−1 to 15)**[b]	−3.1	.002	−0.467	**−12 (−14 to −3)**	**−7 (−10 to 4)**[b]	−3.36	.001	−0.507
Ta	**−1 (−6 to 11)**	**4 (−1 to 11)**[b]	−3.036	.002	−0.458	**−4 (−8 to 5)**	**1 (−3 to 9)**[b]	−2.646	.01	−0.399
Tr	**−3 (−6 to 8)**	**4 (0 to 16)**[b]	−3.912	<.001	−0.59	**−8 (−10 to 2)**	**1 (−5 to 8)**[b]	−3.847	<.001	−0.58
Mo	**−5 (−10 to 3)**	**6 (−1 to 15)**[b]	−3.977	<.001	−0.6	**−8 (−11 to 1)**	**1 (−5 to 13)**[b]	−4.074	<.001	−0.614
Elbow flexion										
Re	**125 (117 to 132)**[b]	**121 (114 to 129)**	−3.36	.001	−0.507	97 (88 to 106)	100 (93 to 109)	−1.51	.13	−0.228
Ta	102 (97 to 112)	105 (98 to 114)	−0.568	.57	−0.086	**99 (89 to 107)**[b]	**98 (88 to 105)**	−2.062	.04	−0.311
Tr	**126 (118 to 133)**[b]	**120 (115 to 129)**	−3.393	.001	−0.512	**101 (91 to 108)**[b]	**99 (84 to 107)**	−2.451	.01	−0.37
Mo	**131 (120 to 136)**[b]	**127 (118 to 133)**	−3.295	.001	−0.497	**125 (115 to 132)**[b]	**119 (113 to 127)**	−3.523	<.001	−0.531
Forearm pronation										

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Re	90 (83 to 101)	99 (86 to 110)	−1.607	.11	−0.242	**26 (18 to 32)**	**34 (24 to 39)**[b]	−2.711	.007	−0.409
Ta	**90 (80 to 101)**	**99 (86 to 108)**[b]	−3.165	.002	−0.477	**77 (73 to 89)**[b]	**75 (55 to 83)**	−2.289	.02	−0.345
Tr	**78 (73 to 88)**[b]	**74 (55 to 83)**	−2.289	.02	−0.345	31 (23 to 37)	32 (23 to 38)	−0.828	.41	−0.125
Mo	31 (23 to 37)	34 (25 to 39)	−0.536	.59	−0.081	22 (16 to 29)	24 (16 to 32)	−0.016	.99	−0.002
Wrist dorsal flexion										
Re	26 (22 to 43)	29 (21 to 42)	−0.243	.81	−0.037	**13 (2 to 24)**	**20 (8 to 34)**[b]	−2.289	.02	−0.345
Ta	22 (9 to 37)	28 (16 to 38)	−1.217	.22	−0.183	16 (4 to 25)	21 (12 to 34)	−1.542	.12	−0.232
Tr	26 (15 to 36)	27 (18 to 39)	−0.243	.81	−0.037	19 (7 to 30)	20 (11 to 31)	−0.925	.36	−0.139
Mo	22 (11 to 33)	23 (11 to 31)	−0.016	.99	−0.002	20 (9 to 30)	20 (7 to 29)	−0.146	.88	−0.022
Wrist ulnar deviation										
Re	**5 (−3 to 10)**	**11 (4 to 16)**[b]	−3.587	<.001	−0.541	−7 (−14 to 2)	−5 (−14 to 5)	−1.347	.18	−0.203
Ta	**1 (−5 to 9)**	**7 (3 to 13)**[b]	−3.328	.001	−0.502	−2 (−10 to 7)	−4 (−12 to 5)	−0.86	.39	−0.13
Tr	**3 (−3 to 12)**	**12 (5 to 16)**[b]	−3.977	<.001	−0.6	**−1 (−7 to 5)**	**5 (1 to 12)**[b]	−3.782	<.001	−0.57
Mo	**2 (−4 to 11)**	**13 (3 to 18)**[b]	−4.074	<.001	−0.614	**1 (−5 to 8)**	**10 (3 to 15)**[b]	−4.042	<.001	−0.609
Neck flexion										
Re	27 (22 to 29)	26 (22 to 28)	−1.445	.15	−0.218	19 (14 to 25)	21 (18 to 26)	−1.25	.21	−0.188
Ta	26 (22 to 30)	25 (23 to 28)	−0.373	.71	−0.056	25 (21 to 29)	24 (22 to 27)	−0.828	.41	−0.125
Tr	25 (21 to 29)	25 (22 to 27)	−0.601	.55	−0.091	21 (18 to 25)	20 (18 to 25)	−1.542	.12	−0.232
Mo	22 (19 to 26)	22 (19 to 28)	−0.73	.47	−0.11	21 (18 to 24)	18 (17 to 24)	−1.899	.06	−0.286
Neck right lateral flexion										
Re	**−1 (−2 to 0)**	**3 (2 to 4)**[b]	−4.107	<.001	−0.619	**−4 (-5 to −3)**	**−2 (−4 to 0)**[b]	−3.036	.002	−0.458
Ta	**−3 (−4 to −2)**	**−2 (−3 to 0)**[b]	−3.036	.002	−0.458	**−3 (−5 to −2)**	**−2 (−4 to 0)**[b]	−2.678	.007	−0.404
Tr	**−2 (-3 to −1)**	**2 (−1 to 3)**[b]	−3.945	<.001	−0.595	**−3 (−4 to −3)**	**−2 (−3 to 0)**[b]	−3.393	.001	−0.512

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Mo	**−2 (−3 to 0)**	**4 (2 to 5)**[b]	−4.107	<.001	−0.619	**−3 (-4 to −1)**	**2 (−1 to 3)**[b]	−3.977	<.001	−0.6
Neck right rotation										
Re	**4 (2 to 6)**[b]	**2 (1 to 3)**	−3.263	.001	−0.492	0 (−1 to 2)	0 (−1 to 2)	−0.211	.83	−0.032
Ta	1 (0 to 2)	1 (0 to 2)	−0.016	.99	−0.002	1 (−1 to 2)	1 (−1 to 1)	−0.179	.86	−0.027
Tr	**4 (1 to 6)**[b]	**3 (0 to 3)**	−3.165	.002	−0.477	1 (−1 to 2)	1 (−1 to 2)	−0.146	.88	−0.022
Mo	**5 (3 to 7)**[b]	**3 (2 to 5)**	−3.393	.001	−0.512	**4 (1 to 5)**[b]	**2 (0 to 3)**	−3.068	.002	−0.463
Hip flexion										
Re	51 (44 to 57)	50 (45 to 54)	−0.73	.47	−0.11	43 (37 to 51)	44 (38 to 50)	−0.016	.99	−0.002
Ta	44 (38 to 51)	46 (38 to 51)	−0.341	.73	−0.051	44 (37 to 51)	45 (38 to 51)	−0.373	.71	−0.056
Tr	50 (43 to 58)	50 (44 to 54)	−0.763	.45	−0.115	44 (38 to 51)	46 (38 to 51)	−0.438	.66	−0.066
Mo	51 (45 to 58)	51 (46 to 56)	−0.179	.86	−0.027	50 (43 to 56)	51 (44 to 54)	−0.308	.76	−0.046

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test ($P<.05$), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b]Larger values from comparisons between chopstick and spoon conditions.

Forearm pronation showed differing trends according to the feeding phase. The minimum forearm pronation angle during the reaching phase and maximum angle during the table phase were smaller with chopsticks than with the spoon. Conversely, the minimum angle during the table phase and maximum angle during the transporting phase were larger with chopsticks. The differences between utensils showed small to large effect sizes. Neck and hip flexion angles did not differ significantly between utensils.

In female participants (Table 2), all shoulder joint angles, nearly all wrist dorsiflexion angles, and neck right lateral flexion angles were smaller with chopsticks than with the spoon. In contrast, most elbow flexion angles were larger when using chopsticks. Forearm pronation and wrist ulnar deviation angles varied inconsistently across feeding phases. In half the forearm pronation angles, chopsticks showed smaller values than did the spoon. However, the minimum angle during the table phase and maximum angle during the transporting phase were larger in the chopstick condition. Minimum wrist ulnar deviation angles during the reaching and table phases were larger with chopsticks, while other angles were smaller. These differences showed small to large effect sizes. Most neck flexion angles and all neck right rotation and hip flexion angles were comparable between chopsticks and spoons.

XSL•FO
RenderX

Table . Comparison of maximum and minimum joint angles between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in female participants[a].

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion										
Re	**20 (14 to 27)**	**37 (31 to 44)**[b]	-4.015	<.001	–0.62	**10 (4 to 12)**	**10 (6 to 22)**[b]	–2.45	.01	–0.378
Ta	**14 (7 to 18)**	**15 (10 to 25)**[b]	–2.694	.007	–0.416	**14 (6 to 17)**	**13 (7 to 23)**[b]	–2.485	.01	–0.383
Tr	**21 (12 to 28)**	**37 (31 to 44)**[b]	–4.015	<.001	–0.62	**13 (6 to 17)**	**14 (9 to 24)**[b]	–2.972	.003	–0.459
Mo	**22 (18 to 31)**	**40 (37 to 49)**[b]	–4.015	<.001	–0.62	**20 (15 to 27)**	**37 (31 to 44)**[b]	–4.015	<.001	–0.62
Shoulder abduction										
Re	**31 (28 to 34)**	**37 (32 to 45)**[b]	–3.98	<.001	–0.614	**21 (19 to 24)**	**23 (19 to 28)**[b]	–2.589	.01	–0.399
Ta	**24 (22 to 28)**	**28 (24 to 36)**[b]	–3.285	.001	–0.507	**23 (21 to 27)**	**28 (23 to 33)**[b]	–3.319	.001	–0.512
Tr	**30 (27 to 35)**	**38 (33 to 47)**[b]	–4.015	<.001	–0.62	**24 (21 to 28)**	**28 (24 to 33)**[b]	–3.493	<.001	–0.539
Mo	**29 (27 to 36)**	**39 (32 to 48)**[b]	–3.98	<.001	–0.614	**28 (26 to 32)**	**36 (31 to 45)**[b]	–3.945	<.001	–0.609
Shoulder internal rotation										
Re	**3 (0 to 8)**	**13 (10 to 18)**[b]	–3.98	<.001	–0.614	**-5 (-8 to -1)**	**5 (-4 to 7)**[b]	–3.493	<.001	–0.539
Ta	**3 (0 to 8)**	**11 (6 to 15)**[b]	–3.25	.001	–0.501	**1 (-2 to 5)**	**9 (5 to 13)**[b]	–3.041	.002	–0.469
Tr	**3 (0 to 6)**	**13 (9 to 17)**[b]	–4.015	<.001	–0.62	**-1 (-5 to 0)**	**7 (2 to 11)**[b]	–3.875	<.001	–0.598
Mo	**4 (-1 to 8)**	**15 (10 to 21)**[b]	–4.015	<.001	–0.62	**0 (-3 to 5)**	**13 (6 to 15)**[b]	–4.015	<.001	–0.62
Elbow flexion										
Re	**135 (130 to 138)**[b]	**128 (120 to 134)**	–3.841	<.001	–0.593	106 (101 to 111)	103 (97 to 112)	–0.504	.61	–0.078
Ta	106 (104 to 111)	103 (97 to 112)	–1.095	.27	–0.169	**103 (97 to 110)**[b]	**97 (90 to 104)**	–3.041	.002	–0.469
Tr	**134 (125 to 137)**[b]	**128 (122 to 131)**	–3.076	.002	–0.475	**104 (97 to 111)**[b]	**97 (90 to 106)**	–3.389	.001	–0.523
Mo	**138 (131 to 141)**[b]	**131 (127 to 137)**	–3.041	.002	–0.469	**134 (126 to 138)**[b]	**127 (120 to 131)**	–3.736	<.001	–0.576
Forearm pronation										
Re	**102 (80 to 116)**	**109 (99 to 123)**[b]	–2.763	.01	–0.426	**17 (12 to 43)**	**32 (19 to 49)**[b]	–3.285	.001	–0.507
Ta	**102 (81 to 119)**	**111 (99 to 122)**[b]	–2.728	.01	–0.421	**93 (68 to 108)**[b]	**80 (64 to 91)**	–3.319	.001	–0.512

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Tr	**93 (68 to 107)**[b]	**80 (64 to 91)**	–3.215	.001	–0.496	26 (16 to 44)	29 (18 to 45)	–0.122	.90	–0.019
Mo	**26 (15 to 41)**	**31 (20 to 48)**[b]	–2.346	.02	–0.362	17 (10 to 35)	27 (11 to 40)	–1.199	.23	–0.185
Wrist dorsal flexion										
Re	**22 (15 to 37)**	**34 (27 to 40)**[b]	–3.563	<.001	–0.55	**9 (-1 to 21)**	**20 (14 to 31)**[b]	–3.875	<.001	–0.598
Ta	**23 (12 to 34)**	**31 (21 to 37)**[b]	–3.18	.001	–0.491	**18 (3 to 25)**	**23 (15 to 31)**[b]	–2.694	.007	–0.416
Tr	**25 (15 to 37)**	**31 (23 to 39)**[b]	–2.381	.02	–0.367	17 (7 to 30)	23 (11 to 31)	–1.616	.11	–0.249
Mo	18 (10 to 34)	25 (15 to 33)	–1.929	.054	–0.298	**17 (7 to 31)**	**22 (12 to 31)**[b]	–2.033	.04	–0.314
Wrist ulnar deviation										
Re	**12 (5 to 17)**	**14 (9 to 18)**[b]	–2.381	.02	–0.367	**4 (-2 to 10)**[b]	**-1 (-10 to 4)**	–3.319	.001	–0.512
Ta	**10 (3 to 15)**	**10 (6 to 18)**[b]	–1.964	.05	–0.303	**6 (0 to 13)**[b]	**-3 (-11 to 4)**	–3.91	<.001	–0.603
Tr	**12 (5 to 15)**	**15 (9 to 19)**[b]	–3.424	.001	–0.528	**9 (1 to 13)**	**10 (5 to 14)**[b]	–2.346	.02	–0.362
Mo	**10 (6 to 16)**	**15 (9 to 20)**[b]	–3.806	<.001	–0.587	**9 (5 to 13)**	**12 (9 to 19)**[b]	–3.597	<.001	–0.555
Neck flexion										
Re	27 (25 to 30)	27 (25 to 29)	–1.199	.23	–0.185	19 (10 to 25)	18 (15 to 23)	–0.678	.498	–0.105
Ta	27 (25 to 30)	27 (25 to 29)	–0.33	.74	–0.051	26 (24 to 29)	26 (25 to 28)	–0.226	.82	–0.035
Tr	26 (24 to 29)	27 (25 to 28)	–0.539	.59	–0.083	18 (11 to 24)	17 (12 to 19)	–1.929	.054	–0.298
Mo	20 (11 to 25)	20 (15 to 22)	–0.678	.50	–0.105	**18 (9 to 23)**[b]	**14 (11 to 18)**	–2.485	.01	–0.383
Neck right lateral flexion										
Re	**2 (–2 to 3)**	**3 (2 to 6)**[b]	–3.736	<.001	–0.576	**-2 (-4 to 0)**	**-1 (-2 to 0)**[b]	–2.52	.01	–0.389
Ta	**–1 (–3 to 2)**	**0 (-1 to 2)**[b]	–2.242	.03	–0.346	–2 (–3 to 1)	-1 (-2 to 0)	–1.825	.07	–0.282
Tr	**2 (–1 to 3)**	**3 (1 to 5)**[b]	–3.667	<.001	–0.566	**–1 (–3 to 1)**	**0 (–1 to 1)**[b]	–3.007	.003	–0.464
Mo	**1 (–1 to 3)**	**4 (3 to 7)**[b]	–3.493	<.001	–0.539	**1 (–2 to 3)**	**2 (1 to 5)**[b]	–3.563	<.001	–0.55
Neck right rotation										
Re	3 (1 to 5)	3 (1 to 5)	–1.025	.31	–0.158	0 (–2 to 1)	0 (–2 to 0)	–0.295	.77	–0.046
Ta	1 (–1 to 1)	0 (–2 to 1)	–0.017	.99	–0.003	0 (–1 to 1)	0 (–2 to 1)	–0.226	0.82	–0.035
Tr	3 (1 to 4)	3 (2 to 4)	–0.469	.64	–0.072	1 (–1 to 1)	0 (–2 to 1)	–0.052	.96	–0.008
Mo	4 (2 to 6)	5 (2 to 6)	–0.226	.82	–0.035	3 (1 to 4)	3 (0 to 4)	–0.608	.54	–0.094

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Hip flexion										
Re	52 (42 to 60)	52 (42 to 58)	–1.095	.27	–0.169	44 (38 to 53)	45 (38 to 52)	–0.052	.96	–0.008
Ta	45 (39 to 54)	45 (38 to 53)	–0.956	.34	–0.148	44 (38 to 54)	44 (38 to 52)	–0.504	.61	–0.078
Tr	53 (43 to 58)	49 (42 to 57)	–1.025	.31	–0.158	45 (39 to 54)	44 (38 to 53)	–1.13	.26	–0.174
Mo	54 (44 to 60)	53 (43 to 58)	–1.06	.29	–0.164	53 (43 to 58)	49 (42 to 57)	–1.616	.11	–0.249

[a]Significant differences in each phase, as determined using the Wilcoxon signed rank test ($P<.05$), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b]Larger values in comparisons between chopstick and spoon conditions.

Changes in Joint Angles

In male participants (Table 3), changes in shoulder flexion, abduction, and neck right lateral flexion angles were generally smaller in the chopstick condition than in the spoon condition across all feeding phases. Additionally, wrist ulnar deviation, neck flexion, and hip flexion showed smaller changes in some phases when using chopsticks. Conversely, neck right rotation in most phases and wrist dorsiflexion during the reaching phase showed increased changes with chopsticks. Elbow flexion showed greater changes in the reaching phase when using chopsticks; however, the changes were smaller in the table and mouth phases. Changes in forearm pronation were smaller during the table phase but larger during the transporting phase. Shoulder internal rotation changes across all phases showed no significant differences between utensils. Effect sizes for these differences were small to large.

XSL•FO
RenderX

Table . Comparison of joint angle changes between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in male participants[a].

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion					
Re	**7 (5 to 12)**	**18 (11 to 21)**[b]	–3.62	<.001	–0.546
Ta	1 (1 to 3)	2 (1 to 3)	–1.217	.22	–0.183
Tr	**6 (4 to 9)**	**13 (7 to 18)**[b]	–3.425	.001	–0.516
Mo	**2 (2 to 3)**	**3 (3 to 5)**[b]	–2.841	.01	–0.428
Shoulder abduction					
Re	**6 (4 to 12)**	**11 (7 to 13)**[b]	–2.419	.02	–0.365
Ta	**1 (0 to 1)**	**1 (1 to 2)**[b]	–2.808	.005	–0.423
Tr	**4 (2 to 6)**	**7 (5 to 11)**[b]	–3.685	<.001	–0.556
Mo	**1 (1 to 1)**	**2 (1 to 2)**[b]	–2.646	.008	–0.399
Shoulder internal rotation					
Re	9 (6 to 15)	9 (6 to 14)	–1.25	.21	–0.188
Ta	2 (1 to 4)	3 (2 to 4)	–0.958	.34	–0.144
Tr	4 (2 to 8)	4 (2 to 8)	–0.308	.76	–0.046
Mo	2 (1 to 4)	3 (2 to 4)	–0.828	.41	–0.125
Elbow flexion					
Re	**26 (22 to 36)**[b]	**22 (13 to 25)**	–3.036	.002	–0.458
Ta	**2 (1 to 4)**	**7 (4 to 13)**[b]	–3.555	<.001	–0.536
Tr	25 (21 to 30)	23 (20 to 28)	–1.055	.29	–0.159
Mo	**4 (3 to 5)**	**6 (4 to 7)**[b]	–2.354	.02	–0.355
Forearm pronation					
Re	68 (57 to 76)	67 (60 to 77)	–0.049	.96	–0.007
Ta	**11 (7 to 13)**	**26 (22 to 34)**[b]	–4.074	<.001	–0.614
Tr	**50 (44 to 55)**[b]	**42 (34 to 52)**	–2.484	.01	–0.374
Mo	8 (5 to 11)	9 (8 to 11)	–1.704	.09	–0.257
Wrist dorsal flexion					
Re	**16 (8 to 29)**[b]	**10 (7 to 14)**	–2.289	.02	–0.345
Ta	6 (4 to 9)	5 (3 to 7)	–1.412	.16	–0.213
Tr	8 (5 to 11)	4 (3 to 8)	–1.899	.06	–0.286
Mo	2 (1 to 4)	2 (1 to 3)	–0.601	.55	–0.091
Wrist ulnar deviation					
Re	**9 (5 to 14)**	**15 (9 to 22)**[b]	–2.386	.02	–0.36
Ta	**4 (2 to 6)**	**11 (6 to 18)**[b]	–3.88	<.001	–0.585
Tr	4 (3 to 6)	4 (2 to 7)	–0.86	.39	–0.13
Mo	2 (1 to 2)	1 (1 to 2)	–0.925	.36	–0.139
Neck flexion					

XSL•FO
RenderX

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Re	4 (2 to 9)	3 (2 to 6)	–1.769	.08	–0.267
Ta	0 (0 to 1)	1 (0 to 1)	–1.088	.28	–0.164
Tr	3 (1 to 6)	4 (2 to 6)	–0.99	.32	–0.149
Mo	**1 (1 to 2)**	**3 (2 to 4)**[b]	–3.1	.002	–0.467
Neck right lateral flexion					
Re	**2 (2 to 4)**	**5 (4 to 7)**[b]	–4.107	<.001	–0.619
Ta	0 (0 to 1)	1 (0 to 1)	–1.347	.18	–0.203
Tr	**1 (1 to 3)**	**2 (1 to 5)**[b]	–2.516	.01	–0.379
Mo	**1 (0 to 1)**	**2 (2 to 3)**[b]	–3.945	<.001	–0.595
Neck right rotation					
Re	**3 (2 to 5)**[b]	**2 (1 to 3)**	–3.717	<.001	–0.56
Ta	0 (0 to 1)	0 (0 to 1)	–0.016	.99	–0.002
Tr	**3 (1 to 4)**[b]	**1 (1 to 2)**	–2.971	.003	–0.448
Mo	**2 (1 to 3)**[b]	**2 (1 to 2)**	–2.354	.02	–0.355
Hip flexion					
Re	7 (3 to 9)	6 (3 to 8)	–1.412	.16	–0.213
Ta	**0 (0 to 0)**	**0 (0 to 1)**[b]	–2.354	.02	–0.355
Tr	5 (2 to 8)	5 (1 to 6)	–1.282	.20	–0.193
Mo	1 (1 to 1)	1 (0 to 2)	–0.308	.76	–0.046

[a]Significant differences in each phase, as determined using the Wilcoxon signed rank test ($P<.05$), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as "Re," the table phase as "Ta," the transporting phase as "Tr," and the mouth phase as "Mo."

[b]Larger values in comparisons between chopstick and spoon conditions.

In female participants (Table 4), changes in shoulder flexion, abduction, and neck right lateral flexion angles across most feeding phases were smaller with chopsticks than with a spoon. Additionally, changes in shoulder internal rotation, wrist ulnar deviation, and neck and hip flexion were smaller in some phases when using chopsticks. Elbow flexion changes during the reaching phase and forearm pronation changes during the transporting phase were larger in the chopstick condition, contrasting with smaller changes in elbow flexion during the table and mouth phases and smaller forearm pronation changes during the table phase. Wrist dorsiflexion and neck right rotation changes were similar across all feeding phases. Significant differences were found in small to large effect sizes.

XSL•FO
RenderX

Table . Comparison of joint angle changes between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in female participants[a].

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion					
Re	**11 (7 to 17)**	**25 (15 to 33)**[b]	–3.98	<.001	–0.614
Ta	**1 (1 to 1)**	**2 (1 to 3)**[b]	–2.485	.01	–0.383
Tr	**7 (5 to 13)**	**19 (15 to 28)**[b]	–3.875	<.001	–0.598
Mo	**3 (2 to 4)**	**5 (4 to 7)**[b]	–3.528	<.001	–0.544
Shoulder abduction					
Re	**8 (6 to 14)**	**13 (6 to 18)**[b]	–2.624	.01	–0.405
Ta	1 (0 to 2)	1 (1 to 2)	–1.581	.11	–0.244
Tr	**5 (3 to 8)**	**8 (5 to 17)**[b]	–2.589	.01	–0.399
Mo	**1 (1 to 2)**	**2 (1 to 3)**[b]	–2.45	.01	–0.378
Shoulder internal rotation					
Re	9 (6 to 12)	9 (6 to 16)	–1.651	.10	–0.255
Ta	**1 (1 to 2)**	**3 (2 to 3)**[b]	to 2.311	.02	–0.357
Tr	5 (3 to 6)	5 (4 to 10)	–1.894	.06	–0.292
Mo	**2 (2 to 4)**	**4 (3 to 5)**[b]	–3.076	.002	–0.475
Elbow flexion					
Re	**29 (22 to 32)**[b]	**21 (14 to 33)**	–2.172	.03	–0.335
Ta	**3 (2 to 4)**	**8 (6 to 12)**[b]	–3.702	<.001	–0.571
Tr	28 (22 to 32)	30 (24 to 34)	–0.782	.43	–0.121
Mo	**4 (3 to 6)**	**7 (5 to 8)**[b]	–3.702	<.001	–0.571
Forearm pronation					
Re	74 (63 to 87)	78 (67 to 82)	–0.678	.50	–0.105
Ta	**8 (5 to 15)**	**33 (23 to 45)**[b]	–4.015	<.001	–0.62
Tr	**58 (52 to 70)**[b]	**49 (38 to 53)**	–3.424	.001	–0.528
Mo	7 (3 to 11)	10 (6 to 12)	–1.616	.11	–0.249
Wrist dorsal flexion					
Re	12 (8 to 20)	11 (6 to 16)	–1.86	.06	–0.287
Ta	8 (5 to 12)	8 (3 to 13)	–0.504	.61	–0.078
Tr	5 (3 to 7)	7 (4 to 11)	–1.825	.07	–0.282
Mo	3 (2 to 4)	3 (2 to 3)	–1.13	.26	–0.174
Wrist ulnar deviation					
Re	**7 (3 to 9)**	**14 (10 to 25)**[b]	–3.354	.001	–0.518
Ta	**3 (1 to 3)**	**14 (6 to 24)**[b]	–4.015	<.001	–0.62
Tr	4 (3 to 5)	5 (3 to 6)	–1.269	.21	–0.196
Mo	1 (1 to 2)	1 (1 to 3)	–1.442	.15	–0.223

XSL•FO
RenderX

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Neck flexion					
Re	8 (3 to 13)	7 (5 to 10)	–1.408	.16	–0.217
Ta	1 (0 to 1)	1 (0 to 1)	–0.365	.72	–0.056
Tr	**8 (3 to 11)**	**10 (8 to 13)**[b]	–2.416	.02	–0.373
Mo	**2 (1 to -2)**	**3 (2 to 5)**[b]	–3.98	<.001	–0.614
Neck right lateral flexion					
Re	**2 (1 to 3)**	**4 (2 to 8)**[b]	–3.215	.001	–0.496
Ta	**0 (0 to 1)**	**1 (0 to 1)**[b]	–1.964	.05	–0.303
Tr	**2 (1 to 3)**	**3 (1 to 6)**[b]	–2.589	.01	–0.399
Mo	**1 (1 to 1)**	**2 (1 to 2)**[b]	–3.18	.001	–0.491
Neck right rotation					
Re	4 (2 to 5)	3 (1 to 5)	–1.303	.19	–0.201
Ta	0 (0 to 1)	0 (0 to 1)	–1.442	.15	–0.223
Tr	3 (2 to 4)	3 (2 to 4)	–0.956	.34	–0.148
Mo	2 (1 to 2)	2 (1 to 2)	–0.713	.48	–0.11
Hip flexion					
Re	6 (5 to 8)	4 (2 to 7)	–1.408	.16	–0.217
Ta	0 (0 to 0)	0 (0 to 0)	–1.512	.13	–0.233
Tr	4 (2 to 7)	4 (1 to 7)	–0.365	.72	–0.056
Mo	**1 (0 to 1)**	**2 (1 to 2)**[b]	–3.076	.002	–0.475

[a]Significant differences in each phase, as determined using the Wilcoxon signed rank test ($P<.05$), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”
[b]Larger values in comparisons between chopstick and spoon conditions.

Hand Spatiotemporal Parameters

In both sexes (Table 5 for males and Table 6 for females), the reaching phase duration was longer with chopsticks than with the spoon. In contrast, the mouth phase duration was shorter with chopsticks. Additionally, female participants showed a shorter transporting phase duration when using chopsticks. These differences showed small to large effect sizes.

Table . Comparison of movement times between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in male participants[a].

Variable and Phase	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Movement time (seconds)					
Re	**4.9 (2.4 to 9.6)**[b]	**2.9 (1.8 to 3.5)**	–3.393	.001	–0.512
Ta	1 (0.7 to 2.2)	1 (0.7 to 1.4)	–0.438	.66	–0.066
Tr	0.8 (0.6 to 1.4)	0.9 (0.7 to 1.2)	–1.136	.26	–0.171
Mo	**0.4 (0.3 to 0.5)**	**0.7 (0.6 to 0.8)**[b]	–3.979	<.001	–0.600

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test ($P<.05$), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”
[b]Larger values in comparisons between chopstick and spoon conditions.

XSL•FO
RenderX

Table . Comparison of movement times between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in female participants[a].

Variable an Phase	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Movement time (seconds)					
Re	**4.8 (3.2 to 7.3)**[b]	**2.8 (1.6 to 4.7)**	–3.702	<.001	–0.571
Ta	1.1 (0.9 to 1.8)	1.5 (1.3 to 1.7)	–0.852	.39	–0.131
Tr	**0.9 (0.8 to 1.1)**	**1.1 (1 to 1.2)**[b]	–2.103	.04	–0.324
Mo	**0.5 (0.4 to 0.7)**	**0.9 (0.8 to 1.2)**[b]	–3.98	<.001	–0.614

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test ($P<.05$), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b]Larger values in comparisons between chopstick and spoon conditions.

During the transporting phase in both sexes (Table 7 for males and Table 8 for females), the actual distance traveled was shorter with chopsticks; however, the relative distance traveled was greater than that with a spoon. Velocity and maximum velocity were lower with chopsticks than with spoon. The timing of maximum velocity and the number of movement units used were comparable between utensils. Small to large effect sizes were observed for significant differences between utensil conditions.

Table . Comparison of hand spatiotemporal parameters between chopstick use for eating solid foods and spoon use for eating semisolid foods during the transporting phase in male participants[a].

Variables	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Actual distance traveled (meters)	**0.11 (0.08 to 0.12)**	**0.14 (0.11 to 0.18)**[b]	–3.782	<.001	–0.570
Relative distance traveled	**1.17 (1.04 to 1.35)**[b]	**1.07 (1.04 to 1.11)**	–2.062	.04	–0.311
Velocity (meters per second)	**0.11 (0.08 to -0.15)**	**0.16 (0.14 to 0.19)**[b]	–3.393	.001	–0.512
Maximum velocity (meters per second)	**0.18 (0.16 to 0.28)**	**0.25 (0.22 to 0.32)**[b]	–3.068	.002	–0.463
Timing of maximum velocity (%)	70.6 (61.7 to 83.2)	60.2 (45.1 to 75.8)	–1.704	.09	–0.257
Number of movement units	5 (3.8 to 8)	4 (3 to 7)	–1.325	.19	–0.200

[a] Significant differences, as determined using the Wilcoxon signed rank test ($P<.05$), are shown as values in boldface. "*r*" denotes the effect size.

[b]Larger values in comparisons between chopstick and spoon conditions.

XSL•FO
RenderX

Table . Comparison of hand spatiotemporal parameters between chopstick use for eating solid foods and spoon use for eating semisolid foods during the transporting phase in female participants[a].

Variables	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Actual distance traveled (meters)	**0.12 (0.11 to 0.14)**	**0.2 (0.17 to 0.22)**[b]	–3.91	<.001	–0.603
Relative distance traveled	**1.15 (1.07 to 1.2)**[b]	**1.05 (1.04 to 1.09)**	–3.424	.001	–0.528
Velocity (meters per second)	**0.12 (0.1 to 0.16)**	**0.18 (0.14 to 0.2)**[b]	–3.702	<.001	–0.571
Maximum velocity (meters per second)	**0.22 (0.19 to 0.29)**	**0.35 (0.28 to 0.37)**[b]	–3.875	<.001	–0.598
Timing of maximum velocity (%)	70.3 (63.4 to 75.9)	63.9 (56.4 to 69.2)	–1.338	.18	–0.206
Number of movement units	5 (4 to 6)	6 (4.5 to 6.5)	–1.171	.24	–0.181

[a] Significant differences, as determined using the Wilcoxon signed rank test (P<.05), are shown as values in boldface. "*r*" denotes the effect size.

[b]Larger values in comparisons between chopstick and spoon conditions.

Discussion

Principal Findings in Joint Kinematics

The aim of this study was to determine upper body joint angles and hand spatiotemporal parameters during the use of chopsticks for solid foods and a spoon for semisolid foods across all feeding phases by comparing these conditions. In addition to these findings observed in both sexes, using subgroup analysis, male participants demonstrated greater neck rotation motions, while female participants exhibited fewer wrist dorsiflexion angles under the chopstick condition. This comparison is crucial for clinical assessments of individuals with upper body sensorimotor dysfunctions, considering food customs in East Asia.

Various joint angle positions of the upper limb and neck were lower in the chopstick condition than in the spoon condition, except for most elbow flexion angles and some forearm pronation, wrist ulnar deviation (in females), and neck right rotation (in males) angles being higher in the chopstick condition. In addition, female participants showed smaller wrist dorsiflexion positions with chopstick use than with spoon use. However, neck and hip flexion angles were generally comparable between the conditions.

The changes in upper body joint angles across phases were smaller in the chopstick condition than in the spoon condition, except for greater or fewer changes in elbow and forearm joint angles with chopsticks, and greater wrist dorsiflexion and neck right rotation angles in male participants. Based on the comparison between the conditions, chopsticks required smaller joint angle positions and changes in shoulder and neck flexion and right lateral flexion across feeding phases and sexes.

These findings align with our hypotheses and the results from previous reports on shoulder and neck joint angles [6,12]. Chopstick use requires less control of shoulder joint motions due to how food is held compared with spoon use. This rationale is supported by the fact that chopsticks may securely hold food through a “scissor-pinching” motion, which reduces trapezius muscle activity compared with “pincer pinching” [26]. The scissor-pinching technique used with chopsticks allows for a stronger grip on the food [27]. The reduction in muscle activity may explain the smaller shoulder joint motions observed in the chopstick condition. In addition, the finding of fewer neck joint angle positions and changes in our chopstick condition is supported by findings in studies showing smaller head or trunk region motions while eating solid foods with a fork or a spoon (not-easy-to-drop-foods condition), compared with eating liquid foods with a spoon (easy-to-drop-foods condition) [8,11]. The actual differences in shoulder joint motions to control utensils and compensatory neck joint motions between chopstick use for solid foods and spoon use for semisolid foods should be clinically highlighted as normative information. However, the current experimental conditions did not clarify whether food was prone to being dropped, indicating the need for further studies.

Notably, in male participants, neck right rotation angles and angle changes during some phases were greater in the chopstick condition than in the spoon condition. Additionally, hip flexion angles and changes, which indicate trunk positioning and motion, did not differ significantly between conditions in either sex group. Proximal body part stability or movements are considered the same. Although they were not anticipated at the start of the study, these postures and movements could be important for clinical assessments. Proximal body part motions are often more pronounced when eating foods prone to being dropped [8,11], which was likely the case for the spoon condition in this study. However, this remains unclear and warrants further investigation under adjusted conditions.

Contrary to our hypothesis, motion extent at the middle and distal joints using chopsticks was not consistently reduced. For instance, when transporting food to the mouth, forearm pronation angle changes were greater in the chopstick condition than in the spoon condition. However, during the table phase (when taking food from the bowl), the changes were not significant in the chopstick condition. These joint motions are

XSL•FO
RenderX

thought to be essential for approaching and manipulating food directly.

Principal Findings in Hand Spatiotemporal Parameters

Hand spatiotemporal parameters suggest that, in contrast with spoon use, chopstick use required more time during the reaching phase and less time during the mouth phase in both sex groups, with female participants also showing shorter times during the transporting phase. These utensil and food type–related differences contradicted our hypothesis. Furthermore, this finding was not recognized in a previous study [6]. A possible reason for the contrasting results is that the previous study used pickles as the food item under both the chopstick and spoon conditions, whereas this study used pickles (solid food) with chopsticks and yogurt (semisolid food) with a spoon to align with the study aim. Yogurt may be easier to consume without chewing (requiring less time during the reaching phase) than pickles [28]. However, yogurt may be more prone to spilling than pickles (requiring more time during the transporting and mouth phases). Accordingly, the present findings related to duration differences in conditions lead to the assumption that this was due to utensil use along with respective food characteristics such as viscosity and not just dependent on utensil types.

Hand spatiotemporal parameters during the reaching movement have been reported [13,14], with clearly defined start and end positions and movement not disrupted. Accordingly, this study analyzed parameters related to distance, speed, motor control strategy, and smoothness during the transporting phase, which showed no disruption during reaching or transporting. The results were interpreted in accordance with the common concept of spatiotemporal parameters to contribute to clinical assessments.

First, the actual distance traveled by the hand from the bowl to the mouth was shorter with chopsticks than with a spoon. Furthermore, the hand holding the chopsticks moved at a slower speed. This slowness might have been caused by the shorter transfer distance [29] observed in the chopstick condition than in the spoon condition. These shorter and slower hand transfers might be due to the larger forearm rotations and smaller shoulder flexion and abduction motions associated with chopstick use, as noted earlier. In contrast, when transporting food with a spoon, the larger hand transfer distance and higher speed might be due to increased shoulder motions rather than forearm movements. Thereby, the effective hand spatiotemporal transfers during the transporting phase might be mainly caused by shoulder joint motion and not forearm joint motion. Contrastingly, forearm joint motion might affect well transfer of the utensil tip rather than the hand.

The relative distance (ie, the ratio between the actual and theoretical shortest distance) for the hand with chopsticks was greater than for the hand with the spoon, indicating lower efficiency in the hand spatial transfer. This also suggests that forearm rotation might make the movement in the tip of chopsticks efficient, not in the hand holding chopsticks. However, the relationship between hand spatiotemporal parameters as well as utensil tip motion and joint angles is beyond the scope of this study and considered an area for future research.

Additionally, we analyzed the timing of maximum velocity (motor control strategy) and number of movement units (smoothness) and found no obvious differences between utensil conditions. The comparable reaching strategies and movement smoothness between conditions might indicate that the healthy participants in this study had well-learned motor control for these modes of feeding. These characteristics are generally assessed in patients [13]; this study provides valuable data on comparable motor control and smoothness between 2 utensil-food type conditions in normal hand motions regarding such assessments.

Clinical Implications

Clinicians might determine appropriate eating utensils used by patients based on their ability to use them, considering factors such as paresis and intellectual ability [2]. Rheumatoid arthritis [30] and Parkinson disease [31] may also be included in target populations as they have altered upper body kinematics during functional tasks. The assessment can refer to the findings of this study regarding joint angle positions, ranges of joint motion, and hand spatiotemporal parameters as fundamental normal movement knowledge, which could support clinical judgments when selecting chopsticks for eating solid foods or a spoon for eating semisolid foods in East Asian countries. For instance, individuals with limited shoulder flexion or abduction or increased elbow flexion angles might be suited to using chopsticks to eat solid foods, which include small-sized pickles used in this study, bite-sized meats and vegetables, or sticky rice; conversely, those who demonstrate significant forearm rotation during the table phase might benefit from using a spoon to eat semisolid foods, which include yogurt, jelly dessert, or rice porridge. Movements for using a spoon to eat semisolid foods may also be consistent with those for using a spoon to eat liquid foods such as soup or solid foods such as separated rice and finely chopped vegetables. These combinations of utensils and foods, alongside the typical movements currently revealed, would be realistic about food customs in East Asia. Additionally, differences in hip flexion positions and motions (ie, trunk inclination) may not be related to utensil selection.

In the rehabilitation of eating difficulties in East Asian countries, when changing food forms from semisolid to solid, patients' forearm supination should be well confirmed for using chopsticks. Compared with the spoon condition, chopstick use typically involves shorter and slower hand movements during the transporting phase, which can be considered normal. If the nondominant hand is trained for chopstick use with solid foods, improvements as spoon use with semisolid foods in control strategy and movement smoothness—similar to those observed in this study—could be assessed, as recently reported [32,33]. Hand spatiotemporal parameters involving nondominant or impaired hands using utensils could also be evaluated based on the current findings. Although male participants showed greater neck right rotation positions and motions and female participants showed smaller wrist dorsiflexion positions in chopstick condition than in spoon condition should be assessed as individuals with normal state in each disease, these sex-related

XSL•FO
RenderX

characteristics would be emphasized particularly in diseases with ununiform sex ratio such as collagen disease and cervical spondylotic myelopathy.

Limitations and Future Research

The present results were derived from measurements obtained during specific movement tasks using an inertial motion-capturing system, which may limit the generalizability of the findings to feeding movements observed in clinical settings. Future studies should investigate whether certain foods could be more prone to being dropped by examining the effects of different eating utensils (eg, chopsticks with supportive features, forks, and spoons) and various types of foods (eg, viscosity, size, and forms) separately, as these may be confounding factors in the current experiments; moreover, these studies should define these effects on feeding movements. This approach might help verify and expand upon the current findings. Drawing connections between joint motions and the hand spatiotemporal parameters and utensil tip motions, despite being out of this study objectives but discussed, might also be a valuable method for more deeply understanding the differences between the current 2 utensil-food type conditions; thus, eating movements would be assessed well. Moreover, future research should also involve populations with upper body dysfunctions.

Conclusions

This study investigated upper body joint angles and hand spatiotemporal parameters during feeding movements involving chopsticks for solid food and a spoon for semisolid food across 4 feeding phases and both sexes. The results indicate that chopstick use, regardless of sex, generally requires smaller shoulder and neck joint angle positions and motions across phases but involves larger elbow and forearm joint motions based on the phase. Additionally, when transporting food from the bowl to the mouth using chopsticks, the hand travels a shorter distance and moves slowly. Spoon use, in contrast, involves larger elbow, forearm, and wrist ulnar deviation motions when taking up semisolid food from the bowl. Regarding sex differences, male participants exhibit increased right neck rotation positions and motions when using chopsticks compared with using a spoon, whereas female individuals demonstrate reduced wrist dorsiflexion positions under the chopstick condition relative to the spoon condition. These movement characteristics, influenced by the combination of utensils and food types specific to East Asia countries, could facilitate clinical assessments of feeding behaviors specific to food customs in individuals with upper body dysfunction and diseases specific to each sex and aid in determining adaptive utensils.

Funding

This study was supported by a clinical research support grant from the University of Miyazaki Hospital.

Authors' Contributions

JN contributed to study conceptualization and design, data acquisition, formal analysis and interpretation of data, writing and editing the manuscript, and approval of the final version. SM and SM contributed formal analysis and interpretation of data, reviewing and editing the manuscript, and approval of the final version. HA and EC contributed to study conceptualization and design, reviewing and editing the manuscript, and approval of the final version.

Conflicts of Interest

None declared.

References

1. Clavé P, Terré R, de Kraa M, Serra M. Approaching oropharyngeal dysphagia. Rev Esp Enferm Dig 2004 Feb;96(2):119-131. [doi: 10.4321/s1130-01082004000200005] [Medline: 15255021]
2. Iokawa K, Sone T, Fujita T, Tsukada T, Kaneda M, Hasegawa K. Functional and cognitive variables predicting successful use of chopsticks or a spoon by the paretic upper extremity in patients following stroke: a cross-sectional study. Top Stroke Rehabil 2019 Jan;26(1):1-5. [doi: 10.1080/10749357.2018.1536021] [Medline: 30355062]
3. Hill SW, Mong S, Vo Q. Three-dimensional motion analysis for occupational therapy upper extremity assessment and rehabilitation: a scoping review. The Open Journal of Occupational Therapy 2022;10(4):1-14. [doi: 10.15453/2168-6408.1901]
4. Nakatake J, Totoribe K, Chosa E. Whole-body joint angles necessary for eating with chopsticks. Jpn Occup Ther Res 2019;38:163-170. [doi: 10.32178/jotr.38.2_163]
5. Nakatake J, Totoribe K, Arakawa H, Chosa E. Exploring whole-body kinematics when eating real foods with the dominant hand in healthy adults. PLoS One 2021;16(10):e0259184. [doi: 10.1371/journal.pone.0259184] [Medline: 34710151]
6. Nagao T. Joint motion analysis of the upper extremity required for eating activities. Bull Health Sci Kobe 2004;19:13-31. [doi: 10.24546/00391929]
7. Chang BC, Huang BS, Chen CK, Wang SJ. The pincer chopsticks: the investigation of a new utensil in pinching function. Appl Ergon 2007 May;38(3):385-390. [doi: 10.1016/j.apergo.2006.03.009] [Medline: 16765903]
8. Chinju K, Yamamoto Y, Inada E, Iwashita Y, Sato H. Analysis of head motions during food intake in Japanese adults using a new motion capture system. Arch Oral Biol 2024 Apr;160:105908. [doi: 10.1016/j.archoralbio.2024.105908] [Medline: 38335700]

XSL•FO
RenderX

9. Inada E, Saitoh I, Nakakura-Ohshima K, et al. Association between mouth opening and upper body movement with intake of different-size food pieces during eating. Arch Oral Biol 2012 Mar;57(3):307-313. [doi: 10.1016/j.archoralbio.2011.08.023] [Medline: 21975117]
10. Doğan M, Koçak M, Onursal Kılınç Ö, et al. Functional range of motion in the upper extremity and trunk joints: nine functional everyday tasks with inertial sensors. Gait Posture 2019 May;70:141-147. [doi: 10.1016/j.gaitpost.2019.02.024] [Medline: 30875600]
11. van der Kamp J, Steenbergen B. The kinematics of eating with a spoon: bringing the food to the mouth, or the mouth to the food? Exp Brain Res 1999 Nov;129(1):68-76. [doi: 10.1007/s002210050937] [Medline: 10550504]
12. Inagaki Y, Ishida T, Sugimori H, et al. Functional range of motion for basic seated activities of daily living tasks. Front Sports Act Living 2025;7:1646326. [doi: 10.3389/fspor.2025.1646326]
13. de los Reyes-Guzmán A, Dimbwadyo-Terrer I, Trincado-Alonso F, Monasterio-Huelin F, Torricelli D, Gil-Agudo A. Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: a review. Clin Biomech (Bristol, Avon) 2014 Aug;29(7):719-727. [doi: 10.1016/j.clinbiomech.2014.06.013]
14. Schwarz A, Kanzler CM, Lambercy O, Luft AR, Veerbeek JM. Systematic review on kinematic assessments of upper limb movements after stroke. Stroke 2019 Mar;50(3):718-727. [doi: 10.1161/STROKEAHA.118.023531] [Medline: 30776997]
15. Robert-Lachaine X, Mecheri H, Larue C, Plamondon A. Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis. Med Biol Eng Comput 2017 Apr;55(4):609-619. [doi: 10.1007/s11517-016-1537-2] [Medline: 27379397]
16. Al-Amri M, Nicholas K, Button K, Sparkes V, Sheeran L, Davies JL. Inertial measurement units for clinical movement analysis: reliability and concurrent validity. Sensors (Basel) 2018 Feb 28;18(3):719. [doi: 10.3390/s18030719] [Medline: 29495600]
17. Zhang JT, Novak AC, Brouwer B, Li Q. Concurrent validation of Xsens MVN measurement of lower limb joint angular kinematics. Physiol Meas 2013 Aug;34(8):N63-N69. [doi: 10.1088/0967-3334/34/8/N63] [Medline: 23893094]
18. Kim S, Nussbaum MA. Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks. Ergonomics 2013;56(2):314-326. [doi: 10.1080/00140139.2012.742932] [Medline: 23231730]
19. Pedro B, Cabral S, Veloso AP. Concurrent validity of an inertial measurement system in tennis forehand drive. J Biomech 2021 May 24;121:110410. [doi: 10.1016/j.jbiomech.2021.110410] [Medline: 33852942]
20. Kwakkel G, van Wegen EEH, Burridge JH, et al. Standardized measurement of quality of upper limb movement after stroke: consensus-based core recommendations from the second stroke recovery and rehabilitation roundtable. Neurorehabil Neural Repair 2019 Nov;33(11):951-958. [doi: 10.1177/1545968319886477] [Medline: 31660781]
21. Nakatake J, Totoribe K, Chosa E, Yamako G, Miyazaki S. Influence of gender differences on range of motion and joint angles during eating in young, healthy Japanese adults. Prog Rehabil Med 2017;2:20170011. [doi: 10.2490/prm.20170011] [Medline: 32789218]
22. Mesquita IA, Fonseca PFPD, Borgonovo-Santos M, et al. Comparison of upper limb kinematics in two activities of daily living with different handling requirements. Hum Mov Sci 2020 Aug;72:102632. [doi: 10.1016/j.humov.2020.102632] [Medline: 32452388]
23. Nakatake J, Arakawa H, Tajima T, Miyazaki S, Chosa E. Age- and sex-related differences in upper-body joint and endpoint kinematics during a drinking task in healthy adults. PeerJ 2023;11:e16571. [doi: 10.7717/peerj.16571] [Medline: 38144196]
24. Waslen A, Friesen KB, Lang AE. Do sex and age influence scapular and thoracohumeral kinematics during a functional task protocol? J Appl Biomech 2024 Feb 1;40(1):29-39. [doi: 10.1123/jab.2023-0085] [Medline: 37917968]
25. Zieliński G. Effect size guidelines for individual and group differences in physiotherapy. Arch Phys Med Rehabil 2025 Dec;106(12):1844-1849. [doi: 10.1016/j.apmr.2025.05.013]
26. Yoo IG, Yoo WG. Comparison of effects of pincer- and scissor-pinching modes of chopstick operation on shoulder and forearm muscle activation during a simulated eating task. J Phys Ther Sci 2012;24(10):953-954. [doi: 10.1589/jpts.24.953]
27. Chen YL. Effects of shape and operation of chopsticks on food-serving performance. Appl Ergon 1998 Aug;29(4):233-238. [doi: 10.1016/s0003-6870(97)00046-x] [Medline: 9701536]
28. Aguayo-Mendoza MG, Ketel EC, van der Linden E, Forde CG, Piqueras-Fiszman B, Stieger M. Oral processing behavior of drinkable, spoonable and chewable foods is primarily determined by rheological and mechanical food properties. Food Qual Prefer 2019 Jan;71:87-95. [doi: 10.1016/j.foodqual.2018.06.006]
29. Gordon J, Ghilardi MF, Cooper SE, Ghez C. Accuracy of planar reaching movements. II. Systematic extent errors resulting from inertial anisotropy. Exp Brain Res 1994;99(1):112-130. [doi: 10.1007/BF00241416] [Medline: 7925785]
30. Gur Kabul E, Unver F, Alptekin A, et al. The effect of rheumatoid arthritis on upper extremity functions: a kinematic perspective. Int J Rheum Dis 2022 Nov;25(11):1279-1287. [doi: 10.1111/1756-185X.14421] [Medline: 35965381]
31. Romano P, Pournajaf S, Ottaviani M, et al. Sensor network for analyzing upper body strategies in Parkinson's disease versus normative kinematic patterns. Sensors (Basel) 2021 May 31;21(11):3823. [doi: 10.3390/s21113823] [Medline: 34073123]

XSL•FO
RenderX

32. Sawamura D, Sakuraba S, Suzuki Y, et al. Acquisition of chopstick-operation skills with the non-dominant hand and concomitant changes in brain activity. Sci Rep 2019 Dec 31;9(1):20397. [doi: 10.1038/s41598-019-56956-0] [Medline: 31892724]
33. Sawamura D, Sakuraba S, Yoshida K, et al. Chopstick operation training with the left non-dominant hand. Transl Neurosci 2021 Jan 1;12(1):385-395. [doi: 10.1515/tnsci-2020-0189] [Medline: 34721894]

Edited by S Munce; submitted 19.Apr.2025; peer-reviewed by P Gawda, S Hill; revised version received 09.Nov.2025; accepted 13.Jan.2026; published 06.Mar.2026.

Please cite as:
Nakatake J, Maeda S, Miyazaki S, Arakawa H, Chosa E
Comparison of Upper Body Joint and Hand Motions in Eating Solid Foods With Chopsticks and Semisolid Foods With a Spoon in Healthy Males and Females: Observational Study
JMIR Rehabil Assist Technol 2026;13:e76239
URL: https://rehab.jmir.org/2026/1/e76239
doi:10.2196/76239

© Jun Nakatake, Shogo Maeda, Shigeaki Miyazaki, Hideki Arakawa, Etsuo Chosa. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Barriers and Enablers for the Intention to Use Assistive Robotics Among People With Spinal Cord Injury and Those Involved in Their Care: Qualitative Study

Susanne Frennert[1], PhD; Johanna Persson[1], PhD; Eva Díez-Rodríguez[2], MSc; Monica Alcobendas-Maestro[2], MSc; Fátima Villamayor Vega[2], MSc; Antonio Oliviero[2], PhD

[1]Department of Design Sciences, Division of Ergonomics and Aerosol Technology, Ingvar Kamprad Designcentrum (IKDC), Lund University, FormStråket 12, Lund, Sweden

[2]FENNSI Group, Hospital Nacional de Parapléjicos and "Instituto de Investigación Sanitaria de Castilla-La Mancha" (IDISCAM), SESCAM, Toledo, Spain

Corresponding Author:
Susanne Frennert, PhD
Department of Design Sciences, Division of Ergonomics and Aerosol Technology, Ingvar Kamprad Designcentrum (IKDC), Lund University, FormStråket 12, Lund, Sweden

Abstract

Background: Spinal cord injury (SCI) may, and often does, profoundly reshape daily life, altering physical abilities, social roles, and personal identities. While assistive technologies, including assistive robotics, are often framed as solutions to re-establish independence, their adoption is shaped by practical, emotional, and social considerations as well as functional qualities. Individuals with SCI, their relatives, and health care professionals need to navigate complex dynamics when encountering assistive robotics. Understanding how assistive technologies are perceived and positioned in everyday life may help developers and designers create assistive robotics that are meaningful and useful for intended users.

Objective: The aim of this qualitative study was to explore how individuals with SCI, relatives, and health care professionals working with patients with SCI perceive and describe the possibilities and limitations of assistive robotics. The study sought to understand the factors that influence the intention to use assistive robotics among individuals with SCIs.

Methods: We used a qualitative approach, conducting semi-structured interviews and participatory workshops in Sweden and Spain. In total, the study involved 18 interview participants with SCI, 21 workshop participants with SCI, 12 relatives, and 26 health care professionals. The interviews and workshops elicited reflections on participants' experiences, expectations, and concerns regarding assistive robotics in general and supernumerary robotic limbs in particular. Data were analyzed using reflexive thematic analysis, with a focus on interpreting the meanings embedded in participants' narratives.

Results: The analysis showed that participants' engagement with assistive robotics was influenced by expectations of technological benefits and by practical constraints in everyday life. The main barriers identified were practical constraints, including the subthemes "navigating a changing reality," "difficulties with awareness and access" and "concerns about costs"; and interaction with robots, including "doubts about meaningfulness," "uncertainty regarding reliability and safety," "uneasiness about competence" and "apprehension of social norms." Participants' visions of enhanced self-efficacy through assistive robotics were described as important enablers of the intention to use and motivation to try assistive robotics. Shared expectations and concerns about future technologies (technological imaginaries) also influenced how participants talked about assistive robotics.

Conclusions: Rather than presenting assistive robotics as an inevitable progression toward greater autonomy, this study highlights the complexities and contingencies that shape how individuals relate to assistive robotics in general and supernumerary robotic limbs in particular. Participants' responses illustrate that robotic assistance is not merely a question of technological feasibility but is deeply entangled with embodied experiences, shifting identities, and evolving social relations. While visions of independence through assistive robotics remain compelling among participants, sociotechnical imaginaries coexist with concerns about meaningful engagement, reliability, safety, competence, and social norms, as well as challenges related to transition periods, costs, and limited awareness and access to assistive robotics.

(JMIR Rehabil Assist Technol 2026;13:e72080) doi:10.2196/72080

KEYWORDS

barriers; enabler; assistive robotics; spinal cord injury; spinal cord injury care

XSL•FO
RenderX

Introduction

Spinal cord injury (SCI) has a substantial impact on quality of life, affecting physical, psychological, social, and economic aspects of an individual's life. The severity and level of the injury influence the extent of this impact, often resulting in paralysis or loss of sensation and a need for extensive assistance in everyday life [1]. This, in turn, negatively affects the ability of individuals with SCI to independently perform everyday tasks, such as hygiene, eating, and dressing. Frequently, they depend on assistance from caregivers, relatives, and various assistive aids [2].

In robotics labs around the world, several projects are being conducted to develop valuable robotic assistance for people with SCI, such as mobile manipulators [3] and robot-assisted feeding [4]. In this paper, we use the term "assistive robotics" to refer to robotic devices that provide physical assistance for everyday activities to people with upper body-limb disabilities (eg, robotic arms, powered gloves, and exoskeletons that support feeding, reaching, dressing, or grooming).

However, despite the potential benefits of assistive technologies, abandonment rates remain high [5], often due to technology-driven design that overlooks the social and cultural factors affecting user acceptance [6]. A key factor influencing user abandonment is the intended users' perception of the technology in question [7-9].

Conceptual Framework and Research Gap

The perception of technology and its relation to the intention to use are explained by the Technology Acceptance Model (TAM) [10] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [11]. TAM focuses on personal beliefs about usefulness and ease of use [10], while UTAUT adds social influence and the broader conditions surrounding the use of technology [11]. In the context of individuals with SCI, the models suggest that the intentions to use assistive technologies are influenced not by personal beliefs but also by the views and practices of caregivers, relatives, and health care professionals (HCPs) who organize care and everyday use [12].

At the same time, TAM and UTAUT were developed for information systems and "able-bodied individuals" [10,11], and may therefore not fully capture the situated, embodied, and relational aspects of technology adoption in a disability context. Building on these models, work in assistive technologies has emphasized the importance of perceived meaningfulness, identity, and the organization of support around assistive technologies [7,9,13,14]. Concepts from Science and Technology Studies such as sociotechnical imaginaries draw attention to shared visions and narratives about how emerging technologies might reshape care and independence [15-19]. Sociotechnical imaginaries highlight how hopes, fears, and expectations about technologies are collectively produced.

Existing research on assistive robotics and people with SCI has primarily addressed rehabilitation and technical feasibility. Studies have investigated robots for rehabilitation [20-22], work participation [23,24], technology use and priorities [25,26], outcome measures [27] and user-centered design guidelines [28]. Several recent systematic reviews and state-of-the-art papers synthesize the development of upper-limb assistive and rehabilitation robots, including exoskeletons, end-effector devices, soft exosuits for people with cervical SCI and other neuromuscular conditions [29-33]. The reviews describe control strategies, clinical indications, motor outcomes, and identify challenges, such as comfort, wearability, and robustness. Furthermore, a systematic review has examined the barriers and facilitators to exoskeleton use [34] and broader acceptance of assistive technologies among people with motor disabilities [9].

Across the existing literature, user perspectives are most often operationalized in terms of usability, comfort, and short-term feasibility, focusing on the person undergoing rehabilitation [9,29,32]. While some studies included HCPs and, more rarely, relatives as respondents, they are not systematically treated as stakeholders, even though they are often responsible for procuring assistive technologies, providing training, and embedding them in care routines [34]. There are also very few empirical studies on how emerging technologies such as supernumerary robotic limbs (SRLs) are imagined before they reach the market [30,33], which leaves a gap in understanding how assistive robotics are positioned within the everyday lives of people with SCI and those involved in their care.

To address the identified gap, this study is conducted within the HARIA project, which develops SRLs to augment upper-limb function in people with substantial arm and hand impairments [35]. SRLs are wearable robotic arms and fingers, designed to enhance human sensorimotor abilities, particularly in tasks involving movement and coordination [36]. For individuals with SCI, SRLs are proposed as aids to compensate for lost functions in impaired limbs, provide necessary assistance, and substitute for impaired limbs in terms of functionality [37].

Aim and Research Question

The objective of this qualitative study was to explore how individuals with SCI, their relatives, and HCPs conceptualize, engage with, and negotiate the possibilities and limitations of assistive robotics in general, with a particular focus on SRLs. More specifically, we examined (1) intentions to use emerging assistive robotics that are not yet widely available, (2) how these intentions are shaped by experiences of both used and abandoned assistive technologies, and (3) how perspectives differ or converge across stakeholders, including individuals with SCI, relatives, and HCPs in Sweden and Spain.

We asked: What barriers and enablers shape the intention to use assistive robotics among people with SCI and those involved in their everyday care in Sweden and Spain? By exploring this question, we aimed to clarify how assistive robotics are positioned within the everyday realities of potential users and their caregivers and to inform the design and implementation of emerging assistive robotics.

Methods

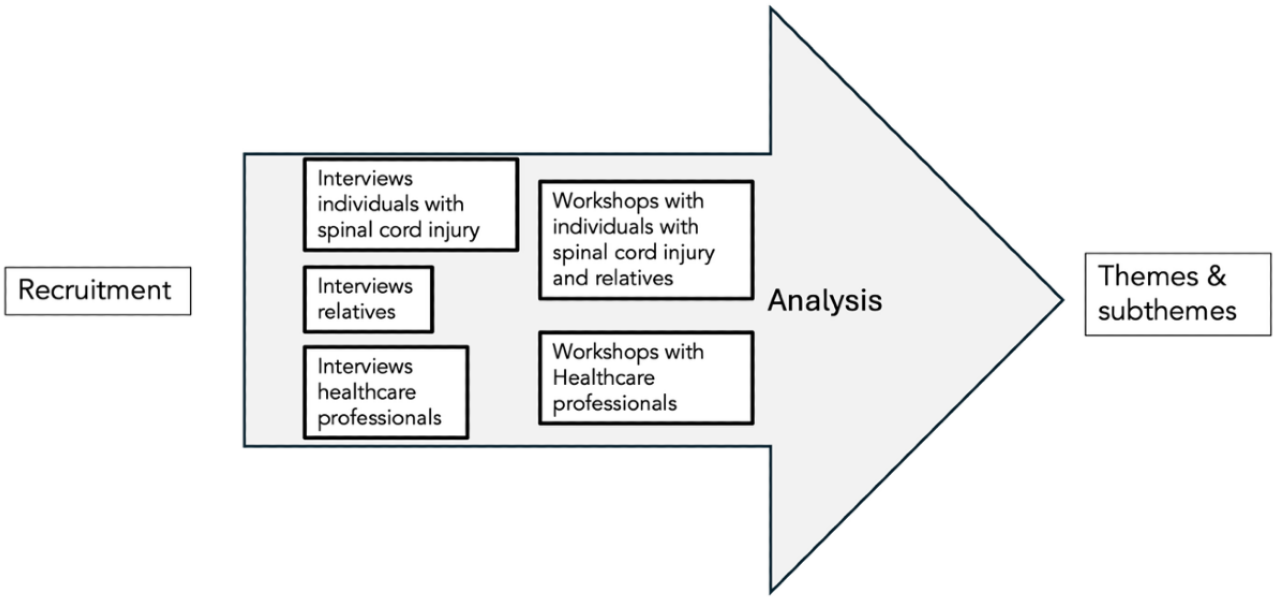
Study Design

This qualitative exploratory study is part of a broader research project centered around SRLs [35,37-39]. We adopted an

XSL•FO
RenderX

interpretive, constructivist approach [40] and used reflexive thematic analysis [41] to explore how different stakeholders make sense of assistive robotics in the context of SCI. Within this study, our focus is on individuals with SCIs and their intention to use assistive robotics, with a particular emphasis on SRLs. We combined individual semi-structured interviews [42] with participatory workshops [43] to capture both in-depth personal narratives and group-based discussions and co-imaginings of future assistive robotics (Figure 1).

Figure 1. Study design.



The multidisciplinary research team included researchers in rehabilitation engineering, design science and technology studies, and clinicians, several of whom had prior experience of SCI rehabilitation and assistive technology research. The diversity of the research team informed the development of the interview and workshop guides and the interpretation of the data.

Setting and Participants

To understand the social dynamics and contextual factors pertinent to assistive robotics, our participant pool included not only individuals with SCI but also HCPs specializing in SCIs, professionals working closely with individuals affected by SCIs, as well as relatives and caregivers. The study was conducted in both Spain and Sweden. In Sweden, participant recruitment involved collaboration with Paraplegic Associations, using their networks to disseminate information about the study to potential participants. Additionally, outreach efforts were extended to various health care facilities specializing in SCIs. In Spain, recruitment targeted HCPs, individuals with SCIs, and relatives through collaboration with Sescam - Gobierno de Castilla-La Mancha, a hospital and rehabilitation facility specializing in SCIs, which facilitated the identification and inclusion of relevant participants.

We focused on adults with cervical SCI because SRLs in the HARIA project aim to augment arm and hand function [35]. We also included relatives and HCPs, based on the assumption that the intention to use assistive robotics is affected by those who support and care for patients with SCI.

Eligibility and Sample

We contacted potential participants through email and phone. From this pool, we selected a purposive sample by applying specific criteria. To qualify for participation, individuals needed to fulfill the following criteria: (1) be 18 years or older, (2) have acquired a cervical SCI as an adult, (3) live in the community (not in long-term institutional care), and (4) be able to participate in an interview or workshop in Swedish or Spanish. Eligibility extended to those who were either a relative of an individual with a SCI or HCPs working with patients with SCI.

We aimed for heterogeneity to capture a wide range of lived experiences, needs, and expectations. Therefore, we did not set inclusion and exclusion criteria regarding time since injury, lesion level within the cervical region, functional level, or previous experience with assistive technologies or robotics. None of the participants had prior hands-on experience with SRLs, as SRLs were introduced by the researchers during the workshops (Multimedia Appendix 1).

Participants were recruited between January 2023 and September 2023. Recruitment was discontinued when we considered that data saturation was reached, indicating that further interviews produced repetitive information and no new themes or responses emerged [41].

Data Collection

Data were gathered through audio-recorded qualitative semistructured interviews [42,44] and participatory workshops [43]. The first author formulated the interview and workshop guides (Multimedia Appendices 1-5), subsequently refining them in collaboration with the rest of the research team [42,45].

XSL•FO
RenderX

The interview questions probed into the experiences of individuals living with SCIs, exploring their current use of assistive technologies and those received but not used (Multimedia Appendix 2). Such questions were included to situate participants' views on robotics within their broader histories of adopting, adapting to, and sometimes abandoning assistive technologies. Later during the interview, the focus shifted to assistive robotics technology, with questions aimed at understanding the interviewees' experiences with assistive technologies. During the latter part of the interview, the interviewer provided a brief introduction to SRLs, as the interviewees lacked prior knowledge, and SRLs are not yet commonly available on the market but are currently under development in research laboratories around the world [36,37,46]. Preferences and ideas regarding the use of assistive robotics technology were explored by asking questions about the functions requested, the required design for actual use, and how the interviewees would describe their needs and wishes concerning assistive robotics.

Interview questions and workshop activities for relatives and HCPs followed the same overall structure as for individuals with SCIs but were designed to explore their insights into the needs of the end-users (individuals with SCI) from their relational and professional standpoints (Multimedia Appendices 3-5). Our questions did not delve into the requirements of relatives and HCPs to acquire, use, or prescribe robotic devices; instead, we focused on their perceptions of the needs of end-users (individuals with SCI). Throughout the sessions, the researchers maintained a field diary. The diary served as a repository for initial analytical reflections, interpretations, and first impressions.

Qualitative semi-structured interviews were conducted either over the phone or digitally by the first [SF] and second [JP] authors in Sweden and by the third [ED], fourth [MA], and fifth [FV] authors in Spain, each lasting approximately 45 to 75 minutes. In total, 18 individuals with SCI, 10 relatives, and 13 HCPs were interviewed.

In the participatory workshops [43], the initial part focused on the lived experiences of individuals with SCIs, identifying both opportunities and challenges (Multimedia Appendix 1). The initial part of the workshop also covered currently used assistive technologies and those left unused. The subsequent part of the workshops centered on SRLs. Here, we showed both illustrations of different scenarios in which SRLs were used and videos of users interacting with SRLs (Multimedia Appendix 1). The same visual material was used across workshops to support comparability between groups of stakeholders and countries. The participants were asked to openly discuss their perceptions of the different scenarios, including perceived barriers and enablers associated with using SRLs. The participatory workshops, with an average duration of 2 hours, included one researcher as the workshop moderator and another responsible for comprehensive notetaking [43]. Upon conclusion, all materials and participant notes were collected.

Two in-person workshops were held in Sweden at Lund University by the first [SF] and second [JP] authors, each lasting around 2 hours. In Spain, 3 in-person workshops were conducted by the third [ED], fourth [MA], and fifth [FV] authors at Sescam - Gobierno de Castilla-La Mancha, a specialized hospital and rehabilitation facility for patients with SCI. Workshops were conducted in small groups and included individuals with SCI and relatives and, in separate workshops, HCPs working in SCI rehabilitation. In total, 21 individuals with SCI, 2 relatives, and 13 HCPs took part in participatory workshops in Sweden and Spain (Table 1).

Table . Participants in the qualitative semistructured interviews and participatory workshops.

Category	Qualitative semistructured interviews	Male	Female	Age (years), mean	Participatory workshops	Male	Female	Age (years), mean
Participants with spinal cord injury	18	14	4	61	21	16	5	51
Relatives	10	3	7	N/A[a]	2	1	1	N/A
HCPs	13	1	12	N/A	13	2	11	N/A

[a]Not available.

Data Analysis

The interviews were transcribed verbatim and subjected to thematic analysis following Clarke and Brown [41]. Initially, the first and the second author individually reviewed the transcripts, immersing themselves in the data through multiple readings. During the initial stage, the researchers systematically assigned codes to expressions and narratives related to barriers and enablers to the intention to use assistive robotics technology. Subsequently, the first and the second author compared codes and resolved any disagreements. The codes were organized into potential key themes and subthemes [41].

Fieldnotes and material from the workshops were analyzed using the same coding framework as the interview transcripts [41]. The shared coding framework allowed us to triangulate across data sources and explore how the same issues were discussed in one-to-one interviews and in group settings (ie, during workshops). Where workshop data introduced new nuances, the coding framework was adjusted and reapplied to interview transcripts.

During the interpretive coding phase, barriers and enabler themes were developed inductively. Initial codes captured a broad range of practical hindrances, concerns, and hopes related to assistive robotics. The codes were gradually organized into

XSL•FO
RenderX

categories that reflected patterns across the data (Table 2 presents the final set of themes and subthemes).

For the purpose of verification, workshops dedicated to member checking were organized. Member checking is a qualitative research practice used to ensure data accuracy and credibility by actively involving participants in the data analysis process [47-49]. During the member checking workshops, research findings were presented to individuals with SCIs, their relatives, and HCPs for feedback. The member checking workshops facilitated the refinement of themes to ensure close alignment with participants' lived experiences of SCI. Quotations are used to illuminate the themes.

Table . Identified barriers and enablers.

Themes	Subthemes
Barriers	
Practical constraints	• Navigating a changing reality • Difficulties with awareness and access • Concerns about costs
Interactions with assistive robotics	• Doubts about meaningfulness • Uncertainty regarding reliability and safety • Uneasiness about competence • Apprehension of social norms
Enablers	
Sociotechnical imaginaries	• Imaginings of greater self-efficacy • Imaginings of fulfillment of functional requirements

Ethical Considerations

Ethical approval for this study was obtained from the Swedish Ethical Review Authority (Dnr 2022-03154-01) and Comité Ético CHT Eic-chto-secretario at Toledo Hospital in Spain (CEIC-961). The study was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent before their participation. Participants were informed in advance that they would not receive any monetary compensation for their time and inconvenience and offered no direct personal benefits. Confidentiality and anonymity of the data provided were maintained throughout the research and publication process.

Results

Sample Characteristics

In total, 41 individuals participated in interviews and 36 in participatory workshops (Table 1). Participants were recruited from 2 countries (Sweden and Spain) and 3 stakeholders' groups (individuals with SCI, relatives, and HCPs). Among participants with SCI, most were men in their fifties and sixties (mean age 61 y in interviews and 51 y in workshops), all with cervical injuries and varying functional levels (ie, most used wheelchairs and required different degrees of assistance with everyday activities such as eating and dressing). Relatives were mainly women, typically partners and close family members. The HCPs were predominantly women and represented a range of professions working with SCI rehabilitation (eg, physiotherapists, rehabilitation physicians, and biomedical engineers).

While the national health care systems differed, for example, in Sweden, participants relied to a greater extent on personal assistants who are paid by the municipalities [50], whereas in Spain, participants with SCIs often depended on their nonpaid relatives [51], similar themes emerged across the 2 country contexts. Such contextual differences, therefore, mainly function as a backdrop for the findings rather than as a primary focus of comparison.

Overview Themes

In our interpretive reading of the material, the data revealed that participants frequently described barriers impeding their intention to use assistive robotics, both generally and specifically in the context of SRLs. Through reflexive thematic analysis, we organized the data into 2 main barrier themes, "practical constraints" and "interactions with assistive robotics," and 1 main theme capturing enablers of intention to use "imaginings of self-efficay and functional fulfilment." Practical constraints include (1) navigating a changing reality, (2) difficulties with awareness and access, and (3) concerns about costs. Interactions with assistive robotics include (1) doubts about meaningfulness, (2) uncertainty regarding reliability and safety, (3) uneasiness about competence, and (4) apprehension of social norms. The main enabler theme encompassed two subthemes: (1) imaginings of greater self-efficacy and (2) imaginings of fulfillment of functional requirements. Key themes and subthemes are represented in Table 2.

Barrier – Practical Constraints

Overall, 3 subthemes related to barriers associated with practical constraints were identified: "navigating a changing reality," "difficulties with awareness and access," and "concerns about costs." The barrier themes were voiced across all stakeholder groups, although the specific experiences they referred to sometimes differed between individuals with SCI, relatives, and HCPs.

Navigating a Changing Reality

The subtheme, "navigating a changing reality," captures how participants described ongoing changes over time as a result of

XSL•FO
RenderX

the injury, aging, fluctuating symptoms, and evolving technologies as a barrier to engaging with assistive robotics. Most participants, including individuals with SCIs, relatives, and HCPs, described living with a SCI as a constant experience of transition periods. Transitions encompassed both the long-term adjustment to altered functionalities and needs and short-term fluctuations in pain and fatigue. Participants illustrated these experiences in expressions, such as

> *It was a brutal impact … from white to black. It shocked me in absolutely everything; it changed my life overnight. Everything has changed.* [SCI_6]

> *As a person with a SCI, you experience neurogenic pains that persist. Some days, it's really tough.* [SCI_15]

Those with SCIs also felt that immediately after the accident, they received considerable rehabilitation and assistance at the hospital, but upon discharge, they were largely on their own:

> *But now, after being discharged, it's difficult... no one is keeping track of me...* [SCI_4]

Such transition periods became a barrier to the intention to use assistive technology as needs, energy levels, and abilities fluctuated throughout the day and changed over time. Consequently, participants were less willing to invest time and energy in assistive technology due to the constant changes, along with fatigue and pain.

Within this subtheme, HCPs frequently cited a barrier to their intention to adopt assistive robotics. They pointed out that rapid technological developments impacted their work, for example, by requiring training in new devices, adapting clinical routines to new technology, and troubleshooting technology during already time-constrained consultations, which they perceived as increasing workload. The continuous changes and advancements placed a strain on them and their engagement with advanced technology. HCPs referred to transition periods in adapting to new technology and the continuous need to learn how to use it. As 1 HCP stated:

> *As a disadvantage, it might seem that we would have to learn and train ourselves on these new devices.* [HCP_18]

Overall, the experience of transitions and constant changes acted as barriers to the intention to use assistive robotics technology.

Difficulties With Awareness and Access

The subtheme, "difficulties with awareness and access," illustrates the barriers individuals with SCIs encounter when interacting with HCPs and others, both in understanding their injury's impact on their lives and in discovering available assistive aids. The subtheme was mainly articulated by individuals with SCI and their relatives, with HCPs offering a partly contrasting perspective on the availability of assistive robotics.

The obstacles outlined by individuals with SCIs included the lack of knowledge among municipality workers, reliance on specialized care professionals, and the challenge of self-identifying their needs:

> *Municipal workers lack knowledge about SCIs…we must have availability to direct contact with neurorehabilitation for assistive device support. We as patients typically need to identify our problems and propose potential aids, which may then be either adapted or developed for our individual needs. However, there is no predefined 'catalog' for selection.* [SCI_3]

Many participants emphasized the lack of support and assistance regarding assistive aids and knowledgeable personnel. They highlighted this by narrating

> *Above all, there is a lack of support on the issue of rehabilitation, there should be more means and more availability, both in terms of aids and personnel.* [SCI_5]

> *…I actually don't know who I would turn to if I wanted something [aids].* [SCI_12]

> *I don't know of any aid. As for support, there isn't much…* [SCI_8]

In contrast, HCPs perceived the availability of numerous assistive technologies. However, many also stressed that they did not recommend or prescribe them:

> *There are a lot of aids. For example, to eat and drink there are devices that I don't remember the name of because I have never recommended them.* [HCP_11]

In general, most participants with SCIs did not rely on advanced assistive technology but on other humans such as relatives and personal assistants. Their dependence on assistive technology mainly centered on wheelchairs and some specialized aids for the bathroom and kitchen. When it came to more advanced assistive technology and robotics, only a limited number of the participants (n=8) had experiences through participating in trials, testing at rehabilitation facilities, or using smart home technology with voice commands to control lighting, TV, and radio.

Concerns About Costs

Participants with SCI commonly faced financial pressure due to the inability to work or being limited to part-time employment. The subtheme, "concerns about costs," emphasizes a significant barrier to the intention to use assistive robotics technology tied to financial constraints. Concerns about the affordability of assistive technologies were raised by individuals with SCI, relatives, and HCPs.

In Sweden, known for subsidized health care [50], individuals with SCIs highlighted, like their Spanish counterparts, that only some assistive aids are subsidized [51]. Obtaining advanced assistive technology for free was perceived as challenging and almost impossible.

Some participants were aware of assistive robotics technology through the media but mentioned that trying it out required paying for it themselves. They perceived themselves as a disadvantaged group due to a lack of income and expressed the opinion that having money would make everything easier. For example, 1 participant stated,

XSL•FO
RenderX

> *Well, the truth is that I do not have much. Around here, we have a saying that "with money" everything is easier. When you have money, you can afford everything, but if you can't pay, you will have nothing.* [SCI_11]

Another participant expressed concerns, stating,

> *In terms of support and aids, I think much is missing. Initially, you receive rehabilitation and physiotherapy support, but there comes a time when you either pay out of pocket or receive nothing.* [SCI_18]

HCPs and relatives also shared concerns about costs, viewing them as a barrier to accessing more advanced assistive technology and human support. When discussing SRLs, participants perceived them as expensive and beyond the scope of ordinary prescribed assistive technologies. As 1 HCP noted,

> *The aids must be very affordable so that patients can easily access them.*
>
> *[HCP_25]*

The cost of assistive technology was consistently perceived as a major barrier affecting the intention to use assistive robotics technology among most of the participants.

Barrier – Interaction With Assistive Robotics

While practical constraints are linked to tangible barriers in the daily lives of individuals with SCIs, impacting their intention to use assistive technology in general, a prominent theme emerged focusing on the specific challenges of interacting with assistive robotics. The main theme comprises 4 subthemes: doubts about meaningfulness, uncertainty regarding reliability and safety, uneasiness about competence, and apprehension of social norms.

Doubts About Meaningfulness

The participants reported several doubts about the meaningfulness of assistive robotics technology. Doubts about the meaningfulness of assistive robotics were expressed primarily by individuals with SCI, drawing on their own previous experiences of technologies that had promised a lot but delivered little, and were echoed by some HCPs.

A few of the participants had prior experiences of taking part in user testing of robotic devices and some had experiences with advanced technologies such as speech synthesis and Alexa. They emphasized that these technologies had not been very meaningful to use in everyday life:

> *The hospital's robotic arm I tried was large, with numerous cords and a complex setup, making it impractical for daily use. To be suitable for everyday life, I envision a more discreet, easily deployable, and user-friendly design with fewer machinery and electronic components to make it useful.* [SCI_9]

Another participant said:

> *I've tried using speech synthesis, but for two reasons, I've given up on it. Firstly, sometimes it goes well for a couple of sentences and then it turns into complete gibberish. Secondly, when constructing a text, you often don't have all the answers in your head at the beginning. You start writing, go back, and make changes, rearrange and so on. This results in a significant amount of editing work. Reversing, erasing, and adjusting are not as straightforward with voice. Dynamically building a text doesn't always mean having it fully clear in your head from the start.* [SCI_1]

These quotes illustrate that the participants questioned the usefulness of advanced technologies based on their prior experience. There were also participants who doubted the usefulness of assistive robotic technologies based on what they learned from other people and through media. They highlighted that they often hear about life-changing technologies from others but doubt the truth in these stories:

> *The media offer people like us [individuals with spinal cord injury] a thousand stories, and then 95% of the things are worthless.* [SCI_3]

They did not believe what they heard due to the lack of evidence. Also, HCPs questioned the lack of evidence regarding the benefits and usefulness of assistive robotic technologies:

> *I don't think there is much scientific evidence in this regard…* [HCP_23]

Uncertainty Regarding Reliability and Safety

Uncertainty about reliability and safety was mentioned mainly by relatives and health care professionals, while participants with SCI themselves did not express these concerns as frequently. Relatives stressed that assistive aids need not be overly sophisticated; from their perspective, low-tech solutions appeared safer. While assistive robotics technology may offer extensive functionality, it also seemed to heighten the apprehension about its reliability and safety. One relative emphasized,

> *...before it [a robot] performs a thousand functions, it has to be able to do a few basic ones in a safe and reliable way...* [REL_9]

HCPs also underscored the importance of reliability and safety, envisioning potential mishaps that could jeopardize the well-being of individuals with SCIs. One HCP illustrated,

> *Imagine it's [a robotic arm] feeding you, it turns off and leaves you halfway with the fork in your mouth, or it's combing your hair and leaves you with the comb tucked into the hair.* [HCP_13]

The uncertainty surrounding reliability and safety emerged as a significant barrier to the intention to use for both relatives and HCPs.

Uneasiness About Competence

Uneasiness about competence referred both to individuals with SCI doubting their own ability to manage complex technologies and to relatives and HCPs questioning whether potential users would be able to operate assistive robotics independently. While the participants with SCI did not raise a lot of concern about reliability and safety, as opposed to relatives and HCPs, they were instead uneasy about their own competence rather than the technologies. Almost all the participants perceived robotics

XSL•FO
RenderX

as advanced and complex technology and worried that it was not for them due to their limited technological competence but also their dependence on others for most things in life. This was illustrated through their expressions, as demonstrated by

> *I am no longer autonomous at all. I went from being the one who solved problems for others to not being able to do anything for anyone; now, they have to do everything for me.* [SCI_1]

> *I think that I've never used anything besides the wheelchair, which is manual and the crutches.* [SCI_12]

> *I don't have any experience, I don't use technology.* [SCI_7]

Relatives and health care professionals were also uneasy about the competence of individuals with SCI in using assistive robotics. They drew on their own experience of helping them to handle everyday life activities as well as technology. As one of the relatives said:

> *He will need another person to place or start the robotic arm... after all that would not be a solution because the person who is there would end up assisting him more than the robotic arm.* [REL_2]

This quote illustrates how the subthemes of doubts about meaningfulness and uneasiness about competence are interrelated. If the intended user does not have the competence to use the technology by themselves, it decreases the meaningfulness of the technology.

Apprehension of Social Norms

The subtheme, "apprehension of social norms," highlights participants' awareness of societal expectations and how others perceive them. Apprehension of social norms was voiced predominantly by individuals with SCI, with relatives and HCPs recognizing and sometimes sharing these concerns. Specifically, when discussing robotic limbs, many participants expressed concerns about being viewed as abnormal or strange.

They were already accustomed to being seen as "different" with 1 participant noting,

> *… people may see you as if you were a robot or something like that…people already look at us with a strange face.*
>
> *[SCI_13]*

The fear was that adding a robotic finger or robotic arm would further set them apart from what is considered "normal."

HCPs and relatives shared similar reservations, expressing doubts about the societal acceptance of robotics and robotic limbs. One HCP remarked,

> *I see the exoskeleton, and they look like 'Robocop.' Besides, who would put this device to work for them every day?* [HCP_8]

The comment not only points to worries about social appearance but also doubts about whether the everyday effort required to set up and use assistive robotics would be perceived as worthwhile.

As the last quotation illustrates, participants reported that thoughts about robotics and robotic limbs often evoke images from movies depicting humans becoming part-robot. However, their concerns also extend to the practical aspects of using such devices, tying back to the subtheme of the perception of meaningfulness.

Enablers

Alongside barriers, our analysis of the data revealed enablers for the intention to use assistive robotics, such as the imaginings of greater self-efficacy and independence. We interpreted the enablers as "imaginings" because they appear to be rooted in hopes and wishes for a different future. The imaginings were not specifically related to assistive robotics but rather to assistive aids in general or life in general. While the barriers described hindered participants' intentions to use assistive robotics significantly, the participants also emphasized enablers or imagined solutions. The 2 main enabler themes: "imaginings of greater self-efficacy" and "imaginings of fulfillment of functional requirements" capture how participants envisioned assistive robotics as potentially transforming everyday life, even while simultaneously doubting whether such visions would materialize.

Imaginings of Greater Self-Efficacy

A significant number of participants expressed a desire for increased independence and the ability to resume their pre-SCI lifestyle. While many participants had come to terms with their injury, its emotional impact remained, as they found themselves unable to lead the life they once did. Participants expressed a keen appreciation for any opportunity for improvement and assistance:

> *All aids that could potentially enhance my function would be very welcome.* [SCI_5]

Participants mostly discussed aids in general that could enhance their independence, rather than making specific reference to assistive robotics or robotic limbs. They did not specify the type of aids but rather conveyed their desires to carry out daily activities independently:

> *I would like to be able to use it [a robotic limb] mainly for cleaning myself, eating, drinking, putting on shoes, dressing, that kind of thing that allows me to be more autonomous. I don't know the possibilities, the more functions, the better. Give me as much autonomy as possible.* [SCI_2]

The majority of the participants with SCIs expressed imaginings of having greater self-efficacy. Such imaginings could be understood as a desire to explore all possible technologies or aids for increased independence. However, the empirical data showed that they mostly used low-tech technology and perceived much technology as lacking meaningfulness to them (as illustrated above).

Imaginings of Fulfilment of Functional Requirements

When discussing factors influencing the intention to use, "imaginings of fulfilment of requirements" emerged as one of the most significant enablers. Aids or assistive technology need to be, according to the participants, "user-friendly" and easy to

XSL•FO
RenderX

learn. Participants emphasised various requirements, with statements such as

> *..it is crucial to consider affordability, esthetic appeal and ease of application and maneuvering.* [HCP_3]
>
> *Resistance to wear and tear, absence of daily problems, ease of repair and reliable technical support are essential.* [REL_4]

If the technology met all these requirements, then it could serve as an enabler for intention to use, according to the participants.

Discussion

Principal Findings

This study explored how individuals with SCI, their relatives, and HCPs in Sweden and Spain conceptualize assistive robotics in general and SRLs in particular, and which factors shape their intention to use assistive robotics.

Our findings complement previous research that has focused on specific robotic devices and their usability [52,53] and on general patterns of assistive technology adoption and abandonment [7,13,14,54-56]. Coherent with previous research, participants in our study were interested in emerging technologies that could support activities of daily living; nevertheless, they also highlighted usability, cost, and stigma as key challenges. Our contribution lies in showing how such issues play out for emerging assistive robotics such as SRLs and in demonstrating how expectations and doubts are distributed across multiple stakeholder groups.

Relating the findings to TAM and UTAUT

Viewing our findings through the UTAUT and TAM lens [10,11], our barriers and enablers can be mapped onto core constructs while also extending them. The theme “doubts about meaningfulness” corresponds closely with performance expectancy or perceived usefulness. Thus, the participants questioned whether assistive robotics would make everyday life easier or add complexity. Subthemes of uncertainty regarding reliability, safety, and uneasiness about competence relate to effort expectancy and facilitating conditions. Hence, relatives and HCPs worried about malfunctions and safety, while individuals with SCI doubted their ability to operate assistive robotics, which were perceived as complex technologies. Apprehension of social norms directly reflects the UTUAT construct on social influence, as individuals with SCI anticipated being perceived as “too robotic” or “strange” when using visible robotic limbs in public.

The practical constraint subthemes: “navigating a changing reality,” “difficulties with awareness and access,” and “concerns about costs” speak directly to facilitating conditions. Participants highlighted limited information pathways, lack of knowledge in municipal services and financial barriers, even in subsidized health care systems like the Swedish one [50]. The findings underscore that the intention to use assistive robots is not only based on the individual attitude but shaped by broader health care system infrastructures and funding arrangements.

The enabler subtheme, “imaginaries of greater self-efficacy” and “imaginings of fulfilment of functional requirements,” align with the behavioral intention construct in TAM and UTUAT [10,11]. Participants imagined assistive robotics that were reliable, aesthetically acceptable, and well supported, and that assistive robotics would allow them to reclaim valued activities such as self-care, eating, and dressing. However, many participants also doubted that currently available or near-future assistive robotics would fully meet these requirements, suggesting a gap between positive behavioral intention and realistic expectations about actual use.

Sociotechnical Imaginaries

Building on our reference to “sociotechnical imaginaries” in the results, we recognized the interplay between expectations and broader sociotechnical narratives, echoing discussions of the myth of technology and its advancement in the participants’ perceptions of assistive robotics. As Mayor describes in her book *“Gods and Robots: Myths, Machines and Ancient Dreams of Technology”* [57] new technologies often elicit both hope and apprehension [58,59] and can be viewed as modern myths [60,61]. The imaginaries or enablers we identified in our analysis presented a paradox. We term the tension the “meaningfulness paradox,” as participants expressed doubt about the meaningfulness of assistive robotics, their competency in handling the technology, its safety, and reliability. At the same time, they also spoke about how assistive technology could enhance self-efficacy if it met all functional requirements (which they doubted).

Thus, while TAM and UTUAT focus primarily on individual beliefs and intentions [10,11], drawing on the concepts of sociotechnical imaginaries [15-19] enables us to situate the participants' beliefs within broader collective visions of what assistive robotics could or should do. Participants' imaginings of independence, reduced caregiver burden, and improved quality of life echo existing imaginaries of robotics as a solution to workforce shortages and increasing care demands [62-64]. At the same time, the participants' narratives revealed a persistent gap between idealized expectations and lived realities, including pain, fatigue, cognitive load, and bureaucratic hurdles. For example, some participants contrasted media portrayals of robots with their own experiences of bulky robotic arms that were too demanding to set up and operate.

The tension was especially visible when participants contrasted the promise of SRLs with their experiences of earlier assistive robotic devices they had tried, which turned out to be cumbersome and unreliable. As such, sociotechnical imaginaries therefore function both as enablers for adoption, motivating interest in trying out emerging technologies, and as a potential source of disappointment if the assistive robotic device cannot live up to the expectations. Our findings suggest that managing the expectation gap is important in order to avoid future abandonment of assistive robotics, echoing previous calls for more realistic and relational approaches to social robots and care technologies [22,65].

Implications for Research and Practice

Our findings suggest considerations for the future development of assistive robotics. They indicate that the participants’ perceptions and conceptualizations of assistive robotics were

XSL•FO
RenderX

informed by ideas and sociotechnical imaginaries. The analysis underscores how idealized visions intersect with intention to use, where individuals may express a desire to use assistive robotics while simultaneously encountering barriers that complicate the translation of intent into action. Rather than framing identified barriers solely as obstacles to adoption, they can be understood as part of a broader negotiation between technical expectations and lived experiences [65]. This perspective shifts the focus from "non-compliant users" to assistive robotic systems and design that may not yet be well aligned with everyday realities of individuals with SCI and those who care for them.

For designers and engineers, our results highlight the importance of focusing on everyday meaningfulness rather than technical sophistication alone. Thus, prioritizing functions that users, relatives, and HCPs perceive as most valuable (such as self-care tasks), minimizing setup time and cognitive load. Furthermore, the subtheme "apprehension of social norms" indicates aesthetics and visibility matter. Hence, the participants did not want to look "different" or "too robotic." As such, the design of assistive robotic devices needs to be compatible with the user's self-perception, for example, by allowing robotic components to be concealed under clothing.

For clinicians and HCPs, our findings point to a need for clear pathways for identifying suitable candidates for assistive robotics, providing realistic information about benefits and limitations, as well as training and follow-up support. Structured introduction protocols, where assistive robotics are tried out in supervised sessions and then revisited after a trial period at home, may help individuals with SCI, relatives, and HCPs jointly assess whether the technology is useful or not. HCPs themselves also need opportunities and time to familiarize themselves with emerging technologies. Without having the time to familiarize themselves with emerging technologies, the rapid technological change of health care technologies may be perceived as additional workload instead of as support.

For policymakers and funders, ongoing concerns about costs and access underscore the need for developing reimbursement models and procurement strategies that make assistive robotics financially accessible, while also ensuring that funding covers training and maintenance. Models that support periodic reassessment may be needed to match the fluctuating needs of people with SCI. Involving people with SCI, their relatives, and HCPs in priority-setting and evaluation processes may help align investments with what is perceived as useful in everyday life.

As evidenced by numerous studies, the process of adopting new technology is far from linear [16,66,67]. Individuals' perceptions and intentions to either embrace or reject a technology are not static; rather, they are dynamic, continually shaped and negotiated as people interact with the technology [68]. Our results suggest that the ways individuals with SCI engage with assistive technologies are contingent on their lived experiences, evolving needs, and changing technological landscape. Overhyping the capabilities of assistive robotics risks fostering unrealistic expectations that could shape disengagement and disappointment [69]. Rather than focusing on promotional narratives, there is value in exploring how assistive robotics are experienced and imagined within participants' everyday lives. As researchers, we can contribute by facilitating discussions that reflect both possibilities and limitations, offering a grounded perspective on what assistive robotics can achieve and the type of support they can provide [65].

Limitations of the Study

A limitation of this study is that the inclusion and exclusion criteria did not take into account participants' time since injury, prior use of assistive technology, or whether they had completed comprehensive rehabilitation. The sampling strategy was a deliberate choice intended to invite a diverse range of perspectives and lived experiences, rather than predefining the sample based on rehabilitation history or technological exposure. However, we acknowledge that prior rehabilitation experiences and the use of assistive technologies may have shaped participants' attitudes toward adopting new assistive robotics solutions. Thus, the findings reported in the paper pertain to individuals with SCIs, constituting a highly heterogeneous group. Rather than viewing heterogeneity as a limitation alone, heterogeneity among the participants hopefully brought a more layered understanding of varying expectations, needs, and concerns surrounding assistive robotics. Nonetheless, the study has been strengthened by involving individuals from 2 different countries, as well as including relatives and HCPs. Therefore, the findings are likely to be beneficial for developers, practitioners, and researchers working on projects with similar technologies and characteristics for the target group.

Another limitation, but also an interpretive choice, was the use of the UTAUT as a lens when discussing the findings. The conceptual framings were used to help interpret and frame overarching patterns from the findings in the discussion. The data were first analyzed inductively, seeking all themes without being led by the need to fit participants' perceptions into a certain framework. As such, the interplay between emergent interpretation and structured theory did not impose a rigid framework but instead facilitated a dynamic engagement when discussing the findings, allowing both participant-driven insights and established conceptual framings to emerge. The interpretive approach was conducted with attention to maintaining the rigor of qualitative research. The recurring patterns observed in the data suggest not only coherence but also shared meanings among participants. Even so, using TAM and UTAUT may foreground certain aspects of technology acceptance while overlooking others. Future research could draw on alternative frameworks.

An additional limitation of the study was that the participants expressed gratitude for participating in the interviews and workshops, which may result in overly positive attitudes toward assistive robotics. However, as seen in the results, the data revealed multiple barriers to the intention to use, as interpreted in our analysis, which suggests that expressions of gratitude did not necessarily equate to uncritical acceptance of assistive robotics but coexisted with articulated concerns and constraints.

Conclusions

This study shows that intentions to use assistive robotics among individuals with SCI, their relatives, and HCPs are shaped by

XSL•FO
RenderX

a combination of barriers and enablers. The findings revealed that the barriers were clustered around practical constraints (such as navigating new realities, fluctuating health, limited awareness and access, and concerns about costs) and interaction with assistive robotics (including doubts about meaningfulness, uncertainty regarding reliability and safety, uneasiness about competence, and apprehension of social norms). Individuals with SCI described fluctuating pain, fatigue, and transition periods after discharge as making it difficult to invest energy in learning and using assistive robotics, while HCPs highlighted how continuous technological change increased their workload and required repeated training. Difficulties with awareness and access were mainly voiced by individuals with SCI and their relatives, who experienced fragmented information and limited support, whereas healthcare professionals perceived a wide range of available aids that were nevertheless seldom prescribed. Concerns about costs were shared by all 3 stakeholder groups and were seen as a major barrier to the intention to use assistive robotics.

The theme of "interaction with assistive robotics" revealed doubts about meaningfulness, uncertainty about reliability and safety, uneasiness about competence, and apprehension of social norms. Individuals with SCI tended to question their own ability to manage complex technologies such as assistive robotics in everyday life, while relatives and HCPs questioned whether assistive robotics would reduce dependence. Apprehension of social norms was voiced predominantly by individuals with SCI, who feared that visible robotic limbs would reinforce perceptions of them as "different." At the same time, participants' imaginings of greater self-efficacy and fulfillment of functional requirements, often expressed as hopes of regaining pre-injury abilities and reducing dependence on others, functioned as important enablers across groups. We conceptualized the pattern of ambivalence as a meaningfulness paradox, as participants expressed doubts about the meaningfulness of assistive robotics in everyday life, their own competence in handling assistive robotics, and their safety and reliability. Nonetheless, they also spoke about how assistive robotics could enhance self-efficacy if assistive robotics met all functional requirements, a circumstance they often doubted. The results suggest that assistive robotics should be understood in relation to the complexities of everyday life with SCI, caring relationships, and the imaginaries that shape how emerging technologies are perceived.

Acknowledgments

We thank all the participants.

Funding

This work was partially supported by the European Union (EU) project HARIA—Human–Robot Sensorimotor Augmentation—Wearable Sensorimotor Interfaces and Supernumerary Robotic Limbs for Humans with upper-limb disabilities (grant agreement No. 101070292). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the EU or CINEA. Neither the EU nor the granting authority can be held responsible for them.

Data Availability

Data are not publicly available due to ethical restrictions concerning participant privacy, in accordance with Swedish and Spanish regulations and ethics committee approvals but may be accessed by approved researchers after a data use agreement.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guidelines for in-depth workshop with individuals with a spinal cord injury.
[PDF File, 61 KB - rehab_v13i1e72080_app1.pdf]

Multimedia Appendix 2
Interview guide for patients.
[PDF File, 56 KB - rehab_v13i1e72080_app2.pdf]

Multimedia Appendix 3
Interview guide for relatives.
[PDF File, 57 KB - rehab_v13i1e72080_app3.pdf]

Multimedia Appendix 4
Interview guide for health care professionals.
[PDF File, 65 KB - rehab_v13i1e72080_app4.pdf]

XSL•FO
RenderX

Multimedia Appendix 5
Guidelines workshops with health care professionals.
[PDF File, 61 KB - rehab_v13i1e72080_app5.pdf]

References

1. Hauwe LVD, Sundgren PC, Flanders AE. Spinal trauma and spinal cord injury (SCI). In: Diseases of the Brain, Head and Neck, Spine 2020–2023: Diagnostic Imaging: Springer; 2020:231-240. [doi: 10.1007/978-3-030-38490-6_19] [Medline: 32119240]
2. Wagner A, Schweizer C, Ronca E, Gemperli A. The most important assistive devices for persons with spinal cord injury in Switzerland: a cross-sectional study. Disabilities 2023;3(3):367-378. [doi: 10.3390/disabilities3030024]
3. Kemp CC, Edsinger A, Clever HM, Matulevich B. The design of stretch: a compact, lightweight mobile manipulator for indoor human environments. Presented at: 2022 IEEE International Conference on Robotics and Automation (ICRA); May 23-27, 2022; Philadelphia, PA, USA p. 3150-3157. [doi: 10.1109/ICRA46639.2022.9811922]
4. Schultz JR, Slifkin AB, Yu H, Schearer EM. Proof-of-concept: a hands-free interface for robot-assisted self-feeding. Presented at: 2022 International Conference on Rehabilitation Robotics (ICORR); Jul 25-29, 2022; Rotterdam, Netherlands p. 1-6. [doi: 10.1109/ICORR55369.2022.9896535]
5. Cruz DM, Emmel MLG, Manzini MG, Braga Mendes PV. Assistive technology accessibility and abandonment: challenges for occupational therapists. Open J Occup Ther 2016;4(1):10. [doi: 10.15453/2168-6408.1166]
6. Lusardi R, Tomelleri S, Wherton J. Living with assistive robotics: exploring the everyday use of exoskeleton for persons with spinal cord injury. Front Med Technol 2021;3:747632. [doi: 10.3389/fmedt.2021.747632] [Medline: 35047959]
7. Sugawara AT, Ramos VD, Alfieri FM, Battistella LR. Abandonment of assistive products: assessing abandonment levels and factors that impact on it. Disabil Rehabil Assist Technol 2018 Oct;13(7):716-723. [doi: 10.1080/17483107.2018.1425748] [Medline: 29334475]
8. Jafar MR, Nagesh DS. Literature review on assistive devices available for quadriplegic people: Indian context. Disabil Rehabil Assist Technol 2023 Aug;18(6):929-941. [doi: 10.1080/17483107.2021.1938708] [Medline: 34176416]
9. Ventura S, Ottoboni G, Pappadà A, Tessari A. Acceptance of assistive technology by users with motor disabilities due to spinal cord or acquired brain injuries: a systematic review. J Clin Med 2023 Apr 19;12(8):2962. [doi: 10.3390/jcm12082962] [Medline: 37109297]
10. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
11. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
12. Pekkarinen S, Hennala L, Tuisku O, et al. Embedding care robots into society and practice: socio-technical considerations. Futures 2020 Sep;122:102593. [doi: 10.1016/j.futures.2020.102593]
13. Dos Santos ADP, Ferrari ALM, Medola FO, Sandnes FE. Aesthetics and the perceived stigma of assistive technology for visual impairment. Disabil Rehabil Assist Technol 2022 Feb;17(2):152-158. [doi: 10.1080/17483107.2020.1768308] [Medline: 32501732]
14. Barbareschi G, Carew MT, Johnson EA, Kopi N, Holloway C. “When They See a Wheelchair, They’ve Not Even Seen Me”-factors shaping the experience of disability stigma and discrimination in Kenya. Int J Environ Res Public Health 2021 Apr 17;18(8):4272. [doi: 10.3390/ijerph18084272] [Medline: 33920601]
15. Jasanoff S, Kim SH. Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power: University of Chicago Press; 2015. URL: https://doi.org/10.7208/chicago/9780226276663.001.0001 [accessed 2026-01-18]
16. Latour B. Reassembling the Social: An Introduction to Actor-Network-Theory: Oxford University Press; 2007. URL: https://doi.org/10.1093/oso/9780199256044.001.0001 [accessed 2026-01-18]
17. Bijker WE, Hughes TP, Pinch T. The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology: MIT press; 1987. URL: https://monoskop.org/images/1/1f/Bijker_Hughes_Pinch_eds_The_Social_Construction_of_Technological_Systems._New_Directions_in_the_Sociology_and_History_of_Technology_no_OCR.pdf [accessed 2026-01-13]
18. Akrich M. The de-scription of technical objects. In: Bijker WE, Law J, editors. Shaping Technology/Building Society: Studies in Sociotechnical Change: MIT press; 1992:205-224. [doi: 10.4000/tc.863]
19. Sand M, Schneider C. Visioneering socio-technical innovations — a missing piece of the puzzle. Nanoethics 2017 Apr;11(1):19-29. [doi: 10.1007/s11569-017-0293-6]
20. Mekki M, Delgado AD, Fry A, Putrino D, Huang V. Robotic rehabilitation and spinal cord injury: a narrative review. Neurotherapeutics 2018 Jul;15(3):604-617. [doi: 10.1007/s13311-018-0642-3] [Medline: 29987763]
21. Holanda LJ, Silva PMM, Amorim TC, Lacerda MO, Simão CR, Morya E. Robotic assisted gait as a tool for rehabilitation of individuals with spinal cord injury: a systematic review. J Neuroeng Rehabil 2017 Dec 4;14(1):126. [doi: 10.1186/s12984-017-0338-7] [Medline: 29202845]

XSL•FO
RenderX

22. Langer A, Feingold-Polak R, Mueller O, Kellmeyer P, Levy-Tzedek S. Trust in socially assistive robots: considerations for use in rehabilitation. Neurosci Biobehav Rev 2019 Sep;104:231-239. [doi: 10.1016/j.neubiorev.2019.07.014] [Medline: 31348963]
23. Trenaman LM, Miller WC, Escorpizo R, SCIRE Research Team. Interventions for improving employment outcomes among individuals with spinal cord injury: a systematic review. Spinal Cord 2014 Nov;52(11):788-794. [doi: 10.1038/sc.2014.149] [Medline: 25179659]
24. Lidal IB, Huynh TK, Biering-Sørensen F. Return to work following spinal cord injury: a review. Disabil Rehabil 2007 Sep 15;29(17):1341-1375. [doi: 10.1080/09638280701320839] [Medline: 17729082]
25. Baldassin V, Shimizu HE, Fachin-Martins E. Computer assistive technology and associations with quality of life for individuals with spinal cord injury: a systematic review. Qual Life Res 2018 Mar;27(3):597-607. [doi: 10.1007/s11136-018-1804-9] [Medline: 29417427]
26. Collinger JL, Boninger ML, Bruns TM, Curley K, Wang W, Weber DJ. Functional priorities, assistive technology, and brain-computer interfaces after spinal cord injury. J Rehabil Res Dev 2013;50(2):145-160. [doi: 10.1682/jrrd.2011.11.0213] [Medline: 23760996]
27. Scivoletto G, Galli G, Torre M, Molinari M, Pazzaglia M. The overlooked outcome measure for spinal cord injury: use of assistive devices. Front Neurol 2019;10:272. [doi: 10.3389/fneur.2019.00272] [Medline: 30967836]
28. Kabir KS, Alsaleem A, Wiese J. The Impact of Spinal Cord Injury on Participation in Human-Centered Research. 2021 Jun 28 Presented at: DIS '21; Jun 28 to Jul 2, 2021; Virtual Event USA p. 1902-1914. [doi: 10.1145/3461778.3462122]
29. Khalid S, Alnajjar F, Gochoo M, Renawi A, Shimoda S. Robotic assistive and rehabilitation devices leading to motor recovery in upper limb: a systematic review. Disabil Rehabil Assist Technol 2023 Jul;18(5):658-672. [doi: 10.1080/17483107.2021.1906960] [Medline: 33861684]
30. Proietti T, Ambrosini E, Pedrocchi A, Micera S. Wearable robotics for impaired upper-limb assistance and rehabilitation: state of the art and future perspectives. IEEE Access 2022;10:106117-106134. [doi: 10.1109/ACCESS.2022.3210514]
31. Guatibonza A, Solaque L, Velasco A, Peñuela L. Assistive robotics for upper limb physical rehabilitation: a systematic review and future prospects. Chin J Mech Eng 2024;37(1):69. [doi: 10.1186/s10033-024-01056-y]
32. Ho JSW, Ko KSY, Law SW, Man GCW. The effectiveness of robotic-assisted upper limb rehabilitation to improve upper limb function in patients with cervical spinal cord injuries: a systematic literature review. Front Neurol 2023;14:1126755. [doi: 10.3389/fneur.2023.1126755] [Medline: 37621855]
33. Martinez-Hernandez U, Metcalfe B, Assaf T, Jabban L, Male J, Zhang D. Wearable assistive robotics: a perspective on current challenges and future trends. Sensors (Basel) 2021 Oct 12;21(20):6751. [doi: 10.3390/s21206751] [Medline: 34695964]
34. Pinelli E, Zinno R, Barone G, Bragonzoni L. Barriers and facilitators to exoskeleton use in persons with spinal cord injury: a systematic review. Disabil Rehabil Assist Technol 2024 Aug;19(6):2355-2363. [doi: 10.1080/17483107.2023.2287153] [Medline: 38009458]
35. HARIA. HARIA Human-Robot Sensorimotor Augmentation. URL: https://clem.diism.unisi.it/~haria [accessed 2026-01-13]
36. Yang B, Huang J, Chen X, Xiong C, Hasegawa Y. Supernumerary robotic limbs: a review and future outlook. IEEE Trans Med Robot Bionics 2021;3(3):623-639. [doi: 10.1109/TMRB.2021.3086016]
37. Prattichizzo D, Pozzi M, Lisini Baldi T, et al. Human augmentation by wearable supernumerary robotic limbs: review and perspectives. Prog Biomed Eng (Bristol) 2021 Sep 17;3(4):042005. [doi: 10.1088/2516-1091/ac2294] [Medline: 41367033]
38. Pozzi M, Malvezzi M, Prattichizzo D, Salvietti G. Actuated palms for soft robotic hands: review and perspectives. IEEE/ASME Trans Mechatron 2023;29(2):902-912. [doi: 10.1109/TMECH.2023.3328944]
39. Bo V, Turco E, Pozzi M, Malvezzi M, Prattichizzo D. A data-driven topology optimization framework for designing robotic grippers. Presented at: 2023 IEEE International Conference on Soft Robotics (RoboSoft); Apr 3-7, 2023; Singapore p. 1-6. [doi: 10.1109/RoboSoft55895.2023.10122000]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Clarke V, Braun V. Thematic analysis. J Posit Psychol 2017 May 4;12(3):297-298. [doi: 10.1080/17439760.2016.1262613]
42. Adams WC. Conducting semi‐structured interviews. In: Handbook of Practical Program Evaluation: Wiley; 2015:492-505. [doi: 10.1002/9781119171386]
43. Ørngreen R, Levinsen KT. Workshops as a research methodology. Electron J E-Learn 2017;15(1):70-81 [FREE Full text]
44. Kallio H, Pietilä AM, Johnson M, Kangasniemi M. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. J Adv Nurs 2016 Dec;72(12):2954-2965. [doi: 10.1111/jan.13031] [Medline: 27221824]
45. Kvale S, Brinkmann S. Interviews: Learning the Craft of Qualitative Research Interviewing: sage; 2009. URL: https://libris.kb.se/bib/10957812 [accessed 2026-01-18]
46. Hussain I, Spagnoletti G, Salvietti G, Prattichizzo D. An EMG interface for the control of motion and compliance of a supernumerary robotic finger. Front Neurorobot 2016;10:18. [doi: 10.3389/fnbot.2016.00018] [Medline: 27891088]
47. Candela AG. Exploring the function of member checking. TQR 2019;24(3):619-628. [doi: 10.46743/2160-3715/2019.3726]

XSL•FO
RenderX

48. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member checking: a tool to enhance trustworthiness or merely a nod to validation? Qual Health Res 2016 Nov;26(13):1802-1811. [doi: 10.1177/1049732316654870] [Medline: 27340178]
49. Motulsky SL. Is member checking the gold standard of quality in qualitative research? Qualitative Psychology 2021;8(3):389-406. [doi: 10.1037/qup0000215]
50. Ludvigsson JF, Bergman D, Lundgren CI, et al. The healthcare system in Sweden. Eur J Epidemiol 2025 May;40(5):563-579. [doi: 10.1007/s10654-025-01226-9] [Medline: 40383868]
51. Bernal-Delgado E, Angulo-Pueyo E, Ridao-López M, et al. Spain: health system review. Health Syst Transit 2024 Sep;26(3):1-187. [Medline: 40094303]
52. Jardón A, Gil Á, de la Peña AI, Monje CA, Balaguer C. Usability assessment of ASIBOT: a portable robot to aid patients with spinal cord injury. Disabil Rehabil Assist Technol 2011;6(4):320-330. [doi: 10.3109/17483107.2010.528144] [Medline: 20969432]
53. Laubacher M, Perret C, Hunt KJ. Work-rate-guided exercise testing in patients with incomplete spinal cord injury using a robotics-assisted tilt-table. Disabil Rehabil Assist Technol 2015;10(5):433-438. [doi: 10.3109/17483107.2014.908246] [Medline: 24712412]
54. Almenara M, Cempini M, Gómez C, et al. Usability test of a hand exoskeleton for activities of daily living: an example of user-centered design. Disabil Rehabil Assist Technol 2017 Jan;12(1):84-96. [doi: 10.3109/17483107.2015.1079653] [Medline: 26376019]
55. Riemer-Reiss ML. Applying Rogers' Diffusion of Innovations Theory to assistive technology discontinuance. J Appl Rehabil Couns 1999 Dec 1;30(4):16-21. [doi: 10.1891/0047-2220.30.4.16]
56. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: a systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
57. Mayor A. Gods and Robots: Myths, Machines, and Ancient Dreams of Technology: Princeton University Press; 2019. URL: https://journeytothewestresearch.com/wp-content/uploads/2022/09/Gods-and-Robots-Myths-Machines-and-Ancient-Dreams-of-Technology-2018.pdf [accessed 2026-01-13]
58. Corn JJ. Imagining Tomorrow: History, Technology, and the American Future 1986. URL: https://libris.kb.se/bib/4738000 [accessed 2026-01-18]
59. Natale S. Introduction: new media and the imagination of the future. : WI; 2014 URL: https://iris.unito.it/bitstream/2318/1768946/1/PUBLISHED_Natale_introduction-new-media-and-the-imagination-of-the-future.pdf [accessed 2026-01-18]
60. Mosco V. The Digital Sublime: Myth, Power and Cyberspace (Cambridge, MA and London, UK): MIT Press; 2004. URL: https://doi.org/10.7551/mitpress/2433.001.0001 [accessed 2026-01-18]
61. Dourish P, Bell G. Divining a Digital Future: Mess and Mythology in Ubiquitous Computing: Mit Press; 2011. URL: https://www.mondotheque.be/wiki/images/9/92/Paul_Dourish_Genevieve_Bell_Divining_a_Digital.pdf [accessed 2026-01-13]
62. Breuer S, Braun M, Tigard D, Buyx A, Müller R. How engineers' imaginaries of healthcare shape design and user engagement: a case study of a robotics initiative for geriatric healthcare AI applications. ACM Trans Comput-Hum Interact 2023 Apr 30;30(2):1-33. [doi: 10.1145/3577010]
63. Vallès-Peris N, Domènech M. Roboticists' imaginaries of robots for care: the radical imaginary as a tool for an ethical discussion. Engineering Studies 2020 Sep 1;12(3):157-176. [doi: 10.1080/19378629.2020.1821695]
64. Krings BJ, Weinberger N. Assistant without master? Some conceptual implications of assistive robotics in health care. Technologies (Basel) 2018;6(1):13. [doi: 10.3390/technologies6010013]
65. Rosén J. What Did You Expect?: A Human-Centered Approach to Investigating and Reducing the Social Robot Expectation Gap: University of Skövde; 2024. [doi: 10.13140/RG.2.2.36413.44001]
66. Shove E, Pantzar M, Watson M. The Dynamics of Social Practice: Everyday Life and How It Changes: SAGE Publications Ltd; 2012. [doi: 10.4135/9781446250655]
67. Hirsch E, Silverstone R. Consuming Technologies: Media and Information in Domestic: Routledge (Taylor & Francis Group); 2003. [doi: 10.4324/9780203401491]
68. Borup M, Brown N, Konrad K, Van Lente H. The sociology of expectations in science and technology. Technol Anal Strateg Manag 2006 Jul;18(3-4):285-298. [doi: 10.1080/09537320600777002]
69. Geels FW, Schot J. Typology of sociotechnical transition pathways. Res Policy 2007 Apr;36(3):399-417. [doi: 10.1016/j.respol.2007.01.003]

Abbreviations

SCI: spinal cord injury
SRL: supernumerary robotic limb
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

XSL•FO
RenderX

Edited by A Scano; submitted 03.Feb.2025; peer-reviewed by J Wangdell, J Chu; revised version received 11.Dec.2025; accepted 02.Jan.2026; published 17.Feb.2026.

Please cite as:
Frennert S, Persson J, Díez-Rodríguez E, Alcobendas-Maestro M, Villamayor Vega F, Oliviero A
Exploring Barriers and Enablers for the Intention to Use Assistive Robotics Among People With Spinal Cord Injury and Those Involved in Their Care: Qualitative Study
JMIR Rehabil Assist Technol 2026;13:e72080
URL: https://rehab.jmir.org/2026/1/e72080
doi:10.2196/72080

© Susanne Frennert, Johanna Persson, Eva Díez-Rodríguez, Monica Alcobendas-Maestro, Fátima Villamayor Vega, Antonio Oliviero. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Blockchain-Based Mobile App for Digital Identification of Older Adults in Rural Peru: Design and Usability Evaluation Study

Wilver Arana-Ramos*, BSc; Aldo Francisco Pastrana-Leon*, BSc; Juan Carlos Morales-Arevalo*, MBA, MSc

Department of Software Engineering, Faculty of Engineering, Peruvian University of Applied Sciences, Av. Alonso de Molina Nro. 1611, Lima Polo And Hunt Club (Av. Primavera 2390), Lima, Peru

*all authors contributed equally

Corresponding Author:
Juan Carlos Morales-Arevalo, MBA, MSc
Department of Software Engineering, Faculty of Engineering, Peruvian University of Applied Sciences, Av. Alonso de Molina Nro. 1611, Lima Polo And Hunt Club (Av. Primavera 2390), Lima, Peru

Abstract

Background: Older adults in rural areas of Peru encounter many challenges in accessing critical public services, such as health care, education, and social assistance, due to low levels of digital literacy, limited access to technology, and the lack of formalized, secure ID. This inhibits entry into digital health, education, and social assistance systems and increases their risk of vulnerability and social exclusion.

Objective: This study aimed to design a blockchain technology–based mobile app architecture that helps facilitate secure and inclusive digital ID for older adults in rural areas of Peru, enabling access to vital services through a decentralized, privacy-preserving solution.

Methods: This study followed the design thinking process, which consists of five phases: empathize, define, ideate, prototype, and evaluate. A total of 16 adults (aged 61-85 years) were interviewed to determine the usability barriers and security and privacy concerns with mobile technology, which was used to define functional and nonfunctional requirements. These requirements were developed based on the interviews. The primary features the target population valued included blockchain authentication, assisted registration, multilingual functionality, and a user-friendly interface. The features were prioritized and prototyped using the Figma web-based app. The architecture of the app was developed using the C4 model and accounted for sequential development while ensuring scalability, modularity, and decentralization. Usability was assessed quantitatively by administering the System Usability Scale to the same 16 participants after they had interacted with the prototype.

Results: The mean System Usability Scale score was 60.78 (SD 13.68), indicating acceptable usability. The main issues identified were a lack of skills to navigate digital interfaces, concerns regarding data security, and accessibility challenges for people with disabilities. Participants provided high ratings for the assisted registration system and notifications. The modular, blockchain-based system architecture showed substantial potential for scalability and broader inclusion. The prioritization matrix identified that, for adoption, features must incorporate good design, be multilingual, and require secure authentication.

Conclusions: The proposed blockchain-based mobile app offers a viable technical and socially inclusive model for secure digital ID of older adults in underserved contexts. Usability testing suggested that the solution was perceived as secure, usable, and appropriate for the target population. Although not fully deployed, our prototypes and system architecture provide a good starting point for future implementation. The findings in this study can contribute to efforts to facilitate digital inclusion, access to services, and respect for people's autonomy in identity management systems for vulnerable people.

(JMIR Rehabil Assist Technol 2026;13:e79553) doi:10.2196/79553

KEYWORDS

blockchain technology; data privacy; digital divide; essential public service; portable digital identity; older people; user-centric design

Introduction

Digital technologies have become increasingly important in identity management across all sectors of the economy. The proliferation of mobile apps designed to improve access to services has further deepened this relevance [1]. Older adults living in rural areas of Peru face unique challenges, such as higher levels of digital illiteracy and limited public service coverage. These challenges hinder access to essential services such as health care, education, and social support and mirror broader barriers reported for digital health innovations in low- and middle-income health care systems in South and Southeast Asia [2-4]. Advances in digital identity models, such as self-sovereign identities (SSIs), have opened up new possibilities

XSL•FO
RenderX

for improving security, privacy, and access to services [5]. The emergence of identity verification systems that leverage blockchain infrastructure and protocols exemplifies this change, as they ensure there are no gaps in identity verification and authentication processes and allow users to exercise holistic control over their personal information [6].

Nonetheless, traditional identity management systems that typically use centralized data formats and 2-factor authentication often fail older adults in rural settings [7]. Centralized databases are limited, as they rely on dependable connectivity, technical skills, and trust in institutions to function as intended, even though there isn't adequate infrastructure in many rural contexts. In addition, centralized database models present vulnerabilities by creating a single point of failure that makes any district-based database susceptible to unauthorized access or misuse of information. Furthermore, the use of passwords, tokens, and SMS codes presents usability challenges, as older adults may be less versed in the use of mobile devices and may have difficulty remembering and entering credentials [8]. Conversely, blockchain protocols decentralize validation across identifiable participants and enable user control over personal data by creating a single entry point to the data. Collectively, these properties make blockchain protocols safer, more transparent, and more empowering for this population.

These advances in technology go well beyond data protection. They are fundamental to improving efficiency in the delivery of public and private services. For example, health credential verification and the interoperability of electronic health records are relevant in the health care context. The same is true in education and public participation [9], where the use of mobile apps that integrate SSIs could help reduce the digital divide and provide access to vital resources for all [2,3]. However, their implementation remains limited due to the fragmentation of existing platforms and the lack of a widely accepted interoperable standard. In addition, challenges remain related to centralized data sets, transparency in verification, adherence to strong data privacy policies, and the centralized control of information [10].

In this context, the focus of this study is on designing a blockchain-based mobile app architecture to support digital identity verification and access to necessary services. The goal is to create a system that can be implemented quickly and efficiently while respecting appropriate privacy, security, and decentralization practices in relation to structural data [11].

This study follows a user-centered design thinking (DT) approach to ensure that the proposed solutions are accessible, secure, and appropriate for the target population, rural older adults [12]. The research will include a comprehensive review of current solutions, such as blockchain and smart contracts applied to identity systems, in order to identify the most effective methods for improving user interaction and building trust in digital service platforms [13]. This research is qualitative and will combine a detailed literature review with the design and creation of system architecture models that will eventually be applied in the real world. The goal will be to establish a robust framework for mobile services that use advanced identity verification technologies, highlighting the potential of blockchain to ensure data permanence and integrity. The results of this study provide useful approaches for building mobile systems that prioritize accessibility, security, and transparency, with a special focus on improving essential services [13]. This strategy is even more significant for rural communities, as it allows older adults in these communities to access digital ID solutions that respect their privacy and give them control of their personal information.

Therefore, the objective of this study was to design and evaluate the usability of a blockchain-based mobile app architecture that supports secure digital ID and access to essential services for older adults living in rural areas of Peru.

Methods

Ethical Considerations

This research involved 16 older adults (aged 61-85 years) living in rural areas of Peru who voluntarily took part in interviews and usability tests. All participants signed an informed consent form. No sensitive data were collected, and personal information was anonymized and managed in accordance with the principles of confidentiality and data protection.

Formal ethics board (IRB) approval was not required according to institutional guidelines [14], as the study focused on the design and usability evaluation of a digital prototype. The research was non-interventional in nature, involved voluntary adult participants, and did not include clinical procedures or the collection of sensitive personal or health data. All participants provided informed consent prior to participation. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Research Methodology

This research adopted the DT methodology to develop a mobile app. The user-centered design approach is highly effective in digital product development because it supports users' needs at every stage while also ensuring a product is functional and accessible to the target audience [15-18]. Among the several established design frameworks, such as user-centered design, participatory design, and agile user experience, DT was selected because it provides a structured yet flexible process that begins with empathy as its core principle. This characteristic aligns with the study's focus on understanding the lived experiences of older adults in rural Peru, whose limited digital literacy and accessibility challenges require solutions grounded in human context rather than purely technical efficiency. By explicitly situating this work within a lineage of research that views empathy as the foundation for inquiry and intervention, the DT framework allows the research team to translate qualitative insights into practical, socially responsive design decisions [15-17].

The design process was broken into five phases: empathize, define, ideate, prototype, and test, and each phase was important for our design of a blockchain-based mobile product for digital ID and safe access to services. A complete overview of these activities is provided below.

XSL•FO
RenderX

Empathize Phase

A total of 16 older adults aged 61 to 85 years were interviewed to identify some of the key barriers when trying to access digital services or engage with new technologies. The interviews revealed a number of key issues directly related to participants' comfort levels and hesitancy with digital technology, and their anxiety surrounding cybersecurity, which often impaired their use of multiple digital platforms. These findings illustrate that solutions must be technically feasible, but also usable, secure, and easy to navigate [15].

Define Phase

The research team identified important findings from the interviews that influenced the project [19]. Based on the interview findings, the research team identified key user priorities, specifically the importance of a user-friendly interface, transparent privacy control, and user management (input). The study also highlighted the need for ID systems that are achievable and meaningful to users (at all digital literacy levels) [17].

Ideate Phase

At this stage of the research and development process, the team specified some core features of the app, such as authentication methods incorporating blockchain technologies, a security alerts feature, and personalization options for service access that the user sets according to their needs and preferences. The goal was to create a well-integrated solution in which users could access and manage strategic services relevant to them in a secure and efficient manner [18].

Prototype Phase

Prototypes and low- and medium-fidelity mockups were developed on the Figma platform. The prototypes helped to visualize the interface design of the app and observe users' interactions with the app's functions [19]. The technical system architecture was developed in accordance with the C4 model (this study included only "context," "containers," and "components," excluding "code," as we did not model the code level) to allow the system to be scalable and adaptable within a blockchain structure [20-22].

Testing Phase

To quantitatively evaluate prototype app usability, the study applied the System Usability Scale (SUS) developed by Brooke [23] in 1996. The SUS is a well-respected instrument used in usability research, due in part to its simplicity and its provision of a straightforward numerical metric of users' perceptions of system usability [23].

A survey was created using the standard 10-item SUS, in which respondents indicated whether they agreed or disagreed with statements using a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree"). The full set of questions used is in Textbox 1.

The launch survey was conducted on 16 older adults that were representative of the targeted users, all of whom had previous interactions with the prototype. The sample size matched the guidelines recommended for initial usability testing using the SUS. Scoring was based on subtracting 1 from the participants' response for the odd-numbered items and subtracting the even-numbered item responses from 5. The scores were summed and multiplied by 2.5 to create a final score with a range of 0 to 100, with higher scores representing better usability. User feedback was an important part of the process for improving interface accessibility and ease of use [24]. This cyclical, user-centered approach ensures that the app meets technical specifications while being accessible, safe, and easy to use for older adults with various levels of technological capabilities [16,17].

Textbox 1. Questions on the System Usability Scale.

1. I would like to use this application frequently.
2. I found the application unnecessarily complex.
3. I found the application easy to use.
4. I would need the help of a person with technical knowledge to be able to use this application.
5. The various functions of the application are well integrated.
6. I found too much inconsistency in the application.
7. I imagine most people could learn to use this application quickly.
8. I found the application cumbersome to use.
9. I felt very confident using this application.
10. I learned to use the application quickly.

Results

This section elaborates on the results from the various design phases and the process of developing a mobile app designed for digital identity and secure access to vital services based on blockchain technology, which was achieved using the DT approach. This allowed the team to better understand the needs of the targeted end users, older adults, and ensured that the prototype genuinely reflects the expectations and needs of this group. The major findings and responses presented in each phase are described below.

Empathize Phase: Interview Findings

During the empathize phase, interviews were conducted with 16 older adults, ranging in age from 61 to 85 years, to identify

XSL•FO
RenderX

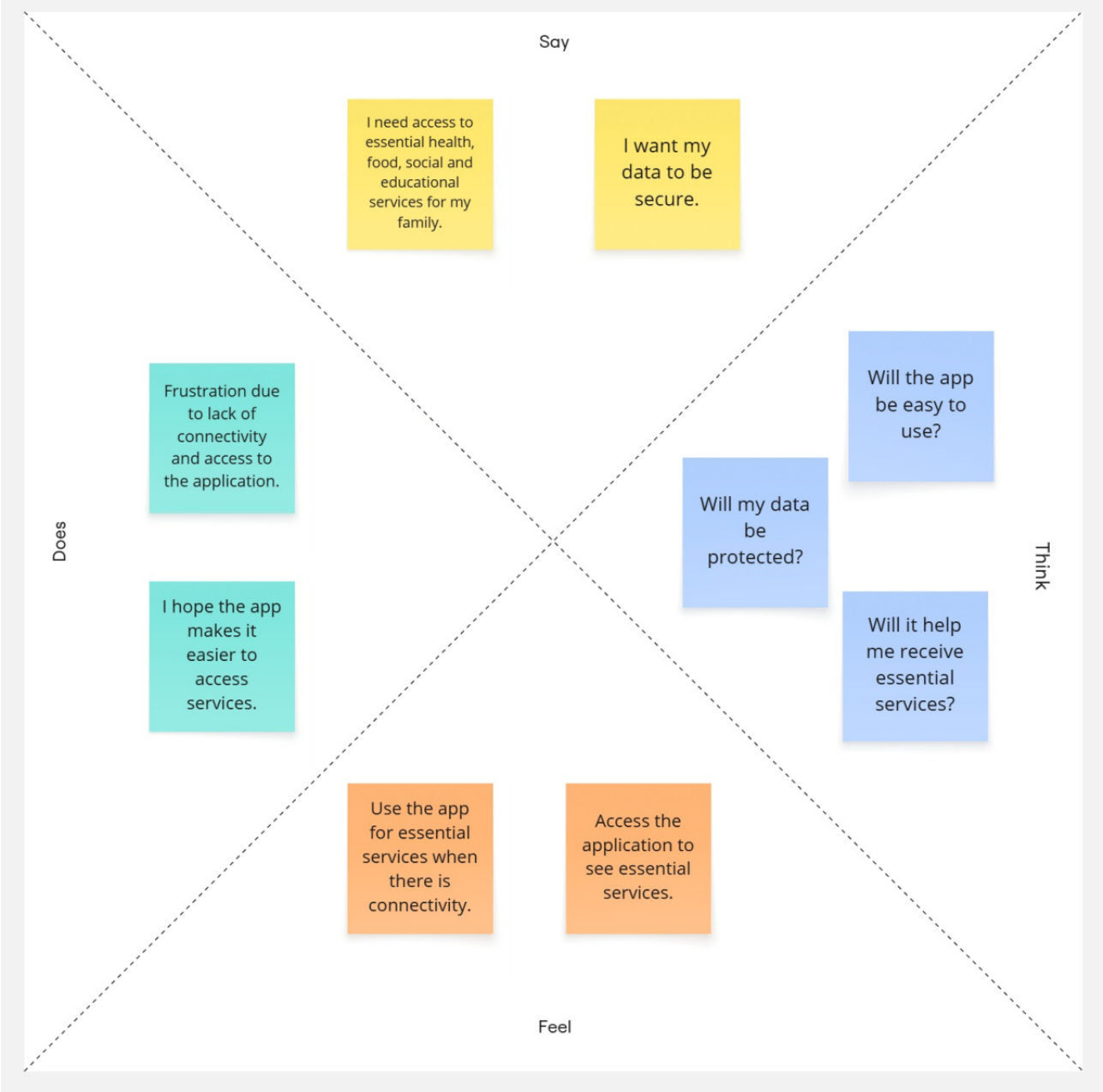
the barriers and difficulties they face when interacting with mobile apps. The most common problems encountered are presented in Textbox 2.

Based on the findings obtained, two empathy maps were developed to reflect the needs and emotions of users in relation to technology [25,26]. These are presented in Figures 1 and 2, where the first corresponds to users and the second to the local authorities.

Textbox 2. Identified problems.

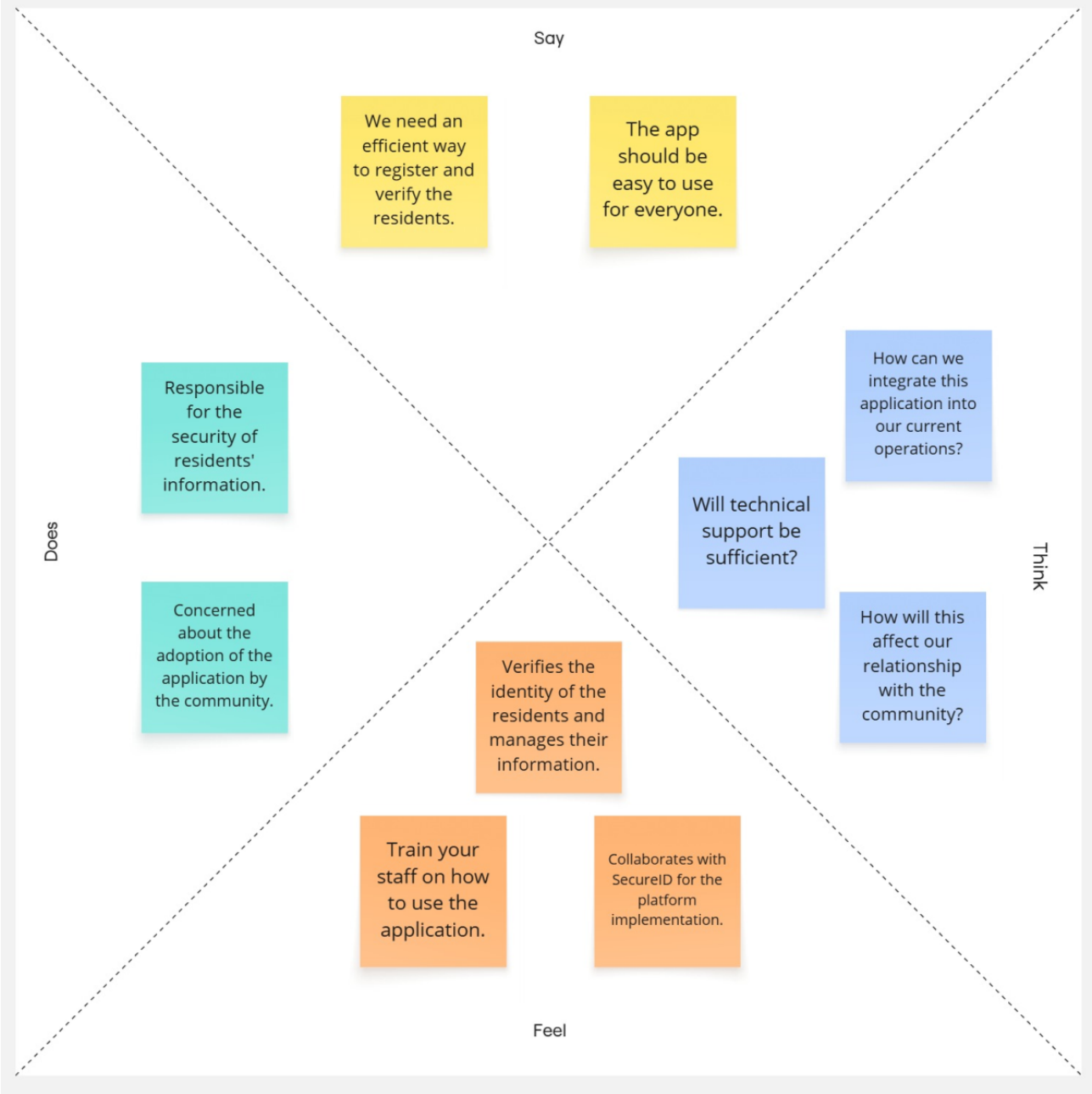
- Difficulty navigating apps: users indicated that the app interfaces were complex and difficult to understand.
- Data security concerns: older adults expressed distrust about the handling of their personal data on digital platforms.
- Lack of interface customization: some users reported that the apps were not tailored to their individual needs, making adoption difficult.
- Lack of digital literacy: many older adults have difficulty understanding apps due to their unfamiliarity with digital technologies.
- Limited accessibility: the lack of accessibility options for people with visual or motor disabilities made interaction complicated for a segment of the population.

Figure 1. Empathy map: user perspective.



XSL•FO
RenderX

Figure 2. Empathy map: local authorities' perspective.



Define Phase: Functional and Nonfunctional Requirements

Based on the findings obtained in the empathize phase, the functional and nonfunctional requirements for the mobile app were defined. These were established based on the specific needs of the users. The most relevant requirements are presented in Table 1 below.

The following are the most relevant user statements, selected from those collected for their significant contribution to understanding users' expectations regarding interaction with the app and ensuring that their needs are adequately addressed. These user stories are presented in Table 2.

XSL•FO
RenderX

Table . Functional and nonfunctional requirements.

Requirement	Description	Type
Registration in the app	Allows the user to register in the mobile app with the support of a caregiver, enabling access to their digital identification and essential services.	Functional
Digital identification display	Allows the user to securely access their digital identification based on blockchain technology, guaranteeing the authenticity and protection of their personal data.	Functional
Selection of essential health service	Provides the user with the ability to view and select health services available in their locality, facilitating timely access to vital information.	Functional
Information security	The app must ensure the integrity, confidentiality, and authenticity of user data through the use of blockchain technology and robust security protocols.	Nonfunctional
Multilingual support	The app must support multiple languages to facilitate access to services for users who do not speak the platform's main language.	Nonfunctional
Notification system	The system must have a notification controller that manages reminders, alerts, and updates, sending them to users and local authorities.	Functional

Table . Relevant user stories.

User ID	User statement
US-01	As an older adult user, I would like to view my digital identification in order to access my personal data securely.
US-02	As a local authority, I want to filter the list of registered seniors by name or digital identification to quickly locate a specific user and facilitate records management.
US-03	As an older adult user, I would like to register in the application in order to access the available services.
US-04	As an older adult user, I wish to attach and upload images such as a signature, face, or identification to be registered in my account.
US-05	As a local authority, I wish to modify the distribution of essential services.
US-06	As an older adult user, I want to visualize the services provided by the application to easily access the available options.
US-07	As a local authority, I want to view the list of registered seniors to properly track their enrollment and access to available services.
US-08	As a local authority, I want to visualize the services provided by the application to easily access the available options and also have the possibility to add new broadcasts to each selected service.

Ideate Phase: Key Functionalities and Prioritization Matrix

In this phase, the key functionalities needed for the mobile app were identified. The functionalities were prioritized according to their impact on user experience and the effort required for their implementation. These results are presented in Table 3.

Table 4 shows the prioritization matrix, which ranks the key functionalities according to their impact and implementation effort.

XSL•FO
RenderX

Table . Selected key functionalities.

Key functionality	Description
Blockchain authentication	A decentralized authentication system to ensure the security and privacy of user data.
Intuitive interface	Simple and accessible design, with customization options to enhance the experience for older adults.
Assisted registration system	Provides personalized assistance during the registration process to ensure the inclusion of older adults who are not fully technologically autonomous.
Multilingual support	The app must support multiple languages to facilitate access to services for users who do not speak the platform's main language.
Notification system	The system should send notifications to remind users about important events or updates related to services, such as health care, education, and social assistance.

Table . Functionality prioritization matrix.

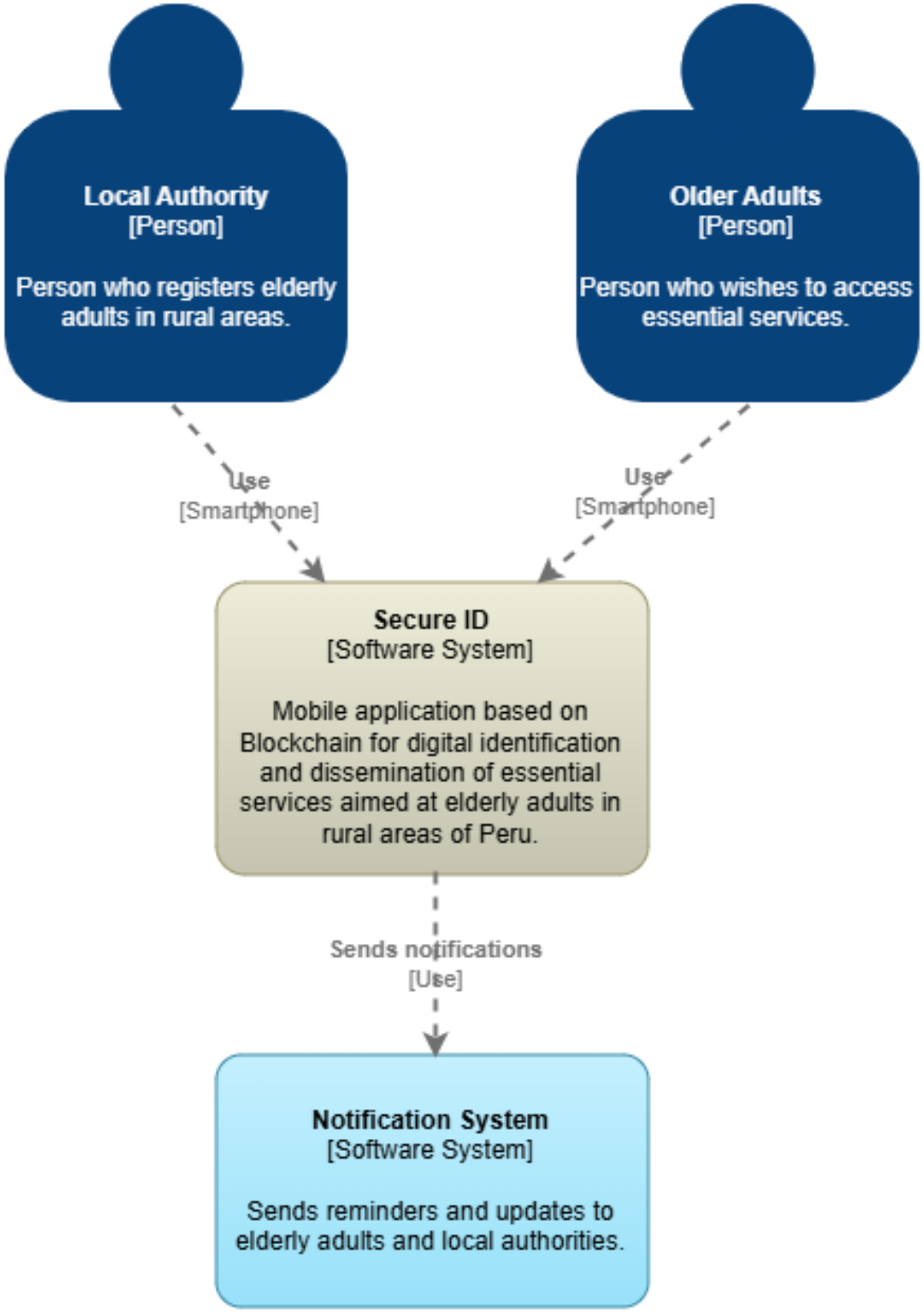
Functionality	Impact	Effort	Priority
Blockchain authentication	High	Medium	High
Intuitive Interface	High	High	Medium
Assisted registration system	High	Medium	High
Multilingual support	High	Medium	High
Notification system	High	Medium	High

Prototyping Phase: C4 Model, User Flow, and Mockups

During the prototyping phase, a comprehensive representation of the system was developed, covering both the technical architecture and the user experience. For this purpose, the C4 model was applied, which allows for structuring the system architecture at different levels of abstraction. Figure 3 shows the context diagram, which illustrates the general interaction between the mobile app, the users (older adults and local authorities), and the external identity verification services. Figure 4 presents the container diagram, which decomposes the system into its main modules: mobile interface, cloud backend, and decentralized database, along with the technologies used. Multimedia Appendix 1 presents the component diagram, which describes the internal elements of the back end, such as the authentication manager, the service controller, and the notification module.

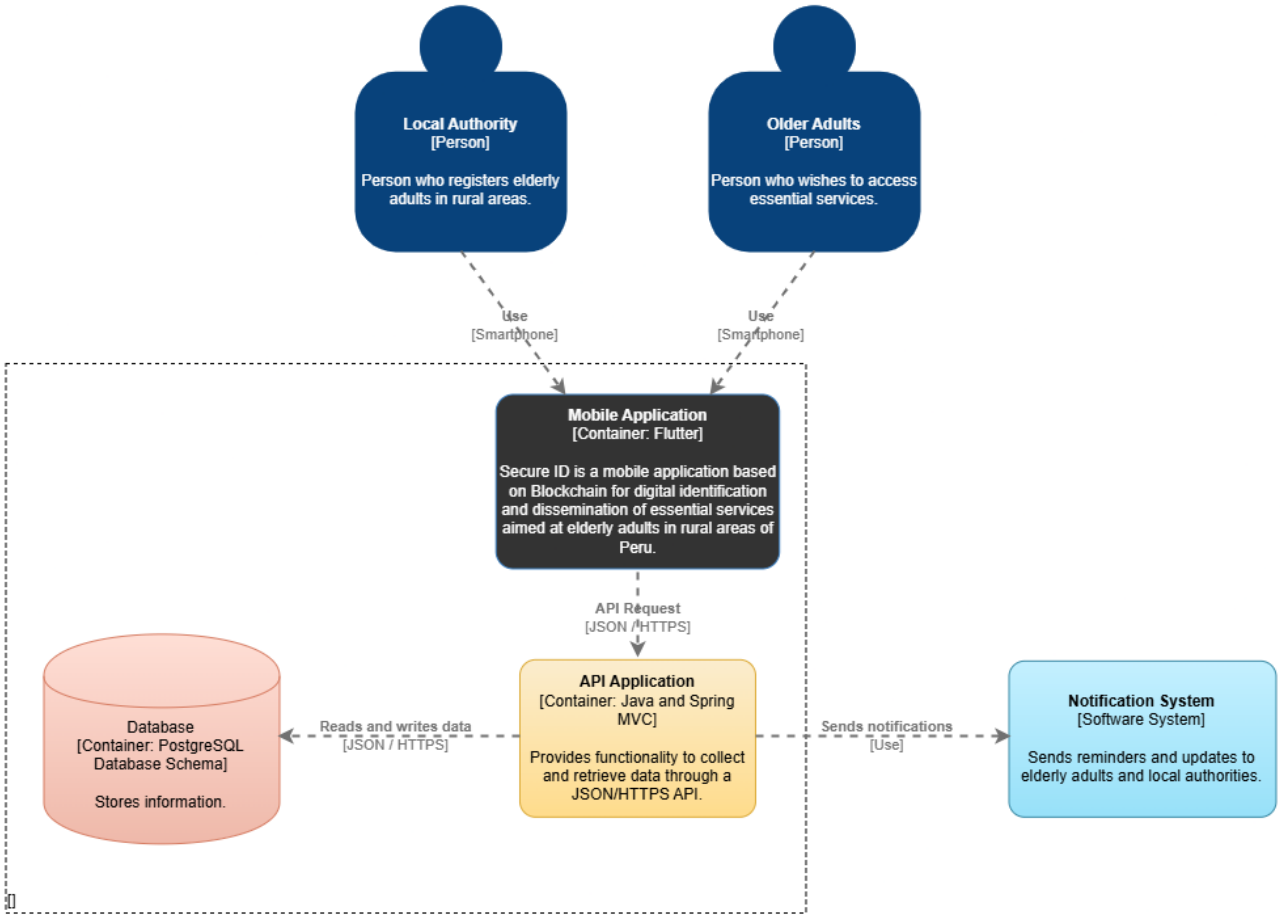
XSL•FO
RenderX

Figure 3. C4 model context diagram.



XSL•FO
RenderX

Figure 4. C4 model container diagram.



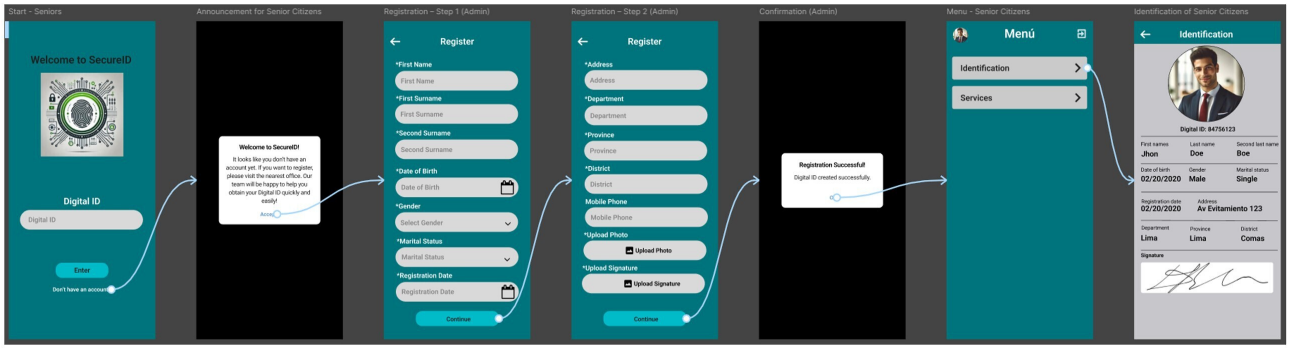
To validate the user experience, navigation flows were developed that represent common user steps from login to service access. Figure 5 shows one of these flows, focusing on digital ID.

Finally, low and medium fidelity mockups were designed in Figma. Figure 6 illustrates the most representative screens of the prototype: user registration, ID display, and service dissemination. These prototypes were essential to validate the suitability of the design to the capabilities and preferences of the target audience.

Figure 5 illustrates the proposed navigation flow for the digital ID process of older adults. This sequential design optimizes the user experience through a clear interface and guided steps from the initial registration to the display of the digital ID, ensuring accessibility, usability, and efficiency in identity management.

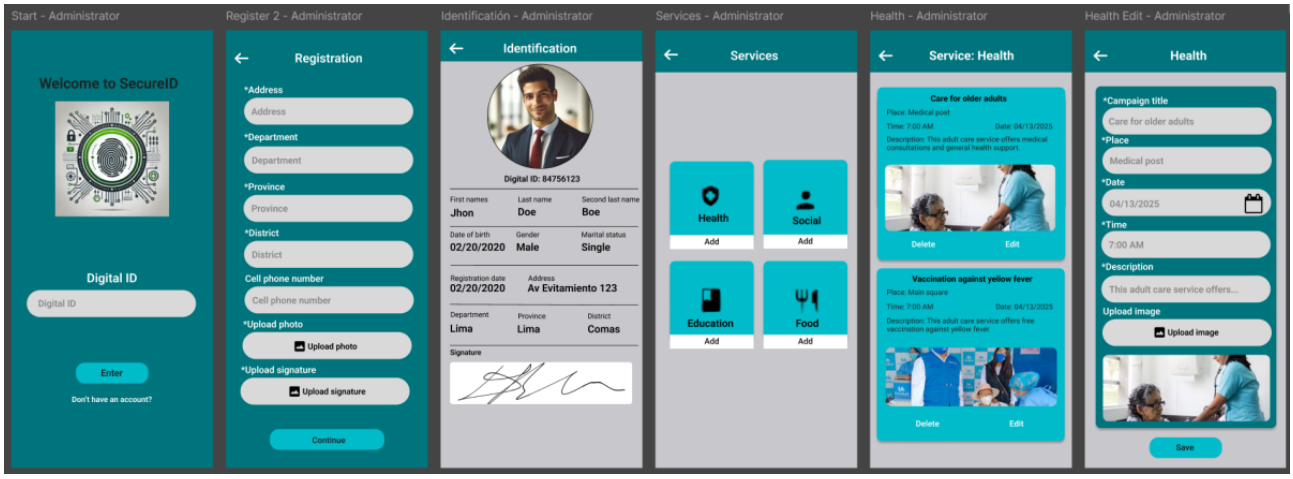
Figure 6 presents a series of mockups that illustrate key system functionalities from the administrator's perspective. These include user registration, display of ID information, and management of essential services such as health care, social support, education, and food. The design promotes efficient administration and clear dissemination of campaigns targeting vulnerable populations.

Figure 5. Ideal navigation flow for identification.



XSL•FO
RenderX

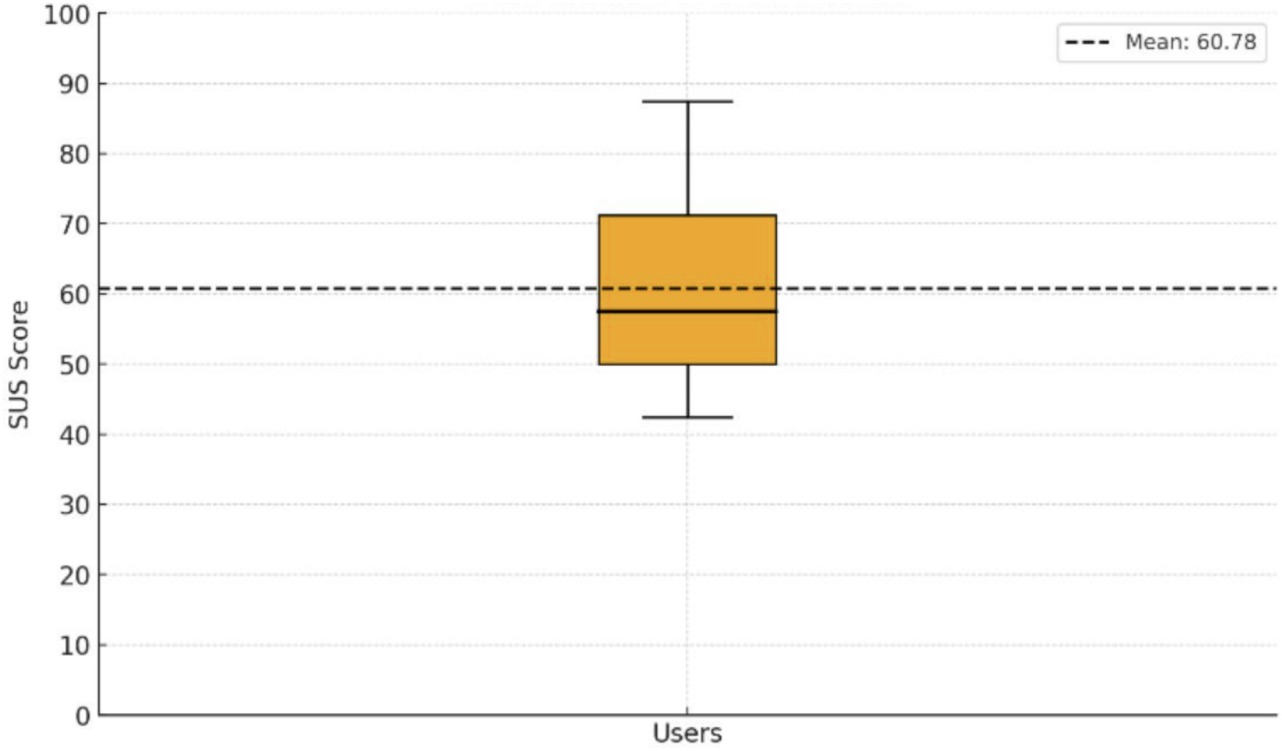
Figure 6. Mockups: registration, identification, and dissemination of essential services.



Testing Phase: Usability Evaluation

In this phase, the results obtained by applying the SUS to a sample of 16 older adults, representative users of the target population who interacted with the mobile app prototype, were analyzed. The mean SUS score was 60.78 (SD 13.68) on a scale from 0 to 100, where higher values indicate better perceived usability, as can be seen in Figure 7. This score is interpreted as an acceptable usability according to the standards established in the literature, which indicates that the designed prototype is perceived as easy to use, safe, and suitable for the target population.

Figure 7. Distribution of System Usability Scale (SUS) scores using a boxplot.



Discussion

Outline

In this study, we designed and evaluated the usability of a blockchain-based mobile app architecture intended to support secure digital ID and access to essential services for older adults in rural Peru. Usability testing with 16 older adults showed that the mean SUS score was 60.78 (SD 13.68), indicating acceptable usability. Older adults viewed the app as secure and mostly easy to use and specifically highlighted the intuitive interface, assisted registration process, and personalized notifications as positive features [25,27], although there were ongoing issues related to digital literacy, trust in data security, and accessibility for users with disabilities.

In terms of accessibility, the design targeted low digital literacy and disability-related barriers through large high-contrast fonts, simplified menus, intuitive iconography, and an assisted registration flow [19]. In our sample, most participants self-reported no visual or motor limitations; those who did generally required more time and occasional assistance during

XSL•FO
RenderX

formative sessions, which directly informed these adaptations [8]. The confirmatory evaluation will quantify effectiveness using prespecified metrics, including completion without assistance across core tasks, median task time, error rate, and SUS, stratified by limitation status.

The prototypes and C4 models developed demonstrate a clear and scalable architecture, since they are structured in modular layers that allow incorporating new functionalities without affecting the stability of the system [22]. This scalability is reflected in the use of the C4 model, which facilitates the expansion of the system both in terms of components and cloud services, adapting to different regions or user groups [28]. The solution presented is accessible and reliable, as evidenced by the results obtained in the usability tests with older adults [29].

Unlike the conceptual framework proposed by Tan et al [30], which focuses on a governance taxonomy for blockchain-based systems in the public sector at macro, meso, and micro levels, our proposal is based on a practical and applied approach, which directly implements such principles in a functional architecture with tested prototypes. This framework includes concrete technical decisions on decentralized authentication, user privacy, and accessibility. The designed app facilitates secure and efficient access to essential services, promoting digital inclusion and trust in the handling of personal information [11]. As a limitation, while the study initially focused on design and prototyping, the app is currently being evaluated in real-world settings with older adults from the target population to validate its effectiveness, usability, and adoption in practical environments.

Related Work

In recent years, academic studies have begun to examine how blockchain technology can reinforce governance structures, support the decentralization of public service delivery, enhance security for mobile apps, and improve the user experience in distributed digital environments. Below are four thematic categories that frame the most relevant work.

Application of Blockchain in Public Sector Governance Models

Tan et al [30] address the implementation of blockchain's potential to transform public services by improving transparency, efficiency, and security, key aspects for the design of blockchain-based mobile apps. Through a conceptual framework, the authors explore key governance decisions at three levels: micro, meso, and macro. These governance elements, illustrated in Figure 8, affect the design and implementation of blockchain-based systems in the public sector [31]. However, they identify limitations such as the lack of interoperable infrastructure and the need for effective governance models. This study contributes to these limitations by proposing a solution based on a more flexible approach to digital identity management and access to essential services, especially through mobile apps that enable better interaction with public services [30].

XSL•FO
RenderX

Figure 8. Key elements in the economic design of governance.



Specifically, the proposed architecture directly operationalizes Tan et al's [30] microlevel governance principle through user-centered control of digital credentials. For instance, the assisted registration system and blockchain-based authentication module enable users to manage their identity data without relying on centralized authorities. Likewise, the notification and verification features provide transparency and consent management, ensuring that users are aware of and approve any data exchanges. At the mesolevel, interoperability is supported by modular components that facilitate integration with local authority systems, while the macrolevel implications relate to potential scalability within national digital identity strategies. In this way, our design translates Tan et al's [30] theoretical governance taxonomy into practical app features that promote individual autonomy and trust in digital identity management.

For his part, Ibrahim [32] explores the impact of decentralization on improving public services, highlighting the need to optimize government efficiency and accountability by delegating authority to local governments. This approach, which highlights the importance of tailoring services to local needs, is essential when considering blockchain-based mobile apps for public services. However, Ibrahim notes that disparities in local resources and capabilities limit the effectiveness of decentralization. Our research complements this analysis by integrating blockchain-based technologies, providing a solution that improves security, accessibility, and efficiency in digital ID and access to essential services, overcoming local resource barriers through a decentralized and accessible infrastructure [32].

Innovations in Blockchain-Based Mobile Apps for Security and Education

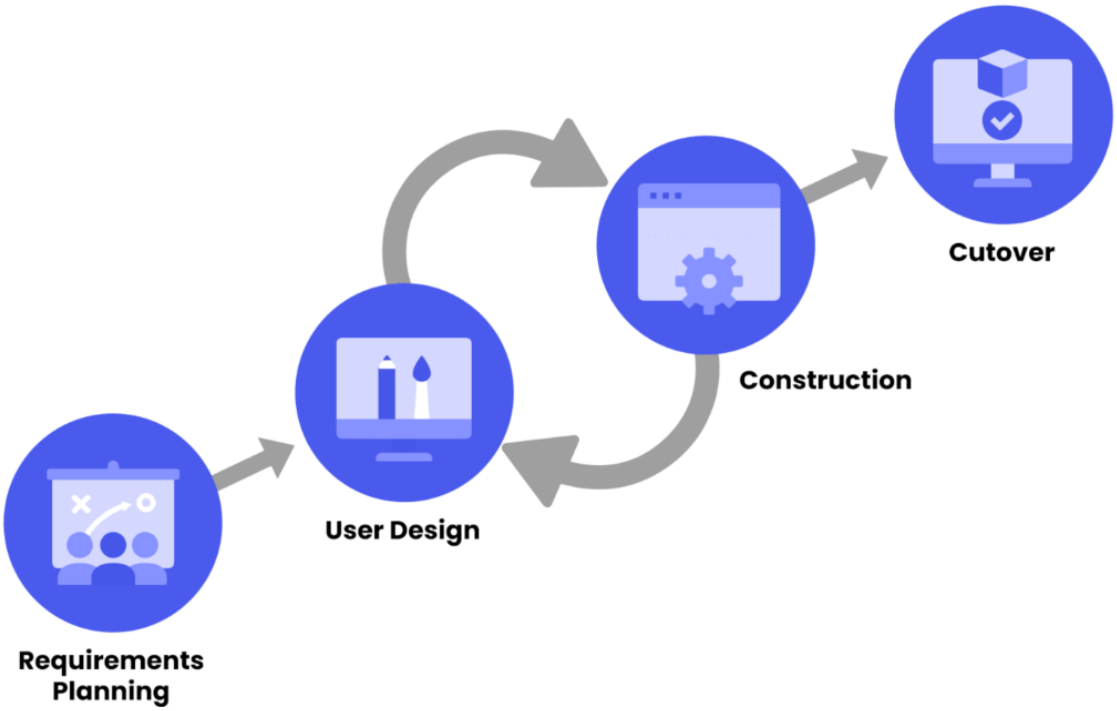
Rizky et al [33] propose a blockchain-based mobile app for decentralization of information management in the field of e-journals, addressing security and reliability issues in

XSL•FO
RenderX

centralized systems. Although their focus is on data security in academic platforms, their findings on decentralization and security are highly relevant to the design of mobile apps that manage digital ID in essential public services. Using a SWOT (strengths, weaknesses, opportunities, and threats) analysis and a waterfall development approach, the study implements blockchain to ensure a network resilient to external interference. However, scalability in larger environments remains a challenge. Our research advances this by proposing technical solutions that improve the efficiency and accessibility of essential services, overcoming these limitations by integrating blockchain for secure and decentralized digital ID.

Similarly, Asmawi et al [34] developed BlockScholar, a blockchain-based mobile educational app to facilitate the understanding of blockchain. Their research addresses the gap in interactive and accessible educational resources, highlighting the need for platforms that offer immersive and accessible learning on blockchain. They used the rapid application development model to create interactive and gamified content, as illustrated in Figure 9. Although the study made progress in accessibility and comprehension, it still faces challenges in adaptability to diverse educational contexts. Therefore, our research proposes solutions that enhance learning personalization and expand the coverage of blockchain-related topics.

Figure 9. Rapid application development model phases.



Security, Identity, and Digital Inclusion with Blockchain

Gumilar et al [35] explore the integration of financial technology in digital inclusion, with an emphasis on digital financial literacy as a catalyst for reducing economic and social disparities [35]. Through a systematic literature review, they used Scopus databases (2020-2024) to examine how the adoption of digital financial services optimizes inclusion. The results highlighted the importance of improving digital financial literacy but noted limitations in consumer use and protection. In line with these findings, our research extends the analysis by proposing innovative technological solutions and a more accessible design to improve adoption in underserved populations.

On the other hand, Musa et al [36] propose a blockchain-based approach to improve the security of data storage in Android mobile apps. They address the problem of vulnerability and unauthorized access to sensitive data by using blockchain to provide decentralized and secure storage. Their methodology includes the implementation of a 6-layer framework called BSADS (blockchain-based secure android data storage), which optimizes efficiency and security, as detailed in their proposed framework. Despite advances, scalability and costs remain limitations. This research contributes to overcoming these challenges by proposing lightweight node solutions and cost optimization techniques.

Identity and User Experience Management on Blockchain

Alanzi and Alkhatib's [37] study presents blockchain-based identity management solutions aimed at improving privacy and security in traditional centralized systems. Using blockchain technologies such as Ethereum and smart contracts, they address issues of third-party control, single point of failure, and vulnerability to data manipulation. However, limitations in scalability and weak authentication of some proposed systems are identified. In this context, our research extends these approaches, proposing improved decentralized infrastructure and optimizing the design of smart contracts to increase security and efficiency in digital identity management systems.

On the other hand, Jang and Han [38] developed a user experience framework for blockchain-based services, specifically focused on improving user interaction in contexts

XSL•FO
RenderX

such as finance and health care [38]. Through an analysis of active services, they identified both general and blockchain-specific functions, highlighting improved efficiency and security. However, their study faces limitations in general applicability due to a lack of standardization. Consequently, our research extends this framework, addressing the implementation of blockchain in digital identity and essential services, improving accessibility and user experience through advanced technological solutions [38]. Table 5 summarizes the key approaches and inputs underlying the development of a blockchain-based mobile app for digital ID and access to essential services.

Table . Fundamental approaches for blockchain-based mobile apps.

Approach	Contribution	Referenced works
Taxonomy of governance	• Rationale for the governance approach	• Tan et al [30]
Decentralization	• Decentralization analysis	• Ibrahim [32]
Patterns of mobile apps	• Mobile app design patterns	• Rizky et al [33] • Asmawi et al [34]
Inclusion, security, identity, and user experience	• Design and accessibility considerations • Blockchain security and authentication • Digital identity management • User experience and accessible design	• Gumilar et al [35] • Musa et al [36] • Alanzi and Alkhatib [37] • Jang and Han [38]

Conclusions

This study designed a functional architectural proposal for a blockchain-based mobile app that facilitates digital ID and access to essential services, with a focus on the inclusion and data security of older adults. The solution addresses usability, privacy, and decentralization issues, overcoming the limitations of traditional centralized and vulnerable systems. In addition to its technical and usability benefits, the proposed system entails ethical challenges related to data governance. If local authorities or institutions act as blockchain nodes, they could potentially exert undue control over users' identity data. To prevent this, future implementations should ensure transparent node management, community oversight, and independent auditing. These measures are key to promoting genuine digital inclusion while safeguarding autonomy and trust in rural contexts. Nevertheless, usability tests applied using the SUS revealed high acceptance and perceived ease of use by older adults, demonstrating that the proposal represents a solid basis for the comprehensive development of the app, which has been implemented in a functional prototype currently undergoing real-world testing with older adult participants in rural areas.

Acknowledgments

The authors would like to thank the Dirección de Investigación de la Universidad Peruana de Ciencias Aplicadas (UPC) for the support provided for the publication of this article.

Funding

This work received financial support from the Dirección de Investigación de la Universidad Peruana de Ciencias Aplicadas (UPC) exclusively for the article processing fee (APF). No additional funding was provided for the research activities.

Data Availability

The data collected in this study consist of usability evaluation responses (eg, survey data such as System Usability Scale scores and qualitative feedback) obtained from voluntary participants during prototype testing. No clinical, laboratory, or medical record data were collected. The dataset is not publicly available due to privacy and ethical considerations, but can be provided by the corresponding author upon reasonable request.

Authors' Contributions

WA-R and AFP-L contributed equally to this work. They led the study execution, including participant recruitment, data collection, usability evaluation, analysis of results, and drafting of the manuscript. JCM-A provided research supervision and leadership, contributed to the study planning and methodological guidance, ensured alignment with journal requirements, coordinated the submission and revision process, and performed critical review and editing of the manuscript.
All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
C4 model component diagram.
[PNG File, 134 KB - rehab_v13i1e79553_app1.png]

References

1. Cai S, Du J, Chen X, Li E, Chen Y. The relationship between e-health literacy and educational participation motivation among elderly individuals: the chained mediating effects of self-identity and social capital. Br J Hosp Med (Lond) 2024 Sep 30;85(9):1-13. [doi: 10.12968/hmed.2024.0261] [Medline: 39347685]
2. Farao A, Paparis G, Panda S, Panaousis E, Zarras A, Xenakis C. INCHAIN: a cyber insurance architecture with smart contracts and self-sovereign identity on top of blockchain. Int J Inf Secur 2024 Feb;23(1):347-371. [doi: 10.1007/s10207-023-00741-8]
3. Morte-Nadal T, Esteban-Navarro M. Recommendations for digital inclusion in the use of European digital public services. Humanit Soc Sci Commun 2025;12(1):273. [doi: 10.1057/s41599-025-04576-7]
4. Yi S, Yam ELY, Cheruvettolil K, et al. Perspectives of digital health innovations in low- and middle-income health care systems from South and Southeast Asia. J Med Internet Res 2024 Nov 25;26:e57612. [doi: 10.2196/57612] [Medline: 39586089]
5. Song W, Nokhbeh Zaeem R, Liau D, et al. Self-sovereign identity and user control for privacy-preserving contact tracing. Presented at: WI-IAT '21; Dec 14-17, 2021; Melbourne, Australia p. 438-445 URL: https://dl.acm.org/doi/proceedings/10.1145/3486622 [doi: 10.1145/3486622.3493914]
6. George M, Chacko AM. Health passport: a blockchain-based PHR-integrated self-sovereign identity system. Front Blockchain 2023;6:1075083. [doi: 10.3389/fbloc.2023.1075083]
7. Ji H, Dong J, Pan W, Yu Y. Associations between digital literacy, health literacy, and digital health behaviors among rural residents: evidence from Zhejiang, China. Int J Equity Health 2024 Apr 9;23(1):68. [doi: 10.1186/s12939-024-02150-2] [Medline: 38594723]
8. Lu SY, Yoon S, Yee WQ, Heng Wen Ngiam N, Ng KYY, Low LL. Experiences of a community-based digital intervention among older people living in a low-income neighborhood: qualitative study. JMIR Aging 2024 Apr 25;7:e52292. [doi: 10.2196/52292] [Medline: 38662423]
9. Prieto-Egido I, Sanchez-Chaparro T, Urquijo-Reguera J. Impacts of information and communication technologies on the SDGs: the case of Mayu Telecomunicaciones in rural areas of Peru. Inf Technol Dev 2023 Jan 2;29(1):103-127. [doi: 10.1080/02681102.2022.2073581]
10. Li J, Liang N, Wu Y, Ma X, Liao Z. Effect of internet usage on the life satisfaction of the elderly spousal caregivers of disabled elderly. Sci Rep 2024;14(1):22294. [doi: 10.1038/s41598-024-73298-8]
11. Konopik J, Blunck D. Development of an evidence-based conceptual model of the health care sector under digital transformation: integrative review. J Med Internet Res 2023;25:e41512. [doi: 10.2196/41512]
12. Husain L, Greenhalgh T. Examining intersectionality and barriers to the uptake of video consultations among older adults from disadvantaged backgrounds with limited English proficiency: qualitative narrative interview study. J Med Internet Res 2025 Jan 6;27:e65690. [doi: 10.2196/65690] [Medline: 39761566]
13. Zuñe Chero L, Capuñay-Uceda OE, Idrogo Burga E, Capuñay-Uceda CE. Inequalities in learning performance caused by the digital divide during the health emergency. Presented at: 22nd LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI 2024; Jul 17-19, 2024 URL: https://laccei.org/LACCEI2024-CostaRica [doi: 10.18687/LACCEI2024.1.1.1042]
14. UPC Code of Ethics in Scientific Research of the Peruvian University of Applied Sciences (UPC). URL: https://investigacion.upc.edu.pe/assets/files/C%C3%B3digo-de-%C3%A9tica-en-la-investigacion-cientifica-de-la-UPC.pdf [accessed 2026-02-02]
15. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e29512. [doi: 10.2196/29512] [Medline: 35029535]
16. Sobrinho ADS, Gomes GDO, Bueno Júnior CR. Developing a multiprofessional mobile app to enhance health habits in older adults: user-centered approach. JMIR Form Res 2024 Apr 15;8(8):e54214. [doi: 10.2196/54214] [Medline: 38619865]
17. Essop H, Kekana R, Smuts H. Co-designing of a prototype mobile application for fetal radiation dose monitoring among pregnant radiographers using a design thinking approach. Health Informatics J 2024;30(3):14604582241284960. [doi: 10.1177/14604582241284960] [Medline: 39348214]
18. Chen E, Bishop J, Guge Cozon L, et al. Integrating human-centered design methods Into a health promotion project: supplemental nutrition assistance program education case study for intervention design. JMIR Form Res 2023 Apr 21;7:e37515. [doi: 10.2196/37515] [Medline: 37083485]
19. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023 Sep 21;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
20. Rashidi H, Rashidi Z. Software architecture tools: a classification and survey with recommendation for an organization. J Comput Secur 2023(10):61-81. [doi: 10.22108/jcs.2023.137862.1131]

XSL•FO
RenderX

21. Voorheis P, Zhao A, Kuluski K, et al. Integrating behavioral science and design thinking to develop mobile health interventions: systematic scoping review. JMIR Mhealth Uhealth 2022 Mar 16;10(3):e35799. [doi: 10.2196/35799] [Medline: 35293871]
22. Liang X, Zhao J, Chen Y, Bandara E, Shetty S. Architectural design of a blockchain-enabled, federated learning platform for algorithmic fairness in predictive health care: design science study. J Med Internet Res 2023 Oct 30;25:e46547. [doi: 10.2196/46547] [Medline: 37902833]
23. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194. [doi: 10.1201/9781498710411]
24. Hill JR, Brown JC, Campbell NL, Holden RJ. Usability-in-place-remote usability testing methods for homebound older adults: rapid literature review. JMIR Form Res 2021 Nov 2;5(11):e26181. [doi: 10.2196/26181] [Medline: 34726604]
25. Sweeney M, Barton W, Nebeker C. Evaluating mobile apps targeting older adults: descriptive study. JMIR Form Res 2023 Apr 27;7:e37329. [doi: 10.2196/37329] [Medline: 37103995]
26. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults’ intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
27. Kim S, Park C, Park S, et al. Measuring digital health literacy in older adults: development and validation study. J Med Internet Res 2025 Feb 5;27:e65492. [doi: 10.2196/65492] [Medline: 39908081]
28. Liang X, Alam N, Sultana T, Bandara E, Shetty S. Designing a blockchain-empowered telehealth artifact for decentralized identity management and trustworthy communication: interdisciplinary approach. J Med Internet Res 2024 Sep 25;26:e46556. [doi: 10.2196/46556] [Medline: 39320943]
29. Koo JH, Park YH, Kang DR. Factors predicting older people’s acceptance of a personalized health care service app and the effect of chronic disease: cross-sectional questionnaire study. JMIR Aging 2023 Jun 21;6:e41429. [doi: 10.2196/41429] [Medline: 37342076]
30. Tan E, Mahula S, Crompvoets J. Blockchain governance in the public sector: a conceptual framework for public management. Gov Inf Q 2022 Jan;39(1):101625. [doi: 10.1016/j.giq.2021.101625]
31. Barrera C. A framework for blockchain governance design: the Prysm group wheel.: Medium; 2019. URL: https://medium.com/prysmeconomics/a-framework-for-blockchain-governance-design-the-prysm-group-wheel-703279c1b0dd [accessed 2025-06-01]
32. Ibrahim AHH. Decentralization and its impact on improving public services. Int J Soc Sci 2024;7(2):45-53. [doi: 10.21744/ijss.v7n2.2278]
33. Rizky A, Lutfiani N, Sri Mariyati W, Atika Sari A, Febrianto KR. Decentralization of information using blockchain technology on mobile apps e-journal. B-FronT 2022;1(2):1-10 [FREE Full text] [doi: 10.34306/bfront.v1i2.37]
34. Asmawi A, Saifulbahri EM, Mohd Ariffin NA. Development of BlockScholar as an educational mobile application on blockchain technology. ARASET 2023;34(1):15-23 [FREE Full text] [doi: 10.37934/araset.34.1.1523]
35. Gumilar DWA, Sangka KB, Totalia SA. Digital financial literacy and digital financial inclusion in the era of digital disruption: systematic literature review. FJMR 2024;3(5):1563-1576 [FREE Full text] [doi: 10.55927/fjmr.v3i5.9213]
36. Musa HS, Krichen M, Altun AA, Ammi M. Survey on blockchain-based data storage security for android mobile applications. Sensors (Basel) 2023 Oct 26;23(21):8749. [doi: 10.3390/s23218749] [Medline: 37960449]
37. Alanzi H, Alkhatib M. Towards improving privacy and security of identity management systems using blockchain technology: a systematic review. Appl Sci (Basel) 2022;12(23):12415. [doi: 10.3390/app122312415]
38. Jang H, Han SH. User experience framework for understanding user experience in blockchain services. Int J Hum Comput Stud 2022 Feb;158:102733. [doi: 10.1016/j.ijhcs.2021.102733]

Abbreviations

DT: design thinking
SSI: self-sovereign identity
SUS: System Usability Scale

Edited by S Munce; submitted 24.Jun.2025; peer-reviewed by C Wrigley, SB Guo; revised version received 20.Nov.2025; accepted 03.Dec.2025; published 02.Feb.2026.

Please cite as:
Arana-Ramos W, Pastrana-Leon AF, Morales-Arevalo JC
Blockchain-Based Mobile App for Digital Identification of Older Adults in Rural Peru: Design and Usability Evaluation Study
JMIR Rehabil Assist Technol 2026;13:e79553
URL: https://rehab.jmir.org/2026/1/e79553
doi:10.2196/79553

XSL•FO
RenderX

© Wilver Arana-Ramos, Aldo Francisco Pastrana-Leon, Juan Carlos Morales-Arevalo. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 2.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Role of Technology Acceptance in the Telerehabilitation of Patients With Metabolic Syndrome: Longitudinal Study

Zsanett Tesch[1*], MA, MSci; Tamás Ujházi[1*], MA; István Kósa[2], MD, PhD; Norbert Buzás[1,3], PhD

[1]Department of Health Economics, Albert Szent-Györgyi Medical School, University of Szeged, Szeged, Hungary

[2]Department of Preventive Medicine, Albert Szent-Györgyi Medical School, University of Szeged, Szeged, Hungary

[3]Department of Theoretical Health Sciences and Health Management, Faculty of Health Sciences and Social Studies, University of Szeged, Szeged, Hungary

[*]these authors contributed equally

Corresponding Author:
Norbert Buzás, PhD
Department of Health Economics
Albert Szent-Györgyi Medical School
University of Szeged
35 Kossuth Lajos Sugárút
Szeged, 6724
Hungary
Phone: 36 70 439 9209
Email: buzas.norbert@med.u-szeged.hu

Abstract

Background: The advent of telerehabilitation has created new opportunities for the care of patients with metabolic syndrome. In distant rehabilitation, technology acceptance is particularly important because home-based projects are based on digital devices, and many patients are less familiar with their use.

Objective: Our aim was to explore technology acceptance among patients undergoing a 3-month complex, telemedicine-supported metabolic rehabilitation. We were curious to see how different factors influence the intention to use rehabilitation technologies and how this changes through the telerehabilitation process.

Methods: Participants were selected from patients in the metabolic telerehabilitation program at the university. Our model was based on the unified theory of acceptance and use of technology 2, which we supplemented with various other constructs. A paper-pencil questionnaire survey was administered on the last day of the preparatory week of the rehabilitation program (T1, n=145) and at the follow-up visit after the closing (T2, n=139). We used structural equation modeling with the least squares method to explore the relationships between model variables. Respondent segments were also identified by performing a hierarchical cluster analysis using Ward's method.

Results: Facilitating conditions (FC) have the greatest impact (0.366) on the behavioral intention (BI) to use technology. Effort expectancy has no direct effect on BI; it operates only through performance expectancy (PE), which may be because, in telerehabilitation settings, patients are more goal-driven than experience-driven. The analyses of the T2 data show that the direct impact of social influence on BI has disappeared by the end of the rehabilitation process. This can be explained by the fact that during device use, it becomes clear that the devices are secure and the data are safe, making this factor implicit in the patient's behavior. Only 2 constructs appeared in both the T1 and T2 models: PE and FC. By comparing the 2 datasets, we have provided empirical support for an old hypothesis: the experience of using the tool for a time has led to a significant reduction in the impact of FC and a corresponding increase in the dominance of PE, which has "absorbed" the impact of some other constructs. Based on respondents' attitudes, we found 3 clusters. The telerehabilitation program itself has a significant impact on patients' BI, as the relative share of "enthusiastic users" (73/145, 50.3%) increased by about 20%, while the share of "distrustful reluctants" (25/145, 17.2%) decreased to a quarter by the end of the program.

Conclusions: This behavior-based functional approach enables treatments to be tailored to actual technology-use demands rather than to presumptive societal features. This means that before beginning rehabilitation, attempts should be undertaken to identify patients' clusters in clinical practice, and rehabilitation should be planned according to the individual's attitude toward technology.

XSL•FO
RenderX

(JMIR Rehabil Assist Technol 2026;13:e82161) doi:10.2196/82161

KEYWORDS

metabolic syndrome; modified UTAUT2; unified theory of acceptance and use of technology 2; patient behavior; technology acceptance; telerehabilitation

Introduction

Background

Obesity and metabolic syndrome, as a consequence, are serious public health challenges worldwide, significantly raising the risk of cardiovascular disease, type 2 diabetes, and other chronic conditions [1,2]. The prevalence of obesity has increased steadily in recent decades, not only in affluent countries but also globally, including in middle- and low-income regions [3]. Structured rehabilitation programs play a key role in the care of patients with metabolic syndrome, but access to these programs is limited. Traditional, location-based rehabilitation programs often face capacity constraints, long waiting lists, and geographical or mobility limitations of patients [4-7]. Telerehabilitation offers an alternative to these problems, enabling the therapeutic process to be carried out in the home environment using mobile technology and digital platforms [5].

The 2 most common uses of telerehabilitation are cardiac and poststroke rehabilitation. Mobile technology has been well accepted in rehabilitation programs for patients with ischemic heart disease and congestive heart failure, and digital self-confidence was a significant predictor of acceptance [8]. A comparison of a technology-based, home-based, remotely monitored program and a traditional cardiac rehabilitation program showed that the home program is feasible and acceptable to patients, but that home rehabilitation presents challenges, including maintaining motivation over the long term, which makes its widespread implementation difficult [9]. A focus group study of 188 patients examining telemedicine rehabilitation care for people with chronic diseases [10] found that the possibility of exercising at home, working on one's own recovery, and receiving better-quality exercise instruction thanks to exercise videos on the portal were among the factors that facilitated the acceptance of mobile health (mHealth) rehabilitation. The majority of participants also viewed the benefits of the telerehabilitation program for survivors of stroke as positive [11]. Widespread mobile phone use has helped to make rehabilitation strategies involving the families of survivors of stroke and remote monitoring via telemedicine feasible and cost-effective even in low- and middle-income countries. Participants in another stroke telerehabilitation program also rated regular contact positively, particularly the provision of care-related information and exercise reminders, which helped them to maintain their commitment to the exercise program [12].

Telemedicine has become a key tool in digital health care over the past decade, with the COVID-19 pandemic accelerating the adoption of digital solutions [13]. However, the success of telemedicine-supported programs varies a lot around the world. The main barriers are technology-specific [14], as the success of a program depends on its user-friendly interface and ease of use, especially for older or less technologically experienced users [15]. On the other hand, the effectiveness of programs strongly depends on patients' general attitudes toward technology use, digital competence, socialization background, and perceived privacy in their personal sphere [16,17].

Prior Work

The basis for predicting technology acceptance behaviors can be traced back to the theory of reasoned action of Fishbein and Ajzen [18], which posits that the adoption of a given behavior depends on an individual's attitude and subjective norms. Davis [19] further developed this framework specifically for the study of technology acceptance, creating the technology acceptance model (TAM), in which perceived usefulness and perceived ease of use determine the intention to use. As a next development of the model, Venkatesh and Davis [20] identified factors influencing perceived usefulness in the TAM 2 model, followed by Venkatesh and Bala [21] identifying factors influencing perceived ease of use in the TAM 3 model. The consecutive step was the unified theory of acceptance and use of technology (UTAUT), developed by Venkatesh et al [22]. A further development of this, the unified theory of acceptance and use of technology 2 (UTAUT2) model [23], specifically examines technology adoption among everyday consumers. The research frameworks for exploring technology acceptance listed here share the commonality that they examine use intention and actual use as target variables. Intent to use reflects the extent to which consumers are open to adopting a new technology based on their existing attitudes, while actual use reveals how much time they spend using the device once they have the opportunity.

In their comprehensive methodological study, Holden and Karsh [24] pointed out more than a decade ago that an important future direction for technology acceptance is to adapt the model explicitly to the health context. Notwithstanding, a decade later, Niknejad et al [25] still identified in telerehabilitation far fewer studies than expected that had shown a quantitative association between promoting factors or inhibitors of telerehabilitation adoption and intention to use. Patient-side adoption of telemedicine is influenced by a number of factors that are not part of classical models explaining technology acceptance but are more specific, contextual determinants [10]. Liu et al [26] showed that expected performance was the strongest predictor of intention to use technology for rehabilitation, whereas effort expectancy (EE) and social impact were not significant. In contrast, facilitating conditions (FC) significantly influenced not only intention but also the actual use. Koo et al [27], analyzing the adoption of personalized health management mobile apps by older patients, highlighted the role of performance expectancy (PE), EE, FC, social influence (SI), habit, hedonic motivation (HM), and trust. Huang et al [28] also confirmed the influence of TAM variables for electronic health recording and sharing apps, while Fezzi et al [29] identified perceived usefulness, perceived ease of use, skill level, task

XSL•FO
RenderX

complexity, and the nature of the technology as influencing factors for functional insulin therapy. Cooks et al [30] associated the acceptance of virtual health care professionals (HPs) with a sense of belonging, local identity, self-concept, similarity, and attractiveness, while Kauttonen et al [31] found that the acceptance of artificial intelligence–based health care apps is determined by trust, data security, and the role of the vendor. Schomakers et al [32] and Siebert et al [33] noted that, for therapeutic mobile apps, trust is an independent explanatory factor alongside the UTAUT construct. Rajak and Shaw [16] attempted to generalize the TAM model to mHealth users and demonstrated its high predictive power. Zhu et al [34] used the UTAUT2 model for a similar purpose and demonstrated that the model's original constructs have a meaningful impact on the intention to use. Schretzlmaier et al [35] also extended the UTAUT2 model with their own constructs and found that the factors trust and received threat of illness are also relevant for the acceptance of mHealth technologies. Yang et al [17], building on the UTAUT theory, have recently demonstrated the impact of the convenience, accuracy, and privacy protection factors on technology intention to use. Examining the adoption of mHealth technologies in a national sample, it was found that, in addition to the classic variables of the UTAUT2 model, privacy, lifestyle, self-efficacy, and trust in Bangladesh [36], while resistance to use, data privacy issues, and HM in the Philippines [37], also influence behavioral intention (BI) to use.

In conclusion, in addition to the application of classical TAMs such as TAM and UTAUT, there is a growing emphasis in practice on the inclusion of psychological and technology-specific constructs that make the explanation of technology acceptance in health care more complex and viable. It is also evident from decades of practice in the widespread use of models that constructs influencing target variables can be freely modified by dropping some constructs or adding new ones to suit specific research contexts or objectives [38]. In line with this, AlQudah et al [39] also emphasize the scope to integrate new constructs to enhance the robustness of models of user acceptance of new technologies in health care. This is also confirmed by Ramachandran et al [40], who argue that although the usefulness of home-based telemedicine-supported rehabilitation programs is clear to patients, adoption is also influenced by a number of external factors.

Study Objectives and Proposed Hypotheses

Overview

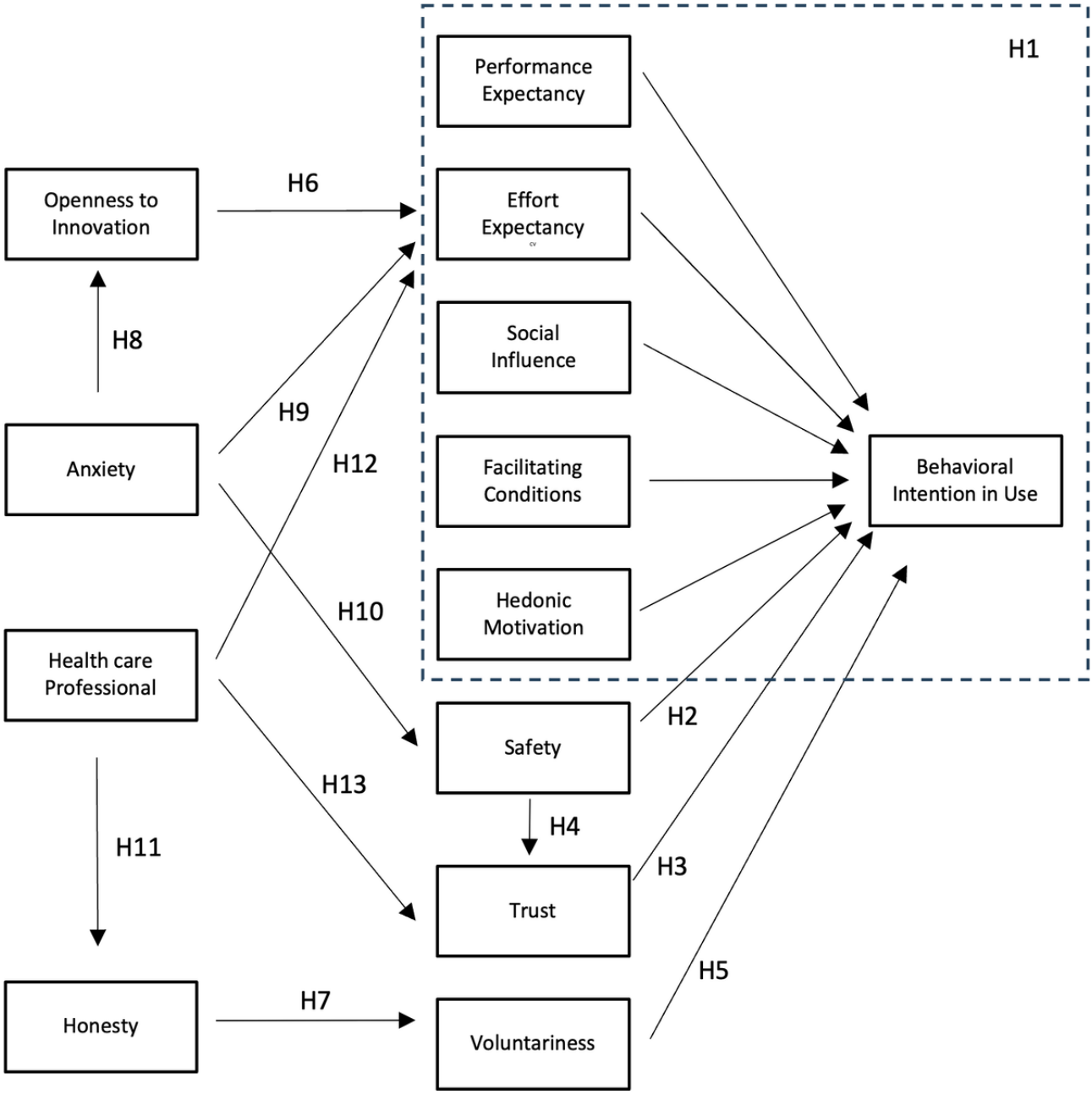
The aim of this study is to investigate the factors influencing technology acceptance in a telemedicine-assisted rehabilitation program for patients with metabolic syndrome. Our aim was to investigate how constructs influence the intention to use technology and, in turn, to develop an acceptance model for the telerehabilitation process for patients with metabolic syndrome. On the other hand, we sought to answer how the experience of the telerehabilitation process influences patients' technology acceptance parameters, that is, how participation in the process shapes their intention to use.

Our modeling was based on the UTAUT2 model [23]. Since the technologies provided to users were free, we omitted the price construct from our model. We did the same with habit, as the participants in the rehabilitation program did not have a routine in using the devices. For the remaining constructs in the UTAUT2 model, we assumed the specified impacts [23] on the intention to use rehabilitation technologies (hypothesis 1).

To further describe the factors influencing technology adoption in metabolic rehabilitation, we also conducted an exploratory qualitative study. The aim of this was to identify additional relevant constructs that may influence the intention to use the technologies, based on the experiences of professionals directly involved in the program, in line with the suggestion of Kapser and Abdelrahman [38], and to incorporate them into our baseline model. Data collection consisted of 1-hour semistructured interviews with a total of 3 physicians, 2 nurses, 2 dietitians, and 2 physiotherapists from different specialized groups within the clinic. The participants all had direct contact with patients undergoing metabolic rehabilitation and were thus able to provide meaningful insights into the manifestation of attitudes, barriers, and supportive factors related to technology use. The information gathered was subjected to a content analysis, which revealed several factors that are less or differently emphasized in models in the literature. Therefore, we have included these constructs (safety, trust, voluntariness, openness to innovation [OP], honesty, anxiety, and HPs) in our conceptual model (Figure 1), and we explain their use in detail below.

XSL•FO
RenderX

Figure 1. The extended UTAUT2 model as research hypotheses. H: hypothesis; UTAUT2: unified theory of acceptance and use of technology 2.



Safety

A sense of safety plays a key role in the acceptance of health technologies, especially for patients who use digital or invasive devices regularly. In this context, the safety factor refers to the extent to which the user perceives the technology as physically harmless, psychologically reassuring, technologically reliable, and privacy-safe [41]. Kitsiou et al [42] also showed that perceptions of safety have a significant positive impact on intention to use, especially in sensitive application areas such as continuous glucose monitoring or remote monitoring systems. Examining the use of home telehealth services among older people, Cimperman et al [43] also found that perceived safety positively influences the intention to use. Lack of safety (eg, fear of device malfunction or data leakage), which is effectively equivalent to the perceived threat variable [34,44], can significantly reduce trust in technology, which plays a key mediating role in acceptance models. On this basis, we hypothesized that the safety construct would positively affect the BI to use rehabilitation technologies (hypothesis 2).

Trust

Trust plays a key role in technology adoption, as patients are in a vulnerable position in health care. Trust, in this context, reflects the user's belief that the technology is reliable, predictable, and safe to use during the treatment process [45]. In the case of telemedicine applications, this belief extends to the system's technical stability, data security, and the belief that HPs will use the device appropriately and ethically [46]. Higher levels of trust can significantly increase the intention to use the technology [34-36,47]. Accordingly, in our modeling, we hypothesized that trust would have a positive effect on intention to use for rehabilitation technologies (hypothesis 3). However, we also hypothesized that safety would have a positive effect

XSL•FO
RenderX

on trust, also because safety enhances overall trust (hypothesis 4).

Voluntariness

Voluntariness is a key factor in technology adoption models, especially in environments where users are sensitive to issues affecting their autonomy, such as health care. This latent variable measures the extent to which a user feels that the use of technology is based on free choice rather than on external constraints or expectations [22]. A sense of voluntariness contributes significantly to the development of positive attitudes, increases willingness to accept, and reduces resistance, especially when new technologies (eg, telehealth systems or wearable sensors) are introduced to populations vulnerable to a loss of control [48]. In a clinical setting, it is particularly important that patients perceive technology as a choice rather than an instruction from the treating physician or institution, as this indirectly influences the likelihood of engagement and long-term use. Voluntariness, as a latent variable, appears as a moderating variable in the UTAUT1 model but is no longer included in the UTAUT2 model [23], as it is implicit in the consumer context that consumption is voluntary. Since in health care, a patient may always feel "pressure" from health care staff, so voluntariness is not implicit, we reintegrated this variable into the model and assumed that voluntariness positively influences intention to use rehabilitation technologies (hypothesis 5).

Openness to Innovation

OP reflects an individual's willingness and ability to try new technologies, solutions, or services, even if they initially have little information about how they work. This personality dimension often serves as a moderator or indirect variable in models of technology adoption, as it significantly impacts attitudes toward technological innovations and, in turn, the intention to use them [49]. Individuals with high OP tend to be more willing to use digital health solutions, such as telemedicine applications, and are less likely to be apprehensive about novelty, making it easier for them to turn initial experimentation into regular use [50]. In the context of metabolic rehabilitation, this is particularly relevant, as patients need to regularly monitor their own condition and provide data, which requires the use of digital tools, the acceptance of which can be greatly facilitated by a positive attitude toward technological novelty. As openness is primarily associated with technology-use efforts, we hypothesized that OP would have a positive effect on EE within the rehabilitation technology adoption model (hypothesis 6).

Honesty

Based on the experience of qualitative in-depth interviews with health care staff, we included social effects such as honesty. This is the willingness of users to share sensitive personal information when using the tools provided to them [51]. This depends on the provider's transparent data management and open communication with the user [52]. These suggest that user honesty, as a specific form of social interaction, can be incorporated into the TAM in digital tool–assisted rehabilitation, where accurate, fair, and voluntary data disclosure is key to effectiveness (hypothesis 7).

Anxiety

Anxiety describes the negative anticipatory reactions and emotional responses associated with technology use, arising from fear of use, fear of making mistakes, or experiencing the complexity and unpredictability of technology. This psychological factor is particularly important in digital health settings, where users often perceive improper use as a health risk [53]. Technology anxiety often negatively impacts perceived ease of use and perceived usefulness, thereby weakening the intention to use [54]. This is particularly relevant for patients undergoing metabolic rehabilitation, as most have limited digital experience and may be more sensitive to the risk of digital system failure due to their health conditions. Anxiety may therefore also be a barrier to OP (hypothesis 8), EE (hypothesis 9), and sense of safety (hypothesis 10).

Health Care Professionals

From the construct called SI in previous models, we have separated the influence of HPs to more accurately describe the effects, as their role is multifaceted and more actively helping than the encouraging role of family or friends. In particular, doctors, nurses, physiotherapists, and dieticians have a significant impact on patients' willingness to use technology in rehabilitation when they introduce new digital solutions such as telemedicine systems or mHealth apps. For patients, health care staff are a key source of information and a point of trust, so their positive attitudes may facilitate technology acceptance, while their skepticism or passivity may prevent it [55]. Accordingly, we hypothesized that HP would positively impact the honesty (hypothesis 11), EE (hypothesis 12), and trust (hypothesis 13) factors.

Methods

Sample Size Estimation

Given the finite population of eligible program participants (N=170), we calculated the minimum required sample size using a conservative proportion-based approach (P=.50), a 95% confidence level, and a 5% margin of error with a finite population correction, resulting in a minimum of 118 participants. Our achieved sample size exceeded this threshold (T1: n=145; T2: n=139). We also included a power-based verification aligned with the primary partial least squares-structural equation modeling model, confirming that the baseline sample provides adequate power (α=.05; 80% power) to predict BI from the number of model predictors.

Recruitment

Participants in the study were selected from patients in the metabolic telerehabilitation program at the Cardiometabolic Rehabilitation Unit of the Internal Medicine Clinic, University of Szeged, using convenience sampling. The study included individuals between the ages of 25 and 70 years who had low levels of physical activity (less than 30 minutes of self-reported exercise per week) and lived with metabolic syndrome. To define metabolic syndrome, we used the Adult Treatment Panel-III criteria [56], according to which, participants must have at least 3 of the 5 risk factors listed therein, as determined from a review of their medical records.

XSL•FO
RenderX

Individuals with any of the following conditions were excluded from selection: planned invasive cardiological intervention, uncontrolled hypertension, type 1 diabetes mellitus, type 2 diabetes treated with more than 1 daily dose of insulin, heart failure, renal failure, cancer, severe cognitive impairment, lack of cooperation, known conditions preventing physical training, and mental conditions severely limiting judgment or capacity to act.

Questionnaire Development

To measure the variables in the model shown in Figure 1, a self-completion questionnaire was developed through a multistep procedure to ensure adequate content and face validity. The selection of constructs was grounded in the theoretical framework and the qualitative semistructured interviews described earlier. Content validity was ensured by aligning questionnaire items with both established literature-based measurement scales and the expert-derived thematic findings. The preliminary questionnaire was reviewed by members of the research team with expertise in health economics, preventive medicine, and digital health research to evaluate item relevance, clarity, and domain coverage. Face validity was assessed through a pilot evaluation involving representatives of the target population participating in the rehabilitation preparation phase, focusing on the comprehensibility, clarity of wording, and interpretability of the items. Based on feedback received, minor linguistic and formatting adjustments were implemented before administering the final questionnaire. All items were measured using a 5-point Likert scale, reflecting the level of agreement with each statement. The constructs and corresponding measurement items are presented in Table 1.

The translation of the questions was carried out using the following method: 2 independent experts translated them, and their translations were then reconciled. The reconciled version was back-translated and approved by the corresponding author (NB).

XSL•FO
RenderX

Table 1. Variables and measurement items in the UTAUT2[a] model for telerehabilitation.

Constructs and items	Content	Loadings	Cronbach α	CR[b]	AVE[c]
Anxiety (AX)			0.932	0.950	0.879
AX1	Nervousness	0.945			
AX2	Confusion	0.945			
AX3	Uncomfortability	0.922			
Behavioral intention (BI)			0.921	0.924	0.864
BI1	Intention	0.934			
BI2	Frequency	0.954			
BI3	Continuity	0.899			
Effort expectancy (EE)			0.913	0.937	0.796
EE1	Easy to learn	0.919			
EE2	User-friendliness	0.944			
EE3	Easy to use	0.928			
EE4	No special skills needed	0.767			
Facilitating conditions (FC)			0.858	0.868	0.701
FC1	Equipment conditions	0.847			
FC2	Troubleshooting	0.871			
FC3	Compatibility	0.791			
FC4	Technical support	0.838			
Hedonic motivation (HM)			0.873	0.969	0.792
HM1	Fun	0.871			
HM2	Joy	0.94			
HM3	Entertainment	0.857			
Health care professionals (HPs)			0.908	0.882	0.787
HP1	Doctors	0.896			
HP2	Dietitians	0.843			
HP3	Physiotherapists	0.925			
HP4	Nurses	0.876			
Honesty (HO)			0.744	0.911	0.784
HO1	True data	0.942			
HO2	Embarrassment	0.828			
Openness to innovation (OP)			0.815	0.868	0.722
OP1	Openness to new products	0.875			
OP2	Up-to-date in technology	0.875			
OP3	First user of a new technology	0.797			
Performance expectancy (PE)			0.853	0.867	0.695
PE1	Usefulness	0.784			
PE2	Accomplishments	0.891			
PE3	Productivity	0.872			
PE4	Data tracking	0.782			
Safety (SA)			0.8	0.798	0.718
SA1	Security	0.75			

XSL•FO
RenderX

Constructs and items	Content	Loadings	Cronbach α	CR[b]	AVE[c]
SA2	Safety of data	0.905			
SA3	Confidentiality	0.879			
Social influence (SI)			0.789	0.797	0.704
SI1	Influence of close people	0.873			
SI2	Proudness	0.794			
SI3	Friends' encouragement	0.847			
Trust (TR)			0.863	0.874	0.708
TR1	Trustworthiness	0.806			
TR2	Privacy	0.868			
TR3	Reliability	0.845			
TR4	Safety of device	0.845			
Voluntariness (VO)					
VO1	Free will	1	—[d]	—	—

[a]UTAUT2: unified theory of acceptance and use of technology 2.

[b]CR: composite reliability.

[c]AVE: average variance extracted.

[d]Not applicable.

Data Collection

In this study, a paper-pencil questionnaire was administered to patients with metabolic syndrome who had participated in a 3-month complex rehabilitation program on 2 occasions: on the last day of the preparatory week prior to the rehabilitation program (T1) and at the follow-up visit after the program closed (T2). Patients completed the questionnaires in a separate, quiet room. All questionnaires were recorded in an electronic database by one of the researchers (ZT).

Statistical Analysis

The first step in analyzing the collected data was to use descriptive statistics to describe the sample population's sociodemographic characteristics. To answer our research question, we used structural equation modeling with the least squares method to explore the relationships among model variables using SmartPLS software (version 4; SmartPLS GmbH). The path analysis allowed us to investigate the extent to which each variable directly influences the intention to use the devices dispensed to patients during metabolic telerehabilitation and to identify interactions between these variables, thereby identifying which variables have an indirect effect on intention to use. Finally, to provide a deeper interpretation of the results, respondent segments were identified through a hierarchical cluster analysis using Ward's method. The procedure uses the participants' responses to the indicators as criteria and, based on these, classifies those with similar responses into similar inwardly homogeneous clusters, while the system ensures that the different clusters are heterogeneous along their criteria, that is, well distinguishable. The output of the cluster analysis is to identify clusters by saving the clusters as variables and comparing averages along the indicator variables to identify the respondent groups.

Ethical Considerations

The study was conducted in accordance with the guidelines of the Declaration of Helsinki and was approved by the Medical Research Council of Hungary (approval BM/25016-1/2024). The research was based on prior written informed consent from the participants. The questionnaire was completed in a separate room, with participants seated at a distance from one another to ensure privacy. After completing the questionnaire, participants could not be identified, thus ensuring confidentiality. Participants did not receive any compensation for their participation in the research.

Results

Demographic Characteristics

The demographic profile of respondents (n=145) is shown in Table 2. During the study period, 58.5% (83/145) of the participants in the telemedicine-supported metabolic rehabilitation program were female, and 41.5% (59/145) were male. Patients aged 25 to 82 years (mean age 52.6, SD 12.65 years) completed the questionnaire. In terms of age distribution, 19.6% (28/145) were younger than 40 years, almost half were between 51 and 70 years, and only 9.8% (14/145) were older than 70 years. In total, 28.7% (41/145) of respondents came from a large city, 37.8% (54/145) from a small town, and 1 in 3 from a village. In terms of education, nearly two-thirds of respondents had a secondary education, one-third had a higher education degree, and 2.1% (3/145) had only a primary education. In terms of marital status, 63.9% of respondents are in a couple, 43.7% have children, and 29.9% declared themselves completely single. Regarding their financial means, 45.5% (66/145) of respondents said that everyday expenses are not a problem, that they can afford some extra expenses in their monthly budget, but that they cannot save. A total of 36.6%

XSL•FO
RenderX

(53/145) of respondents can afford extra spending beyond living expenses and can save at the end of the month. Only 2.8% (4/145) of respondents have serious financial difficulties.

Table 2. Participants' characteristics.

Demographic factors	Values, n (%)
Sex	
Male	59 (41.5)
Female	83 (58.5)
Age (years)	
≤40	28 (19.6)
41-50	31 (21.7)
51-60	36 (25.2)
61-70	34 (23.8)
≥70	14 (9.8)
Marital status	
Married or relationship without children	38 (26.4)
Married or relationship with children	54 (37.5)
Single without children	43 (29.9)
Single with children	9 (6.3)
Place of residence	
Bigger city	41 (28.7)
Small city	54 (37.8)
Village	48 (33.6)
Highest level of education	
Primary school	3 (2.1)
Vocational college	45 (31)
Secondary school	46 (31.7)
College or university	47 (32.4)
Other	4 (2.8)
Financial situation	
My monthly income is not enough to buy basic necessities	4 (2.8)
I can buy basic necessities, but I have no money for anything else; it just about lasts until the end of the month	21 (14.5)
I can buy basic necessities and occasionally allow myself 1 or 2 extra expenses, but I cannot save any money	66 (45.5)
I can buy what I need and even put some money aside, and I can afford occasional extra expenses	53 (36.6)
I can buy what I need and even put some money aside, and I can afford regular extra expenses	1 (0.7)

Partial Least Squares-Structural Equation Modeling Model

The results confirming the reliability and validity of the constructs are also summarized in Table 1. Construct reliability was assessed using the Cronbach α coefficient and the composite reliability (CR) coefficient [57]. The CR is more accurate because it accounts for the model structure and covariances. Both coefficients above 0.70 are considered satisfactory. The results showed that Cronbach α and CR were all above 0.7, with the lowest values being 0.744 and 0.797, respectively. However, none of the values exceeded the redundancy threshold of 0.95, indicating that the constructs are reliable. The convergent validity of the constructs was determined using average variance extracted. The results showed that all average variance extracted values were above 0.5, indicating adequate convergent validity for all constructs, reflecting their 1D nature.

XSL•FO
RenderX

Table 3 presents the hypothesis testing results for the path analysis based on the questionnaires collected in T1, and Figure 2 shows their graphical representation. Table S1 in Multimedia Appendix 1 contains the more detailed statistical parameters. It can be seen that the intention to use (BI) is directly and positively influenced by the PE, SI, and FC factors, all of which are included in the UTAUT2 model. No direct effect of EE was detected, and HM only has an indirect effect on intention to use through PE. Thus, we have only partially confirmed hypothesis 1.

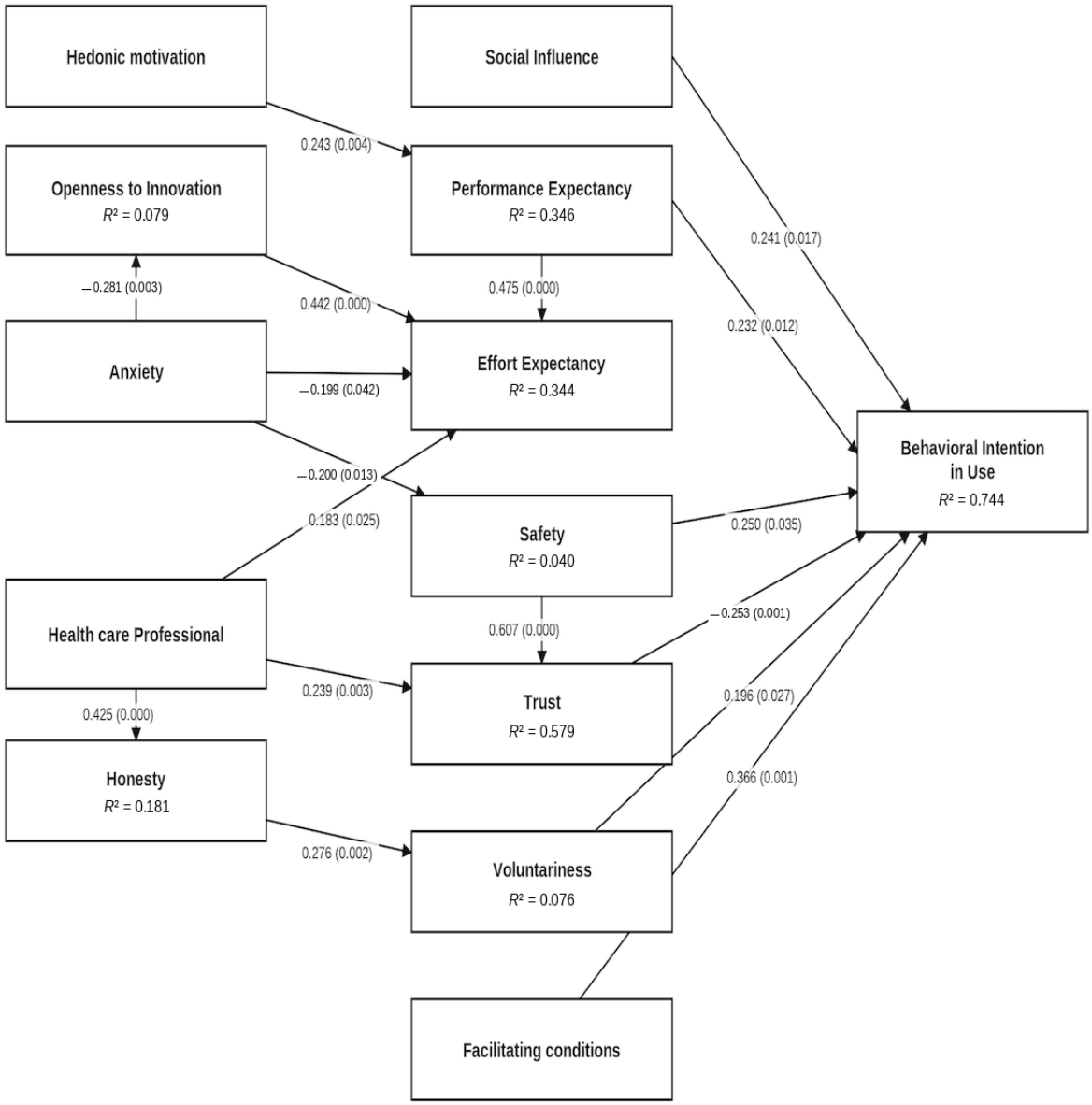
Table 3. Hypothesis testing summary for T1 and T2 surveys.

Path	Hypothesis (H)	T1		T2	
		Original sample	*P* value	Original sample	*P* value
PE[a]→BI[b]	H1	0.232	.01	0.403	<.001
EE[c]→BI	H1	Not confirmed	—[d]	Not confirmed	—
SI[e]→BI	H1	0.241	.02	Not confirmed	—
FC[f]→BI	H1	0.366	.001	0.266	.01
HM[g]→BI	H1	Not confirmed	—	Not confirmed	—
EE→PE	Not presumed	0.475	<.001	0.315	.02
HM→PE	Not presumed	0.243	.004	0.334	.003
SA[h]→BI	H2	0.25	.04	Not confirmed	—
TR[i]→BI	H3	–0.253	.001	Not confirmed	—
SA→TR	H4	0.607	<.001	0.411	<.001
VO[j]→BI	H5	0.196	.03	Not confirmed	—
OP[k]→EE	H6	0.442	<.001	Not confirmed	—
HO[l]→VO	H7	0.276	.002	Not confirmed	—
AX[m]→OP	H8	–0.281	.003	–0.416	<.001
AX→EE	H9	–0.199	.04	–0.53	<.001
AX→SA	H10	–0.2	.01	–0.462	<.001
HP[n]→HO	H11	0.425	<.001	Not confirmed	—
HP→EE	H12	0.183	.03	0.283	.001
HP→TR	H13	0.239	.003	0.268	.004

[a]PE: performance expectancy.
[b]BI: behavioral intention.
[c]EE: effort expectancy.
[d]Not available.
[e]SI: social influence.
[f]FC: facilitating condition.
[g]HM: hedonic motivation.
[h]SA: safety.
[i]TR: trust.
[j]VO: voluntariness.
[k]OP: openness to innovation.
[l]HO: honesty.
[m]AX: anxiety.
[n]HP: health care professional.

XSL•FO
RenderX

Figure 2. Extended UTAUT2 model with latent variables and path coefficients in T1. UTAUT2: unified theory of acceptance and use of technology 2.



In addition to the above, BI is also directly positively affected by safety and voluntariness. Therefore, hypotheses 2 and 5 are accepted. Trust also has a direct negative effect on BI, so hypothesis 3 is not confirmed. Safety, on the other hand, has a very strong effect (0.607) on trust, confirming hypothesis 4. Although EE does not have a direct effect on BI, it has a strong effect on PE (0.475), indicating that the tools used in the telerehabilitation program are easy and user-friendly without prior training.

In terms of the strength of constructs directly affecting intention to use, the largest influence (0.366) is FC, which includes 4 items: the conditions in the rehabilitation participant's home, the problem-solving ability required to use rehabilitation tools, the compatibility of the tools used with those the user is currently using, and the expected assistance with problems that would arise during use. Nearly equally influential on BI is PE (0.232), that is, the perception of how useful the rehabilitation program and the tools used in it will be in achieving the goals set. Just like SI is 0.241, which indicates the influence of close people to the patient; and safety is 0.250, which indicates that the perception of security in the use of the devices and the inaccessibility of the data generated by the use of the devices to unauthorized persons. Voluntariness, that is, the intention to use the system despite the fact that its use is not compulsory, is less significant (0.196) than the former ones.

Examining which other constructs moderate the direct influence on intention to use, we see that FC and SI are not related to other factors. Trust is significantly influenced by HP (0.239), that is, the expectation of active assistance from health staff is a clear contributor to the perceived safety. HP also has a significant effect on honesty (0.425), implying that the attitude of doctors, nurses, dieticians, and physiotherapists toward

XSL•FO
RenderX

patients contributes to the honesty of the data. HM, which has a direct effect on BI in the UTAUT2 model, has a direct effect on PE here, indicating that performance expectations are also proportional to the enjoyment of technology. Anxiety, as originally hypothesized, also affects 3 other factors: OP, EE, and safety. The effect is negative in all 3 cases: lower innovation anxiety is beneficial for openness to new technologies, expectations of the effort required to operate, and perceptions of safety. Figure 2 also shows that all hypotheses (hypotheses 6-13) related to constructs with a moderating effect were confirmed.

To investigate how participation in the rehabilitation program influences intention to use technology, we also conducted a survey in T2 among patients (n=139) returning to the final medical control at the end of the program. The results of this path analysis are also shown in Table 3. Additional statistical parameters are provided in Table S2 in Multimedia Appendix 1. The model appears to have been simplified. The effect of SI on intention to use was eliminated, so the variable was eliminated from the model. In addition, the effects of voluntariness on BI and of honesty on voluntariness have been discontinued, so voluntariness has disappeared from the model. The direct effects of trust and safety on BI have also ceased, so that only FC and PE have an effect on the intention to use after the rehabilitation program. The effect relationships of anxiety and their sign have not changed, but their effects have increased significantly. The effects of HP also remained unchanged, but while its influence on honesty was weakened, its influence on EE was significantly strengthened.

Cluster Analysis

Overview

TAMs provide a useful framework for understanding how patients perceive digital health solutions in the context of telerehabilitation. However, these models assume linear relationships and cannot capture heterogeneity within patient populations or the diverse patterns of technology use. Based on this consideration, we also apply cluster analysis to our research to identify distinct groups of patients with metabolic syndrome characterized by different attitudes and behavior patterns.

By performing hierarchical cluster analysis on the questionnaire data recorded at T1 (using a dendrogram created with a combination of Ward's minimum variance clustering algorithm and squared Euclidean distance), 3 distinct clusters were identified. The groups are heterogeneous among themselves and homogeneous within the cluster. The clusters are not segregated by demographic characteristics; no category of age, education, residence, or income is predominant in any cluster. Consequently, clusters were further analyzed by comparing the means of responses to questions across variables, that is, functional behavioral patterns, attitudes, emotional reactions, and social support related to digital technology use.

Based on the attitudes of the respondents, the 3 distinct groups and their characteristics are described.

Enthusiastic Users

About half of the participants (73/145, 50.3%) belong to this group. They are open and active in their attitudes toward the use of technology. They enjoy using rehabilitation tools and are eager to try new technologies. Their characteristics are as follows: they most enjoy using rehabilitation tools, they are most likely to try new technologies, and they find those who influence their opinions most encouraging in their use of devices.

This group is the engine of rehabilitation programs: their commitment and positive attitude can serve as a model for others.

Liberated Adopters

Members of the second-largest group (32.4%) are open to using tools but are less active and, in some respects, more moderately engaged than enthusiastic users. The following attributes define this group: technology anxiety is the lowest, they rely most on family and friends for help, they have above-average device use, and they learn and use tools more easily than the average person.

This group is a stable base for the program: if given support and encouragement, they can easily become enthusiastic users.

Distrustful Reluctants

This is the smallest, but no less important, group of people (25/145, 17.2%) who are reticent and critical of rehabilitation technologies. This distrust is often due to anxiety, insecurity, or poor social support. They are the ones who enjoy using the devices less, are less attracted to trying new technologies, have above-average anxiety about using devices, are less likely to rely on family support, have more difficulty than average in learning to use new devices, and have the least belief in the role of device use in rehabilitation.

They are the group most in need of attention: personalized support, confidence-building, and mentoring can help them develop a more positive relationship with rehabilitation tools.

The cluster analysis was also performed on the questionnaire data collected in T2. We found that cluster blocks with practically identical characteristics can be distinguished, but their relative numbers of elements have changed. Among those who completed the program, the proportion of enthusiastic users increased from 50.3% to 61%, and the proportion of liberated adopters was almost the same (32.4% and 34.6%, respectively), while the proportion of distrustful reluctants decreased from 17.3% to 4.4%.

Due to the different cluster sizes and the varied distribution of participants, we cannot statistically say that different age groups behave differently. At the same time, the distribution of participants across clusters is not uniform: the group aged 60 years and older is concentrated in the "liberated adopters" cluster, whereas there are no participants aged 50 years and older in the "distrustful reluctants" cluster (although this cluster has the fewest members). Even after adjusting for unequal cluster sizes and the sample's underlying age distribution, age

XSL•FO
RenderX

remained associated with cluster membership, suggesting that age was a significant structuring variable of the clusters.

The association between sex and cluster membership was examined using a chi-square test of independence. The analysis indicated that sex distribution did not differ significantly among clusters (chi-square test, $P>.05$), suggesting that cluster membership was independent of sex.

Discussion

Principal Findings

As the results show, EE in the T1 study has no direct effect on intention to use, only through PE. This seems surprising at first sight, because EE usually has a direct positive effect on BI in UTAUT2 models. Although Venkatesh et al [23] suggested that if BI is highly dependent on PE, it may "absorb" the effect of EE, this has not been quantified. Several studies have found a significant relationship between EE and PE [43,58,59], but in no case did this prevent both constructs from also impacting BI. However, Kim et al [60] observed no direct effect of EE on BI when they used a UTAUT model to examine the adoption of mHealth tools and found that EE only affected BI as a moderating variable of trust. The lack of a direct effect of EE may be due to the fact that the experience of ease of use (EE) is not important per se, but because it increases the perception of usefulness (PE). In telerehabilitation settings, patients are not experience-driven but rather goal-driven, that is, they place more emphasis on whether the technology actually helps their recovery than on how easy it is to use.

In addition, as a result of the study in T1, although HM is included among the influencing constructs in UTAUT2, we did not measure its direct effect on intention to use (BI). Instead, HM moderates only through PE, that is, an enjoyable experience increases willingness to use only if it also positively influences perceived usefulness. Similar results are reported in the mHealth adoption study by Schretzlmaier et al [35], where the influence of PE and fear of illness on BI is clear, while the effect of HM is only conditional. The moderating effect of HM has also been demonstrated by Sitar-Tăut [61], who, in an m-learning setting, found that HM strongly influences PE within the HM-PE-BI sequence.

In addition, the analysis of the T1 study yielded the surprising result that trust has a negative effect on BI. This means that the more users trust the system, the less likely they are to want to use it. This seems to be a logical contradiction, as one would expect trust to have a positive effect on willingness to use, which is confirmed by several previous studies [34,35]. However, this phenomenon is consistent with a kind of safe bubble effect described in the literature [62], whereby users, due to excessive technological trust, feel less motivated to actively use technologies that require more interaction. Excessive technological confidence reduces user participation, and when confidence in the system exceeds perceived need, the active use of technology may even be abandoned [63]. In our case, this may have been reinforced by the technology's actual trial, as data collection for the prerehabilitation analysis was conducted at the end of the preparation week, by which time, patients had learned to use the required tools in practice. In contrast, in T2, we found that the direct effect of trust on BI had disappeared, remaining relevant in the model only as a moderator. The experience gained during use and active participation in the operation is already reflected in the PE, resulting in a PE dominance in T2.

The analysis of the data recorded in T2 shows that the direct influence of SI, voluntariness, and safety on BI has also disappeared by the end of the telerehabilitation process. SI plays an important role in initiating device use, but later, this driving force is replaced by utility and effectiveness, embodied in PE. By examining the development of long-term device use habits, Peng et al [64] also found that SI before use is important, but that later regular use is no longer fueled by SI. The disappearance of voluntariness has no antecedent in the literature, as it is not included in the UTAUT2 model; it was reintroduced only under possible medical and eminence-based pressures, and its validity was confirmed by the T1 analysis. Sharif Abbasi et al [65], using the UTAUT model, found a result similar to ours when examining internet adoption over a longer period: while in T1, the effect of voluntariness on BI was measurable, in T2, it no longer remained significant, as experience with use increased. The disappearance of the effect of safety on BI can be explained by the fact that during device use, it becomes clear that the devices are secure and the data are safe, making this factor implicit in the patient's behavior.

Of the constructs that directly influence technology use intention, only 2 remain: PE and FC, which appear in both the T1 and T2 models. While in the previous measurement, FC is by far the strongest (0.366) and PE is almost the weakest (0.232) direct effect on BI, in the latter measurement, FC's influence is significantly weaker (0.258). One of the most important extensions of UTAUT2 was the introduction of user experience as a moderator on the constructs, including FC [23]. It was argued that, over time, as the routine of using the technology increases, FC becomes less crucial, sometimes even acting as a pullback force. By comparing the 2 datasets, we have now been able to provide empirical support for this hypothesis: the experience of using the tool for 3 months has led to a significant reduction in the impact of FC and a corresponding increase in the dominance of PE, which has "absorbed" the impact of SI, trust, and safety as well.

While clustering approaches commonly used in telemedicine are typically based on static categories such as socioeconomic status [66] or health literacy [67], we investigated how patients behave in a technological environment, what barriers they experience during use, and what motivational dynamics characterize them. The 3 groups explored represent different levels of technological engagement. This concept is more complex than technology acceptance because it includes emotional involvement (enjoyment or anxiety), motivation and volunteering, openness to novelty, willingness to learn, and self-efficacy, as well as the role of social support in technology use.

The cluster analysis of questionnaires collected before and after the telerehabilitation program shows that the program has a significant impact on patients' intention to use. The relative

XSL•FO
RenderX

share of enthusiastic users increased by about 20%, while the share of distrustful reluctants decreased to a quarter. Our results show that 2 factors are crucial for willingness to use: on the one hand, a meaningful trial of the technology (not just a demonstration, but an actual data collection simulating real-life situations at home) is the best way to dispel fears and anxieties about it, that is, a project where participants in a rehabilitation program can try out the technologies for a longer time during the preparation period can significantly contribute to their acceptance and willingness to use. On the other hand, the supportive and motivating role of the immediate environment is crucial; that is, where the family is indifferent or unsupportive, it is worthwhile to motivate family members with targeted communication during the preparation period of the telerehabilitation program.

Conclusions

Our study has shown that the enabling environment is a key determinant of technology adoption in a telerehabilitation program. Consequently, when developing such a program, the focus should be on creating user-friendly interfaces, clearly communicating the benefits of technology and its relationship with other devices, and developing patients' technological competencies. For patients more experienced with the technologies, PE will be the determining factor; if patients can experience the benefits of the technology in action during the preparation for the telerehabilitation program (training in the technologies and data delivery), their engagement will be stronger. If the impact of technological competency is to remain strong throughout the duration of a telerehabilitation program, it makes sense to adopt an adaptive strategy, that is, to measure increases in competency and reduce support intensity accordingly. This is both a way to signal user autonomy and to avoid a sense of excessive control. The cluster analysis also confirms that the key factor in the success of such a program is the tailored support and communication to the specific characteristics of each group, thereby strengthening their long-term commitment to participation. Indeed, the groups identified are not merely descriptive categories but operational profiles directly linked to the practical challenges of telerehabilitation. This functional approach, based on behavioral patterns, allows interventions to be personalized not according to assumed social characteristics but to real technology use needs. In clinical practice, using this, efforts should be made to identify clusters before initiating telerehabilitation. The next task is to develop personalized coaching techniques for each identified cluster and then examine whether attitudes toward technology use and the effectiveness of the entire telerehabilitation process can be improved through personalized interventions. We have already initiated developments in this direction, incorporating artificial intelligence support.

Limitations

A limiting factor could be that, assuming a lack of technological experience, we omitted this from the model, even though there were certainly patients who possess such experience. We also did not assess whether the patients had participated in a similar rehabilitation program before or how this influenced their current attitudes. As the study was designed to compare intentions to use, actual use may have been influenced by factors we were unable to fully account for. With a larger sample size, it would also be possible to explain the behavioral differences observed across clusters using structural equation modeling analysis, which could be an exciting direction for future research.

Acknowledgments

The authors would like to thank the staff of the Cardiometabolic Rehabilitation Unit at the Internal Medicine Clinic, University of Szeged, for their assistance with the questionnaire survey.

Funding

The research was financially supported by the Healthcare Development Cluster of the Centre of Excellence for Interdisciplinary Research, Development and Innovation of the University of Szeged.

Data Availability

The datasets generated during this study are available in the Mendeley Data repository [68].

Authors' Contributions

Conceptualization: ZT, TU, IK, NB
Methodology: ZT, TU
Formal analysis: ZT, TU
Resources: IK
Supervision: NB
Visualization: NB
Writing—original draft: ZT, TU, NB
Writing—review and editing: IK, NB

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Path analysis data for the preparatory week (T1) and the follow-up visit after the closing (T2).
[DOCX File , 18 KB - rehab_v13i1e82161_app1.docx]

References

1. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation 2005;112(17):2735-2752 [FREE Full text] [doi: 10.1161/circulationaha.105.169404]
2. GBD 2021 Risk Factor Collaborators (2024). Global burden of 88 risk factors in 204 countries and territories, 1990-2021: a systematic analysis for the Global Burden of Disease study. Lancet 2024;403:2162-2203 [FREE Full text] [doi: 10.1002/hon.2936/v2/review2]
3. Obesity and overweight. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2026-07-16]
4. Dalal HM, Taylor RS. Telehealth technologies could improve suboptimal rates of participation in cardiac rehabilitation. Heart 2016;102:1155-1156 [FREE Full text] [doi: 10.1136/heartjnl-2016-309429]
5. Thomas RJ, Beatty AL, Beckie TM, Brewer LC, Brown TM, Forman DE, et al. Home-Based cardiac rehabilitation: a scientific statement from the American Association of Cardiovascular and Pulmonary Rehabilitation, the American Heart Association, and the American College of Cardiology. Circulation 2019;140(1):133-153 [FREE Full text] [doi: 10.1161/cir.0000000000000663]
6. Evenson KR, Fleury J. Barriers to outpatient cardiac rehabilitation participation and adherence. J Cardiopulm Rehabil 2000;20(4):241-246 [FREE Full text] [doi: 10.1097/00008483-200007000-00005] [Medline: 10955265]
7. Khadanga S, Gaalema DE, Savage P, Ades PA. Underutilization of cardiac rehabilitation in women: barriers and solutions. J Cardiopulm Rehabil Prev 2021;41(4):207-213 [FREE Full text] [doi: 10.1097/hcr.0000000000000629]
8. Bäuerle A, Mallien C, Rassaf T, Jahre L, Rammos C, Skoda EM, et al. Determining the acceptance of digital cardiac rehabilitation and its influencing factors among patients affected by cardiac diseases. J Cardiovasc Dev Dis 2023;10(4):174 [FREE Full text] [doi: 10.3390/jcdd10040174] [Medline: 37103053]
9. Claes J, Cornelissen V, McDermott C, Moyna N, Pattyn N, Cornelis N, et al. Feasibility, acceptability, and clinical effectiveness of a technology-enabled cardiac rehabilitation platform (Physical Activity Toward Health-I): randomized controlled trial. J Med Internet Res 2020;22(2):e14221 [FREE Full text] [doi: 10.2196/14221] [Medline: 32014842]
10. Jansen-Kosterink S, Dekker-van Weering M, van Velsen L. Patient acceptance of a telemedicine service for rehabilitation care: a focus group study. Int J Med Inform 2019;125:22-29 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.01.011] [Medline: 30914177]
11. Sarfo FS, Adamu S, Awuah D, Sarfo-Kantanka O, Ovbiagele B. Potential role of tele-rehabilitation to address barriers to implementation of physical therapy among West African stroke survivors: a cross-sectional survey. J Neurol Sci 2017;381:203-208 [FREE Full text] [doi: 10.1016/j.jns.2017.08.3265] [Medline: 28991682]
12. Saywell NL, Mudge S, Kayes NM, Stavric V, Taylor D. A six-month telerehabilitation programme delivered via readily accessible technology is acceptable to people following stroke: a qualitative study. Physiotherapy 2023;120:1-9 [FREE Full text] [doi: 10.1016/j.physio.2023.05.001] [Medline: 37352776]
13. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. BMC Public Health 2020;20(1):1193 [FREE Full text] [doi: 10.1186/s12889-020-09301-4] [Medline: 32738884]
14. Kruse CS, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2016;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633x16674087]
15. Sanders C, Rogers A, Bowen R, Bower P, Hirani S, Cartwright M, et al. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study. BMC Health Serv Res 2012;12:220 [FREE Full text] [doi: 10.1186/1472-6963-12-220] [Medline: 22834978]
16. Rajak M, Shaw K. An extension of technology acceptance model for mHealth user adoption. Technol Soc 2021;67:101800 [FREE Full text] [doi: 10.1016/j.techsoc.2021.101800]
17. Yang M, Al Mamun A, Gao J, Rahman MK, Salameh AA, Alam SS. Predicting m-health acceptance from the perspective of unified theory of acceptance and use of technology. Sci Rep 2024;14(1):339 [FREE Full text] [doi: 10.1038/s41598-023-50436-2] [Medline: 38172184]
18. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, PA: Addison-Wesley Publishing Company; 1975.
19. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340 [FREE Full text] [doi: 10.2307/249008]

XSL•FO
RenderX

20. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000;46(2):186-204 [FREE Full text] [doi: 10.1287/mnsc.46.2.186.11926]
21. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008;39(2):273-315 [FREE Full text] [doi: 10.1111/j.1540-5915.2008.00192.x]
22. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478 [FREE Full text] [doi: 10.2307/30036540]
23. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178 [FREE Full text] [doi: 10.2307/41410412]
24. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
25. Niknejad N, Ismail W, Bahari M, Nazari B. Understanding telerehabilitation technology to evaluate stakeholders' adoption of telerehabilitation services: a systematic literature review and directions for further research. Arch Phys Med Rehabil 2021;102(7):1390-1403. [doi: 10.1016/j.apmr.2020.12.014] [Medline: 33484693]
26. Liu L, Miguel Cruz A, Rios Rincon A, Buttar V, Ranson Q, Goertzen D. What factors determine therapists' acceptance of new technologies for rehabilitation—a study using the unified theory of acceptance and use of technology (UTAUT). Disabil Rehabil 2015;37(5):447-455 [FREE Full text] [doi: 10.3109/09638288.2014.923529] [Medline: 24901351]
27. Koo JH, Park YH, Kang DR. Factors predicting older people's acceptance of a personalized health care service app and the effect of chronic disease: cross-sectional questionnaire study. JMIR Aging 2023;6:e41429 [FREE Full text] [doi: 10.2196/41429] [Medline: 37342076]
28. Huang J, Pang WS, Wong YY, Mak FY, Chan FSW, Cheung CSK, et al. Factors associated with the acceptance of an eHealth app for electronic health record sharing system: population-based study. J Med Internet Res 2022;24(12):e40370 [FREE Full text] [doi: 10.2196/40370] [Medline: 36382349]
29. Fezzi C, Refahi H, Albuisson E, Feigerlova E. User-centered approach in the development of an eHealth tool for self-management skills in functional insulin therapy to prevent complications of diabetes. Prev Med Rep 2022;29:101968 [FREE Full text] [doi: 10.1016/j.pmedr.2022.101968] [Medline: 36161109]
30. Cooks EJ, Duke KA, Flood-Grady E, Vilaro MJ, Ghosh R, Parker N, et al. Can virtual human clinicians help close the gap in colorectal cancer screening for rural adults in the United States? The influence of rural identity on perceptions of virtual human clinicians. Prev Med Rep 2022;30:102034 [FREE Full text] [doi: 10.1016/j.pmedr.2022.102034] [Medline: 36531088]
31. Kauttonen J, Rousi R, Alamäki A. Trust and acceptance challenges in the adoption of AI applications in health care: quantitative survey analysis. J Med Internet Res 2025;27:e65567 [FREE Full text] [doi: 10.2196/65567] [Medline: 40116853]
32. Schomakers EM, Lidynia C, Vervier LS, Calero Valdez A, Ziefle M. Applying an extended UTAUT2 model to explain user acceptance of lifestyle and therapy mobile health apps: survey study. JMIR Mhealth Uhealth 2022;10(1):e27095 [FREE Full text] [doi: 10.2196/27095] [Medline: 35040801]
33. Siebert J, Joeckel L, Heidrich J, Trendowicz A, Nakamichi K, Ohashi K, et al. Construction of a quality model for machine learning systems. Software Qual J 2021;30(2):307-335 [FREE Full text] [doi: 10.1007/s11219-021-09557-y]
34. Zhu Y, Zhao Z, Guo J, Wang Y, Zhang C, Zheng J, et al. Understanding use intention of mHealth applications based on the unified theory of acceptance and use of technology 2 (UTAUT-2) model in China. Int J Environ Res Public Health 2023;20(4):3139 [FREE Full text] [doi: 10.3390/ijerph20043139] [Medline: 36833830]
35. Schretzlmaier P, Hecker A, Ammenwerth E. Predicting mHealth acceptance using the UTAUT2 technology acceptance model: a mixed-methods approach. Stud Health Technol Inform 2023;2(301):26-32 [FREE Full text] [doi: 10.3233/shti230007]
36. Alam MZ, Hu W, Kaium A, Hoque R, Didarul Alam MM. Understanding the determinants of mHealth apps adoption in Bangladesh: a SEM-neural network approach. Technol Soc 2020;61:101255 [FREE Full text] [doi: 10.1016/j.techsoc.2020.101255]
37. Ong AKS, Kurata YB, Castro SADG, De Leon JPB, Dela Rosa HV, Tomines APJ. Factors influencing the acceptance of telemedicine in the Philippines. Technol Soc 2022;70:102040 [FREE Full text] [doi: 10.1016/j.techsoc.2022.102040]
38. Kapser S, Abdelrahman M. Acceptance of autonomous delivery vehicles for last-mile delivery in Germany—extending UTAUT2 with risk perceptions. Transport Res C Emerg Technol 2020;111:210-225 [FREE Full text] [doi: 10.1016/j.trc.2019.12.016]
39. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. Appl Sci 2021;11(22):10537 [FREE Full text] [doi: 10.3390/app112210537]
40. Ramachandran HJ, Jiang Y, Teo JYC, Yeo TJ, Wang W. Technology acceptance of home-based cardiac telerehabilitation programs in patients with coronary heart disease: systematic scoping review. J Med Internet Res 2022;24(1):e34657 [FREE Full text] [doi: 10.2196/34657] [Medline: 34994711]
41. Van Velthoven MH, Smith J, Wells G, Brindley D. Digital health app development standards: a systematic review protocol. BMJ Open 2018;8(8):e022969 [FREE Full text] [doi: 10.1136/bmjopen-2018-022969] [Medline: 30121614]
42. Kitsiou S, Paré G, Jaana M. Systematic reviews and meta-analyses of home telemonitoring interventions for patients with chronic diseases: a critical assessment of their methodological quality. J Med Internet Res 2013;15(7):e150 [FREE Full text] [doi: 10.2196/jmir.2770] [Medline: 23880072]

XSL•FO
RenderX

43. Cimperman M, Makovec Brenčič M, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an extended UTAUT model. Int J Med Inform 2016;90:22-31 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
44. Raue M, D'Ambrosio LA, Ward C, Lee C, Jacquillat C, Coughlin JF. The influence of feelings while driving regular cars on the perception and acceptance of self-driving cars. Risk Anal 2019;39(2):358-374 [FREE Full text] [doi: 10.1111/risa.13267]
45. Montague E, Asan O. Trust in technology-mediated collaborative health encounters: constructing trust in passive user interactions with technologies. Ergonomics 2012;55(7):752-763 [FREE Full text] [doi: 10.1080/00140139.2012.663002] [Medline: 22506847]
46. Shin DH. The effects of trust, security and privacy in social networking: a security-based approach to understand the pattern of adoption. Interact Comput 2010;22(5):428-438 [FREE Full text] [doi: 10.1016/j.intcom.2010.05.001]
47. Zhou T. Understanding mobile internet continuance usage from the perspectives of UTAUT and flow. Inform Dev 2011;27(3):207-218 [FREE Full text] [doi: 10.1177/0266666911414596]
48. Or CKL, Karash BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.m2888]
49. Low MW, Ramayah T. It Isn't enough to be easy and useful! Combined use of SEM and necessary condition analysis for a better understanding of consumers' acceptance of medical wearable devices. Smart Health 2023 Mar;27:100370. [doi: 10.1016/j.smhl.2022.100370]
50. Agarwal R, Prasad J. A conceptual and operational definition of personal innovativeness in the domain of information technology. Inform Syst Res 1998;9(2):204-215 [FREE Full text] [doi: 10.1287/isre.9.2.204]
51. Grande D, Luna Marti X, Feuerstein-Simon R, Merchant RM, Asch DA, Lewson A, et al. Health policy and privacy challenges associated with digital technology. JAMA Netw Open 2020;3(7):e208285 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.8285] [Medline: 32644138]
52. Karampela M, Ouhbi S, Isomursu M. Connected health user willingness to share personal health data: questionnaire study. J Med Internet Res 2019;21(11):e14537 [FREE Full text] [doi: 10.2196/14537] [Medline: 31774410]
53. Kamal SA, Shafiq M, Kakria P. Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). Technol Soc 2020;60:101212 [FREE Full text] [doi: 10.1016/j.techsoc.2019.101212]
54. Meuter ML, Ostrom AL, Bitner MJ, Roundtree RI. The influence of technology anxiety on consumer use and experiences with self-service technologies. J Bus Res 2003;56(11):899-906 [FREE Full text] [doi: 10.1016/s0148-2963(01)00276-4]
55. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220 [FREE Full text] [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
56. NCEP. Third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002;106(25):3143-3421. [doi: 10.1161/circ.106.25.3143]
57. Dijkstra TK, Henseler J. Consistent partial least squares path modeling. MIS Q 2015;39(2):297-316 [FREE Full text] [doi: 10.25300/misq/2015/39.2.02]
58. Alrawashdeh TA, Muhairat MI, Alqatawnah SM. Factors affecting acceptance of web-based training system: using extended UTAUT and structural equation modeling. ArXiv Preprint posted online on May 9, 2012 [FREE Full text] [doi: 10.5121/ijcseit.2012.2205]
59. Li W, Gui J, Luo X, Yang J, Zhang T, Tang Q. Determinants of intention with remote health management service among urban older adults: a unified theory of acceptance and use of technology perspective. Front Public Health 2023;11:1117518 [FREE Full text] [doi: 10.3389/fpubh.2023.1117518] [Medline: 36778558]
60. Kim S, Zhong Y, Wang J, Kim HS. Exploring technology acceptance of healthcare devices: the moderating role of device type and generation. Sensors (Basel) 2024;24(24):7921 [FREE Full text] [doi: 10.3390/s24247921] [Medline: 39771659]
61. Sitar-Tăut DA. Mobile learning acceptance in social distancing during the COVID-19 outbreak: the mediation effect of hedonic motivation. Hum Behav Emerg Technol 2021;3(3):366-378 [FREE Full text] [doi: 10.1002/hbe2.261] [Medline: 34222833]
62. Lee JD, See J. Trust in automation: designing for appropriate reliance. Human Factors 2004;46(1):50-80 [FREE Full text] [doi: 10.1518/hfes.46.1.50_30392]
63. Parasuraman R, Riley V. Humans and automation: use, misuse, disuse, abuse. Human Factors 1997;39(2):230-253 [FREE Full text] [doi: 10.1518/001872097778543886]
64. Peng W, Li L, Kononova A, Cotten S, Kamp K, Bowen M. Habit formation in wearable activity tracker use among older adults: qualitative study. JMIR Mhealth Uhealth 2021;9(1):e22488 [FREE Full text] [doi: 10.2196/22488] [Medline: 33464216]
65. Sharif Abbasi M, Hussain Chandio F, Fatah Soomro A, Shah F. Social influence, voluntariness, experience and the internet acceptance: an extension of technology acceptance model within a South-Asian country context. J Enterprise Inform Manage 2011;24(1):30-52 [FREE Full text] [doi: 10.1108/17410391111097410]

XSL•FO
RenderX

66. Kontos E, Blake KD, Chou WYS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
67. Cajita MI, Hodgson NA, Budhathoki C, Han HR. Intention to use mHealth in older adults with heart failure. J Cardiovasc Nurs 2017;32(6):E1-E7 [FREE Full text] [doi: 10.1097/jcn.0000000000000401]
68. Technology acceptance in the telerehabilitation of patients with metabolic syndrome. Mendeley Data. URL: https://data.mendeley.com/datasets/xtvmsd84dn/1 [accessed 2026-05-11]

Abbreviations

BI: behavioral intention
CR: composite reliability
EE: effort expectancy
FC: facilitating condition
HM: hedonic motivation
HP: health care professional
mHealth: mobile health
OP: openness to innovation
PE: performance expectancy
SI: social influence
TAM: technology acceptance model
UTAUT: unified theory of acceptance and use of technology
UTAUT2: unified theory of acceptance and use of technology 2

Edited by S Munce; submitted 10.Aug.2025; peer-reviewed by D Al mojabil, L Zhu; comments to author 26.Feb.2026; revised version received 06.Mar.2026; accepted 23.Mar.2026; published 14.May.2026.

Please cite as:
Tesch Z, Ujházi T, Kósa I, Buzás N
Role of Technology Acceptance in the Telerehabilitation of Patients With Metabolic Syndrome: Longitudinal Study
JMIR Rehabil Assist Technol 2026;13:e82161
URL: https://rehab.jmir.org/2026/1/e82161
doi:10.2196/82161
PMID:

©Zsanett Tesch, Tamás Ujházi, István Kósa, Norbert Buzás. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 14.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Introduction of Nurse-Led Rehabilitation Services for Patients With Stroke After Discharge to Improve Self-Care Management in Bangladesh: Pilot Randomized Controlled Trial

Salma Akhter[1], RN, MSN; KATM Ehsanul Huq[1], MBBS, DTM, MMS, PhD; Shota Date[1], OTR, PhD; Mahmudul Islam[2], MBBS, MD; Farida Khatun Chhobi[2], MBBS; SK Jakaria Been Sayeed[2], MBBS, MRCP; Md Mazharul Islam[2], MBBS, MD; Mahabuba Afrin[1], RN, MPH, MSS, PhD; Kana Kazawa[1], RN, MSN, PhD; Michiko Moriyama[1], RN, MSN, PhD

[1]Graduate School of Biomedical and Health Science, Hiroshima University, Hiroshima, Hiroshima, Japan

[2]National Institute of Neurosciences & Hospital, Dhaka, Dhaka Division, Bangladesh

Corresponding Author:
Michiko Moriyama, RN, MSN, PhD
Graduate School of Biomedical and Health Science
Hiroshima University
Kasumi 1-2-3 Minami-ku
Hiroshima, Hiroshima, 734-8553
Japan
Phone: 1 81 82 257 5367
Fax: 1 81 82 257 5367
Email: morimich@hiroshima-u.ac.jp

Abstract

Background: Stroke is a leading cause of disability and death, resulting in a clinical, social, and economic burden upon families and the health system worldwide.

Objective: This study aimed to introduce nurse-led rehabilitation services for patients after stroke to improve functional independence for self-care management.

Methods: A pilot, open-label, 2-arm (1:1), randomized controlled trial was conducted at the National Institute of Neuroscience & Hospital in Bangladesh between March and August 2025. A total of 64 patients with stroke were enrolled using simple randomization. In the intervention group (IG), patients received rehabilitation education and assistive devices to improve activities of daily living. The control group (CG) received usual hospital care for stroke. Data were collected through a structured questionnaire. After discharge from the hospital, study nurses enrolled patients after stroke after a face-to-face assessment and provided rehabilitation education at the hospital and monthly rehabilitative education at home for 3 months. The study outcomes were improvement of functional independence, self-efficacy, social participation and reduction of caregivers' burden.

Results: Among 64 participants, 47 (73%) completed the study. Results depict an increase in self-care capacity, self-efficacy, social involvement and a decrease in care burden in both groups from baseline to end line; nevertheless, there was no statistically significant difference between IG and CG ($F_{2, 44}$=0.113; P=.74). Functional Independence Measure mean scores (IG: mean 78.4, SD 37.6; CG: mean 74.3, SD 35.4; P=.70; 95% CI –25.16 to 18.14), self-efficacy (IG: mean 27.0, SD 6.9; CG: mean 26.6, SD 7.1; P=.83; 95% CI –3.6 to 4.6), social participation (IG: mean 30.4, SD 21.4; CG: mean 28.2, SD 17.2; P=.61; 95% CI –9.33 to 13.61), and care burden (IG: mean 24.9, SD 13.6; CG: mean 25.2, SD 12.6; P=.83; 95% CI –8 to 7). In qualitative analysis, we noticed enhanced self-care capacities of patients with poststroke disabilities after rehabilitative intervention, as indicated by participants' perceptions.

Conclusions: In this study, patients' self-care capability was improved after the intervention from baseline, and improvements were noted by using assistive devices based on patients' perceptions and responses. The study findings demonstrated the importance of early self-care management for patients with poststroke disabilities. Future research should focus on long-term, community-integrated strategies that involve primary care, enhanced technological support, and caregiver-focused disability adjustment programs to achieve better outcomes.

Trial Registration: ClinicalTrials.gov NCT06786559; https://clinicaltrials.gov/search?id=NCT06786559

XSL•FO
RenderX

(JMIR Rehabil Assist Technol 2026;13:e88808) doi:10.2196/88808

KEYWORDS

caregiver burden; post discharge; rehabilitation; self-care management; stroke

Introduction

Stroke is a leading cause of disability and death, particularly in low- and middle-income countries (LMIC) [1]. Notably, about 86% of stroke-related deaths and 89% of stroke-associated disabilities have occurred in LMIC. The causes are mainly a lack of prevention, acute management, and rehabilitation facilities and services [2]. The prevalence of stroke globally was 1.2%; in developed countries, 3%; and in developing countries, in 2021 [3]. In Bangladesh, stroke prevalence was 1.1% in 2018 [4], and the trend is rising due to an increase in hypertension, diabetes, smoking, and other modifiable risk factors. High mortality and disability are driven by poor access to acute interventions and limited rehabilitation [5]. Unfortunately, developing countries lack a national database system; therefore, the occurrence of stroke is often underreported [6].

Stroke rehabilitation is a multidimensional process provided by multidisciplinary teams, including physicians, physical and occupational therapists, nurses, psychologists, and social workers [6-8]. It encompasses physical rehabilitation [6], occupational therapy [7], and speech and language therapy [9]. Rehabilitation also includes cognitive therapy for improvement of memory and executive deficits [10], psychological care for poststroke depression and anxiety [11], and social/vocational programs for reintegration of patients into society [6].

The *International Classification of Functioning, Disability, and Health* (*ICF*) is the World Health Organization (WHO) framework for measuring functioning disabilities in the environment where patients with disabilities act and are involved in social participation [12]. *ICF* emphasizes the importance of environmental factors, including assistive devices, which enable persons with impaired function to conduct daily activities and participate in society [12]. However, assistive devices are often not available and expensive, so patients cannot afford specially in LMIC [13]. Acute and postacute rehabilitation services are crucial for improving the quality of life for stroke survivors and their caregivers [14]. To mitigate stroke-related disability and death, urgent priorities include scaling up prevention strategies, improving acute care capacity, and facilitating community-based rehabilitation to address the growing burden of stroke in situations where acute and postacute rehabilitation services are not accessible in health care facilities [1].

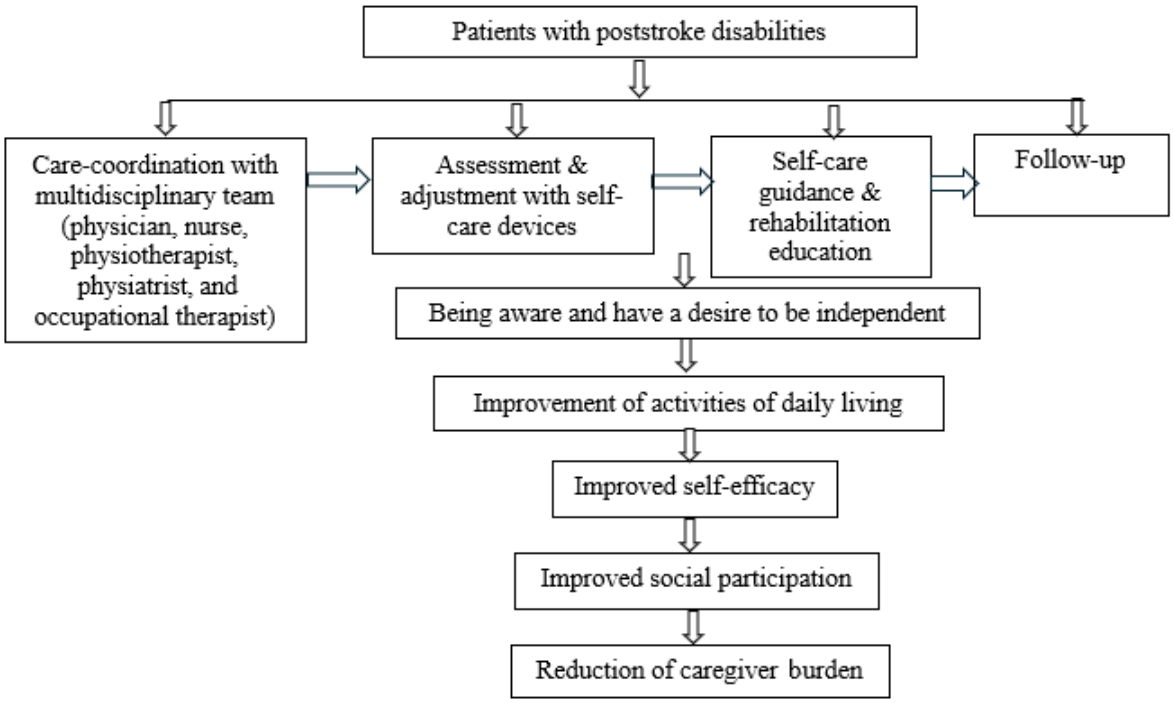
In LMIC, including Bangladesh, there is a lack of institutional acute, postacute, and long-term rehabilitation services available after receiving acute care management from the hospital [15]. Patients with disabilities who need rehabilitation face inadequate health care infrastructure and financial resources [16]. Moreover, poststroke function analysis is a lengthy process that begins with impairment evaluation, rehabilitation training, recovery monitoring, and providing the necessary support to carry out daily tasks [17]. Bangladesh faces a critical shortage of rehabilitation workforce compared to the number of patients with disabilities [18]. Rehabilitation services are limited and fragmented, concentrated in tertiary hospitals, where nongovernment organizations, such as the Centre for the Rehabilitation of the Paralyzed (CRP) and Bangladesh Rural Advancement Committee, provide community-based rehabilitation in some parts of the country [19]. As a result, patients with disabilities have no choice but to be cared for by nonprofessional family members in a nonsupportive and unequipped living environment, which makes it difficult to support their independence and participation in society [20]. Moreover, caregivers face significant challenges, including limited stroke knowledge, financial strain, psychological stress, and cultural expectations [21].

As there are insufficient poststroke institutional services available in Bangladesh, home-based stroke rehabilitation could be an alternative way to play a vital role in ensuring continuity of care and supporting functional recovery within patients' home environments [22]. Globally, it is delivered through multidisciplinary outreach teams, primary care providers, caregiver involvement, and increasingly telerehabilitation, which has proven effective in functional and activities of daily living (ADL) recovery [6,7,23].

Nurse-led rehabilitation is well-practiced globally [24]. They excel at coordinating necessary services and providing rehabilitation care integrated with disease management education for patients and their families. However, in Bangladesh, the role of nurses is limited to acute care nursing, and they do not participate in providing rehabilitation services to patients with poststroke disabilities. Therefore, the purpose of this study was to test the feasibility and efficacy of nurse-led rehabilitation programs focused on improving ADL among patients after stroke in Bangladesh (Figure 1).

XSL•FO
RenderX

Figure 1. Study conceptual framework.



Methods

Study Design and Participants

This was a parallel (1:1), open-label, prospective, randomized controlled trial that adhered to the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Multimedia Appendix 1) [25]. In addition to evaluating the intervention quantitatively, we qualitatively explored the feasibility of home-based rehabilitation and the use of assistive devices for promoting functional independence among patients after stroke and their family caregivers. This study was conducted at the National Institute of Neurosciences & Hospital (NINS&H) in Dhaka, Bangladesh, between March 2025 and August 2025. As NINS&H is the only specialized referral hospital in Bangladesh, most of the patients with stroke are usually discharged within 15 days to 1 month due to overcrowding/high bed occupancy. Therefore, we included patients who were admitted within 5-14 days of the onset of their stroke and discharged within 1 month from the hospital. Patients with stroke and their family caregivers, both male and female individuals, aged ≥18 years, were included. Patients were enrolled irrespective of their time (first or any recurrent) and type (ischemic or hemorrhagic) of stroke, with a modified Rankin scale (mRS) score of 1-4. Patients who had a physician's advice for rehabilitation therapy and needed assistive devices for ADL were included. It is evident that within this time (1-2 weeks), if rehabilitation is initiated, the patient experiences a better prognosis [26]. We conveniently enrolled only participants living in the Dhaka district to ensure follow-up visits with minimal dropout.

Sample Size

Sample size was generated using statistical power of 80% at α=.05 (2-tailed), with an estimated medium effect size of 0.75 [27]. We used this effect size as we had sufficient power to detect a meaningful effect. The sample size was calculated to be 29 for each arm (a total of 58 for both arms). Considering a 10% dropout rate, the estimated total sample size was 64 (32 in each arm). We used G*Power software (version 3.1.9.4; Heinrich Heine University Düsseldorf) for Windows for this calculation.

Randomization and Allocation

Participants were randomized equally into intervention and control groups (CGs; n=32 in each group) using a simple random technique by computer-generated numbers, regardless of their age and sex. Participants were sequentially allocated to the intervention and CGs based on the randomization chart. Randomization was done, and the random allocation sequences were kept by a researcher who was not involved in this study to minimize bias.

Assessment of Eligibility and Recruitment Procedures

At first, patients who were admitted to the hospital were assessed by physicians and the research team to identify any physical impairment. The physiotherapist also assessed the range of motion and notified the physician who required rehabilitation based on their disabilities. After checking the eligibility criteria and verifying with the physician, research assistant (RA) nurses explained the study purpose, ensured home visits for follow-up, and obtained written informed consent from patients and their caregivers. The patient was assessed for functional independence using the Functional Independence Measure (FIM) by RA nurses to evaluate their level of functioning. They also assessed the patient's self-efficacy, social participation, and caregivers' burden.

XSL•FO
RenderX

Interventions

Nurse-Led Intervention

Intervention nurses assessed patients' needs, prepared and supported the use of assistive devices, provided rehabilitation education to patients and carers, used telenursing to monitor patients' recovery remotely, and coordinated multidisciplinary care. This comprehensive nurse-led approach empowered the stroke survivors to participate actively in their recovery, promoted independence in daily activities, reduced complications and carer burden, and ultimately improved self-care management.

Rehabilitation Education

Rehabilitation education was provided according to the researcher-developed health booklet, followed by established guidelines and textbook [28] and a published journal guideline [29]. This booklet includes environmental adjustment, self-care management guidelines with disability (eg, eating and mouth care, grooming, bathing, toileting, and mobility), tips to avoid falling, instructions to use assistive devices, changing dietary habits, health monitoring, strategies to manage stress, exercise instructions, and a daily exercise and medication intake record chart (Multimedia Appendix 2).

Assistive Devices

The research team provided assistive devices to improve ADL (eg, spoon holder, bottle opener, pen holder, and cup/mug holder) using a 3D printer and physical improvement–aiding instruments (eg, special adjustable spoon, walking cane, exercise ball, Thera belt, and massage ball) to help patients with poststroke disabilities with self-care. The 3D printers were used to produce the assistive devices, and these were designed and developed by an occupational therapist and adjusted according to the needs of patients with disabilities.

For the Intervention Group

After assessment, patients and caregivers were trained (for 30 minutes during their discharge from the hospital) in self-care with stroke disabilities (eg, eating, grooming, bathing, dressing, and toileting), physical exercise, lifestyle modification [28,29], and prevention of complications after stroke through a researcher-developed "Health booklet" by the RA nurses. The booklet was provided to patients to follow at home for daily practice and a record of exercise with medication intake. Based on their disabilities, adaptive and exercise devices were provided for use at home to promote functional independence. After discharge from the hospital, they were followed by RA nurses through home visits and teleservices biweekly for 3 months (a total of 5 times). The telehealth services include conversations with patients and caregivers and follow-up on adherence to the use of assistive devices provided by RA nurses. These services follow rehabilitation guidelines and health booklet-guided exercises for improving muscle strength. If needed, guidance on the use of assistive devices was also provided through video calls. During these follow-up visits, they assessed patients' regular self-care activities, adjustment to devices, exercise and adaptive behaviors. During the final follow-up, after 3 months of the initial assessment, researchers and RAs evaluated patient outcomes at the hospital. If the patient was unable to attend the hospital, a home visit was made for follow-up.

For the CG

The patients received the usual care from the RA nurses and physician to which they were entitled. They received hospital advice on treatment adherence for stroke prevention during discharge. During the 3 months, patients were requested to visit the hospital for follow-up visits and data collection. We were provided with a health booklet for all and assistive devices who needed after the final evaluation for ethical purposes.

Preparation for Data Collection

Making a Multidisciplinary Team

Upon starting the study, researchers formed a multidisciplinary team consisting of a physician, RA nurses, a physiotherapist, and an occupational therapist under the supervision of the co–principal investigator (neurologist). They collaborated to assess patients from recruitment through final follow-up for evaluation.

Training the Study Nurses

A total of 7 RA nurses, who were involved in this study, received 1 week of training through lectures and practice on self-care, physical exercise, mobility, and telenursing services by the principal investigator based on the health booklet. They also received protocol-specific hands-on training for data collection, performing study procedures about self-care, behavioral changes, including adaptation and motivation, and assembling assistive devices for patients with disabilities using a 3D printer. They had regular refresher sessions about the enlisted activities with the principal investigator before and during the study period.

Data Collection Procedures

After allocating study participants to either the intervention group (IG) or the CG, the RA nurses conducted a face-to-face interview to collect data about the sociodemographic characteristics and health conditions of stroke survivors and their family caregivers. At baseline (T0), the functional independence measurement, self-efficacy, social participation, and care burden of the family were assessed. The RA nurses periodically followed up with the patients of the IG through home visits (twice a month) and teleservice (twice a month) to check their motivation and cultural values to adjust to the intervention devices and health booklet.

In the community, the nurse visited the home setting to observe the home environment and adjust to patients' needs. If required, they suggested modifications to home settings. Moreover, they assessed the way of implementing assistive devices for performing ADL and rehabilitation for improving muscle strength. Nurses demonstrated the use of assistive devices if needed. They also demonstrated the exercise procedure using the health booklet and motivated patients and family caregivers to maintain the daily record book.

To prevent participant contamination, RA nurses obtained consent and collected data in 2 different rooms. They also maintained different schedules for the IG and CG. As the same

XSL•FO
RenderX

RA nurses collected data from both groups, they strictly maintained confidentiality to share the intervention procedures with the CG. For home-based data collection, RA nurses were divided into the IG and CG.

To evaluate the feasibility and chronological changes of adjustment of home-rehabilitation and assistive devices qualitatively, RA nurses interviewed participants through some open-ended questions and observed participants of the IG every time they contacted them. The interview was conducted for 20-30 minutes by RA nurses based on observation points and interview questions during home-based follow-up. The observation and interview points were as follows: "home environment, adjustment, motivation and use of assistive devices, lifestyle changes pattern with risk factors, level of required assistance after using assistive devices, adherence to rehabilitation regimen, mental status, communication, and social participation" after having a stroke (Multimedia Appendix 3).

Outcome Measures

Primary End Point: Functional Independence

Functional independence was assessed using the FIM scale. The FIM was developed in 1983 by a task force from the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation [30]. The Bengali version of the scale was implemented among the patients with functional disabilities in Bangladesh and found to be valid and reliable (α=.97) [31]. This scale has 18 items from level 1 (total assistance) to level 7 (complete independence). This scale consists of 2 subscales, including physical (13 items) and cognitive (5 items) domains. The physical domain score ranges from 13 to 91, and the cognitive domain score ranges from 5 to 35. The range of the total score is from 18 to 126. The higher the score, the better the level of independence.

Secondary End Points

Self-Efficacy

Self-efficacy was assessed by the Generalized Self-Efficacy Scale [32]. This scale assessed patients' self-efficacy after experiencing poststroke stress and the strength of an individual's beliefs in situations. It was translated into the Bengali version and was found valid and reliable (α=.7 to .91) [33]. It has a total of 10 items with a 4-point Likert scale that ranges from 1=not at all true to 4=exactly true. The lowest score is 10, and the highest score is 40. The higher score indicates better self-efficacy.

Social Participation

Social participation was measured using the Participation Scale based on the WHO *ICF* domain for assessing a person's interaction and participation in community life, which measures the severity of social participation restrictions. They standardized the Bengali version of the scale and found it to be valid and reliable (α=.91) [34]. The score ranges from 0 to 70. The higher the score, the worse the level of participation.

Caregiver Burden

Care burden was assessed by the Zarit Caregiver Burden Assessment (12-item) scale, short version, to assess the subjective burden of informal caregivers of patients with chronic diseases [35]. The Bengali version of the scale was implemented among patients with stroke in Bangladesh and was found valid and reliable (α=.85) [36]. The score ranges from 0 to 48. No cutoff point was assigned for classifying the care burden; a higher score indicated a higher care burden.

Qualitative Observations

We qualitatively analyzed the assistive device-use rate, motivation to use, and unmet needs for physical improvement among participants who received our intervention (IG).

Analysis

To compare data between the 2 groups at different time points, we conducted a per-protocol set analysis to see the actual efficacy of the intervention. For baseline data, descriptive statistics were used after the normality test. The normal data distribution was examined using Shapiro-Wilk's test, histograms, box plots, and normal Q-Q plots for both the control and IGs, with $P>.05$. We used chi-square or the Fisher exact test to compare the IG and the CG using categorical variables, which were reported as frequencies and percentages. Since our data were not normally distributed, we performed the Mann-Whitney *U* test to report the continuous variables as the mean (SD). The mean differences between the 2 groups were compared by an independent *t* test at each time point (T0 and T1). In addition, to assess changes in outcomes within the groups over time, a 2-way repeated-measure ANOVA was used. The level of significance was set at $P<.05$. We conducted generalized linear models considering covariates such as age, residence, and mRS to examine the influence of additional rehabilitation activities (home-based and CRP). We used SPSS (version 31.0; IBM Corp). For the qualitative analysis, after transcribing and translating, we converged and coded the statements. Then we used Dedoose software (version 10.0.59; Socio Cultural Research Consultants, LLC), and the results were categorized into themes and subthemes.

Ethical Considerations

This study was conducted following the ethical principles of the Declaration of Helsinki after getting ethical approval from the institutional review board of NINS&H (approval number IRB/NINS/2024/392), Dhaka, Bangladesh. After an explanation of the study objective, procedure, risk, and benefit, written consent was obtained from participants and their caregivers. Patients' privacy and confidentiality were maintained throughout the study, and withdrawal was ensured without any consequences. Data were deidentified and assigned unique identification numbers for analysis, and no personal information was used in any publication or presentation resulting from this research. Data were stored and accessible by the principal investigator on a password-protected personal computer.

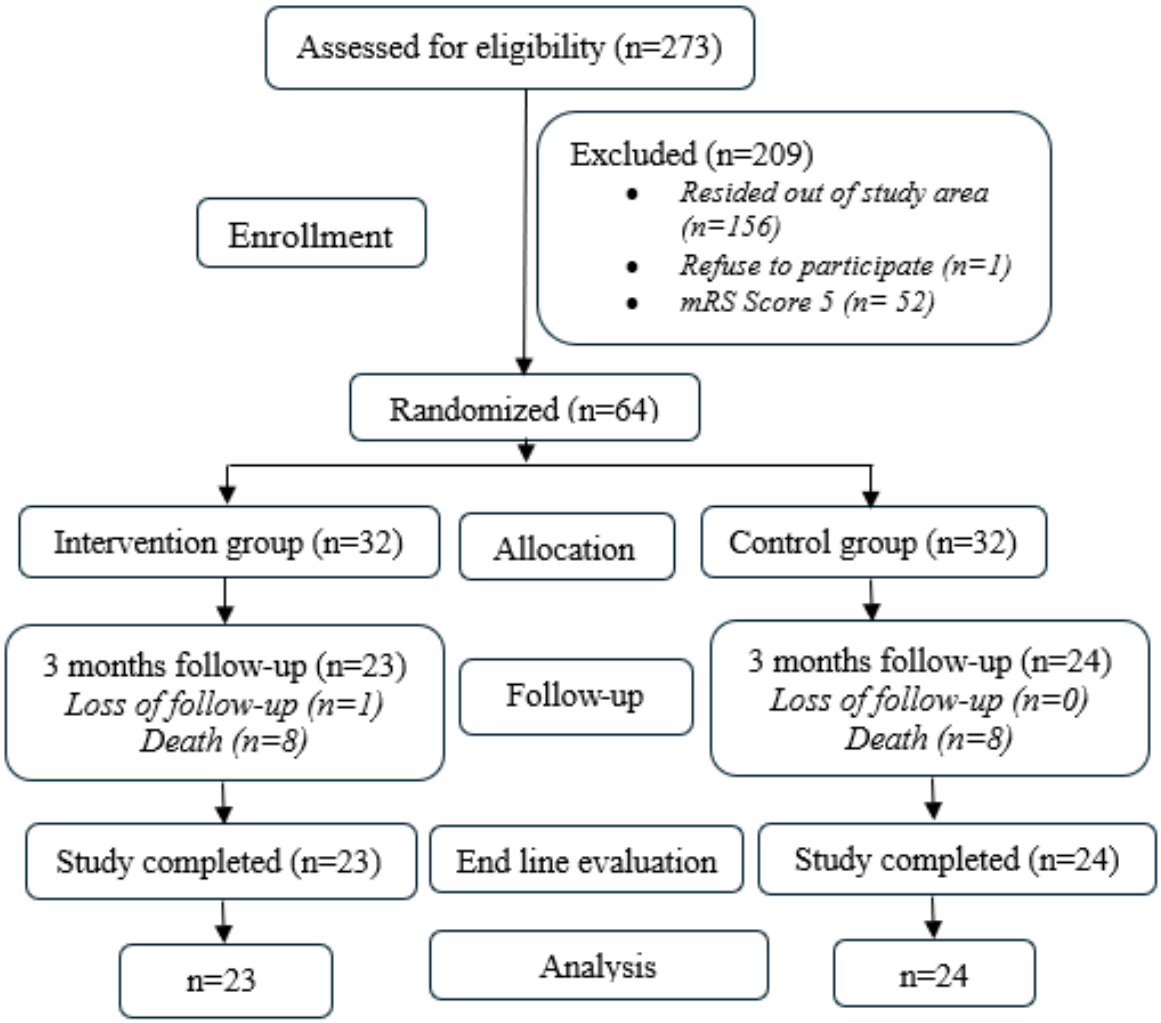
XSL•FO
RenderX

Results

Overview

A total of 273 participants were evaluated to determine their eligibility for the study. Among them, 209 of 273 did not meet the inclusion criteria; 156 of 273 resided out of the study area, 528 of 273 had an mRS score of 5, and 1/273 refused to participate in this study. Therefore, a total of 64 patients were enrolled equally in both groups (IG: n=32 and CG: n=32); 47 (73%) completed (IG: n=23 and CG: n=24) the study, and were analyzed. Among the participants who did not complete the study (n=17), 16/64 (25%) died during the first month of intervention (IG: n=8 and CG: n=8), and 1/64 (1.5%) was lost to follow-up (Figure 2).

Figure 2. Study flowchart based on the CONSORT (Consolidated Standards of Reporting Trials) statement. mRS: modified Rankin scale.



Sociodemographic and Clinical Characteristics of the Participants

The mean age of participants was 55.6 (SD 14.4) years, with a range of 20-82 years. Most of the participants were male (n=52, 81%) and resided in urban areas (n=55, 86%). A higher proportion of rural residents was noted in the CG (n=7, 22% vs n=2, 6%) compared to the IG.

Overall, no significant differences were observed between the IG and CG across sociodemographic and clinical characteristics and for comorbidity, indicating successful randomization (Table 1).

Preintervention FIM scores were compared with sociodemographic and clinical characteristics in Multimedia Appendix 4. The participants' age, religious status, residence, tobacco smoking, and mRS were significantly associated with FIM (all $P<.05$).

XSL•FO
RenderX

Table 1. Sociodemographic and clinical characteristics of the participants at baseline (N=64).

Variable	Total (N=64)	Intervention group (n=32)	Control group (n=32)	*P* value
Age (years), mean (SD; range)	55.6 (14.4; 20-82)	56.3 (15.6; 20-82)	54.9 (12.8; 20-82)	.44[a]
Gender, n (%)				.20[b]
Male	52 (81.3)	28 (87.5)	24 (75)	
Female	12 (18.3)	4 (12.5)	8 (25)	
Employment status, n (%)				.79[b]
Employed	41 (64.1)	20 (62.5)	21 (65.6)	
Unemployed	23 (35.9)	12 (37.5)	11 (34.4)	
Monthly family income (BDT; US $1=BDT 123), mean (SD; range)	N/A[c]	33,250 (24,352.7; 5000-200,000)	33,750 (36,961.2; 5000-200,000)	.49
Education status, n (%)				>.99[d]
Primary (5 years of schooling)	22 (34.4)	11 (34.4)	11 (34.4)	
Secondary (10 years of schooling)	26 (40.6)	13 (40.6)	13 (40.6)	
College and above (>10 years of schooling)	16 (25)	8 (25)	8 (25)	
Religion, n (%)				.30
Muslim	60 (93.8)	31 (96.9)	29 (90.6)	
Hindu	4 (6.3)	1 (3.1)	3 (9.4)	
Residence, n (%)				.07
Urban	55 (85.9)	30 (93.8)	25 (78.1)	
Rural	9 (14.1)	2 (6.3)	7 (21.9)	
BMI, n (%)				.62[d]
Normal weight	25 (39.1)	13 (40.6)	12 (37.5)	
Overweight	35 (54.7)	16 (50)	19 (59.4)	
Obese	4 (6.3)	3 (9.4)	1 (3.1)	
Modified Rankin Scale, n (%)				.44[d]
2	21 (32.8)	8 (25)	13 (40.6)	
3	32 (50)	18 (56.3)	14 (43.8)	
4	11 (17.2)	6 (18.8)	5 (15.6)	
Smoking, n (%)				0.30
Yes	24 (37.5)	10 (31.3)	14 (43.8)	
No	40 (62.5)	22 (68.7)	18 (56.2)	
Stroke type, n (%)				.06
Ischemic	43 (67.2)	25 (78.1)	18 (56.3)	
Hemorrhagic	21 (32.8)	7 (21.9)	14 (43.8)	
Time of stroke, n (%)				.18
First stroke	43 (67.2)	19 (59.4)	24 (75)	
Recurrent stroke	21 (32.8)	13 (40.6)	8 (25)	
Comorbidities, n (%)				
Hypertension	60 (93.8)	30 (93.8)	30 (93.8)	>.99
Diabetes mellitus	50 (78.1)	25 (78.1)	25 (78.1)	>.99
Arrhythmia	19 (29.7)	12 (37.5)	7 (21.9)	.17
Chronic kidney disease	1 (1.6)	0 (0)	1 (3.1)	.31

XSL•FO
RenderX

Variable	Total (N=64)	Intervention group (n=32)	Control group (n=32)	*P* value
Obesity	4 (6.3)	3 (9.4)	1 (3.1)	.30
Others[e]	10 (15.6)	4 (12.5)	6 (18.8)	.49

[a]Mann-Whitney *U* test.

[b]Chi-square test.

[c]N/A: not applicable.

[d]Fisher exact test.

[e]Others: such as meningitis, uterine tumor, and breast cancer.

Outcomes

FIM

At baseline (T0), the mean FIM scores were 43.78 (SD 15.94) for IG and 43.79 (SD 14.10) for CG, with no significant difference (*P*>.99). FIM scores improved chronologically in both IG and CG, showing statistically significant improvements within groups; however, those changes were not statistically significant between groups or in the interaction (Table 2).

Table 2. Comparison of study outcomes between intervention and control groups at 4 time points.

	Baseline (T0), mean (SD)	After 1 month (T1), mean (SD)	After 2 months (T2), mean (SD)	After 3 months (T3), mean (SD)	*F* test (*df*)	*P* value
Functional Independence Measure[a]					37.23 (1)[b]	<.001[b]
Intervention group	43.8 (15.9)	64.4 (24.7)	73.5 (33.5)	78.4 (37.6)		
Control group	43.8 (14.1)	64.3 (30.4)	67.7 (31.8)	74.3 (35.4)		
Self-efficacy[c]					0.046 (5)[d]	.83[d]
Intervention group	23.9 (3.3)	N/A[e]	N/A	27.0 (6.9)		
Control group	24.1 (4.6)	N/A	N/A	26.58 (7.14)		
Social participation[c]					0.272 (6)[d]	.61[d]
Intervention group	39.1 (13.2)	N/A	N/A	30.4 (21.4)		
Control group	36.2 (13.2)	N/A	N/A	28.2 (17.2)		
Care burden[c]					0.046 (5)[d]	.83[d]
Intervention group	31.4 (4.6)	N/A	N/A	24.9 (13.6)		
Control group	29.9 (6.4)	N/A	N/A	25.2 (12.6)		

[a]Functional Independence Measure was analyzed using a 2-way repeated-measures ANOVA. Interaction effect: F_3=0.386; *P*=.60; between-group effect: F_1=0.113; *P*=.74; within-group effect: F_1=37.23; *P*<.001.

[b]2-way repeated measure ANOVA.

[c]Independent *t* test was used to compare groups.

[d]2-way ANOVA.

[e]N/A: not applicable.

Self-Efficacy

At baseline (T0), self-efficacy scores of both groups showed a mean of 23.95 (SD 3.33) for IG and 24.12 (SD 4.59) for CG (*P*=.89). After 3 months (T3), IG showed a higher mean (IG: mean 27.04, SD 6.95; CG: mean 26.58, SD 7.14) than CG; however, the difference was not significant (*P*=.82; Table 2).

Social Participation

At baseline (T0) and end line (T3), the mean restriction in social participation was 39.13 (SD 13.17) and 30.35 (SD 21.44), respectively, for the IG and 36.17 (SD 13.18) and 28.21 (SD 17.19), respectively, for the CG. However, the differences were not significant (*P*=.45 and *P*=.71) at both points (Table 2).

Caregiver Burden

The means of caregiver burden were 31.35 (SD 4.5) for IG and 29.88 (SD 6.39) for CG (*P*=.37) at baseline. At end line (T3), the means were 24.91 (SD 13.58) for IG and 25.21 (SD 12.60) for CG (*P*=.94), and no significant differences were found (Table 2).

FIM Subdomain Scores

FIM domain scores were analyzed at baseline (T0) and end line (T3) between the intervention and CGs. There were no statistically significant differences between the groups at both time points in the following categories: self-care, sphincter control, mobility, locomotion, communication, social cognition, and overall FIM (all *P*>.05; Multimedia Appendix 5).

XSL•FO
RenderX

Regarding FIM scores, the CG demonstrated improvements comparable to those of the IG. To explore the reasons, we conducted further analyses and found that participants in both groups engaged in private rehabilitation services in addition to our intervention. As this was not part of the exclusion criteria, and for ethical reasons, we could not restrict their access to such services. We therefore examined how these additional rehabilitation activities influenced our results. Table 3 presents the FIM score changes by service in each group. The changes were inconsistent and showed no clear relationship with the services received.

To examine which additional rehabilitation services affected the FIM score, we conducted a general linear model with age, mRS, and residence as covariates and found no statistically significant effects**.**

Compared with the CG, participants in the IG showed no significant difference in outcomes (*B*=2.48, SE=10.04, 95% CI –17.21 to 22.16; Wald χ^2_1=0.06; *P*=.81). Similarly, home-based rehabilitation did not significantly predict the outcome relative to the research intervention (*B*=–2.64, SE=10.25, 95% CI –22.72 to 17.44; Wald χ^2_1=0.07; *P*=.80). CRP-based rehabilitation also did not differ significantly from the research intervention condition (*B*=17.22, SE=16.80, 95% CI –15.70 to 50.15; Wald χ^2_1=1.05; *P*=.31; Table 4).

Table 3. Comparison of rehabilitation services from baseline to end line (n=47).

	Partici-pants, n (%)	Baseline (T0)			After 3 months (T3)			Difference (T3 – T0)		
		mean (SD)	*t* test (*df*)	*P* value	mean (SD)	*t* test (*df*)	*P* value	mean (SD)	*t* test (*df*)	*P* value
FIM[a] score (n=47)			–0.002 (45)	>.99		0.388 (45)	.70		0.423 (45)	.67
Intervention group	23 (100)	43.8 (15.9)			78.4 (37.6)			34.6 (34.1)		
Control group	24 (100)	43.8 (14.1)			74.3 (35.4)			30.5 (33.2)		
Study intervention + home-based rehabilitation (n=22)			0.733 (45)	.24		0.174 (45)	.87		–0.189 (45)	.85
Intervention group	8 (34.8)	50.0 (18.8)			79.5 (37.9)			29.5 (32.7)		
Control group	14 (58.3)	44.4 (14.6)			76.6 (35.6)			32.3 (34.4)		
Study intervention + CRP[b] /facility-based (n=5)			0.284 (45)	.80		0.197 (45)	.86		0.089 (45)	.94
Intervention group	4 (17.4)	39.0 (18.9)			83.0 (40.8)			44.0 (30.1)		
Control group	1 (4.2)	33 (N/A[c])			74 (N/A)			41.0 (N/A)		
Study intervention only[d] (n=20)			–0.508 (45)	.62		0.303 (45)	.77		0.517 (45)	.61
Intervention group	11 (47.8)	41.0 (12.5)			75.9 (39.9)			34.9 (38.5)		
Control group	9 (37.5)	44.1 (14.5)			70.6 (38.9)			26.4 (34.6)		

[a]FIM: Functional Independence Measure.

[b]CRP: Centre for the Rehabilitation of the Paralyzed.

[c]N/A: not applicable.

[d]Numbers of “study intervention only” participants were those who did not receive either home-based rehabilitation or CRP/facility-based rehabilitation.

XSL•FO
RenderX

Table 4. Parameter estimates for group and rehabilitation type predicting the outcome variable.

Variable and category	*B* (SE)	95% CI	Wald chi-square (*df*)	*P* value
Group (reference: control group)				
Intervention group	2.48 (10.04)	–17.21 to 22.16	0.061 (1)	.81
Rehabilitation type (reference: standard research intervention)				
Home-based rehabilitation	–2.64 (10.25)	–22.72 to 17.44	0.066 (1)	.80
CRP[a] rehabilitation	17.22 (16.80)	–15.70 to 50.15	1.05 (1)	.31

[a]CRP: Centre for the Rehabilitation of the Paralyzed.

Use of Assistive Devices

Participants in the IG were significantly used in assistive devices (*P*<.001), with 21 (91%) in the IG compared to only 4 (17%) in the CG. These findings indicate that the intervention was associated with significantly greater adoption of assistive devices. In CG, participants purchased assistive devices (exercise ball and TheraBand) with the advice of their physical therapist or outside rehabilitation services.

Extremity Impairments

All (n=47) participants had extremity impairments, and 94% (44/47) had both upper and lower extremity impairments, and 6% (n=3) had only upper extremity impairments. There was no significant difference observed between groups regarding extremity impairment (*P*>.99).

Qualitative Findings

Overview

Among all the participants (n=32) who received the intervention, 23 were interviewed and observed for qualitative assessment. The qualitative data revealed 7 subthemes encompassing three key themes: (1) patients' adherence to rehabilitation, (2) improvement of physical functioning and independence, and (3) perceived improvement by patients and family caregivers. Each theme demonstrated positive transformation between the first month and the third month, supported by participants' perceptions and direct quotes (Multimedia Appendix 6).

Theme 1: Patients' Adherence to Rehabilitation

Acceptance and Motivation to Use Assistive Devices

During the first month, participants were in the process of adapting to assistive devices and integrating them into their daily life activities. By the end of the third month, most had accepted and used these devices independently without any assistance from caregivers, despite emotional and financial challenges.

> *My family spends more on me than they earn* [Male, 48 years, Case #2]

> *These devices feel like blessings, but sometimes I lose hope* [Female, 60 years, Case #6]

Following the Exercise Regimen

Participants initially depended on caregivers to carry out suggested exercises; however, they consistently followed rehabilitation guidelines independently (at least 2 times per day for 30 minutes) in later times. This shift illustrates increased motivation and adherence to recommended regimens of rehabilitation.

Theme 2: Improvement of Physical Functioning and Independence

Mobility and Balance

At baseline, participants required assistance for walking and balance from caregivers. After intervention, most had the ability to ambulate with minimal assistance or modified independence using assistive devices by their own efforts.

> *Nowadays I'm feeling good. I'm going out with a walking cane and walking on my own. My family is relieved* [Male, 20 years, Case #14]

This quote demonstrates regained confidence and functional ability.

ADL

Dependence on daily tasks reduced progressively, from total or moderate dependence to moderate or minimal levels. Initially, they used assistive devices to perform their ADL. Gradually, after gaining the ability, they reduced their use of assistive devices.

> *I do everything by myself (with the help of assistive devices), whatever I can* [Male, 65 years, Case #7]

> *Now I can do my daily activities like eating and going to the toilet on my own. I just need a little help, but not always* [Male, 48 years, Case #22]

Comorbidities and Risk Factors

Participants commonly experienced comorbidities such as diabetes, hypertension, and visual disturbances. While persistent, most reported reduced symptom severity over time.

Participants received health education to combat comorbidities. One participant with diabetes and hypertension felt better after receiving our intervention. He explained:

> *I'm seeing blurry, specially at night... headaches appear when I go to sleep* [Male, 20 years, Case #14]

Theme 3: Perceived Improvement by Patients and Family Caregivers

Physical Improvement

Participants' perceptions shifted from no or moderate improvement to substantial progress and functional independence using assistive devices. Another participant expressed:

XSL•FO
RenderX

> *I feel like I am improving day by day... this progress makes me feel very happy and hopeful* [Male, 48 years, Case #2]

Social Interaction and Participation

Initially, participants were socially withdrawn; they actively reengaged in community and family activities. After our intervention, they engaged more in social gatherings.

> *I go out and meet people. Being active has improved my condition* [Male, 65 years, Case #21]

This reflects enhanced self-efficacy and psychosocial well-being.

Discussion

Principal Findings

This study evaluated the feasibility of a nurse-led rehabilitation education with an assistive device to improve functional independence for self-care management, self-efficacy, social engagement, and reduction of caregiver burden among patients with poststroke disabilities. Over the 3-month study period, there was improvement after intervention; however, no significant differences were observed in any of the outcomes between the intervention and CGs.

Sociodemographic Characteristics

The participants' average age was 50 (range 20-82) years, which is consistent with the demographics of patients with stroke in LMIC, where strokes typically occur in early life compared to high-income countries [2]. Among the participants, male individuals were higher, which indicates that stroke is more common in men worldwide [37]. Therefore, it is essential to explore the underlying factors and comorbidities prevalent among male patients after stroke in that age group.

The majority of individuals came from families with moderate income or were employed, and educational attainment was consistent between groups. According to previous studies, socioeconomic and educational characteristics have an impact on stroke outcomes because decreased likelihood of receiving rehabilitation services and a worse long-term prognosis are associated with lower income and educational attainment [38].

The most prevalent comorbidities were diabetes mellitus and hypertension, which were found in almost a third of both groups, which are recognized risk factor profiles for stroke in Southeast Asian populations [39]. In the IG, we observed more ischemic strokes compared to the CG. As the prognosis of ischemic stroke is limited after intervention, this might affect our intervention, making it nonsignificant. Moreover, compared with ischemic strokes, hemorrhagic strokes are typically associated with higher early morbidity and mortality, resulting in more deaths, as reported in a previous study [40]. Despite persistent comorbidities, participants described adaptation and symptom control over time. This indicates resilience and coping strategies that help patients after stroke to survive with multimorbidity [7,41].

Functional Independence

We did not find any difference in FIM scores over time between the groups. However, if we observe the trend, the FIM scores improved at each time point for both groups. In the IG, we assumed that the FIM scores were increased due to our intervention. We explored the cause of the improvement in FIM scores in the CG and found that they received rehabilitation services from external sources (home-based and CRP), which we could not inhibit due to ethical considerations. CRP provides rehabilitation services with a very highly trained professionals using all conventional therapies (eg, physiotherapy, occupational, speech, and language) in our study areas. Therefore, we can consider our intervention to be effective in 1 arm. This study emphasizes the importance of adequate rehabilitation intervention for patients with poststroke disabilities. Other studies also explained that poststroke disability can be improved after health education and the use of assistive devices [42]. Though our short (3 months) study duration shows FIM improvement in individual groups, one study showed no significant differences in FIM evolution when compared to usual care alone during the short period of time (<10 weeks) [43]. Improvements in mobility, balance, and ADL performance reflect global evidence that structured poststroke rehabilitation fosters early recovery within 3-6 months [44]. The progression from dependence to modified independence, as reported by participants, reinforces the importance of consistent physiotherapy and guided home exercises [9,45]. Our study findings highlight the significance of comprehensive standard care following a poststroke disability.

Self-Efficacy, Social Participation, and Caregiver Burden

With respect to self-efficacy, social participation, and caregiver burden, we did not find any difference between the groups over time. To improve the self-efficacy of patients with stroke after discharge, we provided organized interaction and instruction that helped patients feel more confident and develop the skills they need to manage the aftermath of stroke. As in the CG, some of the patients received group interaction, education on health care and rehabilitation, and a community reintegration program from the CRP; therefore, we found improvement of self-efficacy in both groups. Recent research supports the importance of fostering self-efficacy through community services and standard care in patients, along with natural psychological adaptation [44].

The improvement of social participation and the decrease of caregiver burden were also observed in the IG and in the CG by the influence of CRP services. Despite prominent functional improvements, many patients after stroke report limitations in their ability to participate in the community, which raises concern [46]. Caregivers gradually adjust and become more confident as in the poststroke period, patients regain functioning ability, and families gain experience over time [47]. However, the persistently high levels of burden reported in recent surveys suggest that more focused and direct interventions, such as social support and behavioral education programs, may be effective [47].

The qualitative findings demonstrate how structured therapy promotes progressive recovery of patients after stroke, demonstrating significant gains in functional independence, adherence, and social reintegration. Additionally, patient

XSL•FO
RenderX

narratives enhance comprehension by clarifying contextual and personal factors influencing rehabilitation involvement.

Over time, patients demonstrated more motivation and continued use of assistive devices, which is consistent with earlier studies that focused on gradual adaptation to rehabilitation services [44,48]. However, participants acknowledged that financial and caregiving burdens highlight socioeconomic challenges common in low-resource settings [17,42]. These results are consistent with research indicating that long-term rehabilitation success depends critically on patient motivation and family support [18].

This study shows that nurse-led rehabilitation enhances psychological well-being and quality of life through perceived physical improvement and increased social engagement [44]. Self-efficacy's involvement in fostering recovery is supported by the expressed sense of hope and autonomy [21].

In resource-limited environments, barriers such as inadequate infrastructure, caregiver burden, and financial restraints reduce sustained engagement in rehabilitation [17,49]. These insights highlight the necessity of nurse-led, family-centered, community-based rehabilitation programs that tackle socioeconomic and clinical challenges [21].

Overall, the results show that, even in low-income settings, structured rehabilitation helps stroke survivors recover both physically and psychologically. To improve long-term adherence and social reintegration, future interventions should include systems of social support and monetary assistance.

Strengths and Limitations

This is the first nurse-led randomized controlled trial in Bangladesh to improve patients' ADL with disability by providing rehabilitation intervention. This study represents a proper allocation of participants through successful randomization. We provided assistive devices, education booklets, and follow-up through home visits. We carefully assessed patients and provided the appropriate assistive devices, ensuring that their usage was accurately measured. As this is a feasibility study, we observed that both the device and nurse education were well accepted by patients and family caregivers, indicating that this program could be feasibly implemented in the Bangladesh context.

Nevertheless, limitations must be acknowledged. High early mortality and attrition reflect the acute severity of illness in the study population. It causes underpowering of our study sample size. We conducted our study for a short period (3 months) of time with a small sample, which made it very difficult to interpret the findings. The study was conducted only in urban areas, which did not reflect the scenario of the whole country, which may limit generalizability. Further studies with multiple sites are needed to prevent participant contamination. More accurate measurements, rather than FIM, should be implemented to assess the influencing factors for self-care management. We also need to reconsider the mRS level (mRS-4) as there is a more severe disability that limits physical improvements.

Conclusions

In conclusion, this is the first nurse-led pilot randomized controlled trial aimed at improving the functional abilities of patients with poststroke disabilities in Bangladesh. However, our rehabilitative intervention (3 months) showed no impact on FIM, self-efficacy, social participation, and caregivers' burden for patients after stroke; obvious improvements were observed at each time point (baseline, monthly, and end line) of evaluation for both IG and CG. We assumed that this improvement occurred in the IG due to our intervention, and for the CG, improvement was observed due to the influence of rehabilitation received from home-based and institutional services. The result suggested that a comprehensive rehabilitation intervention is essential for the physical improvement of patients with poststroke disabilities. A future long-term interventional study with a large sample size will be guaranteed for better outcomes, considering all external influences to understand the underlying factors for the improvement of rehabilitative patients after stroke. Strengthening primary care facilities can be a strategic way for the community integration of patients with stroke disability. A government-led, systemic approach is essential to translate policy commitments into equitable access to rehabilitation services for all patients after stroke.

Acknowledgments

I would like to express my sincere gratitude to all the patients and their caregivers for their generous participation in this study. We acknowledge the enormous support from the medical staff at the National Institute of Neurosciences & Hospital (NINS&H) and research assistants for their tremendous effort. Special thanks to the Director of NINS&H for his excellent support of this study. We would like to express our gratitude to Mr Wataru Iwamori, OTR, Ai Reha Home-Visit Nursing Station, Ai Reha Co, Ltd, Hiroshima, Japan, for his guidance on staff training and the development of educational materials for stroke rehabilitation.

Funding

This project was funded by the Japan Society for the Promotion of Science (KAKENHI), Japan (25K24189).

Data Availability

The data supporting the findings of this study are available on request from the corresponding authors. The data are not publicly available due to privacy and ethical restrictions.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: SA, MM
Methodology: SA, MM
Formal analysis: SA
Investigation: SA, MM
Data curation: SA, MI, FKC, SKJBS, MMI, MA, SD
Project administration: SA, MI, FKC, SKJBS, MMI, MA, SD
Supervision: SA, MI, FKC, SKJBS, MMI, MA, SD, MM
Visualization: SA, KATMEH, MM
Funding acquisition: MM
Writing—original draft: SA
Writing—review and editing: KATMEH, MM
All authors reviewed and approved the final manuscript for publication and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT eHEALTH checklist.
[PDF File (Adobe PDF File), 1061 KB - rehab_v13i1e88808_app1.pdf]

Multimedia Appendix 2
Contents of health booklet.
[DOCX File , 17 KB - rehab_v13i1e88808_app2.docx]

Multimedia Appendix 3
Qualitative interview guide.
[DOCX File , 16 KB - rehab_v13i1e88808_app3.docx]

Multimedia Appendix 4
Correlation of sociodemographic and preintervention Functional Independence Measure.
[DOCX File , 18 KB - rehab_v13i1e88808_app4.docx]

Multimedia Appendix 5
Functional Independence Measure subdomain analysis at baseline and endline between groups.
[DOCX File , 28 KB - rehab_v13i1e88808_app5.docx]

Multimedia Appendix 6
Qualitative content analysis.
[DOCX File , 17 KB - rehab_v13i1e88808_app6.docx]

References

1. Feigin VL, Norrving B, Mensah GA. Global burden of stroke. Circ Res 2017;120(3):439-448 [FREE Full text] [doi: 10.1161/CIRCRESAHA.116.308413] [Medline: 28154096]
2. Feigin VL, Brainin M, Norrving B, Martins SO, Pandian J, Lindsay P, et al. World Stroke Organization: Global Stroke Fact Sheet 2025. Int J Stroke 2025;20(2):132-144 [FREE Full text] [doi: 10.1177/17474930241308142] [Medline: 39635884]
3. Aziz M, Bipasha N, Gupta U, Ramnarine IVP, Redgrave J, Ali AN, et al. Stroke in Bangladesh: a narrative review of epidemiology, risk factors and acute stroke services. J Cardiovasc Dev Dis 2025;12(2):58 [FREE Full text] [doi: 10.3390/jcdd12020058] [Medline: 39997493]
4. Cheng Y, Lin Y, Shi H, Cheng M, Zhang B, Liu X, et al. Projections of the stroke burden at the global, regional, and national levels up to 2050 based on the Global Burden of Disease Study 2021. J Am Heart Assoc 2024;13(23):e036142 [FREE Full text] [doi: 10.1161/JAHA.124.036142] [Medline: 39575720]
5. Mondal MBA, Hasan ATMH, Khan N, Mohammad QD. Prevalence and risk factors of stroke in Bangladesh: a nationwide population-based survey. eNeurologicalSci 2022;28:100414 [FREE Full text] [doi: 10.1016/j.ensci.2022.100414] [Medline: 35769921]
6. Shuvo TA, Hosna AU, Hossain K, Hossain S. Prevalence of stroke in Bangladesh: a systematic review and meta-analysis. J Stroke Cerebrovasc Dis 2024;33(12):108017. [doi: 10.1016/j.jstrokecerebrovasdis.2024.108017] [Medline: 39306056]

XSL•FO
RenderX

7. Rahman M, Rana MS, Rahman MM, Khan MN. Healthcare services access challenges and determinants among persons with disabilities in Bangladesh. Sci Rep 2024;14(1):19187 [FREE Full text] [doi: 10.1038/s41598-024-70418-2] [Medline: 39160270]
8. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
9. Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, American Heart Association Stroke Council, Council on CardiovascularStroke Nursing, Council on Clinical Cardiology,Council on Quality of CareOutcomes Research. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016;47(6):e98-e169. [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
10. Brady MC, Kelly H, Godwin J, Enderby P, Campbell P. Speech and language therapy for aphasia following stroke. Cochrane Database Syst Rev 2016;2016(6):CD000425 [FREE Full text] [doi: 10.1002/14651858.CD000425.pub4] [Medline: 27245310]
11. Cicerone KD, Goldin Y, Ganci K, Rosenbaum A, Wethe JV, Langenbahn DM, et al. Evidence-based cognitive rehabilitation: systematic review of the literature from 2009 through 2014. Arch Phys Med Rehabil 2019;100(8):1515-1533 [FREE Full text] [doi: 10.1016/j.apmr.2019.02.011] [Medline: 30926291]
12. Towards a common language for functioning, disabilityhealth: ICF. World Health Organization. 2002. URL: https://www.who.int/classifications/icf/icfbeginnersguide.pdf [accessed 2026-07-04]
13. Khan MA, Saibene M, Das R, Brunner I, Puthusserypady S. Emergence of flexible technology in developing advanced systems for post-stroke rehabilitation: a comprehensive review. J Neural Eng 2021;18(6). [doi: 10.1088/1741-2552/ac36aa] [Medline: 34736239]
14. Ifejika NL, Awosika OO, Black T, Duncan PW, Harvey RL, Katz DI, et al. Improving access to stroke rehabilitation and recovery: a policy statement from the American Heart Association/American Stroke Association. Stroke 2025;56(9):e218-e233 [FREE Full text] [doi: 10.1161/STR.0000000000000493] [Medline: 40740119]
15. Waterworth CJ, Smith F, Kiefel-Johnson F, Pryor W, Marella M. Integration of rehabilitation services in primary, secondary, and tertiary levels of health care systems in low- and middle-income countries: a scoping review. Disabil Rehabil 2024;46(25):5965-5976 [FREE Full text] [doi: 10.1080/09638288.2024.2317422] [Medline: 38376099]
16. Hashemi G, Wickenden M, Bright T, Kuper H. Barriers to accessing primary healthcare services for people with disabilities in low and middle-income countries, a meta-synthesis of qualitative studies. Disabil Rehabil 2022;44(8):1207-1220 [FREE Full text] [doi: 10.1080/09638288.2020.1817984] [Medline: 32956610]
17. Al Imam MH, Jahan I, Das MC, Muhit M, Akbar D, Badawi N, et al. Situation analysis of rehabilitation services for persons with disabilities in Bangladesh: identifying service gaps and scopes for improvement. Disabil Rehabil 2022;44(19):5571-5584. [doi: 10.1080/09638288.2021.1939799] [Medline: 34176400]
18. Hossain MA, Kakuli SA, Numan SM. Barriers of rehabilitation services among the stroke patients in the selected community area of Dhaka, Bangladesh. J Sci Technol Res 2024;5(1):43-48. [doi: 10.59738/jstr.v5i1.23(43-48).tmih4644]
19. Venketasubramanian N, Mannan M. Stroke burden and stroke services in Bangladesh. Cerebrovasc Dis Extra 2021;11(2):69-71 [FREE Full text] [doi: 10.1159/000517234] [Medline: 34247155]
20. Qin H, Reid I, Gorelik A, Ng L. Environmental enrichment for stroke and other non-progressive brain injury. Cochrane Database Syst Rev 2021;11(11):CD011879 [FREE Full text] [doi: 10.1002/14651858.CD011879.pub2] [Medline: 34811724]
21. Kumar A, Yadav AK, Singh VK, Pathak A, Chaurasia RN, Mishra VN, et al. Caregiver burden in caregivers of stroke survivors: a hospital-based study. Ann Indian Acad Neurol 2022;25(6):1092-1098 [FREE Full text] [doi: 10.4103/aian.aian_318_22] [Medline: 36911438]
22. Veronese M, Vellone E, Alvaro R, Pucciarelli G. The transitional care from hospital to home for stroke survivors and their caregivers: a systematic review. J Vasc Nurs 2025;43(2):86-98 [FREE Full text] [doi: 10.1016/j.jvn.2025.03.002] [Medline: 40484579]
23. Ezzell C. Wild chimpanzees upgraded to endangered species. Nature 1988;336(6199):511. [doi: 10.1038/336511a0] [Medline: 3200299]
24. Gutenbrunner C, Stievano A, Nugraha B, Stewart D, Catton H. Nursing - a core element of rehabilitation. Int Nurs Rev 2022;69(1):13-19. [doi: 10.1111/inr.12661] [Medline: 33506550]
25. Hopewell S, Chan A, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. BMJ 2025;389:e081123 [FREE Full text] [doi: 10.1136/bmj-2024-081123] [Medline: 40228833]
26. Coleman ER, Moudgal R, Lang K, Hyacinth HI, Awosika OO, Kissela BM, et al. Early rehabilitation after stroke: a narrative review. Curr Atheroscler Rep 2017;19(12):59 [FREE Full text] [doi: 10.1007/s11883-017-0686-6] [Medline: 29116473]
27. Brydges CR. Effect size guidelines, sample size calculations, and statistical power in gerontology. Innov Aging 2019;3(4):igz036 [FREE Full text] [doi: 10.1093/geroni/igz036] [Medline: 31528719]
28. Tsurumi T. Daily Life Activities and Living Environment Studies (3rd ed, pp. 118-157). Japan: Igaku Shoin Co, Ltd; 2009.
29. Afrin M, Khan SU, Das SC, Huq KATME, Moriyama M. Effectiveness of a health education program for patients who had a stroke and their caregivers by controlling modifiable risk factors to reduce stroke recurrence in a tertiary hospital in

XSL•FO
RenderX

Bangladesh: protocol for a randomized controlled trial. JMIR Res Protoc 2023;12:e51178 [FREE Full text] [doi: 10.2196/51178] [Medline: 38100172]
30. Keith RA, Granger CV, Hamilton BB, Sherwin FS. The Functional Independence Measure: a new tool for rehabilitation. Adv Clin Rehabil 1987;1:6-18. [Medline: 3503663]
31. Rahman DN, Hasan DMM, Moyeenujjaman DM, Shankoor DM. Cross-cultural validation of the Bangla version of the Functional Independence Measure for patients with stroke. Sch J App Med Sci 2023;11(03):679-686. [doi: 10.36347/sjams.2023.v11i04.002]
32. Schwarzer R, Jerusalem M. Causal and control beliefs. 1995. URL: https://www.religiousforums.com/data/attachment [accessed 2025-02-20]
33. Aktar R, Nahar A. Parental acceptance, mental health and self-efficacy of adults in Bangladesh. IOSR J Humanit Soc Sci 2014;19(2):1-7 [FREE Full text]
34. Stevelink SAM, Hoekstra T, Nardi SMT, van der Zee CH, Banstola N, Premkumar R, et al. Development and structural validation of a shortened version of the Participation Scale. Disabil Rehabil 2012;34(19):1596-1607. [doi: 10.3109/09638288.2012.656793] [Medline: 22372970]
35. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980;20(6):649-655. [doi: 10.1093/geront/20.6.649] [Medline: 7203086]
36. Rabin F, Nahar JS, Mullick MSI, Ahmed HU, Farzana N. Adaptation, linguistic and clinometric validation of the Bangla version of Zarit Burden Interview. Bangabandhu Sheikh Mujib Me Univ J 2016;9(4):181-187. [doi: 10.3329/bsmmuj.v9i4.30060]
37. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. Circulation 2019;139(10):e56-e528 [FREE Full text] [doi: 10.1161/CIR.0000000000000659] [Medline: 30700139]
38. Che B, Shen S, Zhu Z, Wang A, Xu T, Peng Y, et al. Education level and long-term mortality, recurrent stroke, and cardiovascular events in patients with ischemic stroke. J Am Heart Assoc 2020;9(16):e016671 [FREE Full text] [doi: 10.1161/JAHA.120.016671] [Medline: 32779506]
39. Sebastian IA, Gandhi DBC, Sylaja PN, Paudel R, Kalkonde YV, Yangchen Y, et al. Stroke systems of care in South-East Asia Region (SEAR): commonalities and diversities. Lancet Reg Health Southeast Asia 2023;17:100289 [FREE Full text] [doi: 10.1016/j.lansea.2023.100289] [Medline: 37849930]
40. Krishnamurthi RV, Ikeda T, Feigin VL. Global, regional and country-specific burden of ischaemic stroke, intracerebral haemorrhage and subarachnoid haemorrhage: a systematic analysis of the Global Burden of Disease Study 2017. Neuroepidemiology 2020;54(2):171-179. [doi: 10.1159/000506396] [Medline: 32079017]
41. Qureshi A, Swain N, Aldabe D, Hale L. Exploring challenges affecting resilience in carers of stroke survivors: a qualitative descriptive study. Disabil Rehabil 2023;45(22):3696-3704. [doi: 10.1080/09638288.2022.2135774] [Medline: 36269117]
42. van Ommeren AL, Smulders LC, Prange-Lasonder GB, Buurke JH, Veltink PH, Rietman JS. Assistive technology for the upper extremities after stroke: systematic review of users' needs. JMIR Rehabil Assist Technol 2018;5(2):e10510 [FREE Full text] [doi: 10.2196/10510] [Medline: 30497993]
43. Li X, He Y, Wang D, Rezaei MJ. Stroke rehabilitation: from diagnosis to therapy. Front Neurol 2024;15:1402729 [FREE Full text] [doi: 10.3389/fneur.2024.1402729] [Medline: 39193145]
44. Li Y, Hu S. Impact of nurse-led interventions on health-related quality of life (HRQOL) in stroke patients: a systematic review and meta-analysis. Clin Nurs Res 2025;34(6):304-315. [doi: 10.1177/10547738251354998] [Medline: 40734648]
45. van Vliet P, Carey LM, Turton A, Kwakkel G, Palazzi K, Oldmeadow C, et al. Task-specific training versus usual care to improve upper limb function after stroke: the "Task-AT Home" randomised controlled trial protocol. Front Neurol 2023;14:1140017 [FREE Full text] [doi: 10.3389/fneur.2023.1140017] [Medline: 37456648]
46. Della Vecchia C, Préau M, Haesebaert J, Viprey M, Rode G, Termoz A, et al. Factors associated with post-stroke social participation: a quantitative study based on the ICF framework. Ann Phys Rehabil Med 2023;66(3):101686 [FREE Full text] [doi: 10.1016/j.rehab.2022.101686] [Medline: 35779831]
47. Alahmadi A, Al-Dubai SAR, Alhelali RA, Alrefaei HA, Alqerafi NM. Assessment of caregiving burden among caregivers of older adults with stroke in Saudi Arabia: analytical cross-sectional study. BMC Public Health 2025;25(1):2605 [FREE Full text] [doi: 10.1186/s12889-025-23895-7] [Medline: 40751162]
48. Lynch EA, Nesbitt K, Gulyani A, Chan RJ, Bidargaddi N, Cadilhac DA, et al. Do self-management interventions improve self-efficacy and health-related quality of life after stroke? A systematic review. Int J Stroke 2025;20(7):786-800 [FREE Full text] [doi: 10.1177/17474930251340286] [Medline: 40275483]
49. Htwe O, Yuliawiratman BS, Tannor AY, Nor Asikin MZ, Soh E, DE Groote W, et al. Barriers and facilitators for increased accessibility to quality rehabilitation services in low- and middle- income countries: a systematic review. Eur J Phys Rehabil Med 2024;60(3):514-522 [FREE Full text] [doi: 10.23736/S1973-9087.24.08154-1] [Medline: 38551518]

XSL•FO
RenderX

Abbreviations

ADL: activities of daily living
CG: control group
CONSORT: Consolidated Standards of Reporting Trials
CRP: Centre for the Rehabilitation of Paralyzed
FIM: Functional Independence Measure
ICF: International Classification of Functioning, Disability, and Health
IG: intervention group
LMIC: low- and middle-income countries
mRS: modified Rankin scale
NINS&H: National Institute of Neurosciences & Hospital
RA: research assistant
WHO: World Health Organization

Edited by S Munce; submitted 02.Dec.2025; peer-reviewed by P Pichayapinyo, EKC Siow; comments to author 12.May.2026; revised version received 17.Jun.2026; accepted 02.Jul.2026; published 17.Jul.2026.

Please cite as:
Akhter S, Huq KATME, Date S, Islam M, Chhobi FK, Sayeed SKJB, Islam MM, Afrin M, Kazawa K, Moriyama M
Introduction of Nurse-Led Rehabilitation Services for Patients With Stroke After Discharge to Improve Self-Care Management in Bangladesh: Pilot Randomized Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e88808
URL: https://rehab.jmir.org/2026/1/e88808
doi:10.2196/88808
PMID:

©Salma Akhter, KATM Ehsanul Huq, Shota Date, Mahmudul Islam, Farida Khatun Chhobi, SK Jakaria Been Sayeed, Md Mazharul Islam, Mahabuba Afrin, Kana Kazawa, Michiko Moriyama. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 17.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experiences of Accessing Physical Rehabilitation Services and Assistive Products Among Orthosis and Wheelchair Users: Qualitative Study

Charlotte Spurway[1], PhD; Andrew Power[2], PhD; Maggie Donovan-Hall[3], PhD; Andrew Amos Channon[4], PhD

[1]Department of Applied Health Sciences, College of Medicine and Health, University of Birmingham, Birmingham, United Kingdom

[2]School of Geography and Environmental Science, University of Southampton, Southampton, United Kingdom

[3]School of Health Sciences, University of Southampton, Southampton, United Kingdom

[4]Department of Social Statistics and Demography, University of Southampton, Southampton, United Kingdom

Corresponding Author:
Charlotte Spurway, PhD
Department of Applied Health Sciences
College of Medicine and Health
University of Birmingham
Health Economics Unit
IOEM Building
Birmingham, B15 2TT
United Kingdom
Phone: 44 121 414 6486
Email: c.m.spurway@bham.ac.uk

Abstract

Background: Access to physical rehabilitation services, including orthotics and wheelchairs, is recognized as a fundamental human right but remains limited in many low- and middle-income countries due to resource constraints, fragmented care, and a shortage of trained rehabilitation professionals. In Cambodia, where the health system still faces challenges from past conflict, these services are delivered through a mix of government, international, and nongovernmental organizations.

Objective: This study investigated the pathways to obtaining orthoses and wheelchairs in Cambodia and their impact on individuals with physical impairments.

Methods: Semistructured interviews were conducted to explore experiences and the impact of accessing physical rehabilitation and assistive products for orthosis and wheelchair users. Interviews were conducted face-to-face with translation support, audio recorded, transcribed, and analyzed using thematic analysis. Interviews were conducted with participants in Phnom Penh and the Kandal province in Cambodia. In total, 17 participants aged 18 to 65 years who were current or previous patients of the Exceed Worldwide physical rehabilitation center located in Phnom Penh took part in the study.

Results: Three themes emerged (early experiences of illness and health care seeking, pathways to assistive products, and the impact of these assistive products). Access to orthoses and wheelchairs often involved lengthy delays and inconsistent experiences with medical care and physiotherapy. Despite these challenges, assistive products played a vital role in improving mobility, independence, and social participation, contributing to greater confidence and emotional well-being. However, issues such as poor environmental accessibility and delays in repair or replacement limited their effectiveness.

Conclusions: The findings demonstrate that, while assistive products offer significant benefits, they do not fully address the systemic barriers faced by people with physical disabilities in Cambodia. Persistent challenges such as stigma, environmental constraints, and emotional strain demonstrate the need for integrated policy reforms and a robust continuum of care that ensures coordinated, ongoing support across all stages of rehabilitation.

(JMIR Rehabil Assist Technol 2026;13:e84735) doi:10.2196/84735

KEYWORDS

Cambodia; orthotic devices; wheelchairs; physical rehabilitation; qualitative; resource-limited settings

XSL•FO
RenderX

Introduction

The United Nations Convention on the Rights of Persons With Disabilities states that access to physical rehabilitation services and assistive products constitutes a fundamental human right [1]. Consequently, individuals with disabilities must be guaranteed equitable access to these services. According to the World Health Organization [1], disability is caused by the interaction between health conditions or impairments and societal or environmental barriers that hinder full participation in society. Within this, impairment refers to issues in body function or structure, including significant deviations or losses. In many low- and middle-income countries (LMICs), physical rehabilitation centers providing prosthetic and orthotic (P&O) services and assistive products are limited, and those services that are available may be inaccessible to potential service users [2].

Assistive products, such as orthoses and wheelchairs, are external products that can be used to prevent impairments and can support individuals in many areas of life, such as education, employment, independent living, and social participation [1,3]. The provision of assistive products can bring socioeconomic benefits for governments and welfare systems, including increased labor force productivity [1,4]. Orthotics and wheelchairs help improve mobility, but they must be appropriate for the user's environment. If an assistive product is not well matched to the setting in which it is used, it may be abandoned [2]. This can result from poor-quality provision, limited user training, and a lack of local services for maintenance and repair [2].

Although population need for wheelchairs and orthoses is lower than the need for other assistive products such as spectacles or walking sticks, reflecting their use among people with more severe functional limitations [1], access remains limited. Only around one-third of those needing basic manual wheelchairs have access to them, with even lower access to more specialized devices. Orthotic provision is similarly low, meeting only 15% to 25% of needs [1].

Orthoses and wheelchairs are included in the World Health Organization Priority Assistive Products List, reflecting their important role in enabling mobility, independence, and participation for large and diverse user groups [5]. However, access to assistive products in LMICs remains limited due to factors such as a lack of political prioritization, insufficient funding and investment, limited awareness of the need for assistive products, market barriers, and shortages of trained rehabilitation workers, resulting in understaffed services and barriers to effective service delivery [6,7]. These challenges can contribute to lengthy delays in accessing physical rehabilitation, such as P&O services, which are further exacerbated by limited referral systems in place and the absence of well-established continuums of care across multiple stages, such as health promotion, preventive care, curative care, and rehabilitation [1,6,8].

This research focused on Cambodia, an LMIC in Southeast Asia. The health system in Cambodia has been heavily impacted by a turbulent past, including conflict and civil war, which has contributed to ongoing challenges [9]. Currently, the delivery of physical rehabilitation services in Cambodia is a joint effort among the Cambodian government, international organizations, and nongovernmental organizations (NGOs) [10]. The People With Disabilities Foundation, a department of the Cambodian Ministry of Health, operates 3 physical rehabilitation clinics and 2 repair workshops [10]. Other organizations, such as Exceed Worldwide, International Committee of the Red Cross, and Humanity & Inclusion, also provide physical rehabilitation and assistive products free of charge [11].

Although there is a growing body of qualitative research on prosthetic services in low-resource settings, including studies on access barriers [12,13], user experiences [11,14], and satisfaction with devices and service delivery [15], fewer studies focus specifically on orthosis or wheelchair users. Where included, these groups are typically examined alongside prosthetic users, limiting evidence on their specific experiences despite the scale and diversity of people who rely on these devices. This study included wheelchair users only, orthosis users only, and users of both to explore shared and divergent experiences across mobility-related assistive product use, enabling analysis of experiences related to rehabilitation access, assistive technology provision, participation, and disability-related stigma, as well as differences between groups. This study therefore explored pathways to accessing physical rehabilitation and obtaining orthoses and wheelchairs in Cambodia and examined how this access shapes the lives of people with physical impairments.

Methods

Design

This qualitative study used semistructured in-depth interviews to explore the pathways to physical rehabilitation services for orthotics and wheelchairs in Cambodia. As this study aimed to understand the pathways taken to receive services, we wanted participants to use their own words to describe their experiences rather than using a fixed survey, leading to the approach undertaken. Multimedia Appendices 1 and 2 contain the completed COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist and associated information about the interviewer.

Setting and Sampling

This study took place in Phnom Penh and the Kandal province, Cambodia, in partnership with a physical rehabilitation center operated by Exceed [16]. Exceed is an international NGO that provides physical rehabilitation services in 3 provinces of Cambodia. Participants included individuals who were currently using or had previously received an orthosis or wheelchair from Exceed's Phnom Penh clinic.

Participants were primarily recruited using a convenience sampling approach involving individuals who attended the clinic either for scheduled appointments or a walk-in visit. Additional recruitment took place during community visits conducted as part of Exceed's outreach activities, including repair workshops and wheelchair delivery events. Some participants were also identified through direct outreach, in which community workers

XSL•FO
RenderX

contacted potential participants to inquire about their interest in taking part in the study. A total of 9 participants were recruited at the clinic, and 8 were recruited either during community visits or through interviews that took place outside of the clinic arranged by community workers. Recruitment was conducted by Exceed staff, who presented the study and emphasized that participation was voluntary.

Interview Procedures

All interviews were conducted face-to-face in Cambodia at Exceed's clinic in Phnom Penh, during community outreach visits or at the homes of participants. For interviews conducted during community visits, a private area within the community space where outreach activities were held was used to ensure confidentiality. The interviews took place from April 2022 to June 2022.

A lecturer from the Cambodian School of Prosthetics and Orthotics in Phnom Penh served as the translator for the interviews. The translator was a native Khmer speaker with English as their second language and professional experience teaching P&O in English. Their clinical and educational expertise supported the translation of technical rehabilitation concepts; however, the absence of a professionally trained interpreter may have resulted in the loss of linguistic nuance or subtle shifts in meaning [17]. Consistent use of a single translator and advance translation of the interview guide were used to support translation consistency [18]. The English version of the interview guide can be found in Multimedia Appendix 3.

During the interviews, questions asked in English by the interviewer were translated into Khmer. Participant responses were then translated back into English using consecutive verbatim interpretation. The quotations presented in the Results section represent the words of the translator and were subsequently edited into readable English for clarity while preserving the original semantic meaning and tone, in line with recommendations within the qualitative research literature regarding translated data [19,20].

The interviews were also recorded using an audio device. Sample size was guided by the aim of generating in-depth accounts rather than achieving data saturation, in line with reflexive thematic analysis practice [21].

Ethical Considerations

All participant documents, such as participant information sheets and consent forms, were produced in English and translated into Khmer by a local team based at Exceed. These documents were provided to the participants before the interviews, and the translator read them to participants to ensure full agreement and explained that they could withdraw at any time and this would not impact the care or treatment they received from the P&O clinic. The potential risks and benefits of taking part and data handling procedures were also explained to participants.

Once the participants had understood and agreed to take part in the study, written informed consent was obtained from all participants prior to data collection. No participants withdrew from the study following consent. All identifiable information was removed from transcripts, and data were stored securely on university servers accessible only to members of the research team.

Ethics approval for this study was granted by the University of Southampton Ethics and Research Governance Online system (68254.A1) and the National Ethics Committee for Health Research in Cambodia (088). All study procedures were conducted in accordance with the ethical standards of the approving committee. Participants received a 12,000 Cambodian riel (US $2.97) token of appreciation for their time.

Data Analysis

The audio files from the interviews were transcribed verbatim by the lead author (CS). The data were analyzed using reflexive thematic analysis [21] to categorize them into themes and subthemes. Thematic analysis was carried out using an interpretivist epistemological framework, which views knowledge as coconstructed and embedded within social, cultural, and contextual settings [22]. From this perspective, findings are not considered objective truths but interpretations shaped through interaction between translated participant accounts and researcher analysis.

Data coding was completed by the lead author (CS) and discussed with the coauthors (AP, MD-H, and AAC). This study formed part of a wider project on barriers to and facilitators of health and rehabilitation services in Cambodia. The wider study informed the initial deductive coding framework; however, coding and theme development for this analysis were primarily inductive and grounded in participants' accounts. The wider study provided a guiding concept rather than a rigid coding structure, allowing additional meanings and interpretations to develop iteratively throughout the analysis through inductive coding [23]. After coding the interview data, codes were reviewed and iteratively developed into broader themes and subthemes through an ongoing process of comparison, refinement, and discussion among the research team. Multimedia Appendix 4 includes all the themes, subthemes, and codes that emerged in this study.

Results

Participant Characteristics

Table 1 presents participant information, including gender, age, type of assistive product, reason for use, and interview location. Seventeen participants were interviewed (n=12, 70.6% women and n=5, 29.4% men), and none dropped out before completing the interview. In total, 52.9% (n=9) of the participants used knee-ankle-foot orthoses, 29.4% (n=5) used ankle-foot orthoses, and 17.6% (n=3) used manual wheelchairs. The average interview length was 43 (SD 12.3) minutes (ranging from 19 to 62 minutes).

XSL•FO
RenderX

Table 1. Participant information.

Pseudonym	Sex	Age (y)	Device	Reason for device use	Location of interview
Bopha	Female	40	KAFO[a]	Polio	Clinic
Sambath	Male	31	AFO[b]	Polio	Clinic
Chaya	Female	55	KAFO	Polio	Clinic
Chhean	Female	30	KAFO	Polio	Clinic
Daevy	Female	33	AFO	Polio	Clinic
Leap	Female	35	AFO	Polio	Community
Sothy	Female	38	KAFO	Polio	Clinic
Samnang	Female	18	AFO	Cerebral palsy	Clinic
Sophal	Female	51	KAFO	Polio	Community
Chamroeun	Female	20	AFO	Cerebral palsy	Community
Rotha	Female	18	KAFO	Polio	Clinic
Choun	Female	18	KAFO	Clubfoot	Clinic
Chea	Female	28	Wheelchair	Spinal condition	Community
Narin	Male	65	Wheelchair	Stroke	Community
Chanthou	Male	51	Wheelchair	Polio	Community
Savy	Male	35	KAFO	Polio	Community
Viseth	Male	32	KAFO	Polio	Community

[a]KAFO: knee-ankle-foot orthosis.

[b]AFO: ankle-foot orthosis.

The interviews were coded into three themes: (1) early experiences of illness and health care seeking, (2) pathways to assistive products, and (3) impact of assistive products.

Early Experiences of Illness and Health Care Seeking

Most participants stated that the reason for orthosis or wheelchair use was fever, with the most common cause of this being polio. Nearly all participants experienced this before they were 5 years old. A minority were uncertain of the cause, with some of their parents being skeptical of medical explanations. For example, one participant reported that her mother believed that antibiotic or antipyretic injections received at a health center were responsible:

> *Before she got ill, she was able to walk right, so her mum doesn't believe that a fever for a night has led to her being unable to stand and walk.* [From interview with Samnang; female; aged 18 years]

As most participants were young at the time of diagnosis, health care decisions were made by parents and grandparents. Health care use varied; while some sought multiple treatments, others received no care despite sudden limb weakness or paralysis. Participants who reported never accessing health care at the time cited several reasons, including parental work demands, financial constraints, and limited health care availability. This was explained by 11.8% (2/17) of the participants, who described the lack of treatment options when they were children due to insecurity and civil unrest in the country:

> *During 1972, she said because the country is insecure or there were problems in the country, there was no medical treatment available, no nothing.* [From interview with Sophal; female; aged 51 years]

Khmer traditional treatments were used exclusively or alongside allopathic medicine, with parents often moving between different types of health care facilities.

Pathways to Assistive Products

Experiences of Medical Treatments or Physiotherapy

Participants learned of orthotic and wheelchair services, including those provided by Exceed and other facilities, through family, neighbors, and outreach by community health workers. Lengthy delays between the onset of impairment and receiving further medical treatment, such as surgery or physiotherapy from hospitals or physical rehabilitation services, were common. Surgery, typically to release lower-limb muscles and tendons to improve mobility, was followed by months of aftercare, although some participants were unable to adhere to it. One participant explained that it was recommended that she undergo surgery on her spine; however, her mother refused due to not believing it was necessary:

> *The physician in Kantha Bopha hospital mentioned that the most important thing is her trunk, her spine. They wanted to do the operation on her spine, but her mum didn't agree with the surgery.* [From interview with Chea; female; aged 28 years]

Although it was not explored in depth, these findings indicate the importance of parental trust in obtaining health care and ensuring that health care professionals can adequately explain information to foster trust between them and patients.

XSL•FO
RenderX

Experiences of Using P&O Services

Participants reported varied experiences with physical rehabilitation services and assistive products. Initial visits typically involved assessment for orthotics or wheelchairs followed by physiotherapy and/or limb casting for orthotic fabrication. Several participants recalled fear and distress during early fittings, particularly as children, due to limited understanding of the process. One participant, who was a child at the time of her first appointment, stated that she believed the orthosis would require cutting into her leg:

> *She said that first time she was scared because she thought that when they brought her to the center, they might cut into her leg or something.* [From interview with Daevy; female; aged 35 years]

Two wheelchair users reported no difficulties adapting; however, one wheelchair user expressed reluctance to use the wheelchair outside the home:

> *He uses the wheelchair just around the house, only indoors. He doesn't use it outdoors because he is afraid of cars and other transport, because they travel very fast, so he is worried about accidents.* [From interview with Narin; male; aged 65 years]

Pain during orthosis fitting and use was commonly reported, particularly by those fitted as children. Gender differences emerged, with men generally reporting minimal discomfort whereas women more often described pain and reported their orthoses as bulky, heavy, and causing blisters in the lower limbs. The 2 quotes below describe some of the experiences that female participants had with their orthoses:

> *She said that when using the device, she feels like it's difficult to walk and is painful. Sometimes she feels like the knee is soft, so she falls over often.* [From interview with Rotha; female; aged 18 years]

> *He said it's not very difficult, like he is able to adapt when he wears the orthosis.* [From interview with Sambath; male; aged 31 years]

Difficulties in adapting to an orthosis sometimes led to device abandonment, particularly when discomfort prevented use. However, many users persisted because they recognized potential functional benefits, although some required multiple orthoses to adapt:

> *She said that after she threw the first device away, she was afraid she would be shy (due to her impairment), so she wanted to try and get a second device and learn to walk with that device.* [From interview with Bopha; female; aged 40 years]

Impact of Assistive Products

Importance of Access to Assistive Products

Several participants reported the positive impact of using a wheelchair or an orthosis. These devices supported their mobility, independence, and inclusion within their communities. Access to assistive products improved participants' confidence and emotional well-being, contributing to a greater sense of happiness and facilitating social interaction. One participant explained that, since she received the orthotic device, she had been able to make friends and work with others:

> *She feels confident and can be herself when she got the new device, before when she didn't have the device, she didn't have friends and worked alone.* [From interview with Chhean; female; aged 30 years]

Independence and the ability to carry out household tasks were highly valued, influencing patterns of assistive product use. Orthoses were generally worn throughout the day and removed only for washing or sleep, although some participants managed short distances without them. This pattern of use contrasted with that of wheelchairs, which were less commonly used within the home environment:

> *For him if he just stays at home, he won't use the wheelchair, but if he travels outside for a short distance, he's going to use his wheelchair.* [From interview with Chanthou; male; aged 51 years]

Despite the significance of assistive products in participants' lives, many reported delays in repairs or replacements of their assistive products. One participant, for example, reported using the same orthotic device for over 10 years, far beyond the typical expected life span of 2 to 3 years. These delays often led participants to resort to temporary self-repairs:

> *He said that his orthosis was broken, so he tried to repair by wrapping plastic around it, but he knows that it is not strong and is only a temporary repair until he gets the new device.* [From interview with Sambath; male; aged 31 years]

Limits in Access to Assistive Products

While assistive products provided significant benefits, they did not eliminate all challenges associated with having a physical impairment in Cambodia. One participant noted that, after discontinuing the use of her orthosis, she felt that her lower limbs had become stronger and became more confident in her mobility:

> *She said that because she has stopped using the device, she feels stronger and more confident because when she was young, she fell over often, but now it's ok, she doesn't fall often.* [From interview with Chamroeun; female; aged 20 years]

Participants noted that environmental conditions affected the functionality of their assistive products. Uneven or sandy terrain increased the risk of falls and wheelchair entrapment, whereas the weather reduced durability and comfort. Flooding and heavy rains during the wet season caused rusting of wheelchair frames, whereas high temperatures affected orthoses by causing the plastic to expand, affecting the fit of the device:

> *Because of the heat right now, the plastic [on the orthosis] increases in size so it makes it loose for him. It doesn't cause pain, but there is less control.* [From interview with Savy; male; aged 35 years]

Despite access to assistive products, daily limitations persisted, varying across individuals. Participants reported avoiding community events such as weddings and religious ceremonies due to discomfort; anxiety; or practical barriers, for example,

XSL•FO
RenderX

concerns about not being able to sit on the floor (customary during temple visits) without removing an orthosis.

The interviews also found challenges with self-perception. Several participants expressed feelings of being a burden on their families due to their reliance on others. One participant, following a stroke, compared his current abilities with his prestroke abilities and described a reduced sense of self:

> *The people around him pity him. He feels like before he could do everything and worked hard, but now he can't do anything, so he feels like he is useless.* [From interview with Narin; male; aged 65 years]

Experiences of negative perceptions were common. Some participants reported being stared at in public, leading to a reluctance to leave their homes:

> *Other problems he faces when he goes out, people look at him like he is not normal and they stare at him.* [From interview with Chanthou; male; aged 51 years]

However, not all participants reported discrimination. Some felt that they were treated equally because they were able to perform daily tasks and contribute economically, aligning with societal perceptions of "normality":

> *She doesn't feel like anyone discriminates or criticizes her about her disability. Her family say that she is ok, she has an impairment, but she is able to get married, have a family and work like other normal people.* [From interview with Leap; female; aged 38 years]

Participants often described employment and contributing to family and community life as helping mitigate the impact of impairment. "Blending in" with people without disabilities was seen as a strategy to avoid discrimination. Some noted improving attitudes toward disability in Cambodia, attributed to greater awareness and education:

> *Nowadays he thinks that information about disability is shared to everyone and people understand more about disabled people, so they don't discriminate against him.* [From interview with Viseth; male; aged 32 years]

However, there was a shared recognition of the ongoing need for more information about disability to be disseminated through schools, television, and other media to reduce stigma.

Discussion

Principal Findings

This study found that early childhood illness, most commonly linked to polio, alongside delayed or absent access to health care and rehabilitation services, shaped long-term disability trajectories and pathways to assistive product use. Assistive products substantially improved mobility, independence, and psychosocial well-being but did not fully remove environmental, social, and attitudinal barriers experienced by users. Pain, fear, and adaptation challenges highlighted the limitations of provision models that focus solely on the device, emphasizing the need for more holistic approaches that address both individual needs and the wider social context.

These broader challenges are rooted in the underlying causes and understandings of impairment. Use of orthoses or wheelchairs in this study was largely attributed to contracting polio in childhood, although other causes included clubfoot, cerebral palsy, and stroke. Some participants with polio were uncertain about their exact diagnosis, with several attributing their impairments to treatments received in hospital. Evidence shows that, in polio-endemic settings, intramuscular injections administered during the incubation period of poliovirus can increase the risk of paralysis, a phenomenon known as provocation poliomyelitis [24-26]. This may help explain why some individuals linked their impairments to medical treatments they received as children.

Many participants reported significant delays in receiving their first assistive product, primarily due to limited awareness of P&O services and restricted service availability. Early rehabilitation is known to improve outcomes for conditions such as stroke and cerebral palsy [27,28]. However, access to physical rehabilitation in many LMIC settings remains constrained by inconsistent service provision and shortages of skilled rehabilitation professionals [29]. Pathways to assistive products can also be difficult to navigate due to fragmented referral systems, particularly where rehabilitation services are delivered by international organizations and NGOs operating separately from wider health systems [1]. Despite this, P&O and other rehabilitation services are essential components of the health care system and play an important role along the continuum of care, particularly in ensuring access during both the postacute and long-term phases of care [30].

Initial use of orthoses was frequently associated with pain, contributing to device abandonment in some cases. This study found that female users of orthotics described experiences of pain and discomfort with their orthoses more frequently than male users. This aligns with existing literature that showed that women's experiences with orthoses were more unpleasant and uncomfortable [31], with another study finding persistent, gendered disparities in access, design, comfort, and function outcomes of assistive products for women compared to men [32]. An analysis of data on orthosis users in Cambodia found that the proportion of current female users was nearly 10% lower than that of male users [33]. This highlights the need for more research specifically examining gendered outcomes of orthoses and wheelchairs. Addressing these gaps will require approaches that account for physiological differences, user experience, and the structural factors shaping women's rehabilitation outcomes [34].

The findings from this study highlight the role of assistive products in the lives of people with physical impairments. Participants reported that these products enhanced independence, enabled work, and supported daily activities, echoing findings from other studies involving P&O users [11,35]. However, access alone does not ensure use, nor do assistive products fully remove barriers, especially in contexts where the broader social and political environment remains restrictive or stigmatizing [3,36]. In such settings, ongoing barriers to education and

XSL•FO
RenderX

employment may constrain the potential of assistive devices to advance inclusion.

Experiences of discrimination or prejudice related to physical disabilities were reported by participants. Misconceptions about the cause of disability, often rooted in traditional or religious understandings, can help perpetuate disability stigma [37,38]. Such stigma can restrict access to employment, education, and community inclusion. However, not all participants experienced discrimination or identified as having a disability, and some described their ability to work and have a family as being the same as that of others. Additionally, it was reported that wearing an assistive product reduced negative treatment from others. Previous research has shown that access to resources that can be used to obtain assistive products, such as finances and knowledge, was an indicator of privilege [39]. This finding was linked to the notion that poverty stigma can be more discriminatory than disability stigma in LMICs. This may help explain why participants had more positive attitudes after receiving an orthosis or wheelchair.

Not all participants wore an orthotic device daily or at all, underlining that access alone does not guarantee use of assistive products or rehabilitation services. In many LMICs, limited design options may not suit users' environments, contributing to nonuse [1,2]. While this study did not examine device abandonment in depth, findings suggest that provided products may not always meet users' needs or preferences. Known contributors to abandonment include pain, discomfort, inadequate training, and personal preference [40,41]. These gaps highlight the need for a comprehensive service approach that goes beyond initial provision to include follow-up, context-appropriate design, user education, and sustained support across the continuum of care.

In Cambodia, P&O service provision has shifted over time, with changes in rehabilitation services, increased government involvement, and reduced international funding. Sustaining a continuum of care for those requiring physical rehabilitation depends on establishing standardized referral pathways across all levels of the health system [30]. This necessitates collaboration among a range of stakeholders, including the government, public and private hospitals and health facilities, and international organizations.

Reflexivity

As White, anglophone researchers based in a high-income country, we recognize that our professional backgrounds, experiences, and prior assumptions may have shaped the study design, data interpretation, and construction of meaning from the data. The absence of Cambodian researchers within the author team is an important limitation and reflects broader structural inequities in global health research and ongoing debates around decolonizing humanitarian and development work. This study was conducted as part of a PhD project, which influenced the research team composition and authorship structure and may have limited the extent to which locally grounded interpretations informed the analysis.

Data collection and interpretation were supported by Cambodian rehabilitation professionals who facilitated recruitment, translation, and contextual understanding. Although their contributions informed the research process and were remunerated, they were not involved in coauthorship or full thematic analysis, primarily due to limited availability alongside clinic responsibilities. This may have influenced how meaning was interpreted and themes were developed.

The interviewer, as an external researcher, also occupied an outsider position [42], which may have shaped participant responses and the coconstruction of interview data. To enhance reflexivity and minimize bias, the findings were discussed within the research team, and contextual input from in-country collaborators informed interpretation. Nevertheless, limited local analytical involvement remains a key limitation, and future research should prioritize more equitable coproduction, including shared analysis and authorship with Cambodian researchers.

Strengths and Limitations

This study has a number of strengths. First, it makes a unique contribution to the literature as there is limited empirical research examining access to physical rehabilitation services specifically among individuals who use orthoses and wheelchairs. While there is a growing body of work on assistive technology access in LMICs, most studies focus on prosthetics or broader disability populations, with relatively few exploring the distinct needs and barriers experienced by orthosis and wheelchair users. Second, the findings are situated within the social, cultural, and environmental contexts of participants. This contextual grounding is important in rehabilitation and assistive technology research as cultural perceptions of disability, social support networks, and environmental barriers shape access, use, and outcomes [1].

This study also has some limitations (beyond those noted above about reflexivity). Participants were exclusively current or former users of Exceed services in Phnom Penh, excluding those who accessed other Exceed centers or services from different rehabilitation organizations, who may have had different pathways to obtaining orthoses and wheelchairs.

The interviewer (CS) did not speak Khmer, requiring the use of a translator. While clarification was sought to ensure accurate interpretation, some loss of meaning may have occurred. The translations were conducted by a local colleague, who is a native speaker of Khmer without formal training in professional interpretation or translation. This is a limitation of the study as translation in qualitative research is an interpretive act and can influence the precision and nuance of participant meaning [17]. To mitigate this, the same translator was used throughout the study to support consistency, and the interview guide was translated in advance to reduce ad hoc interpretation during interviews [18]. The translator also had contextual familiarity with the study setting, which supported comprehension of local and physical rehabilitation terminology. However, it is recognized that these measures do not replace the role of a professionally trained interpreter and that limited formal linguistic training may have influenced the accuracy and richness of some translations. The translator's role at the Cambodian School of Prosthetics and Orthotics in Phnom Penh, which collaborates closely with Exceed, may have also

XSL•FO
RenderX

influenced participants' willingness to voice criticisms, although assurances were given that care would not be affected.

Conclusions

This study investigated the pathways to physical rehabilitation services for users of orthoses and wheelchairs in Cambodia. The experiences of participants offer valuable insights into how individuals find and engage with P&O services for physical rehabilitation and assistive products and how these interventions influence their mobility and livelihood. This study highlights the important role of assistive products in enhancing mobility, independence, and engagement in daily activities. However, it also shows that they cannot fully overcome broader environmental, social, and emotional barriers, including stigma, discrimination, and negative self-perception.

Assistive products must be part of a broader, continuous system of care. A functioning continuum of care requires standardized referral mechanisms, consistent follow-up, and integration across all levels of the health system. Without this, individuals risk falling through gaps in service provision, particularly during critical stages of life such as childhood and adolescence.

Future research should further investigate the causes of assistive product abandonment and explore the complex reasons why individuals may choose not to use assistive products. A holistic, person-centered approach, one that supports individuals across the full continuum of care and acknowledges both physical and sociocultural dimensions, is key for improving quality of life for people with physical impairments in Cambodia.

Acknowledgments

The authors would like to thank Exceed Worldwide for their support with the completion of this work in the facilitation of interviews and guidance. In particular, they want to thank Sisary Kheng and Sophary Met for their support.

Data Availability

The datasets generated and analyzed during this study are publicly available in the University of Southampton institutional research repository [43]. The repository contains the deidentified study data and accompanying documentation required to reproduce the analyses.

Funding

Funding for this project was provided by the UK Economic and Social Research Council through the South Coast Doctoral Training Partnership.

Authors' Contributions

CS, AP, and AAC conceived and designed the study. Data collection was carried out by CS. Data analysis and interpretation were performed by CS, AP, MD-H, and AAC. The manuscript was drafted by CS, with input from AAC, AP, and MD-H. All authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ checklist.
[DOCX File , 19 KB - rehab_v13i1e84735_app1.docx]

Multimedia Appendix 2
Additional reflexivity statement.
[DOCX File , 15 KB - rehab_v13i1e84735_app2.docx]

Multimedia Appendix 3
Interview guide.
[DOCX File , 19 KB - rehab_v13i1e84735_app3.docx]

Multimedia Appendix 4
Themes and codes.
[DOCX File , 20 KB - rehab_v13i1e84735_app4.docx]

References

XSL•FO
RenderX

1. Global report on assistive technology. World Health Organization. 2022 May 15. URL: https://www.who.int/publications/i/item/9789240049451 [accessed 2026-05-15]
2. Borg J, Lindström A, Larsson S. Assistive technology in developing countries: a review from the perspective of the Convention on the Rights of Persons with Disabilities. Prosthet Orthot Int 2011 Mar;35(1):20-29. [doi: 10.1177/0309364610389351] [Medline: 21515886]
3. Tebbutt E, Brodmann R, Borg J, MacLachlan M, Khasnabis C, Horvath R. Assistive products and the Sustainable Development Goals (SDGs). Global Health 2016 Nov 29;12(1):79 [FREE Full text] [doi: 10.1186/s12992-016-0220-6] [Medline: 27899117]
4. The case for investing in assistive technology: the dramatic economic, health and social benefits of assistive a billion people to live fulfilling and dignified lives. ATscale. 2020. URL: https://atscalepartnership.org/atscale-publications/case-investing-assistive-technology [accessed 2026-07-16]
5. Priority assistive products list. World Health Organization. URL: https://iris.who.int/server/api/core/bitstreams/665850a1-6207-4c68-b1d2-b3b507ee77cd/content [accessed 2026-05-15]
6. Product narrative: prostheses. ATscale. 2020 Apr. URL: https://at2030.org/static/at2030_core/outputs/Prostheses_Product_Narrative_a11y_20200827.pdf [accessed 2025-08-23]
7. Metcalf CD, Ostler C, Thor P, Kheng S, Srors S, Sann R, et al. Engaging multisector stakeholders to identify priorities for global health innovation, change and research: an engagement methodology and application to prosthetics service delivery in Cambodia. Disabil Rehabil 2024 Feb;46(4):685-696 [FREE Full text] [doi: 10.1080/09638288.2023.2173313] [Medline: 36823949]
8. Heine M, Hanekom S. Chronic disease in low-resource settings: prevention and management throughout the continuum of care-a call for papers. Int J Environ Res Public Health 2023 Feb 17;20(4):3580 [FREE Full text] [doi: 10.3390/ijerph20043580] [Medline: 36834272]
9. Grundy J, Khut QY, Oum S, Annear P, Ky V. Health system strengthening in Cambodia-a case study of health policy response to social transition. Health Policy 2009 Oct;92(2-3):107-115. [doi: 10.1016/j.healthpol.2009.05.001] [Medline: 19501425]
10. Rehabilitation Strategic Plan for Cambodia 2018-2023. Phnom Penh, Cambodia: Ministry of Health and Ministry of Social Affairs, Veterans and Youth Rehabilitation, Cambodia; 2026.
11. Ramstrand N, Maddock A, Johansson M, Felixon L. The lived experience of people who require prostheses or orthoses in the Kingdom of Cambodia: a qualitative study. Disabil Health J 2021 Jul;14(3):101071 [FREE Full text] [doi: 10.1016/j.dhjo.2021.101071] [Medline: 33583726]
12. Allen AP, Bolton WS, Jalloh MB, Halpin SJ, Jayne DG, Scott JD. Barriers to accessing and providing rehabilitation after a lower limb amputation in Sierra Leone - a multidisciplinary patient and service provider perspective. Disabil Rehabil 2022 Jun;44(11):2392-2399. [doi: 10.1080/09638288.2020.1836043] [Medline: 33261506]
13. Farrar M, Niraula YR, Pryor W. Improving access to prosthetic services in Western Nepal: a local stakeholder perspective. Disabil Rehabil 2023 Apr;45(7):1229-1238. [doi: 10.1080/09638288.2022.2057599] [Medline: 35387522]
14. Mattick K, Oldfrey B, Donovan-Hall M, Magomere G, Gakunga J, Holloway C. Experiences of lower limb prosthesis users in Kenya: a qualitative study to understand motivation to use and satisfaction with prosthetic outcomes. Disabil Rehabil 2023 Dec;45(26):4478-4488 [FREE Full text] [doi: 10.1080/09638288.2022.2152875] [Medline: 36495104]
15. Magnusson L, Ahlström G. Patients' satisfaction with lower-limb prosthetic and orthotic devices and service delivery in Sierra Leone and Malawi. BMC Health Serv Res 2017 Feb 01;17(1):102 [FREE Full text] [doi: 10.1186/s12913-017-2044-3] [Medline: 28143549]
16. Exceed Worldwide. URL: https://www.exceed-worldwide.org/ [accessed 2026-05-10]
17. Squires A. Language barriers and qualitative nursing research: methodological considerations. Int Nurs Rev 2008 Sep;55(3):265-273 [FREE Full text] [doi: 10.1111/j.1466-7657.2008.00652.x] [Medline: 19522941]
18. Yunus NA, olde Hartman T, Lucassen P, Barton C, Russell G, Altun A, et al. Reporting of the translation process in qualitative health research: a neglected importance. Int J Qual Methods 2022 Dec 06;21. [doi: 10.1177/16094069221145282]
19. Temple B, Young A. Qualitative research and translation dilemmas. Qual Res 2004 Aug;4(2):161-178. [doi: 10.1177/1468794104044430]
20. van Nes F, Abma T, Jonsson H, Deeg D. Language differences in qualitative research: is meaning lost in translation? Eur J Ageing 2010 Dec;7(4):313-316 [FREE Full text] [doi: 10.1007/s10433-010-0168-y] [Medline: 21212820]
21. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: SAGE Publications; Oct 13, 2021.
22. Williams V, Boylan AM, Nunan D. Qualitative research as evidence: expanding the paradigm for evidence-based healthcare. BMJ Evid Based Med 2019 Oct;24(5):168-169. [doi: 10.1136/bmjebm-2018-111131] [Medline: 30850372]
23. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006 Mar;5(1):80-92. [doi: 10.1177/160940690600500107]
24. Gromeier M, Wimmer E. Mechanism of injury-provoked poliomyelitis. J Virol 1998 Jun;72(6):5056-5060 [FREE Full text] [doi: 10.1128/JVI.72.6.5056-5060.1998] [Medline: 9573275]

XSL•FO
RenderX

25. Strebel PM, Ion-Nedelcu N, Baughman AL, Sutter RW, Cochi SL. Intramuscular injections within 30 days of immunization with oral poliovirus vaccine--a risk factor for vaccine-associated paralytic poliomyelitis. N Engl J Med 1995 Feb 23;332(8):500-506. [doi: 10.1056/NEJM199502233320804] [Medline: 7830731]
26. Halsey NA. Commentary: poliomyelitis and unnecessary injections. Int J Epidemiol 2003 Apr;32(2):278-279. [doi: 10.1093/ije/dyg071] [Medline: 12714549]
27. Al Imam MH, Jahan I, Muhit M, Hardianto D, Laryea F, Chhetri AB, et al. Predictors of rehabilitation service utilisation among children with cerebral palsy (CP) in low- and middle-income countries (LMIC): findings from the Global LMIC CP Register. Brain Sci 2021 Jun 25;11(7):848 [FREE Full text] [doi: 10.3390/brainsci11070848] [Medline: 34202162]
28. Smythe T, Freeze L, Cuthel A, Flowers M, Seghers F, Zia N, et al. Provision of rehabilitation for congenital conditions. Bull World Health Organ 2022 Nov 01;100(11):717-725 [FREE Full text] [doi: 10.2471/BLT.22.288147] [Medline: 36324557]
29. Bernhardt J, Urimubenshi G, Gandhi DB, Eng JJ. Stroke rehabilitation in low-income and middle-income countries: a call to action. Lancet 2020 Oct 31;396(10260):1452-1462. [doi: 10.1016/S0140-6736(20)31313-1] [Medline: 33129396]
30. Rehabilitation in health systems. World Health Organization. 2017 Jan 01. URL: https://www.who.int/publications/i/item/9789241549974 [accessed 2025-08-20]
31. Lorke M, Stefanou T. Improving research and development of wearable assistive and rehabilitation technologies: a systematic review on diversity factors. J Neuroeng Rehabil 2025 Feb 21;22(1):31 [FREE Full text] [doi: 10.1186/s12984-025-01562-7] [Medline: 39985039]
32. Mc Guinness E, Patel D, Patrick M, Austin V. Outcomes of assistive technology use by sex and gender; a scoping review. Disabil Rehabil Assist Technol 2025 Nov;20(8):2666-2690 [FREE Full text] [doi: 10.1080/17483107.2025.2516615] [Medline: 40522876]
33. Spurway C, Dickinson A, Gates L, Metcalf C, Kheng S, Harte C, et al. Examining the changing demand for orthotic service provision from routinely collected digital patient data: a national retrospective descriptive study across three clinics in Cambodia. PLoS One 2025 Dec 09;20(12):e0338461 [FREE Full text] [doi: 10.1371/journal.pone.0338461] [Medline: 41364686]
34. Barth CA. Mind the gap: sex and gender disparities in mobility assistive technology. Assist Technol 2025;37(4):229-230. [doi: 10.1080/10400435.2025.2498281] [Medline: 40424597]
35. Andregård E, Magnusson L. Experiences of attitudes in Sierra Leone from the perspective of people with poliomyelitis and amputations using orthotics and prosthetics. Disabil Rehabil 2017 Dec;39(26):2619-2625 [FREE Full text] [doi: 10.1080/09638288.2016.1236409] [Medline: 27829289]
36. Trafford Z, van der Westhuizen E, McDonald S, Linegar M, Swartz L. More than just assistive devices: how a South African social enterprise supports an environment of inclusion. Int J Environ Res Public Health 2021 Mar 06;18(5):2655 [FREE Full text] [doi: 10.3390/ijerph18052655] [Medline: 33800783]
37. Nuth M. How context influences the ideologies and strategies of disabled people's organisations: a case study from Cambodia. Disabil Soc 2018 May 31;33(7):1046-1060. [doi: 10.1080/09687599.2018.1464433]
38. Rohwerder B. Disability stigma in developing countries. Institute of Development Studies. 2018. URL: https://assets.publishing.service.gov.uk/media/5b18fe3240f0b634aec30791/Disability_stigma_in_developing_countries.pdf [accessed 2026-05-25]
39. Barbareschi G, Kopi NS, Oldfrey B, Holloway C. What difference does tech make? Conceptualizations of disability and assistive technology among Kenyan youth. In: Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility. 2021 Presented at: ASSETS '21; Oct 18-22, 2021; Virtual Event. [doi: 10.1145/3441852.3471226]
40. Federici S, Meloni F, Borsci S. The abandonment of assistive technology in Italy: a survey of National Health Service users. Eur J Phys Rehabil Med 2016 Aug;52(4):516-526 [FREE Full text] [Medline: 26784731]
41. Sugawara AT, Ramos VD, Alfieri FM, Battistella LR. Abandonment of assistive products: assessing abandonment levels and factors that impact on it. Disabil Rehabil Assist Technol 2018 Oct;13(7):716-723. [doi: 10.1080/17483107.2018.1425748] [Medline: 29334475]
42. Palaganas EC, Sanchez MC, Molintas MP, Caricativo RD. Reflexivity in qualitative research: a journey of learning. Qual Rep 2017 Feb 12;22(2):426-438. [doi: 10.46743/2160-3715/2017.2552]
43. Owen C. Dataset to support the University of Southampton doctoral thesis 'Access to healthcare and assistive technologies for people with physical disabilities in Cambodia'. University of Southampton. 2024. URL: https://eprints.soton.ac.uk/489643/ [accessed 2026-07-21]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
LMIC: low- and middle-income country
NGO: nongovernmental organization
P&O: prosthetic and orthotic

XSL•FO
RenderX

Edited by S Munce; submitted 24.Sep.2025; peer-reviewed by Anonymous, S Oh; comments to author 09.Apr.2026; revised version received 27.May.2026; accepted 13.Jul.2026; published 29.Jul.2026.

Please cite as:
Spurway C, Power A, Donovan-Hall M, Channon AA
Experiences of Accessing Physical Rehabilitation Services and Assistive Products Among Orthosis and Wheelchair Users: Qualitative Study
JMIR Rehabil Assist Technol 2026;13:e84735
URL: https://rehab.jmir.org/2026/1/e84735
doi:10.2196/84735
PMID:

©Charlotte Spurway, Andrew Power, Maggie Donovan-Hall, Andrew Amos Channon. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 29.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Basic Technology to Support Stimulation Access, Activity, and Correct Posture in Individuals With Intellectual and Multiple Disabilities: Development and Usability Study

Giulio E Lancioni[1], PhD; Gloria Alberti[1], MA; Valeria Chiariello[1], MA; Elia Zacchini[1], MA; Nirbhay N Singh[2], PhD; Mark F O'Reilly[3], PhD; Jeff Sigafoos[4], PhD; Serafino Buono[5,6], PhD

[1]Lega F. D'Oro Research Center, Osimo, Marche, Italy

[2]Medical College of Georgia, Augusta University, Augusta, GA, United States

[3]Department of Special Education, The University of Texas at Austin, Austin, TX, United States

[4]Faculty of Education, Health, and Psychological Sciences, Victoria University of Wellington, Wellington, Wellington Region, New Zealand

[5]Oasi Research Institute - Istituto di Ricovero e Cura a Carattere Scientifico, Troina, Sicily, Italy

[6]Department of Medicine and Surgery, Università degli Studi di Enna Kore, Enna, Sicily, Italy

Corresponding Author:
Giulio E Lancioni, PhD
Lega F. D'Oro Research Center
Via Linguetta 3
Osimo, Marche, 60027
Italy
Phone: 39 3408323465
Email: giulio.lancioni@uniba.it

Abstract

Background: People with severe to profound intellectual disability and visual or visual and motor impairments often experience passivity and isolation. This can lead to limited environmental stimulation, reduced physical activity, and the acquisition of incorrect body postures. Approaches to improve the situation have generally involved increasing environmental stimulation, which was regulated by the context or controlled by the people through adaptive responding and assistive technology.

Objective: This study assessed 2 intervention programs serving as extensions and advancements of previous approaches aimed at helping participants with intellectual and multiple disabilities control environmental stimulation through their own responses and assistive technology. The first program was designed to help 4 participants with severe intellectual disability and blindness alternate between periods of preferred music stimulation and mild physical activity. The second program was designed to help 5 participants with severe to profound intellectual disability, visual and motor impairments, and postural problems (ie, a tendency to tilt their torso and head forward). These participants were led to access preferred environmental stimulation, engage in mild physical activity, and increase correct posture.

Methods: Each of the 2 programs was implemented according to a single-case research design. The technology for the first program involved (1) a smartphone with the MacroDroid app, which regulated the intervention conditions; (2) pressure sensors that allowed the participants to activate music stimulation; (3) a proximity sensor that monitored the participants' physical activity responses (ie, placing objects in a container high up in front of them); and (4) a mini speaker. The same technology was also used for the second program. In this program, the participants' physical activity responses consisted of arm-stretching movements, and the smartphone also served for monitoring the participants' correct posture. Within each program, a session included 7 music periods and 6 activity periods.

Results: During programs 1 and 2, the mean percentage of music pieces activated independently of research assistants' guidance was between 94.7% and 99.4%. The mean percentage of objects placed in the carton and of shelf-touching responses performed independently was between 98.7% and 99.7%. The baseline mean percentages were between 0% and 24.5%. The mean percentage of session time in a correct posture (program 2) increased to 71.1%-89.6% from baseline values of 18.8%-36.5%. A total of 8 of the 9 participants involved in the 2 programs showed higher levels of indices of happiness during program sessions than control sessions.

XSL•FO
RenderX

Conclusions: The results suggest that the programs might be a useful resource for supporting people with severe to profound intellectual disability and visual or visual and motor impairments.

(JMIR Rehabil Assist Technol 2026;13:e92873) doi:10.2196/92873

KEYWORDS

blindness; intellectual disability; light sensor; MacroDroid application; motor impairment; pressure sensors; proximity sensor; smartphone

Introduction

Background

People with severe or severe to profound intellectual disability and visual or visual and motor impairments tend to be passive and isolated, with limited or no opportunities to control environmental events (and stimulation) around them [1-5]. Their situation may be largely shaped by their inability to locate, recognize, and operate conventional stimulation sources (eg, television and music devices), their lack of self-determination, their poor motor condition, and a general disinterest in activity engagement [4,6-8]. The negative implications of this situation can include a reduced level of environmental stimulation, limited physical activity, and possibly incorrect body postures, that is, a range of conditions that can further hinder people's rehabilitation prospects, general well-being, and quality of life [9-15].

The literature available suggests that several stimulation approaches may be potentially useful in curbing, at least partially, the impact of those negative implications. Two common stimulation approaches involve the use of environmental enrichment and multisensory rooms, respectively. Environmental enrichment consists of an increase in the level of environmental stimulation regulated by staff or other personnel in charge of the context [10,14,16-18]. This approach is considered functional for (1) countering the most feared of the negative implications mentioned above, that is, the risk of low stimulation input (with consequent personal impoverishment), and (2) improving the participants' quality of life [10,19]. Multisensory rooms are structured areas equipped with multiple stimulation sources that are designed to engage the participants' senses and thus enrich their sensory input and possibly provide them with a feeling of satisfaction [20-23].

A third stimulation approach is characterized by the fact that the stimulation is mediated by direct staff and caregiver involvement (intervention). One strategy used within this approach includes an increase in social (verbal and physical) contact between participants and personnel sharing the same context [24-27]. Another strategy exposes the participants to simple stories whose narration by staff is accompanied by the presentation of stimulation events that underline the content of the stories [28-31].

A fourth stimulation approach, contrary to those mentioned above, ensures that the stimulation is controlled by the participants through their own responses and the use of microswitch (sensor) technology [11,13,32,33]. Specifically, the participants are taught to activate a microswitch device through a specific response (eg, a simple hand movement or a hand-stretching movement involving the use of objects). Such a response becomes instrumental in enabling them to access brief periods of preferred stimulation. The distinctive aspects of this approach are that the participants (1) are actively involved in seeking stimulation (rather than being automatically exposed to it) and (2) are required to increase their physical engagement (ie, through their microswitch activation responses) and thus counter their passivity and detachment.

A final stimulation approach differs from the fourth in that its goal is not limited to bringing the participants to seek environmental stimulation through their own responses but also aims to promote their postural control [34]. For example, participants may be (1) taught to perform specific (functional) arm and hand responses to activate microswitch devices and access preferred stimulation and (2) required to maintain a correct posture during the stimulation to ensure that this continues without interruptions [34,35].

In light of the above, one might argue that the last 2 approaches can be considered advantageous because they allow the participants to have an active role in seeking environmental stimulation and because they promote the participants' physical engagement and possibly postural improvement. In essence, these approaches can be viewed as valuable complements to standard occupational therapy and physiotherapy within daily contexts [4,9,12]. With regard to this point, it may also be noted that these approaches rely on the use of basic technology and are only marginally demanding on staff time [4,36,37].

Objectives

The present study was aimed at assessing 2 intervention programs serving as extensions and advancements of the last 2 approaches described above. The first program was directed at 4 participants who presented with severe intellectual disability and blindness. These participants were led to alternate periods in which they could activate preferred music stimulation pieces with periods of mild physical engagement in which a functional response requiring some physical effort was to be repeated several times. Each music stimulation piece lasted 1 minute. Each physical engagement period required the participants to place 6 to 10 objects (ie, common and age-appropriate objects) in a tall carton, which was in front of them. The performance of every response involved forward and upward arm and shoulder stretching (ie, movements that physiotherapists had recommended as useful and appropriate for the participants).

The second program was directed at 5 participants who presented with severe to profound intellectual disability, visual and motor impairments, and postural problems (ie, a tendency to tilt their torso and head forward). These participants were led to alternate stimulation periods such as those described for the

XSL•FO
RenderX

previous group of participants with physical engagement periods. During the latter periods, the participants were required to perform upward and sideward arm-stretching responses (ie, responses that physiotherapists considered useful because of the nature of the movements involved and because they were antagonistic to the participants' postural problems). The participants' posture was monitored throughout the sessions. Loss of correct posture during a stimulation period made the stimulation inaudible until the correct posture was regained.

The programs were developed by the authors, who have long experience working with and developing technology-aided solutions for people with intellectual and multiple disabilities. The authors operated in continuity with previous research work in the area, as indicated above, and in collaboration with staff personnel and caregivers. Each of the 2 programs was implemented according to a single-case research design. The technology used for the programs included commercial components such as a smartphone, pressure and proximity sensors, an interface to connect the sensors to the smartphone, the MacroDroid app, and a mini speaker.

Methods

Ethical Considerations

Staff considered the participants' involvement in the study to be a beneficial and pleasant experience for them (ie, an experience that extended their daily rehabilitation and care treatment by promoting self-determination, enriching stimulation input, and facilitating physical activity). The participants' condition did not allow staff to ask them for or obtain their clear consent to be involved in the study. In light of this, their legal representatives read and signed a consent form on their behalf. The consent form specified that there was no compensation for the participants' involvement, that participants could be withdrawn from the study at any time, and that their data would be deidentified. The study was approved by the Ethics Committee of the Lega F. D'Oro, Osimo (AN), Italy (P031020252). All procedures applied during the study were in line with internationally recognized ethical standards and complied with the 1964 Helsinki Declaration and its later amendments.

Participants

Table 1 reports the chronological age, Vineland age equivalents for daily living skills (personal subdomain), and sensory-motor impairments for the 9 participants. The first 4 (identified with the pseudonyms of Eloise, Grace, Jackson, and Margot) were involved in program 1, and the last 5 (identified with the pseudonyms of Matias, Colton, Alina, Hazel, and Graham) were involved in program 2. Their chronological age ranged from 27 to 59 years. The Vineland age equivalents for daily living skills (personal subdomain) measured via the second edition of the Vineland Adaptive Behavior Scales [38,39] varied between 1 year and 7 months and 4 years. The participants had a diagnosis of congenital encephalopathy and intellectual disability, which was combined with blindness or severe visual impairment. Four of the last 5 participants also presented with motor impairments that precluded any ambulation or allowed only limited levels of supported ambulation. While no recent IQ scores were available, the psychological records of the rehabilitation and care centers that the participants attended reported the participants' level of intellectual disability to be in the severe range (first 4 participants) or the severe-to-profound range (last 5 participants).

The participants were included in the study based on a number of conditions, which had been preliminarily verified. First, they tended to be passive and dependent on staff to access environmental stimulation and engage in physical activity (eg, using objects or performing functional body movements). Second, they seemed to enjoy different types of music stimulation. Third, they were considered capable of responding to simple verbal instructions concerning basic actions such as touching or putting away objects. Fourth, staff were favorable to the programs and the technology involved, which had been illustrated to them in advance.

XSL•FO
RenderX

Table 1. Participants' chronological age, Vineland age equivalents for daily living skills (personal subdomain), and sensory-motor impairments.

Participant (pseudonym)	Chronological age (years)	Vineland age equivalents[a], years (months)	Sensory-motor impairments
Eloise	59	3 (10)	• Blindness
Grace	41	2 (7)	• Blindness
Jackson	35	3 (3)	• Blindness
Margot	50	4 (0)	• Low vision
Matias	58	2 (4)	• Minimal residual vision • Lack of ambulation
Colton	52	2 (8)	• Blindness • Lack of ambulation
Alina	27	3 (1)	• Blindness
Hazel	45	2 (4)	• Minimal residual vision • Lack of ambulation
Graham	48	1 (7)	• Blindness • Lack of ambulation

[a]Age equivalents are based on the Italian standardization of the Vineland Scales [38].

Setting, Sessions, Research Assistants, and Music Stimulation

A quiet room in the participants' daily contexts, fitted with a desk at which the participants sat, served as the setting for the baseline and the program sessions (see below). The participants' regular occupational room served as the setting for the control sessions (see below). Sessions were typically carried out once per day, 5 or 6 days a week. A total of 4 research assistants implemented the sessions. They had a university education and were experienced in applying technology-aided programs with people with extensive disabilities and using various data recording procedures.

Different music pieces (songs and other music stimuli) recommended by staff were evaluated to determine whether they would be used during the study. The evaluation was carried out via a stimulus preference screening procedure. The procedure required that 3 brief segments of each music piece suggested by staff for the individual participants be presented nonconsecutively (at least 10 times per segment) over different screening occasions. Only the music pieces whose segments were considered (ie, by the research assistant and staff member involved in the screening procedure) to produce positive reactions (eg, alerting and smiling) in at least 50% of their presentations were retained for use in the study [4,13].

Technology System: Components

The technology system used for the programs involved (1) a smartphone operating on Android and fitted with a light sensor, (2) a Bluetooth Encore Plus interface (leonardoausili.com), which was linked to 2 pressure sensors and a proximity sensor, and (3) a Bluetooth mini speaker (JBL Go 3) amplifying the smartphone's volume. The smartphone was connected to the internet and fitted with the Encore Plus and MacroDroid apps. The pressure sensors consisted of small Buddy buttons, circles with a diameter of 6.3 cm (leonardoausili.com). The proximity sensor consisted of a Big Candy Corn, a triangle with a side of about 10 cm (leonardoausili.com). In program 1, the pressure sensors were on the desk while the proximity sensor was fixed to the carton in which the participants were to place common, age-appropriate objects taken from a container (Figure 1). The smartphone was in the desk drawer. In program 2, the pressure sensors were on the desk, the proximity sensor was on a small shelf that the participants were to touch, and the smartphone was attached to the back of their chair (Figure 2) or to their wheelchair's headrest.

In program 2, the smartphone on the back of the chair or wheelchair's headrest served to record whether the participants' torso or head posture was correct or incorrect. The closer the participant's torso or head was to the smartphone's light sensor, the lower the level of light reaching the sensor. Based on this relationship, the smartphone was programmed to record the posture as correct up to a prefixed level of light reaching the sensor. Beyond that level, the posture was considered incorrect. This smartphone and light sensor functioning was successfully tested prior to the study by 2 judges (a staff member and a research assistant), who compared the smartphone's (correct or incorrect) posture recording with video images of the posture collected simultaneously. Testing was conducted for each participant under environmental illumination conditions that remained stable during the baseline and program sessions.

XSL•FO
RenderX

Figure 1. Schematic representation of the 2 pressure sensors placed on the desk, the proximity sensor fixed to the carton, and the container with objects.

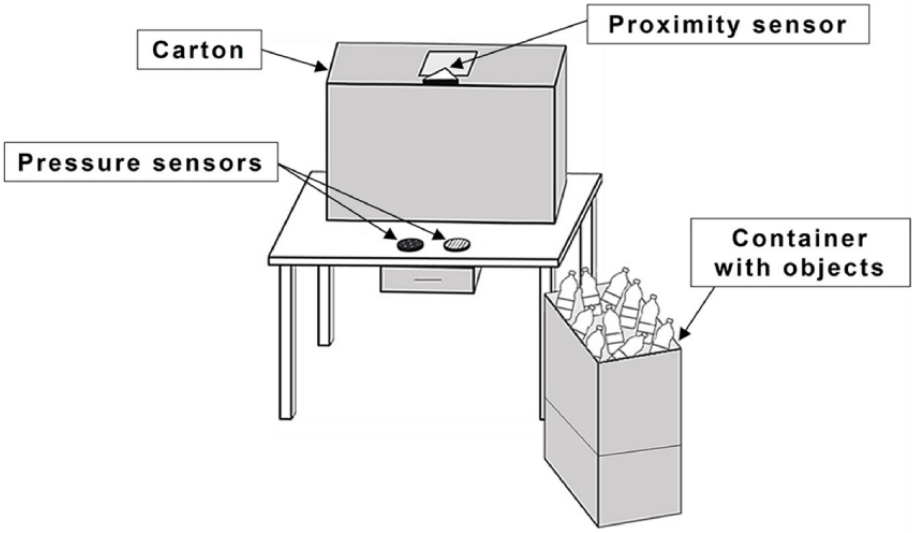
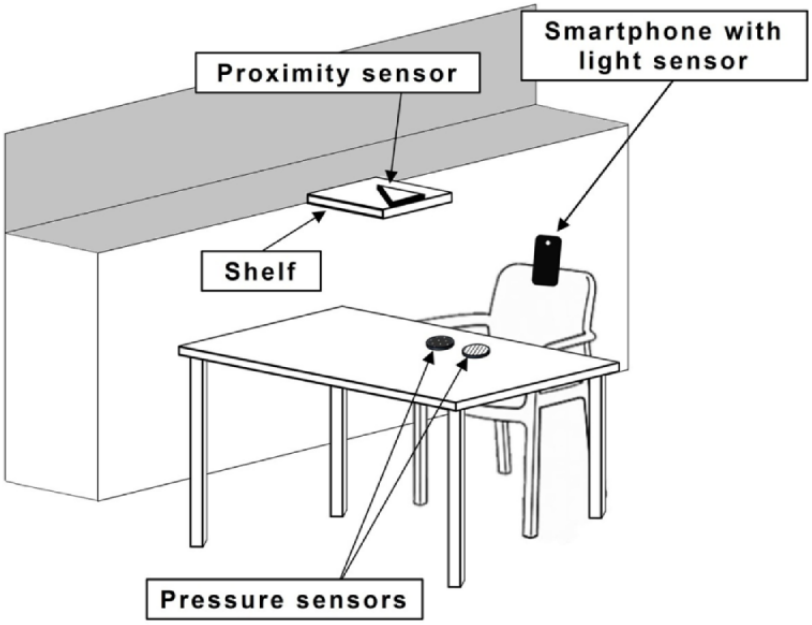


Figure 2. Schematic representation of the 2 pressure sensors placed on the desk, the proximity sensor fixed to the shelf, and the smartphone with the light sensor fixed to the back of the chair.



Technology System: Functioning

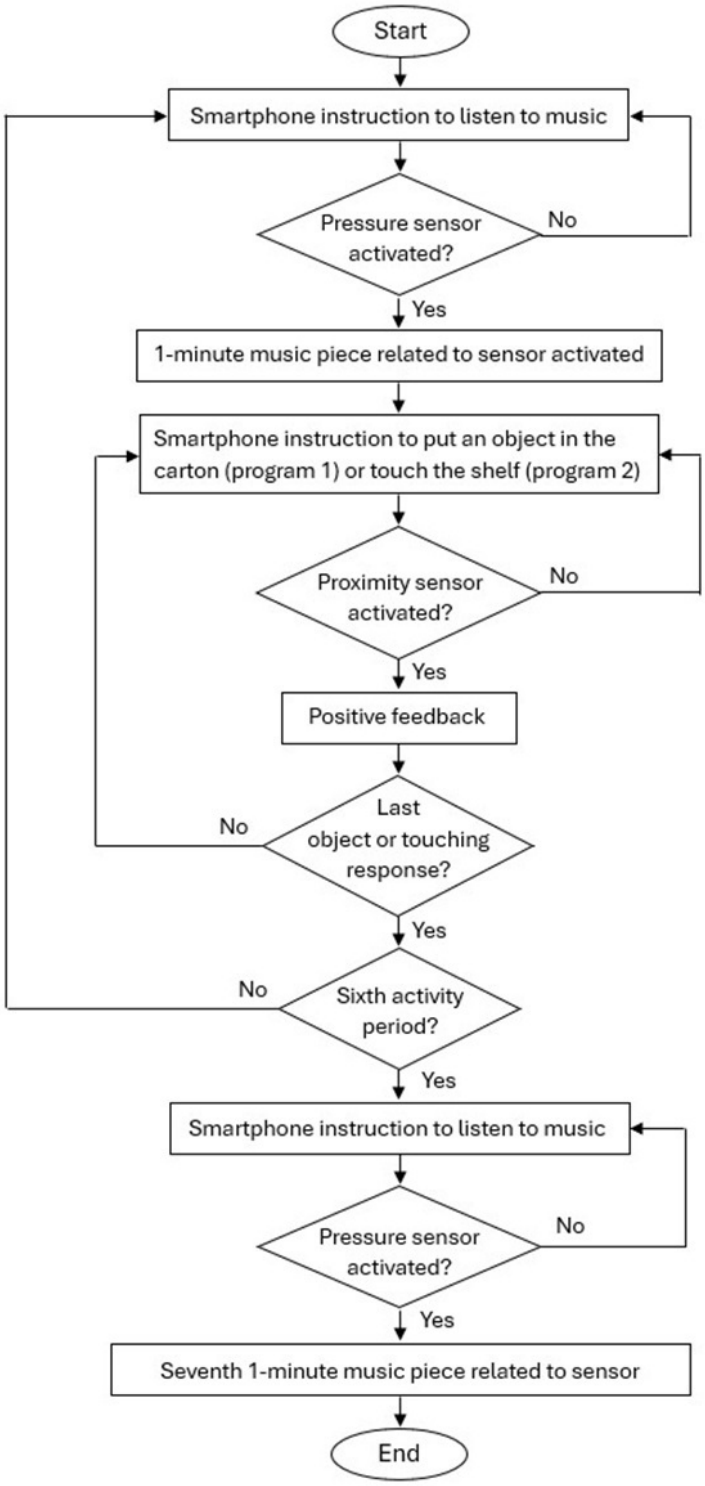
A session with either program started with the smartphone instructing the participants (via the mini speaker) to listen to music (eg, to select the name of a singer or the name of a band). Figure 3 provides a summary of the system's functioning. The participants could touch (activate) one of the pressure (music) sensors available in front of them, thus sending a signal to the smartphone. The smartphone, using the Encore Plus and the MacroDroid apps, (1) detected and discriminated the signal and (2) presented, via the mini speaker, a 1-minute music piece consistent with the sensor activated. In program 2, the piece was audible to the participants if their posture (as determined by the smartphone fixed on the back of their chair or headrest) was correct and inaudible if their posture was incorrect. In the latter case, the smartphone instructed them to straighten their body or raise their head. As the participants regained the correct posture, the music piece became audible again.

Following the end of the music piece, an activity period started during which the music sensors were unresponsive. The smartphone instructed the participants to put an object in the carton (program 1) or touch the shelf (program 2). These responses activated the proximity sensor fixed near the carton's opening (program 1) or on the shelf (program 2). Proximity sensor activation led the smartphone (supported by the Encore Plus and the MacroDroid apps) to (1) count each response (ie, an object placed in the carton or a shelf-touching response), (2) provide a one-word positive feedback statement for each response, (3) follow each feedback statement with the instruction to take another object (program 1) or touch the shelf again (program 2), and (4) determine when the participants had produced the number of responses programmed for the activity period (ie, had placed 6-10 objects in the carton or touched the shelf 6-9 times). Once the programmed number was reached, the smartphone instructed the participants to listen to music (ie, as above) and made the proximity sensor on the carton or shelf

XSL•FO
RenderX

unresponsive. The instruction to listen to music or to place an object in the carton or touch the shelf was repeated at intervals of 10 seconds until the participants responded to it. The session continued with new opportunities to access preferred music pieces followed by activity periods according to the conditions described above. A session was programmed to include 7 music pieces and 6 activity periods. During the 6 activity periods, the participants were expected to place a total of 36 to 60 objects in the carton or to touch the shelf a total of 36 to 54 times.

Figure 3. Flowchart summarizing the system's functioning.



Measures and Data Recording

The four measures recorded during both programs involved (1) the number of music pieces activated independently of research assistants' guidance, (2) the number of objects placed in the carton or of shelf-touching responses performed independently of research assistants' guidance, (3) the length of the sessions, and (4) the participants' indices of happiness (eg, behaviors

XSL•FO
RenderX

such as alerting or orienting to the music, smiling, and making music-based rhythmic movements [4,11,40-43]). The fifth measure, recorded only during program 2, concerned the amount of session time the participants spent in a correct posture. With the exception of indices of happiness, the measures were also recorded during the baseline phase that preceded the programs. Indices of happiness were recorded during 25-30 program sessions and as many control sessions (see below).

The first 2 measures were recorded by the research assistants during the baseline and by the technology system during the programs. During the programs, the system's recorded number of music pieces activated, of objects placed in the carton, and of shelf-touching responses performed would be modified (reduced) if research assistants' guidance had occurred in the process. The length of the sessions and the time spent in a correct posture were recorded by the technology system throughout the study. The indices of happiness were recorded from videos of the sessions using a partial-interval strategy in which 10-second intervals were used for observing and 5-second intervals were used for recording [4].

Interrater agreement was assessed over (1) all baseline sessions for the first 2 measures and (2) 50% of the program and control sessions for indices of happiness with the involvement of a reliability observer. The percentage of agreement on the first 2 measures (computed by dividing the number of sessions in which research assistants and the reliability observer had scores differing by no more than 1 unit by the total number of sessions and multiplying by 100%) was 100% for all participants. The percentage of agreement for indices of happiness (computed for the single sessions by dividing the number of intervals with agreement by the total number of intervals and multiplying by 100%) ranged between 87.5% and 100%, with means exceeding 93% for all participants.

Experimental Conditions and Data Analysis

Overview

The first 4 participants (ie, the group involved in program 1) and the last 5 participants (ie, the group involved in program 2) were initially exposed to a baseline phase without the technology system. Then, they entered the scheduled program during which the technology system described above was used. For each of the 2 groups, the program was implemented according to a nonconcurrent multiple-baseline design across participants (ie, the program started after the group members had received different numbers of baseline sessions) [44,45]. Parallel to 25-30 program sessions, both groups received control sessions. The latter sessions, which were implemented immediately before the program sessions, served to determine whether the program sessions had an impact on the participants' indices of happiness [4,46].

Two basic strategies were used to ensure high levels of accuracy in the research assistants' implementation of the sessions (ie, high levels of procedural fidelity [47,48]). First, the research assistants were involved in preliminary training through simulated sessions in which they were to apply the procedural conditions. Such training continued until they had shown the ability to manage those conditions. Second, the research assistants were provided with feedback about their performance during the study. The feedback (which they received from a study coordinator who had access to videos of the sessions) consisted of being informed as to whether their performance was accurate or included inaccuracies that needed to be amended (with the latter being virtually absent and thus confirming the research assistants' procedural correctness).

The Nonoverlap of All Pairs (NAP) method (ie, a method for the analysis of single-case research data [49,50]) was used to assess the effect size (1) of program 1 on the number of music pieces activated, objects placed in the carton, and indices of happiness, and (2) of program 2 on the number of music pieces activated, shelf-touching responses performed, session time spent in a correct posture, and indices of happiness displayed.

Baseline Sessions

The baseline phase included 5-10 sessions for the first 4 participants (who received program 1) and 5-11 sessions for the last 5 participants (who received program 2). The first four participants sat at a desk with (1) a smartphone that could be used for accessing preferred songs, (2) a carton placed in front of them, and (3) a container with objects placed to their right. The last five participants had (1) a smartphone that could be used for accessing preferred songs; (2) a smartphone with a light sensor and the MacroDroid app, fixed on the back of their chair or their wheelchair's headrest (Figure 2), which served to record the session time spent in a correct posture; and (3) a small shelf with a proximity sensor on it, which was fixed high up to their right.

The smartphone that the participants could use was set up to play music pieces from a YouTube playlist. The playlist had been preliminarily arranged for the individual participants based on their preferred music pieces. The playlist was regulated through a play and pause button on the smartphone's screen. To mark the button area for the participants (given their visual impairment), a small cotton ball was placed just below that area. Touching the smartphone screen above the cotton ball (ie, activating the play and pause button) would start or stop the music.

At the beginning of a baseline session, the research assistants helped the first four participants (1) touch the smartphone, locate the cotton ball, and activate the music briefly; (2) touch the carton's opening; and (3) take an object from the container and place it in the carton. The difference for the last 5 participants was that they were helped to touch the shelf, rather than to touch the carton and place an object in it.

After this introductory help, the research assistants instructed (encouraged) the participants to proceed on their own. If the participants managed to activate a music piece, place objects in the carton, or touch the shelf on their own, their responses were recorded as independent of research assistants' guidance. If the participants did not respond to the instruction (encouragement) for 25 seconds, the research assistants provided them with guidance (ie, verbal and/or physical support) to activate a music piece, which was then recorded as dependent on guidance. Any music piece was interrupted after 1 minute. If the participants remained passive for 25 seconds following

XSL•FO
RenderX

the end of the music piece, the research assistants provided them with guidance to put an object in the carton (first 4 participants) or touch the shelf (last 5 participants). The session continued until (1) the participants had activated 7 music pieces and had placed the objects in the carton or performed the shelf-touching responses programmed for them or (2) 4 consecutive instances of research assistants' guidance had occurred.

Program Sessions

Programs 1 and 2 included 59-73 and 61-80 sessions (with differences due to participants' availability), respectively. The participants were provided with the technology system, which worked as described in the Technology System: Components and Technology System: Functioning sections and Figure 3. The programs were introduced by 2 practice sessions during which the research assistants could repeat the smartphone's instructions and help the participants respond to those instructions. During the following sessions, the research assistants would provide verbal and physical guidance if the participants did not respond to a smartphone instruction after this had been presented 3 times. During 25-30 sessions spread over the central and final part of the programs, data recording also involved the participants' indices of happiness (eg, behaviors such as alerting or orienting to the music, making music-based rhythmic movements, smiling, or vocalizing [40-43]). Recording occurred within the first 10 minutes of the sessions so that the data could be compared to those of 10-minute control sessions.

Control Sessions

Each of the 25-30 program sessions in which indices of happiness were recorded was preceded by a 10-minute control session involving the recording of the same indices. During the control sessions, the participants were not involved in specific activities but had objects available to them and could have some interaction with staff and listen to staff interacting with other individuals with intellectual and visual disabilities sharing the same context. Recording the participants' indices of happiness during program sessions and parallel control sessions served to determine whether the programs had an impact on those indices.

Results

Figure 4 reports the data for the first 4 participants on music pieces activated and objects placed in the carton independently during program 1 and the preceding baseline sessions. Figure 5 reports the data for the last 5 participants on music pieces activated and shelf-touching responses performed independently and for session time spent in a correct posture during program 2 and the preceding baseline sessions. The data points appearing in the figures are mean percentage values computed over blocks of 2 sessions (blocks of 3 sessions are marked with an arrow). Percentage values were computed over (1) the music pieces played, (2) the objects used or shelf-touching responses completed during the activity periods, and (3) the session durations (ie, for correct posture). During the program sessions (which were never interrupted), 7 music pieces were played and 6 activity periods were regularly completed (involving totals of 36-60 objects or 36-54 shelf-touching responses). During the baseline sessions (which were always interrupted), the music pieces played and activity periods started could vary between 2 and 6 (with variable numbers of response opportunities).

During the baseline sessions preceding the introduction of programs 1 and 2, the mean percentage of music pieces activated independently of research assistants' guidance was 0% except for Margot and Hazel. Their mean percentage was 10.7% and 3.5%, respectively. The mean percentage of objects placed in the carton (Figure 4) varied between 0% (Eloise and Margot) and 20.5% (Jackson), with the values of the last 2 or 3 sessions never exceeding the highest values reached in the phase. The mean percentage of shelf-touching responses (Figure 5) ranged between 8.7% (Matias) and 24.5% (Colton). The mean percentage of session time spent in a correct posture (Figure 5) was between about 18.8% (Colton) and 36.5% (Alina). All baseline sessions were interrupted after 4 consecutive instances of research assistants' guidance. Their mean length was always below 9 minutes.

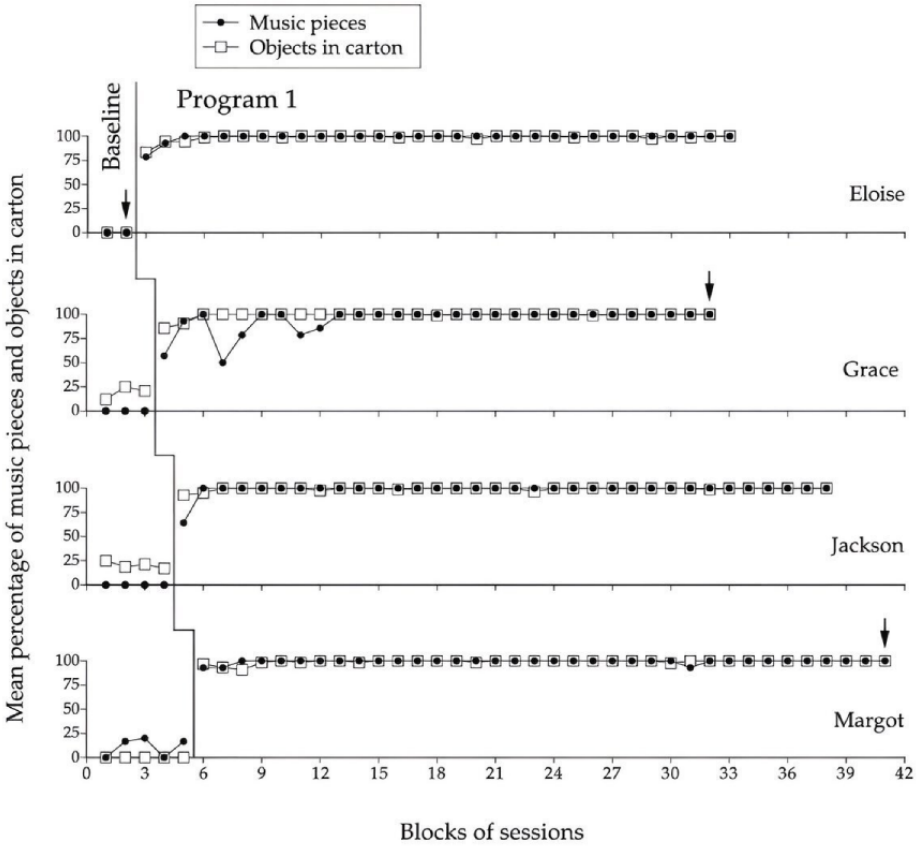
During programs 1 and 2, the mean percentage of music pieces activated independently of research assistants' guidance was between 97% and 99.4%, except in the case of Grace, whose level was 94.7%. The mean percentage of objects placed in the carton and shelf-touching responses performed independently was between 98.7% and 99.7%. It may be noted that the lower percentage values were basically concentrated in the initial sessions of the programs, except in the case of Grace (Figure 4). Her lower values for music pieces activated independently spread beyond the initial sessions as she (1) could continue placing objects in the carton without responding to the smartphone's instruction to listen to music and thus (2) would receive research assistants' guidance. The mean percentage of session time spent in a correct posture (Figure 5) increased to between about 71.1% (Hazel) and 89.6% (Matias). The mean length of the sessions for the 9 participants ranged between about 12 minutes (Jackson) and 16 minutes (Colton and Hazel).

The NAP method used to compare the program and baseline data on the aforementioned measures showed indices of 1 (as the program data on each measure showed percentage values higher than the baseline values for all participants). These indices, together with the fact that the change from the baseline to the program sessions was rapid and clearly visible, may be taken to confirm the large effect size of the programs.

Figures 6 and 7 report the percentage of intervals with indices of happiness during program and control sessions for the first 4 participants and the last 5 participants, respectively. The mean percentage ranged from 7.3% (Eloise) to 64.3% (Margot) during the sessions of program 1 and from 16.8% (Colton) to 37% (Graham) during the sessions of program 2. The mean percentage during the control sessions was below 10% except for Graham (whose mean percentage was 10.2%). The NAP method used to compare the data from the program sessions with the data from the control sessions showed that the programs had a strong effect (with indices of 0.93 to 1) for Jackson, Margot, Matias, Alina, and Graham, a moderate effect (with indices of 0.82 and 0.90) for Colton and Hazel, and no effect (with an index of 0.56) for Eloise.

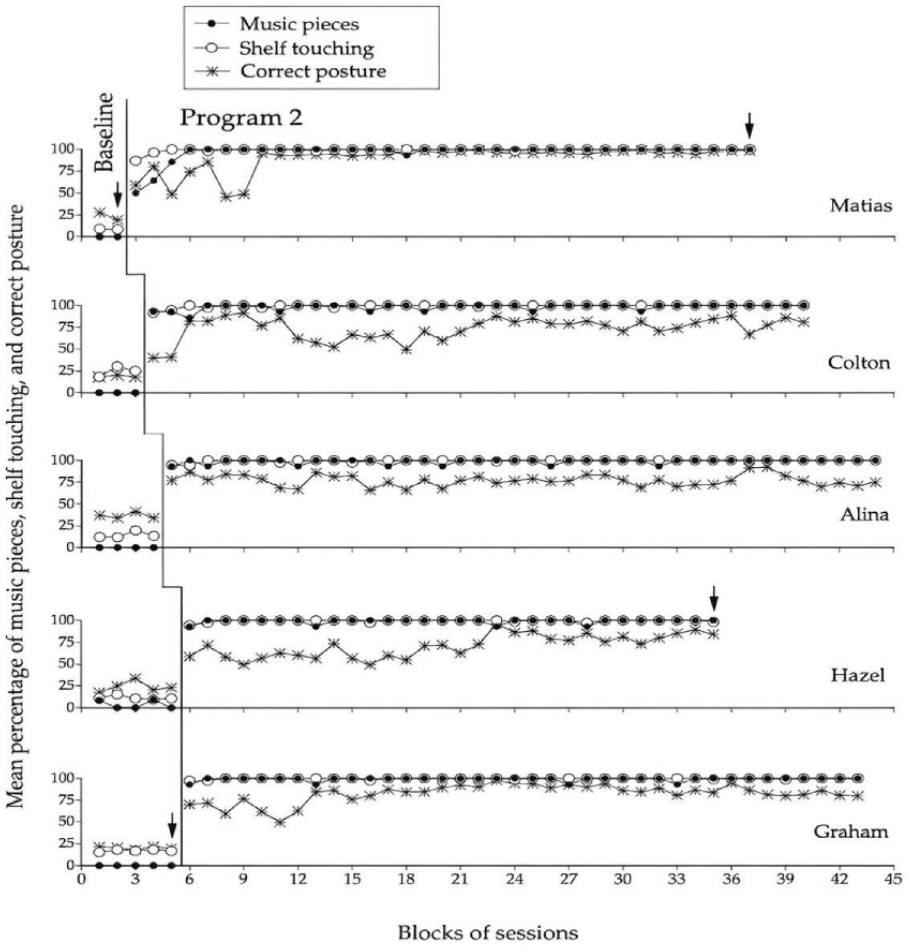
XSL•FO
RenderX

Figure 4. Data for the 4 participants involved in program 1. Dots and open squares represent the mean percentage of music pieces activated and objects placed in the carton independently of research assistants' guidance, respectively, over 2-session blocks. An arrow indicates 3-session blocks.



XSL•FO
RenderX

Figure 5. Data for the 5 participants involved in program 2. Dots, open circles, and asterisks represent the mean percentage of music pieces activated, shelf-touching responses performed independently of research assistants' guidance, and session time spent in a correct posture, respectively. Data are plotted as in Figure 4.



XSL•FO
RenderX

Figure 6. Data for the 4 participants involved in program 1. Black and open triangles indicate the percentage of intervals with indices of happiness during program and control sessions, respectively.

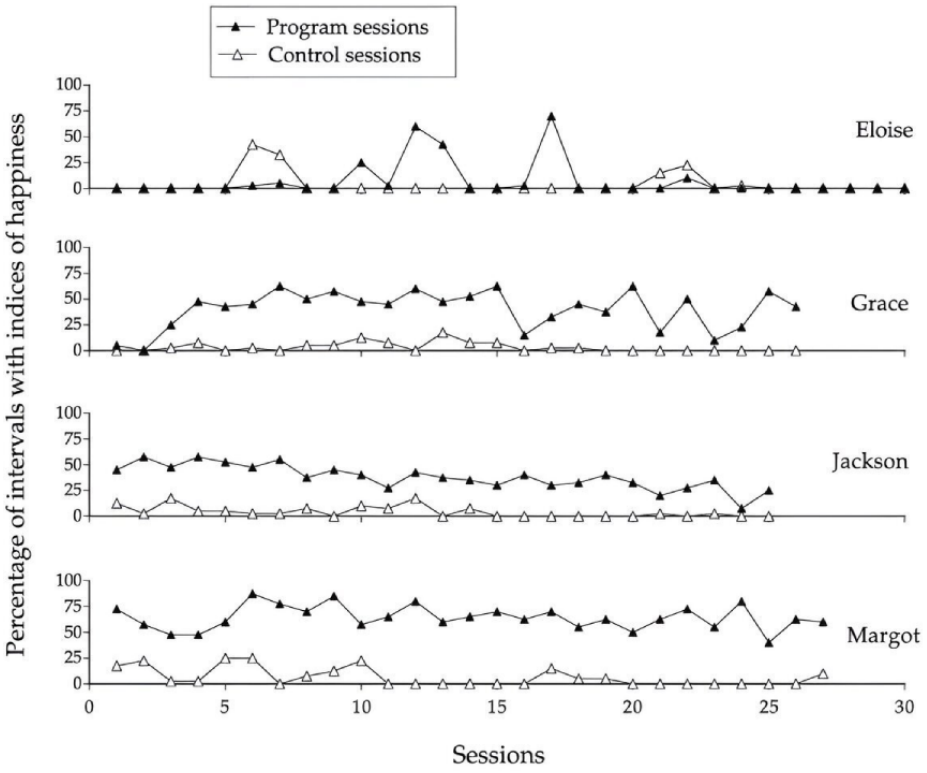
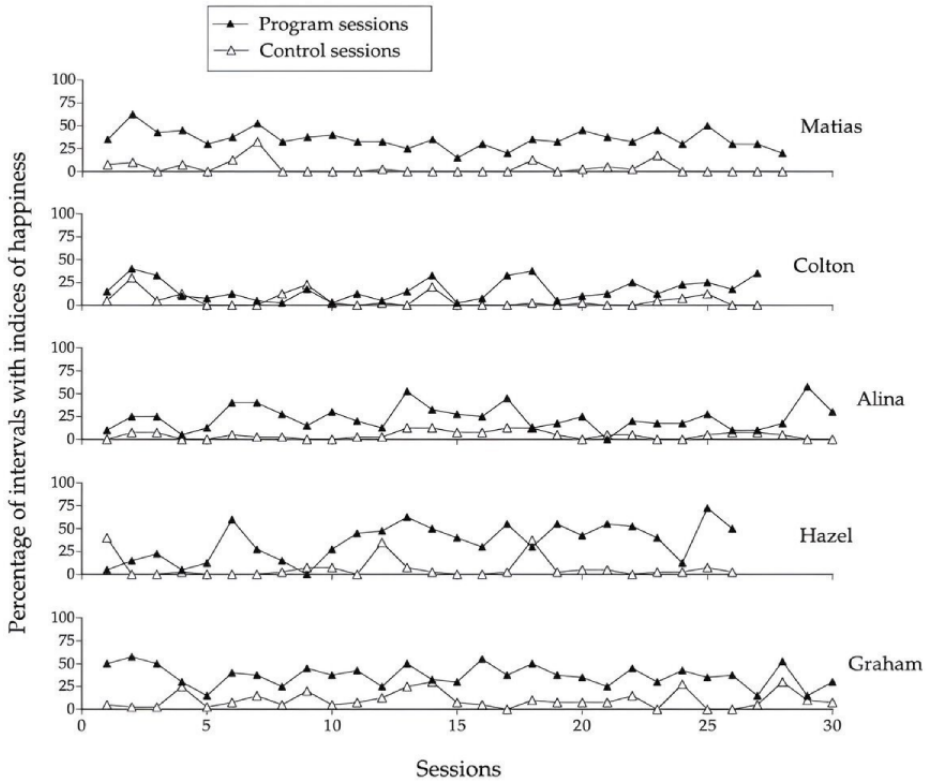


Figure 7. Data for the 5 participants involved in program 2. Black and open triangles indicate the percentage of intervals with indices of happiness during program and control sessions, respectively.



XSL•FO
RenderX

Discussion

Principal Findings

The findings suggest that the programs assessed in this study were helpful in improving the situation of people with intellectual disability and visual or visual and motor impairments. All participants learned to activate music stimulation and to engage in recommended forms of mild physical activity. The participants involved in program 2 also showed an improvement in body posture. Taken as a whole, these findings seem to confirm and extend previous data in the area [33,35,51,52] and to provide new encouraging evidence about the possibility of developing technology-aided programs to help participants with extensive disabilities become responsible for increasing their stimulation input and enhancing their physical activity and postural control. In light of these findings, several considerations are in order.

First, helping people with extensive disabilities manage their access to environmental stimulation and the performance of functional physical actions may be considered relevant from a rehabilitation standpoint and from a practical standpoint. Regarding the rehabilitation aspect, one can argue that the participants' behavioral activation and self-determination not only curb the risks of stimulation impoverishment and limited physical activity but also likely provide a basis for the participants' developmental progress, well-being, and satisfaction [15,53-57]. Regarding the practical aspect, one may consider the findings a demonstration that relevant participants' achievements can be ensured (1) through the support of rather simple and easily accessible technology and (2) with relatively small staff time investment [3,9,36,58-60].

Second, interspersing physical activity with access to preferred music stimulation was thought to be a useful strategy to (1) deal with the participants' low level of interest in and tendency to avoid such activity (ie, a generally nonpreferred type of engagement) and (2) create a link between the activity and the following preferred stimulation with the likelihood that this would become a reinforcing (motivating) factor for the participants' activity engagement [61-64]. Pursuing an improvement in body posture could be considered a useful supplement to the goal of increasing physical activity in participants with motor impairments and postural problems. Bringing participants to control their posture through motivational strategies may be the only way to ensure the independence of such control from external supervision [62,64-67]. The aforementioned hypotheses on promoting physical activity and postural control seem to find support in the participants' performance during the 2 programs. Interpretation of the data reported on postural control may require some caution as the sensor (ie, the smartphone's light sensor) used to record body posture is not a conventional (validated) instrument for use in that specific assessment area.

Third, while the importance of increasing physical activity in people with extensive disabilities is largely acknowledged [9,10,36,37,51,58-60], the evaluation of the beneficial effects of such an increase is difficult or practically impossible when the activity is mild and lasts for relatively short periods of time. In this case, in fact, one cannot rely on any specific assessment criteria and tools. A similar consideration can be extended to the issue of postural control. Identification of functional assessment criteria and tools may be viewed as an important research objective in this area.

Fourth, the data on indices of happiness show that 8 of the participants found their involvement in the program sessions more satisfactory (enjoyable) than their involvement in the control sessions (ie, sessions featuring common conditions in periods of no specific activity or direct supervision). This evidence, together with the findings that all participants had high and stable levels of performance throughout the program sessions, suggests that (1) the program conditions were largely motivating, (2) the positive impact of the stimulation available in the program sessions exceeded the cost (efforts) of the physical activity or postural control pursued within those sessions, and (3) technology-aided programs such as those reported might help promote participants' initiative and functional occupational engagement and foster their developmental progress, with possibly positive implications for their quality of life [4,33,37,40,43,57,61,63].

Fifth, the components of the technology system used in the programs, that is, the smartphone, the Encore Plus interface and app, the pressure sensors, the proximity sensor, the mini speaker, and the MacroDroid app, are commercially available and easily affordable. Commercial availability and affordability make the acquisition of the technology and the application of the programs reported in this study feasible for daily contexts [68,69] as well. One clarification that should be provided at this point is that the acquisition of the technology components does not allow one to have a readily usable system. In fact, the usability of the system depends on the programming of its functioning through the MacroDroid app. This programming, albeit not particularly complex, requires a certain level of familiarity with technology issues. A similar point could be made regarding the implementation of the programs. Although not difficult, it requires staff to acquire a clear understanding of the procedural conditions and the technology arrangement. The authors of this study are willing to provide assistance to any context interested in adopting and using the technology system and programs reported.

Sixth, the fact that the system does not require adjustments from one session to the next and the evidence that the participants can be independent of any external assistance during the sessions can lead one to basic considerations about program applicability and sustainability [70-72]. For example, one might argue that (1) the participants involved in this study could continue to use the programs reported over time with relatively small staff time investment, and (2) the same programs could represent a realistic treatment opportunity for other people with extensive disabilities and limited access to stimulation and physical activity (ie, a treatment opportunity that their daily contexts might be able to afford in terms of time cost and technology requirements).

Limitations and Future Research

While the results appear promising, some caution may be required in interpreting them given the limitations of the study. The most obvious limitations are the relatively small number

XSL•FO
RenderX

of participants, the unavailability of maintenance and generalization data, and the absence of social validation of the programs and the technology system involved. An additional limitation may be the use of an experimental sensor (ie, the smartphone's light sensor) to record the participants' posture. With regard to the small number of participants, 2 points may need to be clarified. The first point is that the single-case research designs used with the participants involved in program 1 and the participants involved in program 2 can be considered adequate to support the internal validity of the data reported [44,45,73]. The second point is that new studies (ie, single-case replication studies and/or group-design studies) would be critically important to determine the external validity of the data and thus their level of representativeness [73-76].

To deal with the issues of maintenance and generalization, more comprehensive studies are needed. Those studies would have to (1) extend the data collection periods and include follow-up checks to verify whether the participants' gains remain stable over time, (2) test the effects of the programs with different groups of people so as to verify whether those effects can be reproduced across participants, and (3) include different contexts in the evaluations (eg, homes, schools, and community and rehabilitation settings) to ascertain the applicability and effectiveness of those programs across settings as well as the sustainability of the programs within those settings [61,63,70,77,78].

To address the social validation question, new studies may need to gather the opinions of staff, family members, and service providers about the acceptability, relevance, and usability of the programs and the technology system on which they rely [79,80]. The process of gathering such opinions may involve 2 steps. The first step would consist of presenting all these personnel with videos of program sessions in which participants use the technology system. The second step would consist of presenting the same personnel with a series of questions concerning the programs and the technology (eg, their impact, social relevance, and applicability) [81-83]. The aforementioned process may be complemented by formal scales assessing the usability of technology (eg, the System Usability Scale) [84,85].

With regard to the recording of body posture, new studies may need to evaluate the usability and practicality of accelerometers, depth cameras, or infrared optical tracking sensors [86-88]. New studies may also consider ways of measuring (1) the impact of mild physical activity such as that adopted in the 2 programs of this study and (2) the level of purposefulness in the participants' selection of the music pieces (ie, in their decision about the pressure sensor to activate at each music opportunity). To address the first issue, one might need to identify evaluation criteria and instruments more sophisticated than those presently available [9,51,58,64]. To address the second issue, one may need to associate highly preferred music pieces with one sensor and moderately preferred music pieces with the other sensor and verify whether the activation of the former sensor is more frequent than the activation of the latter sensor [13,89-91].

Conclusions

The findings suggest that the programs assessed in this study were helpful in improving the situation of people with intellectual disability and visual or visual and motor impairments. All participants learned to activate music stimulation and to engage in recommended forms of physical activity. The participants involved in program 2 also showed an improvement in body posture. These findings provide new evidence of the possibility of developing technology-aided programs to improve the situation of participants with extensive disabilities. Notwithstanding this evidence, caution is required in making general statements and drawing conclusions based on these findings given the limitations of this study, which need to be addressed through new, carefully planned research efforts.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated or analyzed during this study are available in the OneDrive repository [92].

Authors' Contributions

GEL was responsible for setting up the study, acquiring and analyzing the data, and writing the manuscript. GA, VC, and EZ collaborated in setting up the study and the technology system, acquiring and analyzing the data, and editing the manuscript. NNS, MFOR, JS, and SB collaborated in setting up the study, analyzing the data, and writing and editing the manuscript.

Conflicts of Interest

None declared.

References

1. Dijkhuizen A, Hilgenkamp TI, Krijnen WP, van der Schans CP, Waninge A. The impact of visual impairment on the ability to perform activities of daily living for persons with severe/profound intellectual disability. Res Dev Disabil 2016;48:35-42. [doi: 10.1016/j.ridd.2015.10.001] [Medline: 26529065]

XSL•FO
RenderX

2. Evenhuis HM, Sjoukes L, Koot HM, Kooijman AC. Does visual impairment lead to additional disability in adults with intellectual disabilities? J Intellect Disabil Res 2009;53(1):19-28. [doi: 10.1111/j.1365-2788.2008.01114.x] [Medline: 18771511]
3. Hanzen G, van Nispen RMA, Vlaskamp C, Korevaar EL, Waninge A, van der Putten AAJ. Improving the participation of adults with visual and severe or profound intellectual disabilities: a process evaluation of a new intervention. BMC Health Serv Res 2020;20(1):319 [FREE Full text] [doi: 10.1186/s12913-020-05161-1] [Medline: 32299453]
4. Lancioni GE, Singh NN, O'Reilly MF, Sigafoos J, Resta E, Campodonico F, et al. Self-regulated versus staff-regulated stimulation for promoting indices of satisfaction in persons with severe/profound and multiple disabilities. J Dev Phys Disabil 2020;33(1):137-152. [doi: 10.1007/s10882-020-09737-x]
5. Munde VS, Vlaskamp C, Maes B, Ruijssenaars AJJM. Catch the wave! Time-window sequential analysis of alertness stimulation in individuals with profound intellectual and multiple disabilities. Child Care Health Dev 2014;40(1):95-105. [doi: 10.1111/j.1365-2214.2012.01415.x] [Medline: 22774925]
6. Jarjoura W. Disorientation and loss of wayfinding in individuals with congenital blindness and other affecting comorbidities. Br J Vis Impair 2019;37(3):240-247. [doi: 10.1177/0264619619861653]
7. Kroezen M, Van DAN, Van GM, Wolkorte R. Visual impairment and blindness in people with intellectual disabilities: a systematic review. Optom Vis Perform 2020;8:109-121 [FREE Full text]
8. van der Aa H, Jonker F, de Looff P, Didden R. A study into the relationship between adaptive skills and visual impairment in individuals with and without intellectual disabilities. J Vis Impair Blind 2023;117(3):224-232. [doi: 10.1177/0145482x231180615]
9. Burns J, Carter A, Draper S, Foad A. Engaging and sustaining people with intellectual disabilities in physical activity: a narrative review of existing evidence. Int J Dev Disabil 2024;70(5):803-813 [FREE Full text] [doi: 10.1080/20473869.2022.2149096] [Medline: 39131751]
10. De Giorgio A. The roles of motor activity and environmental enrichment in intellectual disability. Somatosens Mot Res 2017;34(1):34-43. [doi: 10.1080/08990220.2016.1278204] [Medline: 28140743]
11. Lancioni GE, Alberti G, Filippini C, Singh NN, O'Reilly MF, Sigafoos J, et al. A technology system to help people with multiple disabilities increase contact with objects and control environmental stimulation: single-case research design. JMIR Rehabil Assist Technol 2025;12:e70378 [FREE Full text] [doi: 10.2196/70378] [Medline: 40117571]
12. Shih CH, Shih CJ, Shih CT. Assisting people with multiple disabilities by actively keeping the head in an upright position with a Nintendo Wii Remote controller through the control of environmental stimulation. Res Dev Disabil 2011;32(5):2005-2010. [doi: 10.1016/j.ridd.2011.04.008] [Medline: 21601420]
13. Tam GM, Phillips KJ, Mudford OC. Teaching individuals with profound multiple disabilities to access preferred stimuli with multiple microswitches. Res Dev Disabil 2011;32(6):2352-2361. [doi: 10.1016/j.ridd.2011.07.027] [Medline: 21824746]
14. Villouta-Gutiérrez O, Pérez-Villalobos C, Rojas-Ponce R, Sáez-Delgado F. Environmental enrichment and intellectual disability: systematic review of neurocognitive effects in children and adolescents. Rev Mex Neuroci 2022;23(1):34-43. [doi: 10.24875/rmn.21000012]
15. Hammell KW. Quality of life, participation and occupational rights: a capabilities perspective. Aust Occup Ther J 2015;62(2):78-85. [doi: 10.1111/1440-1630.12183] [Medline: 25656544]
16. Ball NJ, Mercado E, Orduña I. Enriched environments as a potential treatment for developmental disorders: a critical assessment. Front Psychol 2019;10:466 [FREE Full text] [doi: 10.3389/fpsyg.2019.00466] [Medline: 30894830]
17. Downs J, Rodger J, Li C, Tan X, Hu N, Wong K, et al. Environmental enrichment intervention for Rett syndrome: an individually randomised stepped wedge trial. Orphanet J Rare Dis 2018;13(1):3 [FREE Full text] [doi: 10.1186/s13023-017-0752-8] [Medline: 29321033]
18. Sansevero G, Sale A. Environment as therapy: neuroscience for intellectual disability and dementia. Oncotarget 2017;8(4):5682-5683 [FREE Full text] [doi: 10.18632/oncotarget.14602] [Medline: 28086230]
19. Simões C, Santos S. The impact of personal and environmental characteristics on quality of life of people with intellectual disability. Appl Res Qual Life 2016;12(2):389-408. [doi: 10.1007/s11482-016-9466-7]
20. Carter M, Stephenson J. The use of multi-sensory environments in schools servicing children with severe disabilities. J Dev Phys Disabil 2011;24(1):95-109. [doi: 10.1007/s10882-011-9257-x]
21. Chan SW, Thompson DR, Chau JP, Tam WW, Chiu IW, Lo SH. The effects of multisensory therapy on behaviour of adult clients with developmental disabilities--a systematic review. Int J Nurs Stud 2010 Jan;47(1):108-122. [doi: 10.1016/j.ijnurstu.2009.08.004] [Medline: 19762017]
22. Fava L, Strauss K. Multi-sensory rooms: comparing effects of the snoezelen and the stimulus preference environment on the behavior of adults with profound mental retardation. Res Dev Disabil 2010;31(1):160-171. [doi: 10.1016/j.ridd.2009.08.006] [Medline: 19815373]
23. Stephenson J, Carter M. The use of multisensory environments with individuals with developmental disabilities: a systematic review. J Dev Phys Disabil 2024;37(5):731-757. [doi: 10.1007/s10882-024-09982-4]
24. Hostyn I, Maes B. Interaction with a person with profound intellectual and multiple disabilities: a case study in dialogue with an experienced staff member. J Intellect Dev Disabil 2013;38(3):189-204. [doi: 10.3109/13668250.2013.798400] [Medline: 23984881]

XSL•FO
RenderX

25. Kamstra A, van der Putten AAJ, Post WJ, Vlaskamp C. Informal social networks of people with profound intellectual and multiple disabilities: relationship with age, communicative abilities and current living arrangements. J Appl Res Intellect Disabil 2015;28(2):159-164. [doi: 10.1111/jar.12115] [Medline: 25131857]
26. Kamstra A, van der Putten AAJ, Vlaskamp C. The structure of informal social networks of persons with profound intellectual and multiple disabilities. J Appl Res Intellect Disabil 2015;28(3):249-256. [doi: 10.1111/jar.12134] [Medline: 25431193]
27. Munde V, Vlaskamp C. Initiation of activities and alertness in individuals with profound intellectual and multiple disabilities. J Intellect Disabil Res 2015;59(3):284-292. [doi: 10.1111/jir.12138] [Medline: 24806398]
28. Matos A, Rocha T, Cabral L, Bessa M. Multi-sensory storytelling to support learning for people with intellectual disability: an exploratory didactic study. Procedia Comput Sci 2015;67:12-18. [doi: 10.1016/j.procs.2015.09.244]
29. Penne A, Ten Brug A, Munde V, van der Putten A, Vlaskamp C, Maes B. Staff interactive style during multisensory storytelling with persons with profound intellectual and multiple disabilities. J Intellect Disabil Res 2012;56(2):167-178. [doi: 10.1111/j.1365-2788.2011.01448.x] [Medline: 21726327]
30. Ten Brug A, Van der Putten AA, Penne A, Maes B, Vlaskamp C. Factors influencing attentiveness of people with profound intellectual and multiple disabilities to multisensory storytelling. Policy Practice Intel Disabi 2015;12(3):190-198. [doi: 10.1111/jppi.12128]
31. Ten Brug A, Van der Putten AAJ, Penne A, Maes B, Vlaskamp C. Making a difference? A comparison between multi-sensory and regular storytelling for persons with profound intellectual and multiple disabilities. J Intellect Disabil Res 2016;60(11):1043-1053. [doi: 10.1111/jir.12260] [Medline: 27038012]
32. Lancioni GE, Singh NN, O'Reilly MF, Sigafoos J, Campodonico F, Oliva D, et al. Using microswitch-aided programs for people with multiple disabilities to promote stimulation control and mild physical exercise. J Intellect Dev Disabil 2016;43(2):242-250. [doi: 10.3109/13668250.2016.1253831]
33. Roche L, Sigafoos J, Lancioni GE, O'Reilly MF, Green VA. Microswitch technology for enabling self-determined responding in children with profound and multiple disabilities: a systematic review. Augment Altern Commun 2015;31(3):246-258. [doi: 10.3109/07434618.2015.1024888] [Medline: 25791421]
34. Lancioni GE, Singh NN, O'Reilly MF, Sigafoos J, Perilli V, Campodonico F, et al. Persons with multiple disabilities engage in stimulus choice and postural control with the support of a technology-aided program. Behav Modif 2015;39(3):454-471. [doi: 10.1177/0145445515572187] [Medline: 25733662]
35. Lancioni GE, Singh NN, O'Reilly MF, Sigafoos J, Alberti G, Campodonico F, et al. Everyday technology to help people with intellectual and other disabilities access stimulation via functional motor responses and improved body posture. Dev Neurorehabil 2022;25(1):59-67. [doi: 10.1080/17518423.2021.1989513] [Medline: 34629013]
36. Lancioni GE, Singh NN, O'Reilly M, Sigafoos J, Alberti G, Desideri L. Programs using stimulation-regulating technologies to promote physical activity in people with intellectual and multiple disabilities: scoping review. JMIR Rehabil Assist Technol 2022;9(2):e35217 [FREE Full text] [doi: 10.2196/35217] [Medline: 35389365]
37. Shih C, Shih C, Luo C. Assisting people with disabilities in actively performing physical activities by controlling the preferred environmental stimulation with a gyration air mouse. Res Dev Disabil 2013;34(12):4328-4333. [doi: 10.1016/j.ridd.2013.09.001] [Medline: 24139710]
38. Balboni G, Belacchi C, Bonichini S, Coscarelli A. Vineland-II: Scale Vineland del Comportamento Adattivo. Seconda edizione. Florence, Italy: Giunti Psychometrics; 2016.
39. Sparrow SS, Balla DA, Cicchetti DV. Vineland Adaptive Behavior Scales, Second Edition (Vineland-II): Survey Forms Manual. Circle Pines, MN: AGS Publishing; 2005.
40. Agarwal T, Ramey D. Identifying and validating indices of happiness and unhappiness in autistic children and examining them as quality of life indicators during presession pairing. Adv Neurodev Disord 2023;8(3):429-444. [doi: 10.1007/s41252-023-00352-3]
41. Darling JA, Circo DK. Measuring happiness in individuals with profound multiple disabilities. Res Dev Disabil 2015;47:117-125. [doi: 10.1016/j.ridd.2015.09.005] [Medline: 26476484]
42. Dillon CM, Carr JE. Assessing indices of happiness and unhappiness in individuals with developmental disabilities: a review. Behav Intervent 2007;22(3):229-244. [doi: 10.1002/bin.240]
43. Ramey D, Healy O, McEnaney E. Defining and measuring indices of happiness and unhappiness in children diagnosed with autism spectrum disorder. Behav Anal Pract 2023;16(1):194-209 [FREE Full text] [doi: 10.1007/s40617-022-00710-y] [Medline: 37006433]
44. Slocum TA, Pinkelman SE, Joslyn PR, Nichols B. Threats to internal validity in multiple-baseline design variations. Perspect Behav Sci 2022;45(3):619-638 [FREE Full text] [doi: 10.1007/s40614-022-00326-1] [Medline: 36249165]
45. Slocum TA, Joslyn PR, Nichols B, Pinkelman SE. Revisiting an analysis of threats to internal validity in multiple baseline designs. Perspect Behav Sci 2022;45(3):681-694 [FREE Full text] [doi: 10.1007/s40614-022-00351-0] [Medline: 36249172]
46. Stasolla F, Caffò AO, Perilli V, Boccasini A, Damiani R, D'Amico F. Assistive technology for promoting adaptive skills of children with cerebral palsy: ten cases evaluation. Disabil Rehabil Assist Technol 2019;14(5):489-502. [doi: 10.1080/17483107.2018.1467972] [Medline: 29732901]
47. Morris C, Jones SH, Oliveira JP. A practitioner's guide to measuring procedural fidelity. Behav Anal Pract 2024;17(2):643-655. [doi: 10.1007/s40617-024-00910-8] [Medline: 38966272]

XSL•FO
RenderX

48. Strain P, Fox L, Barton EE. On expanding the definition and use of procedural fidelity. Res Pract Pers Severe Disabil 2021;46(3):173-183. [doi: 10.1177/15407969211036911]
49. Manolov R, Tanious R. Assessing nonoverlap in single-case data: strengths, challenges, and recommendations. J Behav Educ 2024;34(4):869-901. [doi: 10.1007/s10864-024-09552-w]
50. Parker RI, Vannest KJ, Davis JL. Effect size in single-case research: a review of nine nonoverlap techniques. Behav Modif 2011;35(4):303-322. [doi: 10.1177/0145445511399147] [Medline: 21411481]
51. Enkelaar L, Oosterom-Calo R, Zhou D, Nijhof N, Barakova E, Sterkenburg P. The LEDs move pilot study: the light curtain and physical activity and well-being among people with visual and intellectual disabilities. J Intellect Disabil Res 2021;65(11):971-988. [doi: 10.1111/jir.12882] [Medline: 34542212]
52. Lancioni GE, Singh NN, O'Reilly MF, Sigafoos J, Alberti G, Perilli V, et al. Promoting functional activity engagement in people with multiple disabilities through the use of microswitch-aided programs. Front Public Health 2017;5:205 [FREE Full text] [doi: 10.3389/fpubh.2017.00205] [Medline: 28848730]
53. Caouette M, Lachapelle Y, Moreau J, Lussier‐Desrochers D. Descriptive study of caseworkers' practices to support the development of self-determination of adults with intellectual disabilities. J Policy Pract Intellect Disabil 2018;15(1):4-11. [doi: 10.1111/jppi.12217]
54. Cummins RA. Quality of life of adults with an intellectual disability. Curr Dev Disord Rep 2020;7(3):182-187. [doi: 10.1007/s40474-020-00205-x]
55. Friedman C. Natural supports: the impact on people with intellectual and developmental disabilities' quality of life and service expenditures. J Fam Soc Work 2021;24(2):118-135. [doi: 10.1080/10522158.2020.1861158]
56. Gudelytė U, Ruškus J, McCrea KT. "Help me to decide": a study of human rights-based supported decision making with persons with intellectual disabilities. Am J Orthopsychiatry 2024;94(3):297-310. [doi: 10.1037/ort0000724] [Medline: 38512186]
57. Wehmeyer ML. The importance of self-determination to the quality of life of people with intellectual disability: a perspective. Int J Environ Res Public Health 2020;17(19):7121 [FREE Full text] [doi: 10.3390/ijerph17197121] [Medline: 33003321]
58. Bossink L, van der Putten AA, Vlaskamp C. Understanding low levels of physical activity in people with intellectual disabilities: a systematic review to identify barriers and facilitators. Res Dev Disabil 2017;68:95-110. [doi: 10.1016/j.ridd.2017.06.008] [Medline: 28750208]
59. Jacinto M, Vitorino AS, Palmeira D, Antunes R, Matos R, Ferreira JP, et al. Perceived barriers of physical activity participation in individuals with intellectual disability: a systematic review. Healthcare (Basel) 2021;9(11):1521 [FREE Full text] [doi: 10.3390/healthcare9111521] [Medline: 34828567]
60. Jacob US, Pillay J, Johnson E, Omoya O, Adedokun AP. A systematic review of physical activity: benefits and needs for maintenance of quality of life among adults with intellectual disability. Front Sports Act Living 2023;5:1184946 [FREE Full text] [doi: 10.3389/fspor.2023.1184946] [Medline: 37361407]
61. Cooper JO, Heron TE, Heward WL. Applied Behavior Analysis. 3rd Edition. Hoboken, NJ: Pearson; 2019.
62. Michalsen H, Wangberg SC, Anke A, Hartvigsen G, Jaccheri L, Arntzen C. Family members and health care workers' perspectives on motivational factors of participation in physical activity for people with intellectual disability: a qualitative study. J Intellect Disabil Res 2020;64(4):259-270. [doi: 10.1111/jir.12716] [Medline: 31981261]
63. Pierce WD, Cheney CD. Behavior Analysis and Learning. 6th Edition. New York: Routledge; 2017.
64. Rotta K, Raaymakers A, Poling A. Using positive reinforcement to increase the physical activity of individuals with developmental disabilities: a review. J Dev Phys Disabil 2022;35(4):551-575. [doi: 10.1007/s10882-022-09873-6]
65. McGrane N, Galvin R, Cusack T, Stokes E. Addition of motivational interventions to exercise and traditional physiotherapy: a review and meta-analysis. Physiotherapy 2015;101(1):1-12. [doi: 10.1016/j.physio.2014.04.009] [Medline: 25239472]
66. Obrusnikova I, Cavalier AR, Novak HM, Blair AE. The effect of systematic prompting on the acquisition of two muscle-strengthening exercises by adults with moderate disabilities. J Behav Educ 2020;29(3):584-605 [FREE Full text] [doi: 10.1007/s10864-019-09328-7] [Medline: 33737797]
67. Ptomey L, Washburn R, Lee J, Greene J, Szabo-Reed A, Sherman J, et al. Individual and family-based approaches to increase physical activity in adolescents with intellectual and developmental disabilities: rationale and design for an 18-month randomized trial. Contemp Clin Trials 2019;84:105817 [FREE Full text] [doi: 10.1016/j.cct.2019.105817] [Medline: 31344519]
68. Friedman C. Technology first: the impact on technology availability for people with intellectual and developmental disabilities. J Appl Res Intellect Disabil 2023;36(6):1251-1263. [doi: 10.1111/jar.13144] [Medline: 37482893]
69. Reddy KJ. Accessibility and affordability of technologies. In: Innovations in Neurocognitive Rehabilitation: Harnessing Technology for Effective Therapy. Cham, Switzerland: Springer Nature Switzerland; 2025.
70. Bodkin A, Hakimi S. Sustainable by design: a systematic review of factors for health promotion program sustainability. BMC Public Health 2020;20(1):964 [FREE Full text] [doi: 10.1186/s12889-020-09091-9] [Medline: 32560718]
71. Ceptureanu SI, Ceptureanu EG, Luchian CE, Luchian I. Community based programs sustainability. a multidimensional analysis of sustainability factors. Sustainability 2018;10(3):870. [doi: 10.3390/su10030870]

XSL•FO
RenderX

72. Smith EM, Huff S, Wescott H, Daniel R, Ebuenyi ID, O'Donnell J, et al. Assistive technologies are central to the realization of the Convention on the Rights of Persons with Disabilities. Disabil Rehabil Assist Technol 2024;19(2):486-491 [FREE Full text] [doi: 10.1080/17483107.2022.2099987] [Medline: 35900971]
73. Kazdin AE. Single-Case Research Designs: Methods for Clinical and Applied Settings. 3rd Edition. New York: Oxford University Press; 2020.
74. Coiera E, Tong HL. Replication studies in the clinical decision support literature-frequency, fidelity, and impact. J Am Med Inform Assoc 2021;28(9):1815-1825. [doi: 10.1093/jamia/ocab049] [Medline: 34226931]
75. Tanious R, Manolov R, Onghena P, Vlaeyen JWS. Single-case experimental designs: the importance of randomization and replication. Nat Rev Methods Primers 2024;4(1):27. [doi: 10.1038/s43586-024-00312-8]
76. Walker SG, Carr JE. Generality of findings from single-case designs: it's not all about the "N". Behav Anal Pract 2021;14(4):991-995 [FREE Full text] [doi: 10.1007/s40617-020-00547-3] [Medline: 34868812]
77. Taubner H, Tideman M, Staland Nyman C. Employment sustainability for people with intellectual disability: a systematic review. J Occup Rehabil 2022;32(3):353-364 [FREE Full text] [doi: 10.1007/s10926-021-10020-9] [Medline: 34958449]
78. Pennington B, Simacek J, McComas J, McMaster K, Elmquist M. Maintenance and generalization in functional behavior assessment/behavior intervention plan literature. J Behav Educ 2018;28(1):27-53. [doi: 10.1007/s10864-018-9299-6]
79. Philip VS, Santhosh KE. Efficacy and social validity of intervention using speech generating device and tactile symbols to develop communication and language in a child with visual impairment and additional disability. Disabil Rehabil Assist Technol 2025;20(7):2557-2572. [doi: 10.1080/17483107.2025.2539436] [Medline: 40739816]
80. Quigley SP, Romani PW, Field S, Ellwood G. Social validity assessment. In: Luiselli JK, Bird FL, Maguire H, Gardner RM, editors. Behavior Safety and Clinical Practice in Intellectual and Developmental Disabilities: Evidence-Based Practices in Behavioral Health. Cham, Switzerland: Springer Nature Switzerland; 2024.
81. Snodgrass MR, Cook BG, Cook L. Considering social validity in special education research. Learn Disabil Res Pract 2023;38(4):311-319 [FREE Full text] [doi: 10.1111/ldrp.12326]
82. Stasolla F, Caffò AO, Perilli V, Albano V. Supporting locomotion fluency of six children with Cornelia de Lange syndrome: awareness of microswitch responding and social validation. Technol Disabil 2019;30(4):209-220. [doi: 10.3233/tad-180216]
83. Worthen D, Luiselli JK. Comparative effects and social validation of support strategies to promote mindfulness practices among high school students. Child Fam Behav Ther 2019;41(4):221-236. [doi: 10.1080/07317107.2019.1659544]
84. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
85. Vlachogianni P, Tselios N. Perceived usability evaluation of educational technology using the System Usability Scale (SUS): a systematic review. J Res Technol Educ 2021;54(3):392-409. [doi: 10.1080/15391523.2020.1867938]
86. Curto E, Araujo H. An experimental assessment of depth estimation in transparent and translucent scenes for Intel RealSense D415, SR305 and L515. Sensors (Basel) 2022;22(19):7378 [FREE Full text] [doi: 10.3390/s22197378] [Medline: 36236472]
87. Guo Y, Chen Y, Deng J, Li S, Zhou H. Identity-preserved human posture detection in infrared thermal images: a benchmark. Sensors (Basel) 2022;23(1):92 [FREE Full text] [doi: 10.3390/s23010092] [Medline: 36616688]
88. Leone A, Rescio G, Caroppo A, Siciliano P, Manni A. Human postures recognition by accelerometer sensor and ML architecture integrated in embedded platforms: benchmarking and performance evaluation. Sensors (Basel) 2023;23(2):1039 [FREE Full text] [doi: 10.3390/s23021039] [Medline: 36679839]
89. Rajaraman A, Austin JL, Gover HC. A practitioner's guide to emphasizing choice-making opportunities in behavioral services provided to individuals with intellectual and developmental disabilities. Int J Dev Disabil 2023;69(1):101-110 [FREE Full text] [doi: 10.1080/20473869.2022.2117911] [Medline: 36743319]
90. Stancliffe RJ. Choice availability and people with intellectual disability. In: Stancliffe RJ, Wehmeyer ML, Shogren KA, Abery BH, editors. Choice, Preference, and Disability: Promoting Self-Determination Across the Lifespan. Cham, Switzerland: Springer; 2020.
91. Tullis CA, Cannella-Malone HI, Basbigill AR, Yeager A, Fleming CV, Payne D. Review of the choice and preference assessment literature for individuals with severe to profound disabilities. Educ Train Autism Dev Disabil 2011;46(4):576-595 [FREE Full text] [doi: 10.1177/215416471104600409]
92. Program 1 and program 2 participant datasets. Microsoft SharePoint. URL: https://fondazionelfo-my.sharepoint.com/:x:/g/personal/alberti_g_legadelfilodoro_it/IQA20Tz2ViHvS5pOO4tuGjg2AYxP38S4Sr58cH1H4WkIx7o?e=tLvDaU [accessed 2026-07-23]

Abbreviations

NAP: Nonoverlap of All Pairs

XSL•FO
RenderX

Edited by S Munce; submitted 04.Feb.2026; peer-reviewed by N Modugno, K Kascak, M Yamagami; comments to author 01.Jun.2026; revised version received 16.Jul.2026; accepted 22.Jul.2026; published 06.Aug.2026.

Please cite as:
Lancioni GE, Alberti G, Chiariello V, Zacchini E, Singh NN, O'Reilly MF, Sigafoos J, Buono S
Basic Technology to Support Stimulation Access, Activity, and Correct Posture in Individuals With Intellectual and Multiple Disabilities: Development and Usability Study
JMIR Rehabil Assist Technol 2026;13:e92873
URL: https://rehab.jmir.org/2026/1/e92873
doi:10.2196/92873
PMID:42562384

©Giulio E Lancioni, Gloria Alberti, Valeria Chiariello, Elia Zacchini, Nirbhay N Singh, Mark F O'Reilly, Jeff Sigafoos, Serafino Buono. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 06.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Usability and Functional Outcomes of the Obi3 Robotic Feeding Device Among People With Disabilities, Caregivers, and Providers: Mixed Methods Study

Andrea D Fairman[1,2], PhD, MOT, OTR/L, ATP, DRP; Ryan B Osal[2,3], OTD, MS, OTR/L, NZROT; Heather Keeton[1], ATP, CDME; Jonathan P Dekar[1]

[1]DESĪN, LLC, Jacksonville, FL, United States

[2]The Driving Doctor: Adaptive Driving and Rehabilitation Solutions, PC, Glendale, RI, United States

[3]University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Andrea D Fairman, PhD, MOT, OTR/L, ATP, DRP
DESĪN, LLC
7018 A C Skinner Parkway, STE270
Jacksonville, FL, 32256
United States
Phone: 1 401 474 9418
Fax: 1 248 282 9599
Email: andrea@desincorp.com

Abstract

Background: Loss of self-feeding ability due to severe bilateral upper extremity motor impairment (BUEMI) limits functional independence with a critical activity of daily living, increases medical risks, reduces social participation, and affects overall well-being. Robotic feeding devices are an emerging category of assistive technology. Obi3 is a third-generation, robotic feeding device, regulated in the United States as a Class I medical device and CE marked as a Class I device under the European Union Medical Device Regulation. The device is designed to enable people with BUEMI to independently select and safely consume food and liquid using accessible switches or interfaces while maintaining culturally normative dining practices. Although earlier Obi generations showed promise, comprehensive usability data across multiple stakeholder groups remains limited.

Objective: This study evaluated the usability, safety, and clinical relevance of Obi3 in real-world use from the perspectives of people with BUEMI, caregivers, and rehabilitation providers. The findings were intended to generate evidence to inform clinical decision-making and determinations of medical necessity for potential DME coverage, consistent with the device's intended use and clinician-assessed implementation.

Methods: A mixed methods usability study provided a 1-week home trial of Obi3. Quantitative measures included the System Usability Scale (SUS), survey items adapted from the Matching Person and Technology (MPT) framework, and self-feeding ratings based on the *International Classification of Functioning, Disability, and Health* impairment scale. Paired-sample *t* tests, Wilcoxon signed-rank tests, and effect sizes (Cohen *d*) assessed changes in self-feeding scores. Qualitative data from open-ended survey responses and follow-up semistructured interviews were analyzed thematically to complement quantitative findings.

Results: A total of 50 participants were enrolled, of whom 42 completed follow-up surveys. Participants included 15 people with BUEMI (pediatric and adult), 14 caregivers, and 13 providers. Participants' mean self-feeding impairment scores decreased significantly from 3.80 (SD 0.41) at baseline to 0.60 (SD 0.74) post trial (n=15; t_{14}=10.75; P<.001; Cohen d=3.40). All patient participants demonstrated improvement in functional self-feeding ability. Mean SUS scores exceeded the benchmark for acceptable usability (≥68) across stakeholder groups: 82.8 (SD 17.2) for people with BUEMI, 85.2 (SD 12.6) for caregivers, and 85.0 (SD 10.7) for providers. Responses to MPT items indicated strong alignment between user goals and device capabilities, perceived safety, and low complexity. Caregivers reported reduced feeding-related workload and stress, and providers endorsed ease of clinical integration. No adverse events or device malfunctions occurred.

Conclusions: Obi3 demonstrated high usability, safety, and clinical use in restoring self-feeding independence among individuals with severe BUEMI. Findings across people with BUEMI, caregivers, and providers support Obi3 as an effective DME that may

XSL•FO
RenderX

enhance user autonomy and reduce caregiver burden. Lengthier longitudinal studies are needed to establish long-term clinical efficacy and health outcomes.

(JMIR Rehabil Assist Technol 2026;13:e86338) doi:10.2196/86338

KEYWORDS

activities of daily living; assistive technology; caregiver; mixed methods study; rehabilitation; robotics; self-feeding; upper extremity; usability study

Introduction

Background

Self-feeding is a fundamental activity of daily living that supports autonomy, nutritional health, and social participation. Within the *International Classification of Functioning, Disability, and Health* (*ICF*) and the *Occupational Therapy Practice Framework*, self-feeding is recognized as a core activity linked to independence, dignity, and quality of life [1,2]. For clarity, this manuscript uses the term self-feeding to refer specifically to bringing food to the mouth, distinguishing it from eating and swallowing, as defined by the American Occupational Therapy Association [3]. Loss of self-feeding ability due to bilateral upper extremity motor impairment (BUEMI) affects more than nutritional intake, contributing to reduced autonomy, dignity, and social participation [4,5]. BUEMI arises from conditions such as cerebral palsy, amyotrophic lateral sclerosis, muscular dystrophy, spinal cord injury, and traumatic brain injury, all of which may disrupt voluntary arm and hand function [6-11]. Although caregiver-assisted feeding can maintain nutrition, it does not restore independence and may introduce variability in feeding pace, technique, and safety [12-15]. Over time, reliance on caregiver assistance can increase caregiver burden and reduce quality of life for both individuals and their caregivers [16]. In addition to physical limitations, loss of self-feeding reduces participation in the social and cultural aspects of eating. Mealtimes are important social activities, and dependence on others may limit engagement and contribute to a reduced sense of identity and belonging. Assistive and robotic feeding technologies aim not only to restore physical function but also to support social participation and emotional well-being [17-21].

Powered feeding systems are an emerging class of assistive technology (AT) designed to restore self-feeding ability in individuals with BUEMI [22,23]. In the United States, these devices are regulated as Class I medical devices under US Food and Drug Administration (FDA), as a daily activity assist device, indicating low risk when used as intended [24]. Several powered feeding devices are commercially available, including the Mealtime Partner [25] and Neater Eater systems [26]. Earlier devices typically rely on limited motor control and produce less fluid movement. In contrast, Obi incorporates a 6–degree-of-freedom (DOF) robotic arm designed to enable smoother, more adaptive utensil motion and individualized positioning, reflecting advances in human-centered assistive robotics [27].

Evidence from Al-Halimi and Moussa [28] supports the functional validity of such multi-DOF robotic approaches, demonstrating that individuals with severe BUEMI were able to perform complex daily tasks, including self-feeding and drinking, using a 6-DOF robotic arm mounted on a wheelchair. Their study highlighted individualized control strategies and adaptive movement patterns, underscoring how increased robotic dexterity can restore meaningful interaction and autonomy while self-feeding.

Obi is designed for repeated use, serves a medical purpose, and is appropriate for home use. These characteristics are consistent with the Centers for Medicare & Medicaid Services (CMS) definition of durable medical equipment (DME). Its design for home use further supports its use in long-term care settings. Obi has been widely used in clinical and home settings, with no reported adverse events in major regulatory databases. It has been prescribed by clinicians and reimbursed through multiple funding sources, supporting its role as a clinically recognized intervention for self-feeding impairment [31]. Despite these advances, limited research has evaluated the real-world usability of newer robotic feeding systems across multiple stakeholders. Usability is a key determinant of adoption and sustained use, particularly for devices intended for long-term home use.

Obi has been prescribed by physicians, deemed medically necessary by occupational therapists, and reimbursed by multiple sources, including the Veterans Health Administration, state Medicaid programs, private insurers, and international health systems. These patterns of prescription and funding highlight Obi's recognition as a legitimate medical intervention to restore self-feeding independence. A recent descriptive study reported that 19 individuals with BUEMI, ranging in age from 8 to 60 years, successfully used an earlier version of Obi in homes, schools, and clinics, achieving independence in food delivery with positive impacts on nutritional intake, autonomy, and social participation [33]. Obi's third-generation model (Obi3) incorporates user-informed enhancements, including new and expanded utensil options, improved utensil-arm retention, new software features ("reward mode" and "patient power-on"), a 2-inch increase in reach, and faster speeds from scooping to mouth delivery.

The literature and regulatory evidence highlight a critical clinical and societal need for safe, effective, and user-centered technologies that restore self-feeding independence among individuals with BUEMI. The systematic evaluation of Obi3's usability over a 1-week period, including the enhanced features and perspectives of multiple stakeholders, including participants with BUEMI, caregivers, and providers, has not been conducted previously. Addressing this gap is essential because usability is a key determinant of the adoption and sustained use of assistive technologies in real-world contexts. Usability, defined as the effectiveness, efficiency, and satisfaction with which users achieve goals in specific environments [33], remains

XSL•FO
RenderX

particularly critical for devices classified as DME given their intended long-term use.

Study Purpose and Hypothesis

Building on prior descriptive findings [32], this study conducted a mixed methods usability evaluation of the Obi3 robotic feeding device to assess its performance, safety, and clinical relevance across multiple stakeholder groups in real-world settings. We hypothesized that Obi3 would demonstrate high usability, safety, and clinical use, reflected by System Usability Scale (SUS) scores exceeding the industry benchmark of 68 [34,35], significant improvements in self-feeding impairment scores, and positive qualitative feedback related to satisfaction and integration into daily routines. This approach aimed to generate both quantitative and qualitative evidence to inform iterative device refinement, guide clinical implementation, and support broader adoption through real-world clinical evidence.

Methods

Study Design and Oversight

This mixed methods usability study evaluated the Obi3 Robotic Feeding Device across 3 stakeholder groups: people with BUEMI, caregivers, and rehabilitation providers. Quantitative and qualitative methods were combined to determine whether Obi3 fulfills its intended purpose of restoring self-feeding ability for individuals with severe upper-extremity motor impairments and whether it meets formally defined user needs. The study was designed in accordance with the International Organization for Standardization (ISO) 14155:2020 Clinical Investigation of Medical Devices for Human Subjects: Good Clinical Practice and was approved by the WCG Clinical Institutional Review Board (Protocol #20251472).

A target sample of up to 60 participants (20 people with BUEMI, 20 caregivers, and 20 providers) was based on the FDA guidance recommending ≥15 participants per user group for usability or design-validation testing of medical devices [36,37]. The study was coordinated by the device manufacturer, DESĪN LLC, as a single-sponsor remote investigation. All consent procedures, survey administration, and survey data collection were conducted electronically using the HIPAA (Health Insurance Portability and Accountability Act)–compliant Qualtrics platform [38].

Recruitment and Enrollment

Recruitment followed a structured, multistage process to ensure broad representation while maintaining ethical compliance. Recruitment materials, including flyers and digital notices, were distributed nationally. A triadic recruitment model (patient-caregiver-provider) was prioritized but not required. Providers were eligible if they held professional credentials (eg, occupational therapist, speech-language pathologist, or AT professional [ATP]) and had experience recommending or evaluating assistive technologies for individuals with self-feeding limitations. Caregivers were qualified if they directly assisted an individual with mealtime activities. People with BUEMI were eligible if aged ≥5 years, required assistive feeding technology to increase self-feeding independence, and could provide consent or assent.

Exclusion criteria included residency in California for new Obi users (due to state-mandated 30-day device-trial requirements), inability to provide informed consent, cognitive impairments preventing meaningful participation, or non-English fluency. Certified ATPs employed by DESĪN LLC prescreened potential participants using standard operating procedures aligned with routine clinical practice.

Informed consent (and assent when applicable) was obtained electronically via Qualtrics. Adults, caregivers, and providers electronically signed consent forms; pediatric participants (aged <17 years) provided assent with parental or guardian consent. Verbal consent was documented when physical signing was not feasible. Participants were informed of study procedures, potential risks, and withdrawal rights.

Obi3 Trial Procedures

Following consent and baseline screening, providers completed the Obi Medical Device Needs Assessment (MDNA) form [38]. The MDNA captured demographics, diagnosis, and self-feeding impairment at baseline and follow-up using the Self-Feeding Impairment Scale. This scale was not intended to serve as a standardized outcome measure and has not undergone formal psychometric validation. Rather, it is a study-specific clinical rating scale embedded within the Obi MDNA form and is conceptually informed by the *ICF* framework. Specifically, the scale reflects levels of functional self-feeding performance ranging from complete dependence to independent performance and was developed to pragmatically characterize the degree of self-feeding impairment observed during routine clinical assessment and device trials. Because the scale was developed internally for clinical use, there are no published studies on its validity, reliability, or responsiveness.

People with BUEMI and their caregivers received an Obi3 unit for a 1-week home trial. The Obi3 (third generation) device incorporated user-informed hardware and software enhancements, expanded utensil options, increased reach, improved utensil-arm retention, and new features such as reward mode and patient power-on (Figure 1).

XSL•FO
RenderX

Figure 1. Obi3 robotic feeding device and new features. The Obi3 robotic feeding system includes several user-informed enhancements, such as expanded utensil options, stronger utensil attachment, additional 2-inch horizontal reach, faster delivery speed, reward mode, patient power-on mode, in-bowl collision detection, new switch color, and improved labeling and aesthetic design.



The trial was conducted in participants' customary real-world environments, including private homes, outpatient clinics, schools, rehabilitation facilities, and other community-based settings where meals would ordinarily occur, to evaluate how effectively the Obi3 robotic feeding device could be integrated into routine daily activities rather than under controlled laboratory conditions. This approach emphasized ecological validity, allowing researchers to capture authentic user experiences, environmental influences, and contextual factors that affect real-world usability. Each caregiver and patient used the device for a minimum of 5 meals within a 1-week period, providing repeated exposure sufficient to assess learning effects, ease of setup, and consistency of performance across multiple sessions. Providers introduced the Obi3 system using the manufacturer's *Quick Start Guide* [39], ensuring standardized orientation while permitting flexibility for individualized instruction. Remote technical and clinical support was available throughout the trial to mirror the level of assistance users might receive during early adoption in home or community settings.

Participants interacted with the device in its standard configuration, which included power controls, utensil selection options, and the reward mode feature designed to enhance motivation and engagement. Meals were self-selected to align with each participant's typical diet, cultural food preferences, and mealtime routines. This design allowed the study to capture performance across a diverse range of food consistencies, environmental contexts, and social settings, thereby providing a more comprehensive understanding of the device's usability, adaptability, and potential for integration into everyday life.

Outcome Measures

Electronic surveys were administered to people with BUEMI, caregivers, and providers, with developmentally appropriate adaptations for pediatric participants. Surveys combined standardized and nonstandardized items assessing usability, satisfaction, and integration of Obi3 into daily routines. The primary outcome was the SUS, a 10-item instrument scored 0-100, with 68 as the benchmark for average usability [34,35]. The SUS was administered to adult patients, caregivers, and providers; pediatric participants were excluded because the SUS has not been validated for children. Secondary measures included items adapted from the Matching Person and Technology (MPT) framework, which evaluates personal predisposition, psychosocial fit, and environmental support [40,41]. Adults and teen people with BUEMI (aged ≥15 years) completed the SUS and MPT-based items verifying ≥5 meal exposures and answered nonstandardized questions about instruction clarity, ease of operation, utensil satisfaction, and

XSL•FO
RenderX

perceived independence and safety. People with BUEMI aged 15 years and older were grouped within the adult patient category because the survey instruments used in the study included self-report items requiring a higher level of reading comprehension, abstract reasoning, and self-reflection regarding device usability, satisfaction, and independence. These constructs align more closely with adolescent and adult cognitive and communicative abilities than with those of younger children. Grouping patient participants aged 15 years and older with adults allowed for consistent interpretation of survey responses and ensured that the usability and satisfaction data reflected meaningful self-reports.

Younger people with BUEMI (aged 5-14 years) were analyzed separately, as their perspectives were often supplemented or interpreted through caregiver feedback due to differences in literacy, attention span, and cognitive maturity. This stratification thus enhanced the validity and comparability of responses across age groups while maintaining clinical and developmental appropriateness. Pediatric participants with BUEMI (aged <15 years) used simplified, emoji-based Likert items evaluating confidence, enjoyment, and safety.

Caregivers completed survey questions designed to capture both their experience using the Obi device and its impact on daily caregiving demands. These included the SUS [34] to assess perceived ease of use and overall satisfaction; the Zarit Caregiver Burden Scale [42] to evaluate potential changes in emotional, physical, and time-related strain; and a series of custom items addressing device setup, cleaning procedures, and perceived reductions in hands-on feeding time. Together, these instruments provided a comprehensive view of how Obi influenced caregiver workload, stress, and independence for the individual being served.

Providers completed a separate set of survey items, including measures capturing professional background characteristics and impressions of Obi3. They rated the device's usability and clinical use, including the clarity and completeness of its instructional materials, setup guides, and documentation tools. These data were intended to evaluate the device's practicality for integration into routine clinical assessment and intervention processes, as well as its suitability for recommendation and long-term use across diverse care settings.

Participants who completed the survey received a US $25 electronic gift card as approved compensation. Optional semistructured interviews were offered to participants following survey completion, whose responses required clarification or elaboration.

Qualitative Interviews

To complement quantitative data, semistructured virtual interviews were conducted by a researcher independent of DESĪN LLC. Interviews captured lived experiences and stakeholder perspectives, addressing the risk that excluding user voices can yield ineffective technologies. A total of 12 participants were contacted via email, phone, and/or text to elaborate on survey themes. Initial contact with participants who completed the follow-up interviews began the week following survey completion, with recruitment and reminders occurring over a 3-week period. Follow-up interviews were conducted during a limited 4-week period. Early sessions included both provider and pediatric user perspectives, followed by adult user and caregiver interviews later within the 1-month data collection window to ensure accurate recall of the Obi3 trial. Sessions were conducted via Google Meet, with verbal consent to record and transcribe. Transcripts were anonymized and analyzed qualitatively as described in the next section.

Data Analysis

All statistical analyses were conducted using IBM SPSS Statistics for Windows (version 31.0.0). Descriptive and inferential procedures were used to summarize participant characteristics and evaluate changes across study measures. Continuous variables, including SUS scores, MPT alignment ratings, Zarit Caregiver Burden scores, and self-feeding impairment ratings, were analyzed using descriptive statistics, including means, SDs, medians, and 95% CIs to describe central tendency and variability. Categorical variables, such as demographic characteristics, professional discipline, and prior AT experience, were summarized as frequencies and percentages.

Assumptions of normality were evaluated using Shapiro-Wilk tests and visual inspection of score distributions, given the relatively small sample sizes within stakeholder groups. Shapiro-Wilk test statistics (W values) and exact *P* values were calculated and are reported for primary outcome measures. Within-subject changes from baseline to post trial were analyzed using paired-sample *t* tests, with *t* values, degrees of freedom, exact *P* values, and 95% CIs reported. For variables that did not meet assumptions of normality, or because of the ordinal nature of the self-feeding impairment ratings and small sample size, confirmatory nonparametric analyses were conducted using Wilcoxon signed-rank tests, with corresponding *Z* statistics and exact *P* values reported.

Effect sizes (Cohen *d*) were calculated to quantify the magnitude of change, using conventional benchmarks (0.2=small, 0.5=medium, and 0.8=large) [43]. Missing data were reviewed descriptively to confirm whether the missing data appeared to occur at random or if they reflected systematic withdrawal-related patterns. Given the low proportion of missing data and absence of identifiable systematic patterns, analyses were conducted using available-case data without imputation. Data were screened for potential outliers by inspecting distributions and performing range checks; no implausible values or data-entry errors requiring removal were identified.

Open-ended survey responses and interview transcripts were analyzed using thematic content analysis. Two researchers independently coded the data, iteratively refined categories, and resolved discrepancies through consensus. To support consistency in the coding language, a generative AI tool (Google Gemini, version 2.5) was used in a preliminary stage to assist with organizing the raw text and suggesting candidate theme labels. The researchers critically reviewed all AI-assisted outputs, confirmed or revised them through manual coding, and determined the final themes. Triangulation with quantitative findings and weekly audit meetings further enhanced methodological rigor [44].

XSL•FO
RenderX

Ethical Considerations

All procedures adhered to Good Clinical Practice and ethical standards for human research. The study protocol was reviewed and approved by the WCG Clinical Institutional Review Board (IRB; Protocol #20251472). All participants, or their legal guardians in the case of minors, provided informed consent or assent prior to participation. Data collection and handling followed HIPAA-compliant procedures through the Qualtrics electronic platform. Participants who completed the follow-up survey received a US $25 electronic gift card as IRB-approved compensation for their time and effort.

Results

Participant Enrollment and Retention

In total, 50 individuals enrolled in the study across the stakeholder groups, reflecting a balanced distribution among providers, caregivers, and people with BUEMI. Of these, 42 (84%) participants completed posttrial surveys. Overall attrition was at 16% and was limited to posttrial survey completion rather than participation in the Obi trial itself. Participants who did not complete posttrial surveys were considered lost to follow-up after they could not be reached despite multiple contact attempts over a 2- to 3-week period following the expected survey completion date. This time frame was established to ensure that participants' experiences with the device remained recent enough to support accurate recall.

Among providers, 13 (76%) of 17 completed all study requirements, with attrition primarily due to missed posttrial surveys. Caregivers demonstrated a comparable completion rate of 82% (14/17), with 1 early withdrawal prior to trial initiation and 2 incomplete follow-up surveys. Participants with BUEMI aged 15 years and older demonstrated full retention and 100% completion of study procedures, suggesting high engagement and feasibility of participation in this subgroup. Participants aged 5-14 years also demonstrated strong engagement, with 6/7 (86%) completing the surveys. Overall, these results indicate excellent adherence and acceptability across participant groups, particularly among patient stakeholders, with minimal losses attributable to study-related factors. Enrollment, completion, and attrition rates by stakeholder group are summarized in Table 1.

Table 1. Stakeholder enrollment, survey completion, and attrition.

Stakeholder groups	Participants enrolled, n	Surveys completed, n	Total dropouts and reasons for attrition
Providers	17	13	4 (1=no patient available; 3=did not complete posttrial surveys)
Caregivers	17	14	3 (1=dropped before trial; 2=did not complete posttrial surveys)
People with BUEMI[a] (aged ≥15 years)	9	9	None—100% survey completion
Participants aged 5-14 years	7	6	1 (completed trial but no survey)
Totals	50	42	8 lost to attrition

[a]BUEMI: bilateral upper extremity motor impairment.

Patient Demographics

Demographic and descriptive data were obtained from provider-completed Obi MDNA forms. For 1 pediatric participant whose original provider withdrew from the study, the MDNA form was instead completed by a certified ATP. In accordance with standard clinical procedures, the ATP reviewed the participant's existing clinical documentation, including therapy notes, prior evaluations, and progress reports, and conducted a detailed video review of feeding sessions. This approach ensured that the MDNA data accurately reflected the child's functional abilities, feeding performance, and support needs, consistent with how such assessments are typically conducted in clinical practice when direct observation by the evaluating therapist is not feasible.

Among adult and adolescent participants aged ≥15 years (n=9), the mean age was 33.9 (median 18, range 15-67) years. Gender distribution was balanced (4 males and 5 females). The most common diagnosis was cerebral palsy (n=4), followed by other neurological or neuromuscular conditions, as summarized in Table 2.

The pediatric patient cohort (n=6) ranged in age from 7 to 11 (mean 9.2, SD 1.83) years and included 4 females and 2 males. Rett syndrome was most prevalent (n=3), followed by cerebral palsy (n=2) and Charcot-Marie-Tooth disease (n=1) as listed in Table 3.

XSL•FO
RenderX

Table 2. Patient demographics (aged 15 years and older; n=9).

Age (years)	Birth sex	Primary diagnosis code (*ICD-10*[a])	Diagnosis
15	Female	Q87.2	Cornelia de Lange syndrome
15	Female	G80	Cerebral palsy
48	Male	G80.9	Cerebral palsy
50	Male	G80.9	Cerebral palsy
15	Male	G80.0	Spastic quadriplegia cerebral palsy
15	Female	G11.3	Cerebellar ataxia with defective DNA repair
67	Female	G12.21	Amyotrophic lateral sclerosis
18	Male	G71.01	Duchenne muscular dystrophy
62	Female	G12.1	Spinal muscular atrophy

aICD-10: International Classification of Diseases, Tenth Revision.

Table 3. Patient demographics (aged <15 years; n=6).

Age (years)	Birth sex	Primary diagnosis code (*ICD-10*[a])	Diagnosis
7	Female	F84.2	Rett syndrome
10	Male	G80.9	Cerebral palsy
9	Female	F84.2	Rett syndrome
11	Male	G80.3	Athetoid cerebral palsy
7	Female	F84.2	Rett syndrome
11	Female	G60.0	Charcot-Marie-Tooth disease

[a]*ICD-10*: *International Classification of Diseases, Tenth Revision.*

Caregiver Characteristics

Caregiver participants (n=14) were primarily mothers (n=11), with additional representation from spouses/partners (n=2) and 1 paid caregiver. Participants were predominantly middle-aged, with varied household income levels and generally high educational attainment. Burden ratings indicated moderate to high stress, particularly related to balancing caregiving with employment and family responsibilities (9/14, 64%). Half of caregivers reported a high overall burden. Descriptive data related to the caregiver results are summarized in Table 4.

XSL•FO
RenderX

Table 4. Caregivers' descriptive data (n=14).

Category and response	Value, n (%)
Relationship to the patient	
Mother	11 (78.6)
Spouse/partner	2 (14.3)
Paid home care staff	1 (7.1)
Age in years	
18-30	1 (7.1)
31-40	3 (21.4)
41-50	7 (50.0)
51-60	1 (7.1)
61-70	1 (7.1)
Prefer not to answer	1 (7.1)
Annual household income (US $)	
150,000 or more	4 (28.6)
100,000-149,999	3 (21.4)
75,000-99,999	2 (14.3)
50,000-74,999	0
25,000-49,999	1 (7.1)
Less than 25,000	1 (7.1)
Prefer not to answer	3 (21.4)
Highest level of education attained	
Doctoral degree	1 (7.1)
Master's degree	1 (7.1)
Bachelor's degree	6 (42.9)
Associate's degree	4 (28.6)
High school graduate	1 (7.1)
Prefer not to answer	1 (7.1)

Caregiver Burden

Selected items from the Zarit Caregiver Burden Scale were used to descriptively characterize caregiver experiences and caregiving demands among participants. Caregivers reported moderate to high levels of burden across multiple domains. The greatest burden was associated with balancing caregiving responsibilities with work and family obligations (mean 2.64, SD 1.01), followed by insufficient personal time due to caregiving demands (mean 2.57, SD 1.09). Caregivers also reported that their health had been affected by caregiving responsibilities (mean 1.93, SD 1.07). Overall perceived caregiver burden averaged 2.07 (SD 1.14), with 50% (7/14) of caregivers reporting very high levels of burden. These findings suggest that participants providing feeding assistance frequently experienced substantial caregiving demands, highlighting the potential value of interventions that promote greater self-feeding independence.

Provider Characteristics

Of 17 providers enrolled, 13 completed the study. Most were occupational therapy practitioners (n=12), and 1 was a speech-language pathologist. Providers represented outpatient clinics, private practice, and rehabilitation centers, with experience ranging from <3 to >20 years. Half of the providers were recommending Obi for the first time, and most reported limited prior integration of AT into clinical workflows. Descriptive data related to the providers who participated in this study are highlighted in Table 5.

XSL•FO
RenderX

Table 5. Provider demographics (n=13).

Category and response	Value, n (%)
Professional background	
OT[a]/OTA[b]	12 (92.3)
SLP[c]	1 (7.7)
Practice setting	
Outpatient clinics	5 (38.5)
Private practice	3 (23.1)
Schools	2 (15.4)
Home health	1 (7.7)
Hospitals	1 (7.7)
Early intervention	1 (7.7)
Years in professional practice	
20 or more	4 (30.8)
11-20	3 (23.1)
7-10	3 (23.1)
4-6	2 (15.4)
2-3	1 (7.7)
Experience recommending Obi	
First time recommending Obi	8 (61.5)
Recommended 2-4 times	3 (23.1)
Recommended within the past year	2 (15.4)
Education level	
Master's	10 (76.9)
Bachelors	2 (15.4)
Doctoral	1 (7.7)
Experience with AT[d]	
Rare use (<1 day/week)	7 (53.8)
<50% of time with AT	5 (38.5)
>50% of time with AT	1 (7.7)

[a]OT: occupational therapist.

[b]OTA: occupational therapy assistant.

[c]SLP: speech-language pathologist.

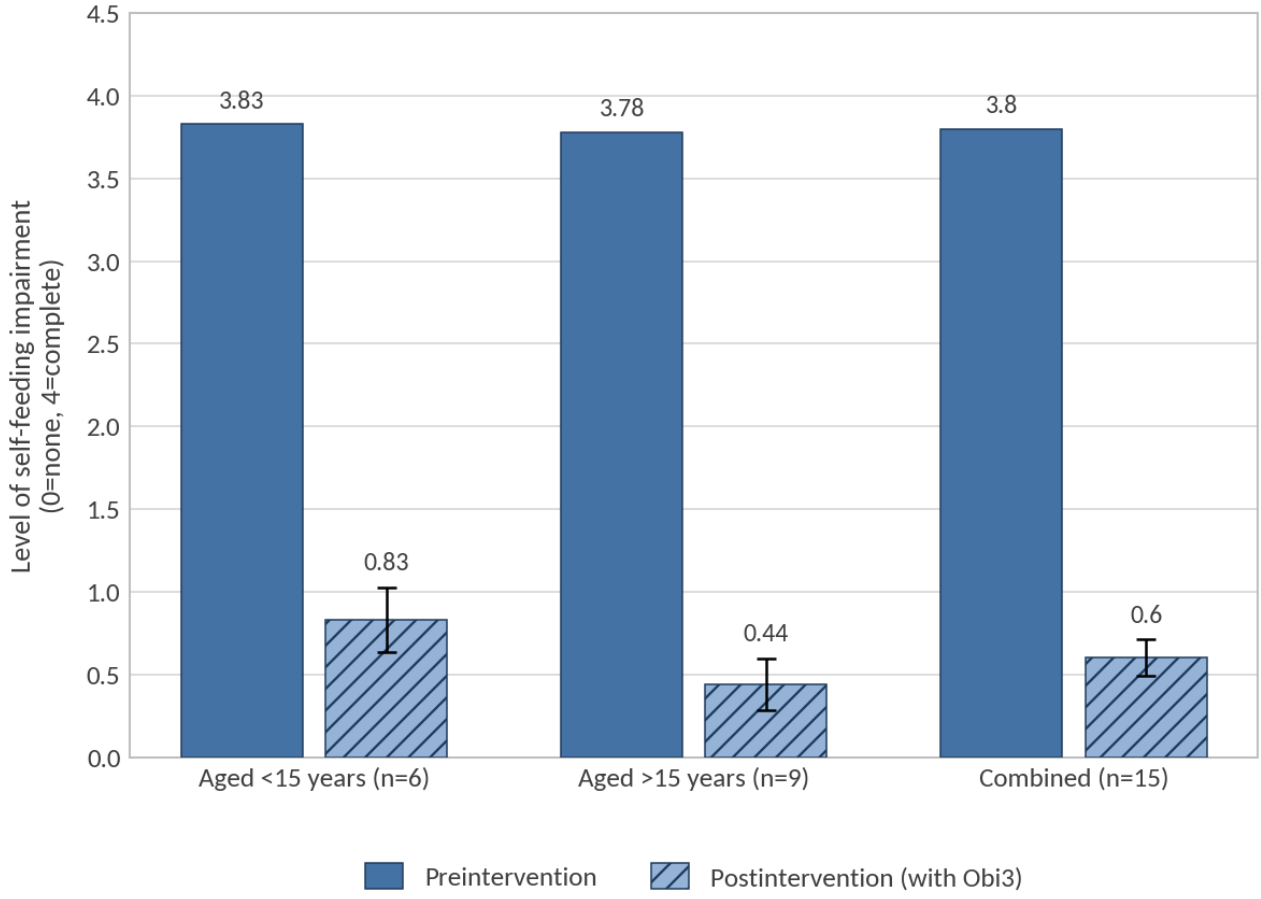
[d]AT: assistive technology.

Participants With BUEMI Self-Feeding Outcomes

Changes in self-feeding ability were measured using the Self-Feeding Impairment Scale (0=no problem to 4=complete impairment). Pediatric participants (n=6) showed significant improvement, with mean impairment decreasing from 3.83 (SD 0.41) at baseline to 0.83 (SD 0.98) post trial (t_5=5.38; P=.002; Cohen d=2.37). Adult participants (n=9) demonstrated a similar pattern, improving from 3.78 (SD 0.44) to 0.44 (SD 0.53; t_8=10.22; P<.001; Cohen d=4.71). Across all people with BUEMI (n=15), mean impairment decreased from 3.80 (SD 0.41) to 0.60 (SD 0.74), representing a mean reduction of 3.20 points (t_{14}=10.75; P<.001; Cohen d=3.40). All participants demonstrated improvement in functional self-feeding ability following the use of Obi3 as represented in Figure 2.

XSL•FO
RenderX

Figure 2. Pre- and postintervention self-feeding impairment scores using Obi3. Bar chart showing baseline and postintervention self-feeding impairment scores. Pediatric group (aged <15 years): 3.83 → 0.83; adult group (aged ≥15 years): 3.78 → 0.44; combined group: 3.80 → 0.60. All groups demonstrated large, statistically significant improvements.



SUS Results

The SUS was administered to adult participants with BUEMI, caregivers, and providers; pediatric participants were excluded because the SUS has not been validated for children. Adult participants with BUEMI reported a mean SUS score of 82.78 (SD 17.16). Caregivers reported a mean score of 85.18 (SD 12.58), and providers reported a mean score of 85.00 (SD 10.73). Detailed results are presented in Table 6.

Table 6. System Usability Scale scores by stakeholder group.

Stakeholder group	n	Mean (SD)	Range	Score ≥68[a], n (%)
Adult participants with BUEMI[b]	9	82.78 (17.16)	47.5-100	8 (88.9)
Caregivers	14	85.18 (12.58)	60-100	13 (92.9)
Providers	13	85.00 (10.73)	72.5-100	13 (100)

[a]System Usability Scale scores ≥68 are considered above the benchmark for acceptable usability. Pediatric participants did not complete the System Usability Scale.

[b]BUEMI: bilateral upper extremity motor impairment.

Modified MPT Findings

Participants With BUEMI Perspectives

Survey items adapted from the MPT framework demonstrated strong alignment between Obi and user goals, competencies, and environmental supports. Adult participants (n=9) rated most items highly (*M*=4.11-4.56 on a 5-point scale) but reported lower comfort integrating Obi into daily routines (*M*=3.44) and social settings (*M*=2.50-3.78). Pediatric participants (n=6) provided consistently higher ratings (*M*=3.50-4.75), with the greatest agreement for the statements “Obi helps me achieve my feeding goals” (*M*=4.17) and “Obi improves my quality of life” (*M*=4.75). Across both groups, participants endorsed ease of use (*M*=4.50-4.56) and environmental support (*M*=4.44-4.67), indicating strong usability and goal alignment. However, additional research is warranted to examine how Obi can further facilitate integration into daily routines and community participation.

XSL•FO
RenderX

Caregiver Outcomes

Caregivers reported reduced burden and stress, with increased time for personal responsibilities and less physical fatigue during meals. Open-ended feedback emphasized Obi's dual benefit, enhancing user independence while easing caregiver workload. Common suggestions included improving portability for travel, stabilizing switch ports, adding visual or auditory cues for users with sensory impairments, and optimizing utensil scooping performance. While "reward mode" was praised by some for motivation, others found it confusing, suggesting clearer instructions or customization options. Overall sentiment was highly positive, with several caregivers describing Obi as "life-changing" for both users and families.

Provider Perspectives

Providers described Obi3 as highly beneficial, easy to use, and clinically valuable. They highlighted smooth operation, intuitive setup, and clear instructional materials, particularly the Quick Start Guide and online tutorials. Suggested refinements included utensil geometry, bite consistency, and environmental adaptability. Several providers emphasized that Obi3 promoted measurable independence and engagement among participants with BUEMI, with one describing it as "a necessary tool for promoting autonomy."

Safety Findings

No adverse events, device deficiencies, or unanticipated problems were reported during the study. The absence of safety incidents supports Obi's favorable safety profile, consistent with its FDA Class I designation and regulatory history.

Follow-Up Qualitative Interviews

Overview

To complement the survey data and gain a deeper understanding of usability experiences, follow-up virtual interviews were conducted. Qualitative data from these semistructured interviews were analyzed to enrich, validate, and extend the quantitative findings. Interview participants included caregivers, health care providers, and clients who had previously used the Obi3 feeding device and completed the corresponding surveys. A total of 7 participants (3 participants with BUEMI, 1 caregiver, and 3 providers) completed follow-up virtual interviews to expand on survey findings.

Participants with BUEMIs ranged in age from 11 to 48 years, with primary diagnoses of spinal cord injury (quadriplegia), cerebral palsy, and Charcot-Marie-Tooth disease. Two participants were adults (aged 35 and 48 years), and 1 was a minor (aged 11 years) who participated with support from her mother; all clients completed follow-up interviews. The caregiver participant was a mother aged 31-40 years with a doctoral degree and an annual household income exceeding US $150,000; she also completed a follow-up interview. Providers included an occupational therapist and a speech-language pathologist, both with master's degrees. Their years of practice ranged from 2-3 years to more than 20 years, and their current settings included an outpatient clinic, private practice, and a state developmental disabilities resource center. Overall, 2 providers worked in Louisiana and 1 in Nebraska; 2 had prior experience recommending Obi. Time dedicated to AT varied, with one provider spending less than 50% but more than 20% of their time on AT, another about 20%, and the speech-language pathologist 50%-75%.

The following themes emerged from this analysis:

Theme 1: Enhanced Independence and Quality of Life

This theme highlights the significant positive impact of Obi3 on the lives of users and their families. Clients, including a client with quadriplegia and a young child, reported a newfound sense of autonomy and freedom during mealtimes. One adult client appreciated being able to eat on their own terms and time, rather than relying on a caregiver for every bite, which provided a mental and emotional boost. A provider noted that their client became tearful at the thought of the device being taken away, indicating a profound improvement in their quality of life. Furthermore, the device facilitated social interaction during meals, as it allowed a client and their caregiver to sit and talk together while eating, transforming mealtimes from a task into an enjoyable, social experience.

> *The timing was perfect for the feeding because I was capable of pressing the button and giving the bite when I wanted it.* [adult patient]

Theme 2: Usability and Ease of Setup

Participants generally found the Obi3 to be user-friendly and easy to set up. A caregiver described the initial setup as "not frustrating" and straightforward after calibrating it with the manual. A provider, who is a therapist, reported that they learned to use the device in under 30 minutes and found it "super easy" and "very user-friendly," even for caregivers who may not have extensive technical knowledge. The *Quick Start Guide* and online resources were also praised for being helpful. However, some participants, including a provider, experienced challenges with certain button functionalities, such as getting the device to start or enter training mode, suggesting a need for clearer instructions.

> *It was very, very user-friendly. I was able to show three of his caregivers...And it took us less than 10 minutes to get it set up.* [provider]

Theme 3: Portability and Design Limitations

Both caregivers and clients identified the device's size and lack of portability as a major limitation. A caregiver expressed a need for a smaller, more portable version for use outside the home, such as at restaurants, church dinners, or school. An adult client shared a similar sentiment, stating they would not feel comfortable taking the large device into a public restaurant. The client likened bringing the device to a "briefcase" that takes up half the table, which they felt detracted from the social experience. A provider also mentioned challenges in finding a suitable table to accommodate the device in a clinical setting.

> *Obi was a little too big for the wheelchair tray. Like, we made it work, but um if it were a little smaller, I think it would have worked better.* [caregiver]

XSL•FO
RenderX

Theme 4: Functional Improvements and Feature Suggestions

Participants offered specific suggestions to improve the device's functionality and versatility. The need for additional utensils, such as a fork, a spoon, and a dull knife for cutting food, was raised by a caregiver. Another client and a caregiver noted the issue of the spoon frequently coming up empty, suggesting that the device should be programmed to always use the bowl's edge to ensure a successful scoop. Different plate sizes, including a larger bowl option, were also recommended. A provider highlighted the benefit of the smaller spoon in the Obi3 version for a client with a small mouth and choking concerns. The device's ability to handle liquids was also a significant benefit for clients with dehydration issues.

> *The machine was easy to use, but the spoon couldn't scoop food from the edge, so I think it needs a larger circumference.* [adult patient]

Theme 5: Motivating Impact and Self-Advocacy

The device was found to be highly motivating for both clients and their families. A pediatric patient loved using the device and was excited to show it to others, even wanting to make a YouTube video about it. A provider noted that the device was motivating for both a child user and their parent, helping them feel that success with the device was achievable at home. One client even used the device to express their food preferences by immediately dumping food they did not want and giving a different choice. This demonstrates that the device empowers users to exercise greater control and self-advocacy over their mealtimes.

> *The Obi device helped reduce my caregiving workload during mealtimes by allowing my child...to self-feed. The Obi device has the potential to be a 'game changer' for our family, as we have been hand-feeding our child three meals a day for ten years.* [caregiver]

Discussion

Principal Findings

Obi3 demonstrated high usability and short-term functional improvements in self-feeding performance under real-world conditions. Statistically and clinically significant gains were observed across all patient participants, who progressed from severe or complete self-feeding impairment at baseline to mild or no impairment after a 1-week home trial. These improvements indicate that Obi3 successfully fulfills its intended medical function, compensating for the loss of voluntary upper-limb movement and enabling users to independently select and consume food and liquid. These findings support the feasibility of the device as an assistive intervention; however, longer-term and controlled studies are needed to establish sustained clinical efficacy, nutritional outcomes, and psychosocial impact.

Quantitative outcomes were reinforced by qualitative and observational data. SUS scores exceeded the industry benchmark of 68 across participants with BUEMI (mean 82.8, SD 17.2), caregivers (mean 85.2, SD 12.6), and providers (mean, SD 10.7), demonstrating a consistent pattern of high perceived usability, reliability, and ease of clinical integration. Complementary survey findings aligned with the MPT framework revealed strong alignment among device capabilities, user goals, and environmental supports, including key predictors of sustained adoption in real-world contexts. Triangulation with interview data further confirmed that participants experienced enhanced autonomy, dignity, and reduced caregiver burden. Collectively, these findings demonstrate that Obi3 meets essential design and performance criteria for DME, addressing not only the physical restoration of self-feeding but also the psychosocial and participatory dimensions of independence.

Participant feedback regarding utensil options and food handling underscores the importance of caregiver and provider education on optimal food presentation when using robotic feeding devices. Although a caregiver suggested adding utensils such as forks or dull knives to facilitate food cutting, this feedback reflects a potential misunderstanding of the device's intended use. As with most adaptive feeding interventions, foods presented for robotic delivery should already be prepared in appropriately sized, manageable portions consistent with each user's swallowing safety and motor abilities. Proper meal preparation, including cutting food into uniform, bite-sized pieces and selecting suitable textures, ensures safe, reliable spoon loading and reduces the risk of choking or aspiration.

The *Instructions for Use* provided with Obi3 include guidance on food size, placement, and consistency to promote optimal performance; however, these directions may not always be sufficiently emphasized or understood by all users. Some caregivers may benefit from more explicit visual examples or in-person demonstrations to reinforce these principles and ensure consistent adherence to safe feeding practices. Integrating this training into initial device setup, clinician-guided instruction, and follow-up sessions could enhance device performance, minimize user frustration, and strengthen caregiver confidence. Ultimately, reinforcing food-preparation education is essential to achieving safe, effective, and independent mealtime participation with robotic feeding systems.

Comparison With Prior Work

Unlike other commercially available powered feeding systems, Obi3 uses a 6-DOF robotic arm that produces naturalistic utensil trajectories and customizable positioning, allowing accommodation for diverse user anatomies and dining environments. These technical advancements correspond with prior evidence from Al-Halimi and Moussa [28], who demonstrated that users with severe motor impairments could complete complex activities of daily living, including self-feeding, using a 6-DOF robotic arm. This study also builds upon earlier work by Burgos [33], which reported improved independence among 19 Obi users in home and clinical contexts, by expanding the sample to a larger, more diverse group and incorporating caregiver and provider perspectives.

The observed reductions in caregiver workload and stress mirror findings from studies of adaptive eating devices and robotic assistance in neurological populations [12,13], underscoring the dual benefit of such technologies for users and care partners alike. In addition, the absence of adverse events aligns with

XSL•FO
RenderX

existing safety records in the FDA's Manufacturer and User Facility Device Experience [31] and the European Database on Medical Devices [32], supporting Obi3's favorable benefit-to-risk profile. Consistent with principles of human-centered design [36], the findings affirm that usability, environmental fit, and psychosocial relevance are critical determinants of long-term adoption and meaningful engagement with assistive technologies.

Limitations

Several limitations should be considered when interpreting these findings. First, data collection primarily relied on self-report surveys, which are well-suited for capturing perceptions of usability, satisfaction, and caregiver burden, but are inherently subject to recall bias, social desirability bias, and individual differences in self-awareness or communication ability. Follow-up interviews were incorporated to contextualize these responses and reduce misinterpretation.

Although 50 participants were enrolled and 42 completed all study requirements, the sample size remains relatively small, limiting the ability to detect subgroup differences or perform multivariate analyses. Although the target enrollment of 60 participants was not achieved, the final sample of 50 participants exceeded typical sample sizes reported in usability studies and was considered sufficient to address the study objectives. While the number of participants aligns with the FDA's guidance for formative and summative usability testing, larger multisite studies would enhance the external validity and representativeness of findings. Future research should include greater demographic, diagnostic, and cultural diversity, as well as participants using alternative switch interfaces or mounting configurations to evaluate device performance across varied clinical and environmental contexts.

The self-feeding impairment rating scale used in this study was a study-specific clinical measure embedded within the MDNA assessment process and has not undergone formal psychometric validation or interrater reliability testing. Although structured scoring anchors, clinician-guided assessment procedures, and repeated ratings by the same provider were used whenever feasible to support consistency, the measure was intended primarily to descriptively characterize functional change during the Obi3 usability trial rather than to serve as a definitive clinical outcome instrument. Consequently, the observed improvements in self-feeding performance should be interpreted as preliminary evidence of functional feasibility rather than conclusive evidence of clinical efficacy. In contrast, the SUS, which served as the primary standardized usability outcome measure, is a well-established, psychometrically validated instrument widely used in health and AT research.

The study was conducted over a short-term trial period (1 week), which may not fully capture the effects of sustained device use on long-term outcomes such as user satisfaction, skill acquisition, changes in muscular strength or tone, or caregiver burden. Extended longitudinal and home-based evaluations are warranted to understand the potential impact on health outcomes, maintenance needs, and adherence over time.

The device manufacturer sponsored this study, and several authors maintain professional or financial relationships with the sponsor. Although these relationships were fully disclosed and reviewed by the IRB, they introduce a potential risk of bias in study design, data interpretation, and reporting. To mitigate this risk, data collection followed standardized protocols. Ratings of participants' self-feeding performance were completed by licensed therapists, who were also study participants, using a structured clinical data collection instrument (Obi MDNA) consistent with routine practice. All survey data were collected remotely via the Qualtrics platform, with participants entering responses independently and without direct interaction with the researchers responsible for data analysis. Qualitative interviews were conducted by a researcher independent of the sponsor, and data analysis incorporated dual independent coding, triangulation, and audit procedures. Additionally, all authors had access to aggregated study data, and statistical analyses were independently verified. Finally, this investigation did not include objective nutritional or biomechanical measures (eg, meal duration, food intake volume, posture, or reach analysis), which could strengthen the interpretation of functional and clinical outcomes. Integrating such measures in future studies would allow for a more comprehensive understanding of the device's impact on independence, health, and participation across stakeholder groups.

Acknowledgments

This study was supported by DESĪN LLC, the manufacturer of the Obi Robotic Feeding Device. The sponsor provided financial support for study coordination, device provision, and participation in the decision to submit this manuscript for publication. DESĪN LLC also contributed to aspects of study design and data interpretation; however, all analyses were conducted in accordance with Good Clinical Practice, and the authors retained full responsibility for the accuracy, integrity, and presentation of the data.

The authors thank the people with disabilities, caregivers, and rehabilitation providers who participated in this study, as well as the assistive technology professionals who assisted with recruitment and device setup.

A generative AI tool (Google Gemini, version 2.5) was used under the researcher's supervision to assist with the organization of qualitative text and preliminary theme labeling during analysis. All final coding, thematic interpretation, and manuscript writing decisions were conducted and verified by the authors.

Authors' Contributions

Conceptualization: ADF, JPD

XSL•FO
RenderX

Methodology: ADF, HK, JPD
Formal analysis: ADF, RBO
Investigation: RBO, HK
Project administration: ADF, RBO, HK
Supervision: ADF
Writing—original draft: ADF, RBO, JPD
Writing—review and editing: ADF, RBO, HK, JPD
ADF served as the principal investigator and oversaw adherence to the institutional review board–approved protocol. To minimize potential bias, ADF did not participate in data collection or have direct contact with research participants. JPD directed activities related to International Organization for Standardization (ISO) 14155 compliance and the Clinical Investigation Plan. Due to his ownership interest in DESĪN LLC, JPD did not have access to identifiable participant data and did not participate in data analysis. All authors reviewed, revised, and approved the final manuscript.

Conflicts of Interest

ADF discloses significant financial interests and relevant financial relationships with DESĪN LLC, the developer and manufacturer of the Obi robotic feeding device. She currently serves as Director of Clinical Affairs at DESĪN LLC, where she oversees clinical research, evidence development, and reimbursement strategy for the device. During the implementation of the study, she was contracted by DESĪN LLC as an independent consultant. RBO reports no direct financial relationship with DESĪN LLC. He received compensation for his contributions to this study through part-time employment with The Driving Doctor: Adaptive Driving and Rehabilitation Solutions under the supervision of ADF. HK discloses significant financial interests and relevant financial relationships with DESĪN LLC. She is a paid employee of DESĪN LLC and served as a subinvestigator on the research activities related to the Obi3 described in this manuscript. JPD discloses a significant financial interest through equity ownership in DESĪN LLC, which constitutes a potential financial conflict of interest. To manage this conflict, he was restricted from access to subject identification, data collection, and analysis activities. All conflicts of interest were reviewed and approved by the WCG Institutional Review Board, which provided independent oversight to ensure compliance with US Food and Drug Administration regulations (21 CFR Part 54) and International Council for Harmonization E6 Good Clinical Practice guidelines.

References

1. World Health Organization. International Classification of Functioning, Disability and Health (ICF). Geneva, Switzerland: World Health Organization; 2001.
2. Occupational therapy practice framework: domain and process. 4th ed. American Occupational Therapy Association. Bethesda, MD: AOTA Press; 2020. URL: https://www.aota.org/practice/domain-and-process/framework [accessed 2025-09-01]
3. American Occupational Therapy Association. The practice of occupational therapy in feeding, eating, and swallowing. Am J Occup Ther 2017;71(Supplement_2):7112410015. [doi: 10.5014/ajot.2017.716S04] [Medline: 29308999]
4. Hammell KW. Engagement in Living: Critical Perspectives on Occupation, Rights, and Wellbeing. In: Canadian Journal of Occupational Therapy. Ottawa: Canadian Association of Occupational Therapists; 2020.
5. Mandy A, Sims T, Stew G, Onions D. Manual feeding device experiences of people with a neurodisability. Am J Occup Ther 2018;72(3):7203345010. [doi: 10.5014/ajot.2018.025353] [Medline: 29689183]
6. Keller JW, Fahr A, Lieber J, Balzer J, van Hedel HJA. Impact of upper extremity impairment and trunk control on self-care independence in children with upper motor neuron lesions. Phys Ther 2021;101(8):pzab112 [FREE Full text] [doi: 10.1093/ptj/pzab112] [Medline: 34464449]
7. Tang L, Chen X, Cao S, Wu D, Zhao G, Zhang X. Assessment of upper limb motor dysfunction for children with cerebral palsy based on muscle synergy analysis. Front Hum Neurosci 2017;11:130 [FREE Full text] [doi: 10.3389/fnhum.2017.00130] [Medline: 28386223]
8. Grad LI, Rouleau GA, Ravits J, Cashman NR. Clinical spectrum of amyotrophic lateral sclerosis (ALS). Cold Spring Harb Perspect Med 2017;7(8):a024117 [FREE Full text] [doi: 10.1101/cshperspect.a024117] [Medline: 28003278]
9. Pierzchlewicz K, Kępa I, Podogrodzki J, Kotulska K. Spinal muscular atrophy: the use of functional motor scales in the era of disease-modifying treatment. Child Neurol Open 2021;8:2329048X211008725 [FREE Full text] [doi: 10.1177/2329048X211008725] [Medline: 33997096]
10. Bhat SG, Shin AY, Kaufman KR. Upper extremity asymmetry due to nerve injuries or central neurologic conditions: a scoping review. J Neuroeng Rehabil 2023;20(1):151 [FREE Full text] [doi: 10.1186/s12984-023-01277-7] [Medline: 37940959]
11. Moritz C, Field-Fote EC, Tefertiller C, van Nes I, Trumbower R, Kalsi-Ryan S, et al. Non-invasive spinal cord electrical stimulation for arm and hand function in chronic tetraplegia: a safety and efficacy trial. Nat Med 2024;30(5):1276-1283 [FREE Full text] [doi: 10.1038/s41591-024-02940-9] [Medline: 38769431]
12. Cavalcanti A, Amaral MF, Silva E Dutra FCM, Santos AVF, Licursi LA, Silveira ZC. Adaptive eating device: performance and satisfaction of a person with Parkinson's disease. Can J Occup Ther 2020;87(3):211-220. [doi: 10.1177/0008417420925995] [Medline: 32410462]

XSL•FO
RenderX

13. Westergren A, Karlsson S, Andersson P, Ohlsson O, Hallberg IR. Eating difficulties, need for assisted eating, nutritional status and pressure ulcers in patients admitted for stroke rehabilitation. J Clin Nurs 2001;10(2):257-269. [Medline: 11820347]
14. Hestevik CH, Molin M, Debesay J, Bergland A, Bye A. Older patients' and their family caregivers' perceptions of food, meals and nutritional care in the transition between hospital and home care: a qualitative study. BMC Nutr 2020;6:11 [FREE Full text] [doi: 10.1186/s40795-020-00335-w] [Medline: 32206325]
15. Miller C, Jones SP, Bangee M, Martinez-Garduno CM, Brady MC, Cadilhac DA, et al. Hydration and nutrition care practices in stroke: findings from the UK and Australia. BMC Nurs 2023;22(1):403 [FREE Full text] [doi: 10.1186/s12912-023-01575-4] [Medline: 37891567]
16. Mori H, Nakane A, Yokota Y, Tohara H, Nakayama T. Caregiver burden and eating-related guilt during dysphagia rehabilitation: a descriptive cross-sectional time series study. JPEN J Parenter Enteral Nutr 2024;48(7):818-826. [doi: 10.1002/jpen.2679] [Medline: 39172047]
17. Nanavati A, Alves-Oliveira P, Schrenk T, Gordon EK, Cakmak M, Srinivasa SS. Design principles for robot-assisted feeding in social contexts. 2023 Presented at: Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction; 2023 March 13–16; Stockholm, Sweden. [doi: 10.1145/3568162.3576988]
18. Chapman K, Dixon A, Kendall E, Clanchy K. Defining dignity at the intersection of disability: a scoping review. Disabil Rehabil 2024;46(23):5404-5414 [FREE Full text] [doi: 10.1080/09638288.2024.2302582] [Medline: 38265032]
19. Albarqi MN. Exploring the Effectiveness of Technology-Assisted Interventions for Promoting Independence in Elderly Patients: A Systematic Review. Healthcare 2024;12(21):2105 [FREE Full text] [doi: 10.3390/healthcare12212105] [Medline: 39517318]
20. Chou C, Tzuang Y, Chen M, Lu C. The impact of caregiver feeding assistance on mealtime behavior and caregiver burden. J Nutr Health Aging 2019;23(9):839-846. [doi: 10.1007/s12603-019-1243-5]
21. Zhang Y, Gong Z, Cai J, Yu W, Dai Y, Wang H. Incidence of dysphagia-related safety incidents in older adults across feeding methods: a systematic review and meta-analysis. J Nutr Health Aging 2025;29(5):100522 [FREE Full text] [doi: 10.1016/j.jnha.2025.100522] [Medline: 39985956]
22. Assistive Technology Act of 1998, 29 U.S.C. §3001. Office of the Law Revision Counsel. URL: https://uscode.house.gov/view.xhtml?path=/prelim@title29/chapter31&edition=prelim [accessed 2026-06-30]
23. Cook AM, Polgar JM. Assistive Technologies: Principles and Practice. Amsterdam: Elsevier; 2020.
24. Product classification database: ILC' daily activity assist device, eating. US Food and Drug Administration. 2025. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpcd/classification.cfm?ID=ILC [accessed 2025-10-03]
25. Mealtime P. Mealtime partner adaptive feeding equipment. Mealtime Partners, Inc. 2025. URL: https://www.mealtimepartners.com [accessed 2025-10-08]
26. Neater eater robotic feeding system. Neater Solutions Ltd.: Neater Solutions Ltd URL: https://www.neater.co.uk [accessed 2025-10-08]
27. Obi robotic feeding device. DESĪN LLC. URL: https://meetobi.com [accessed 2025-10-03]
28. Al-Halimi RK, Moussa M. Performing complex tasks by users with upper-extremity disabilities using a 6-DOF robotic arm: a study. IEEE Trans Neural Syst Rehabil Eng 2017;25(6):686-693. [doi: 10.1109/TNSRE.2016.2603472] [Medline: 28113593]
29. IEC 60601-1-11:2020-Amendment 1 - medical electrical equipment - part 1-11: general requirements for basic safety and essential performance - collateral standard: requirements for medical electrical equipment and medical electrical systems used in the home healthcare environment. International Electrotechnical Commission. Geneva, Switzerland: IEC; 2020. URL: https://webstore.iec.ch/en/publication/67384 [accessed 2025-10-23]
30. Obi: adaptive feeding device [device description and FDA registration]. DESĪN LLC. 2025. URL: https://meetobi.com [accessed 2025-10-03]
31. MAUDE - manufacturer and user facility device experience. US Food and Drug Administration. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm [accessed 2025-10-03]
32. EUDAMED: UDI/device registration. European Commission. Brussels, Belgium: European Commission URL: https://health.ec.europa.eu/medical-devices-eudamed/udidevice-registration_en#udidevices-data-sets [accessed 2025-10-15]
33. Burgos BB. The use of the Obi robot for self-feeding with individuals with upper extremity limitations. Assist Technol Outcomes Benefits 2025;19:1-12 [FREE Full text]
34. Howard J, Fisher Z, Kemp AH, Lindsay S, Tasker LH, Tree JJ. Exploring the barriers to using assistive technology for individuals with chronic conditions: a meta-synthesis review. Disabil Rehabil Assist Technol 2022;17(4):390-408. [doi: 10.1080/17483107.2020.1788181] [Medline: 32663110]
35. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry. London, UK: Taylor & Francis; 1996:189-194.
36. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
37. Applying human factors and usability engineering to medical devices: guidance for industry and Food and Drug Administration staff. US Food and Drug Administration. 2016. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices [accessed 2025-10-20]

XSL•FO
RenderX

38. Borsci S, Macredie RD, Martin JL, Young T. How many testers are needed to assure the usability of medical devices? Expert Rev Med Devices 2014;11(5):513-525. [doi: 10.1586/17434440.2014.940312] [Medline: 25033757]
39. Qualtrics XM [software]. Qualtrics. URL: https://www.qualtrics.com [accessed 2025-10-03]
40. Obi quick start guide. DESĪN LLC. URL: https://meetobi.com/device-documents/ [accessed 2025-10-20]
41. Scherer MJ, Craddock G. Matching person and technology (MPT) assessment process. Technol Disabil 2002;14(3):125-131. [doi: 10.3233/TAD-2002-14308]
42. Soares A, Ribeiro O, Ferreira PL, Sousa L. Psychometric properties of the matching person and technology (MPT) model in assistive technology outcomes. Disabil Rehabil Assist Technol 2024;19(1):22-31 [FREE Full text]
43. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980;20(6):649-655. [doi: 10.1093/geront/20.6.649] [Medline: 7203086]
44. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

Abbreviations

AT: assistive technology
ATP: assistive technology professional
BUEMI: bilateral upper extremity motor impairment
DME: durable medical equipment
DOF: degree of freedom
FDA: US Food and Drug Administration
HIPAA: Health Insurance Portability and Accountability Act
ICF: International Classification of Functioning, Disability, and Health
IRB: institutional review board
ISO: International Organization for Standardization
MDNA: Medical Device Needs Assessment
MPT: Matching Person and Technology
SUS: System Usability Scale

Edited by A Scano; submitted 24.Oct.2025; peer-reviewed by J Fei, H Mahajan, V Palama; comments to author 07.Apr.2026; revised version received 15.Jun.2026; accepted 25.Jun.2026; published 12.Aug.2026.

Please cite as:
Fairman AD, Osal RB, Keeton H, Dekar JP
Usability and Functional Outcomes of the Obi3 Robotic Feeding Device Among People With Disabilities, Caregivers, and Providers: Mixed Methods Study
JMIR Rehabil Assist Technol 2026;13:e86338
URL: https://rehab.jmir.org/2026/1/e86338
doi:10.2196/86338
PMID:

©Andrea D Fairman, Ryan B Osal, Heather Keeton, Jonathan P Dekar. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 12.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

AI Scribe in Specialized Rehabilitation: Usability Study of Perceived Usability and Usefulness Among Physicians and Psychologists

Harald Kåre Engan[1], PhD; Ingeborg Landstad[1], M.Sc; Kari Sand[2], PhD; Ståle Evenshaug[1], M.Sc; Einar Henjum Halsnes[1], M.Sc

[1]Department of Research and Development, Unicare Rehabiliation, Oslo, Oslo, Norway

[2]Department of Research and Development, SINTEF Digital, Trondheim, Trøndelag, Norway

Corresponding Author:
Harald Kåre Engan, PhD
Department of Research and Development
Unicare Rehabiliation
Fyrstikkalléen 19
Oslo, Oslo, 0661
Norway
Phone: 47 99516082
Email: harald.kare.engan@unicare.no

Abstract

Background: Clinical documentation in specialized, interdisciplinary rehabilitation is often time-consuming and complex and may contribute to clinician burnout and reduced time for patient interaction. AI-based scribes transcribe and summarize clinical notes and may reduce documentation burden; however, their usefulness and usability in specialized rehabilitation remain uninvestigated.

Objective: This study aimed to assess the use of an ambient AI scribe tool among physicians and psychologists in specialized rehabilitation, focusing on perceived usefulness, usability, note quality, impact on patient interaction, perceptions of time spent on documentation, and documentation-related fatigue.

Methods: Over approximately 4 weeks, physicians and psychologists at 9 specialized rehabilitation centers used an ambient AI scribe during clinical consultations. Usability and user experience were assessed using the System Usability Scale (SUS) and a modified technology acceptance model (TAM). Participants rated an AI-generated note using a modified 8-item Physician Documentation Quality Instrument (PDQI-8) questionnaire to evaluate documentation quality. Additional questions explored perceived impact on patient interaction, documentation-related fatigue, and general attitudes toward AI technology.

Results: Thirty-one clinicians (18 physicians and 13 psychologists) used the AI scribe over approximately one month and completed evaluation questionnaires. The mean SUS score was 78.3 (SD 10.4), with 86.7% of participants scoring above the usability threshold. Perceived usefulness (PU) and perceived ease of use (PEOU) were relatively high (mean PU score 5.65, SD 1.33; mean PEOU score 5.56, SD 1.17), and 87.1% expressed intention to use the AI scribe in future practice. Most participants (80.6%) agreed that using the AI scribe reduced time spent on documentation, while 66.6% reported a reduced perception of documentation-related fatigue. The quality of AI-generated notes was rated relatively high, with a mean modified PDQI-8 score of 4.3 (SD 0.3) out of 5. Comprehensibility received the highest rating (mean 4.6, SD 0.6), whereas thoroughness received the lowest rating (mean 3.9, SD 1.0) among the note quality variables. Overall, 71.0% of participants expressed some level of agreement that the AI scribe enhanced attentiveness during patient consultations, while a similar proportion (70.9%) agreed that it made active listening easier by reducing distractions associated with note-taking. A few participants reported that some notes were incomplete, lacked biopsychosocial details, and contained inaccuracies.

Conclusions: This study demonstrated broad acceptance of an AI-based ambient scribe among physicians and psychologists working in specialized rehabilitation. The AI scribe was generally perceived as valuable, easy to use, and clinically relevant. Most participants reported a reduced perceived documentation burden and improved attentiveness during consultations. These findings suggest that clinicians perceived the AI scribe as supporting more focused patient interactions. Some concerns were noted regarding completeness and occasional inaccuracies, highlighting areas for refinement. Future research should explore

XSL•FO
RenderX

long-term implementation in routine rehabilitation settings and its effects on workflow, documentation quality, and patient experience.

(JMIR Rehabil Assist Technol 2026;13:e90442) doi:10.2196/90442

KEYWORDS

eHealth; artificial intelligence; AI; rehabilitation; technology; usability; adoption; implementation; scribe

Introduction

Clinical documentation in electronic health records (EHRs) is essential for ensuring the accuracy and comprehensiveness of clinical information, promoting shared decision-making, and enhancing care coordination [1-3]. In specialized rehabilitation, health care professionals must thoroughly document the patient's functional status, therapeutic interventions, and progress over time [4,5]. This requires that medical records accurately reflect the interdisciplinary approach to patient care, often incorporating contributions from various health care professionals. At the same time, fulfilling these documentation requirements often places a heavy burden on clinicians [6,7]. For instance, research indicates that many physicians spend several hours outside regular working hours completing EHRs, leading to career dissatisfaction, negative impacts on work-life balance, and disruption of clinician-patient relationships [8,9].

Manual documentation has been associated with increased cognitive demands and distractions and may contribute to clinician burnout because of the excessive time spent on clinical documentation [10]. This burden may adversely affect health care personnel by reducing productivity and contributing to frustration, emotional exhaustion, and potentially compromised patient care [11,12]. Patients have expressed concerns that the increase in EHR-based administrative tasks leads physicians to spend more time at digital devices, making visits feel impersonal and routine [13]. Recent qualitative research shows that electronic documentation in rehabilitation is perceived as burdensome, limiting time for patient interaction and affecting relational quality [5].

In recent years, advances in AI scribes have offered potential solutions to optimize documentation by reducing the time clinicians spend on manual note-taking, alleviating the administrative burden, and improving the quality and consistency of clinical records [14]. AI scribes use advancements in natural language processing (NLP) and machine learning to record, transcribe, and summarize clinical encounters [15]. Automated digital scribes can capture conversations between clinicians and patients and generate documentation for inclusion in the medical record. By automating the documentation process, AI scribes may potentially reduce the administrative burden on health care providers, allowing them to dedicate more time to direct patient care [14,16]. This could enhance the patient experience and improve clinical outcomes [17]. Recent studies suggest that AI-enhanced medical scribe services can maintain high levels of accuracy and consistency in clinical documentation while minimizing errors associated with manual data entry [17]. Furthermore, integrating AI into health care documentation has shown promise in reducing physician burnout by decreasing the time required for manual medical record documentation [18]. Digital scribes may enable clinicians to engage more with patients, maintain eye contact during consultations, and avoid diverting attention from the patient to a computer for manual documentation, which may ultimately lead to more patient-centered care [19].

Although these developments are promising, evidence from peer-reviewed studies on the benefits of digital scribe technology in rehabilitation settings remains limited. AI scribes may help health care providers in specialized rehabilitation meet documentation requirements and allocate more time to direct patient interactions. In addition, AI scribes may help standardize clinical documentation and improve information flow across interdisciplinary teams, potentially improving care coordination. Despite the potential of AI in clinical documentation, concerns remain about the quality of AI-generated notes [16]. Further research is therefore needed to examine this issue in a rehabilitation setting. To date, no studies have examined clinicians' experiences with the latest version of AI scribe technology in specialized rehabilitation. Unicare Rehabilitation is among the first health care institutions to pilot an ambient AI scribe in this context.

The overall aim of this study was to investigate physicians' and psychologists' experiences using ambient AI-scribing technology as a clinical documentation assistant in specialized rehabilitation. This aim was addressed through 2 specific objectives: first, to evaluate perceived usefulness (PU), perceived ease of use (PEOU), overall usability, perceived interaction support (PIS), the quality of generated clinical notes, and the impact on documentation-related fatigue; and second, to explore the relationship between AI scribe usability, measured using the System Usability Scale (SUS), and key acceptance factors in the modified technology acceptance model (TAM), providing preliminary insight into the relevance of usability in specialized rehabilitation.

Methods

Materials

This exploratory evaluation assessed clinicians' perceptions of the usability and usefulness of the AI scribe Noteless (Noteless AS) following 4 weeks of routine clinical use. The study was reported in accordance with relevant recommendations for observational studies [20].

Product Information and Workflow Overview

The AI scribe evaluated in this study was Noteless, a web-based software application designed to transcribe and summarize medical conversations. The system operates through a series of structured steps to streamline clinical documentation. It uses a transformer-based speech-to-text model, refined to handle

XSL•FO
RenderX

medical terminology, to enhance transcription accuracy. The AI scribe was accessible via a web-based app. At the beginning of patient consultations, the clinician initiated the recording session. The device's voice-recording hardware (Nedis [Nedis BV], model MICCU100BK), in conjunction with the AI scribe app, captured the conversation between the clinician and the patient. Upon completion, the recording was transmitted to the cloud for processing by AI designed to interpret and analyze speech. The AI scribe analyzed the recording and generated a draft derived from the recorded dialogue. For summarization, it used a combination of large language models integrated with a customized prompt structure and supplementary rules [21]. During the study period, the evaluated system was Noteless version 2.0. The system was used to transcribe and generate clinical documentation from consultations conducted in Norwegian.

The platform used a model-agnostic, modular architecture that allowed different commercially available large language models to be used for specific documentation tasks. Noteless was designed as a clinical documentation support tool rather than a clinical decision-support system and was not intended to provide diagnostic, treatment, or other clinical recommendations. Generated notes were copied into the clinician's EHR system, where they were reviewed, edited as needed, and approved by the clinician before finalization. Clinicians retained full responsibility for the final documentation.

Audio recordings were processed in real time through Speechmatics (Speechmatics Inc) speech-recognition services hosted on Microsoft Azure infrastructure located in Ireland and Spain and were not retained after transcription. Audio data were deleted immediately once the corresponding text transcript had been generated. Resulting transcripts and AI-generated notes were temporarily stored within the Noteless platform for a maximum of 24 hours in a Microsoft Azure data center located in Norway (Norway East region) to support document generation and clinician review before being securely deleted. AI-based note generation was performed using services hosted on Microsoft Azure infrastructure in Sweden and Google Cloud Platform infrastructure in Belgium. Neither Noteless nor its subprocessors retained patient data following processing or used such data for model training or any other secondary purposes. Privacy and security safeguards included encryption in transit and at rest, access controls based on the principle of least privilege, audit logging, periodic access reviews, contractual data-processing agreements with subprocessors, and automated deletion procedures to ensure removal of transcripts and generated notes after the retention period.

Setting

This cohort study was conducted at Unicare's specialized academic rehabilitation centers in Norway. Unicare Rehabilitation comprises 10 specialized centers offering both inpatient and outpatient services. Many rehabilitation programs in Norway are organized as outpatient programs at hospitals or rehabilitation centers near hospitals and as inpatient programs at rehabilitation centers primarily located in rural areas. The current study was conducted in inpatient rehabilitation centers. Unicare provides specialized interdisciplinary rehabilitation services, including professionals such as physical medicine physicians, psychologists, physiotherapists, nutritionists, and occupational therapists. The rehabilitation centers annually receive approximately 9000 patients across all major diagnostic groups. The study was conducted across all centers from January to February 2025.

Recruitment and Participants

Ambient scribe technology was offered to all physicians and psychologists for use during inpatient and outpatient medical appointments, with informed consent obtained from all participants. A convenience sample was used, comprising physicians and psychologists who were invited to use the ambient scribe technology and agreed to participate in the evaluation. Potential participants were recruited through a dual approach: informal recruitment emails to engage interested clinicians and proactive efforts by department leaders to disseminate information about the study across 9 distinct rehabilitation centers. This recruitment strategy successfully identified 31 clinicians, comprising 18 physicians and 13 psychologists, all of whom consented to participate in the pilot program. Patients verbally consented to the use of the ambient scribing tool during clinical encounters; however, the project team did not collect or review any patient information at any point during the study. Participants received no reimbursement for participating.

Procedure

Prior to the commencement of the 4-week pilot period, participants attended a 45-minute orientation session via video call, during which the project leader provided information about the study and basic instructions on using the ambient scribe tool. The training session primarily emphasized practical aspects of setting up and configuring the ambient AI scribe technology, including hardware and software components. While the session deliberately limited detailed guidance on usage to allow participants to explore the system's ease of use independently, it nonetheless introduced key principles for practical and safe application of the technology in clinical practice.

Participants were instructed to use their microphone to record the encounter and generate summarized documentation that was automatically uploaded to and stored on a secure cloud, accessible on their desktop computer through secure web interfaces for later incorporation and text editing into the EHR. During the initial briefing, participants were also told how to obtain patients' permission to use the AI medical scribe for in-person or virtual encounters. Throughout the study, participants used ambient scribing for specialized rehabilitation, with continuous technological support from support staff.

Patient Information Resources

Informational brochures were developed and distributed to the rehabilitation centers where the ambient AI scribe was implemented. These brochures contained essential information, which participants could summarize verbally during the consent process. Additionally, laminated posters were offered for placement in visible areas within the clinics to inform patients that physicians and psychologists might use AI scribe technology during their consultations.

XSL•FO
RenderX

Inclusion Criteria

Licensed physicians and psychologists working at one of the participating rehabilitation centers are currently involved in patient care and are willing to use an AI scribe during medical appointments.

Exclusion Criteria

Individuals who had previously used any form of AI-based scribe were not eligible to participate.

Outcomes

PU and Usability

The study used a dual-strategy approach, combining the TAM from Information Systems and the SUS from human-computer interaction [22]. Integrating these 2 complementary approaches was intended to provide a comprehensive understanding of both usability and user experience factors, thereby enhancing the evaluation process. In addition, another purpose was to enable a comparison of the measurement tools themselves, offering insights into their respective contributions to assessing technology in this context.

TAM Methodology

TAM is one of the most widely used and researched models for predicting the adoption and use of information technology by individuals, with its validity in health research affirmed by numerous studies [23]. This model highlights 2 key factors that influence user acceptance of technology: PU and PEOU. PU is defined as "the degree to which a person believes that using a particular system would improve his/her job performance" [24], and PEOU is defined as "the degree to which a person believes that using a particular system would be free of effort" [24].

Initially, TAM items elicited likelihood ratings rather than agreement ratings because the model was designed to predict future product use rather than to rate the experience of actual use [24]. Since the aim of this study was to evaluate PU and PEOU following actual product experience rather than to predict future use, the TAM was adapted to align with the usability metric for user experience (UMUX) format [25]. This modification allowed respondents to express their level of agreement with items that reflect real user experience rather than anticipated usage. The TAM questionnaire used in this study consisted of 12 items: 6 to measure PU and 6 to measure PEOU. In addition, 2 dimensions considered determinants of PU according to the TAM methodology were included, as they are highly relevant to the evaluation of an AI scribe: (1) job relevance (JR), defined as "the degree to which an individual believes that the target system applies to his or her job," and (2) quality output (QO), defined as "the degree to which an individual believes that the system performs his or her job tasks well." Additionally, the construct intention to use (IU) was included, defined as "the degree to which an individual intends to use the system." JR, QO, and IU were each measured with 2 items [26]. TAM items were measured on a 7-point Likert scale, from "1=strongly disagree" to "7=strongly agree." The TAM literature does not define any established cutoff values for interpreting mean scores.

SUS

The SUS is among the most used standardized questionnaires for assessing perceived usability [27]. SUS is a 10-item questionnaire that includes an equal number of positive and negative statements. Each statement is rated on a scale from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating better usability. The SUS is a widely used and validated measure of system usability, and extensive normative research offers various methods for interpreting SUS scores, including the Curved Grading Scale (CGS) approach [28]. To calculate the SUS score, one is subtracted from each odd-numbered item score, and 5 is subtracted from each even-numbered item score. Then, these adjusted scores are summed and multiplied by 2.5 to obtain the overall SUS score, which ranges from 0 (extremely poor usability) to 100 (excellent usability). Usability was classified into 7 descriptive categories based on the adjective rating scale proposed by Bangor et al [3]. The corresponding SUS score ranges were "worst imaginable" (0-12.5), "awful" (12.6-20.3), "poor" (20.4-35.7), "OK" (35.8-50.9), "good" (51.0-71.4), "excellent" (71.5-85.5), and "best imaginable" [3,29].

Perception of Documentation Fatigue

Participants were given a question regarding their perception of reduced tiredness following use of the AI scribe: (1) "The use of the AI tool has reduced the experience of fatigue due to decreased documentation work." Items in the questionnaire were measured on a 7-point Likert scale, from "1=strongly disagree" to "7=strongly agree" (Table 1).

Table 1. Participant characteristics.

	All (n=31)	Physicians (n=18)	Psychologists (n=13)
Age (years), mean (range)	47.5 (28-75)	55.4 (33-75)	37.9 (28-56)
Sex (male), n (%)	12 (39)	10 (56)	2 (15)
University education, n (%)	31 (100)	18 (100)	13 (100)

Perceived Benefits and Challenges of Using the AI Scribe

Participants were asked to provide free-text responses to two open-ended questions: (1) Did you experience any benefits from using the AI tool in your work? If so, what were they? and (2) Were there any challenges or difficulties you encountered while using the AI tool in your work? If so, what were they?

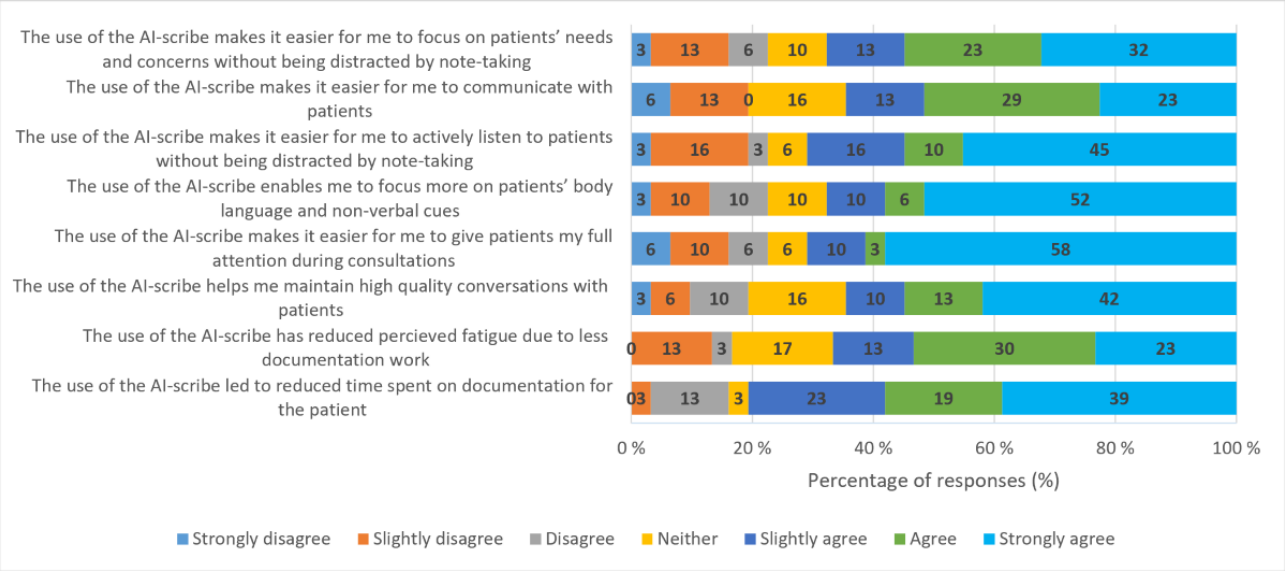
Perceived Impact on Clinician-Patient Interaction

A structured questionnaire, primarily self-developed, with some questions informed by Duggan et al [30], was used to gather data on various aspects of patient interactions. The questions addressed the following dimensions: perceived reduction in time spent on documentation, maintaining high-quality patient conversations, ability to give undivided attention to patients,

XSL•FO
RenderX

focus on nonverbal cues, ease of active listening, overall interaction effectiveness, and understanding of patient needs and concerns. Items were measured on a 7-point Likert scale, from "1=strongly disagree" to "7=strongly agree" (Figure 1).

Figure 1. Distribution of responses to 8 statements assessing the perceived usefulness of the AI scribe in relation to time consumption, patient interaction, and perceived documentation fatigue. Bars represent the percentage of respondents in each Likert category, with the percentages displayed within the segments.



A composite score, referred to as PIS, was created from 6 study-specific items developed for this exploratory study to capture different aspects of how the AI scribe was perceived to influence clinician-patient interaction. The composite score was used as an exploratory summary measure and to facilitate correlation analyses with factors from the TAM framework. This approach is widely used when multiple items reflect different facets of the same underlying concept [31,32]. PIS was designed to capture participants' overall experience of interaction quality when using the AI scribe and should be regarded as a study-specific exploratory measure rather than a validated scale. Higher scores indicate stronger perceived support for interaction. The PIS score was calculated as the mean of the 6 items. The 6 items are provided in Multimedia Appendix 1.

General Attitude Toward AI Technology

Information regarding participants' general attitude toward AI technology was collected using the AI Attitude Scale (AIAS) [33]. The AIAS is a concise self-report instrument designed to evaluate public perceptions of AI technology. Items were measured on a 10-point Likert scale, from "1=not at all" to "10=completely agree."

Quality of Notes

The primary outcome was clinician-rated note quality, assessed using a modified 8-item version of the 9-item Physician Documentation Quality Instrument (PDQI-9) that differed from the original validated instrument by removal of one item and modification of another item. The PDQI-9 has been validated as a reliable tool for evaluating the quality of progress notes and discharge summaries. This tool assesses note quality in 9 domains on a 9-item scale with responses on a 5-point Likert scale ranging from 1 (not at all) to 5 (extremely). The original items are: up-to-date, accurate, thorough, useful, organized, comprehensible, succinct, synthesized, and internally consistent. As in Tierney et al [34], we excluded "up-to-date" because it was deemed not relevant. In addition, we changed the definition of the domain "synthesized" from "reflects the author's understanding of the patient's status and ability to develop a plan of care" to "The note reflects my understanding of the patient's status as it was presented during the consultation." This change was made because, in the context of this study, the author was the AI scribe itself. Thus, the total score for the modified 8-item Physician Documentation Quality Instrument (mPDQI-8) could range from 8 (worst possible note quality) to 40 (best possible quality). Participants were instructed to rate an AI-generated admission note as presented by the AI scribe. This rating was conducted during the final week of the study.

Data Collection Method

The study used a Microsoft Forms survey to assess PU, PEOU, JR, QO, IU, SUS, patient interaction, perceived documentation fatigue, evaluation questions, and AIAS after the 4-week test period. Participants were instructed to complete a paper version of the modified PDQI-8 to evaluate notes generated by the AI scribe after patient consultations and then return them to the project leader.

Analysis

Descriptive statistics were calculated for all variables, including sociodemographic characteristics and scores from the TAM, SUS, AIAS, and custom items related to patient interaction. Group comparisons by profession were conducted using independent-samples Student *t* tests or Mann-Whitney U tests. Pearson correlation coefficients were computed to assess associations between constructs. Following the approach described by Venkatesh and Bala [35], mean item scores were calculated for each instrument. Construct-level scores were then derived by computing the mean of item means, which were used

XSL•FO
RenderX

in subsequent statistical analyses. Given the exploratory nature of the study and the modest sample size, no adjustment for multiple testing was performed. Correlation analyses and subgroup comparisons were therefore interpreted as exploratory and hypothesis-generating. To evaluate the internal consistency and dimensionality of the PIS composite variable, reliability and factor analyses were performed. Internal consistency was assessed using Cronbach alpha, and item quality was examined through corrected item-total correlations and interitem correlations. An exploratory factor analysis (EFA) was conducted to assess the underlying structure of the PIS items, including evaluation of sampling adequacy (Kaiser-Meyer-Olkin [KMO]) and the Bartlett test of sphericity. The PIS score was calculated as the mean of the 6 individual items. Results were considered statistically significant at $P<.05$. Data were analyzed using IBM SPSS Statistics.

Qualitative Analysis of Free-Text Responses

The free-text responses were analyzed by means of qualitative content analysis combining deductive and inductive approaches. The deductive component was based on the 2 open-ended evaluation questions regarding perceived advantages and disadvantages associated with the AI scribe. In the first stage of the analysis, the responses to these 2 questions were analyzed separately. Initially, one researcher (IL) gained an overall understanding of the topics addressed and sorted the responses into preliminary codes for each question. This coding structure was reviewed and discussed by 3 researchers (HKE, IL, and KS), who jointly refined the codes into a final codebook. The names of the codes were based on participants' own wording.

In the next step, the entire dataset was coded. Codes were compared in terms of similarities and differences in content and grouped into preliminary code groups. Through several rounds of discussion and revision among 4 researchers, the code groups were abstracted into categories and subtopics (Table 2). The categories represent broader areas of participants' experiences, while the subtopics describe specific aspects within these areas. The coding process and the development of categories and subtopics were inductive.

During the process of developing categories, the research team merged responses from the 2 open-ended questions into a single dataset since the analysis revealed that the participants described both advantages and disadvantages within the same category. For example, in the category of quality of clinical documentation, it was identified as an advantage that the AI scribe improved summaries and captured clinically relevant information and as a disadvantage that the tool occasionally drew incorrect conclusions or generated text that required manual revision. The predefined dimensions of advantages and disadvantages were retained by including a separate column in Table 2 (inspired by Duggan et al [30]). The organization of codes, the development of categories and subthemes, and the selection of illustrative quotations were discussed and refined by all 4 members of the research team. Any differences in interpretation were resolved through discussion until consensus was reached. The study did not include a member-checking procedure.

The originally Norwegian free-text responses were translated into English at the final stage of the analysis. An AI tool was prompted to translate the responses into language resembling the style used by health care professionals. All translations were subsequently reviewed and, where necessary, revised by 3 researchers. The analysis team consisted of 4 researchers with experience in specialized rehabilitation, technology implementation, health services research, and qualitative and quantitative research methods.

XSL•FO
RenderX

Table 2. Key themes from qualitative comments.

Category and subtopic	Quotes (+ indicates advantages; – indicates disadvantages)
Workload related to documentation task	
Stress and fatigue	• (+) “Less time spent on documentation – reduced fatigue” (K2) • (+) “Time-saving, reduces stress” (K19) • (+) “Easier to go back and write/complete notes later the same day. Noticeably less stress during the workday.” (K22) • (+) “Gives me more time to prepare between sessions” (K29) • (+) “Good summaries of the conversation drastically reduce the time and energy spent on documentation.” (K4)
Documentation burden	• (+) “Because a note was already generated, it was much easier for me to finalize the note right after a session. The threshold for starting documentation was lower since the task felt smaller and already started, which meant I had less backlog.” (K3)
Time spent	
Initial draft/preliminary note	• (+) “Repeats all findings during the examination and explains more to us along the way. Provides a good starting point for documentation if there isn’t time to finish writing between sessions.” (K20) • (+) “Captured the main essence of the conversation, which was transcribed and could be used with only minor edits.” (K11)
Quality of clinical documentation	
Summary	• (+) “Good summary and structure in the notes” (K22) • (+) “More time, more complete notes, better note structure” (K25) • (+) “It summarized longer conversations just fine. Provided a quick starting point for the journal entry.” (K15)
Captures more information	• (+) “It could sometimes pick up things that I didn’t manage to note during the consultation, which was valuable when transcribing the clinical examination.” (K9) • (+) “Reduced risk of missing something important the patient says during the conversation” (K5)
Generates suggestions for follow-up interventions	• (+) “Good summaries of relevant topics in the consultation, along with useful suggestions for the assessment section of the note.” (K10) • (+) “Helpful that AI can assist in generating follow-up interventions after the summary and occasionally provide additional advice.” (K25) • (+) “It did a good job with simple assessments as suicide risk evaluation, and it was helpful for gathering background information/anamnesis.” (K32) • (+) “Suggestions for referrals are useful.” (K4)
Generates its own clinical evaluations	• (–) “It sometimes picked up that the patient was using medications that hadn’t been mentioned at all during the session.” (K32) • (–) “It also added a bit too many interpretations, which means things can easily end up in the note that you don’t fully stand behind if you’re not careful when reviewing.” (K15) • (–) “I’ve also noticed that it occasionally provides incorrect information, which makes it hard to fully trust the tool. The program’s recommendations for next steps are often not very useful for my practice.” (K9)
Selects information	• (+) “Being able to talk with the patient calmly and then get a transcript of the entire conversation, excluding small talk, is a big advantage.” (K17) • (–) “It filters out parts of the conversation about upbringing, conflicts, etc., which can be important to include.” (K18) • (–) “It doesn’t write exactly the way I want, and not everything I’d like to include makes it into the note.” (K11) • (–) “There’s a need for a template, especially to include biopsychosocial context in the note structure and in the summary/assessment section. I feel the wording lacks some holistic perspective.” (K19) • (–) “The software is very selective in? What it listens to, leaving out large parts of information that I find relevant, especially during longer conversations.” (K9)
Proofreading/editing	• (–) “Overall, it took longer to edit the AI-generated note than to write my own notes, and my notes had higher quality and greater relevance to the patient’s course of treatment.” (K32) • (–) “Small errors can slip in and need to be caught during editing/proofreading.” (K23)
Quality of patient consultation	

XSL•FO
RenderX

Category and subtopic	Quotes (+ indicates advantages; – indicates disadvantages)
Maintains focus on the patient	• (+) "I could focus on history-taking and the clinical examination." (K8) • (+) "I was more focused on the conversation with the patient and wrote a good summary at the end." (K5) • (+) "Able to focus on the conversation without having to take too many notes along the way." (K29) • (+) "Being able to talk with the patient calmly and then get a transcript of the entire conversation, excluding small talk, is a big advantage." (K17)
Patient can access the note immediately after the consultation	• (+) "Reading through the AI note together with the patient helps create a neutral summary of the session and invites the patient to correct or share an alternative perspective without feeling contradicted." (K4)
Sensitive topics	• (–) "I feel the recording limits the ability to talk about difficult topics." (K12)

Role of the Vendor

Noteless provided access to the AI scribe software free of charge during the study period and assisted with user registration in the system. The company provided general information about the technology but did not provide training to participants. Noteless had no role in the study design, data collection, data analysis, interpretation of results, manuscript preparation, or the decision to submit the manuscript for publication.

Ethical Considerations

The study was conducted in accordance with the Norwegian Research Ethics Act, the National Guidelines for Research Ethics, and the Declaration of Helsinki. Approval from a Regional Committee for Medical and Health Research Ethics (REK) was not required under Norwegian regulations. Unicare, as the data controller, was responsible for ensuring a lawful basis for processing personal data in accordance with the General Data Protection Regulation (GDPR). All participants provided informed consent prior to participation, and the processing of personal data was based on consent in accordance with Article 6(1)(a) and Article 9(2)(a) of the GDPR. Participation was voluntary, and participants were informed that participation or non-participation would have no consequences for their employment. All collected data were de-identified and stored securely, and only authorized members of the research team had access to the data. The authors declare no conflicts of interest.

Results

Sample Characteristics

A total of 33 physicians and psychologists attended the information meeting and were invited to participate in the study. One physician declined participation, and another did not complete the evaluation questionnaires. The study sample consisted of 31 participants: 18 physicians and 13 psychologists. All participants used the AI scribe for 4 weeks and completed the evaluation questionnaires. Among the physicians, 8 were women, and 10 were men, with a mean age of 55.4 years (SD 12.5). The psychologist group included 11 women and 2 men, with a mean age of 38.3 years (SD 9.0). Among physicians, the mean number of years since graduation was 16.0 (SD 13.8), while among psychologists it was 9.3 years (SD 7.0). Table 1 presents the demographic information of the included participants.

TAM Constructs, Usability, Attitudes to AI Technology, and Patient Interaction Support

All 31 participants completed the questionnaires, although data were missing for 1 participant on the modified PDQI-8 and SUS. The most relevant and statistically meaningful correlations are reported below, along with other observed results. Table 3 shows the correlations between TAM factors, SUS, AIAS, PIS, and categorical variables.

XSL•FO
RenderX

Table 3. Pearson correlations between TAM[a] factors, SUS[b], AIAS[c], PIS[d], and categorical variables.

	PU[e]	PEOU[f]	SUS	BI[g]	JR[h]	QO[i]	PIS	AIAS	Age
PU, *r* (*P* value)	1	0.612 (.001)	0.433 (.017)	0.585 (<.001)	0.731 (.001)	0.843 (.001)	0.778 (.001)	0.541 (.002)	0.168 (.37)
PEOU, *r* (*P* value)	0.612 (.001)	1	0.794 (.001)	0.505 (.004)	0.332 (.07)	0.652 (<.001)	0.519 (.003)	0.389 (.03)	0.231 (.20)
SUS, *r* (*P* value)	0.433 (.017)	0.794 (.001)	1	0.601 (.001)	0.296 (.11)	0.486 (.006)	0.413 (.02)	0.355 (.054)	–0.023 (.90)
BI, *r* (*P* value)	0.585 (<.001)	0.505 (.004)	0.601 (.001)	1	0.766 (.001)	0.490 (.005)	0.481 (.006)	0.718 (.001)	0.021 (.91)
JR, *r* (*P* value)	0.731 (.001)	0.332 (.07)	0.296 (.11)	0.766 (.001)	1	0.552 (.001)	0.560 (.001)	0.651 (.001)	0.091 (.63)
QO, *r* (*P* value)	0.843 (.001)	0.652 (<.001)	0.486 (.006)	0.490 (.005)	0.552 (.001)	1	0.832 (.001)	0.411 (.02)	0.205 (.27)
PIS, *r* (*P* value)	0.778 (.001)	0.519 (.003)	0.413 (.02)	0.481 (.006)	0.560 (.001)	0.832 (.001)	1	0.466 (.008)	0.293 (.11)
AIAS, *r* (*P* value)	0.541 (.002)	0.389 (.03)	0.355 (.054)	0.718 (.001)	0.651 (.001)	0.411 (.02)	0.466 (.008)	1	–0.011 (.95)
Age, *r* (*P* value)	0.168 (.37)	0.231 (.20)	–0.023 (.90)	0.021 (.91)	0.091 (.63)	0.205 (.27)	0.293 (.11)	–0.011 (.95)	1
Profession, *r* (*P* value)	–0.046 (.81)	0.019 (.92)	0.068 (.72)	–0.008 (.97)	–0.112 (.55)	–0.108 (.56)	–0.233 (.21)	0.041 (.82)	0.631 (.001)

[a]TAM: technology acceptance model.

[b]SUS: System Usability Scale.

[c]AIAS: AI attitude scale.

[d]PIS: perceived interaction support.

[e]PU: perceived usefulness.

[f]PEOU: perceived ease of use.

[g]BI: behavioral intention.

[h]JR: job relevance.

[i]QO: quality output.

PU

Twenty-two of 31 participants (71.0%) agreed or strongly agreed that the AI scribe was useful, choosing either “agree” or “strongly agree” on the Likert scale. The mean PU score across all participants was 5.65 (SD 1.33), with physicians scoring 5.71 (SD 1.15) and psychologists 5.59 (SD 1.56). There was no statistically significant difference between groups (*P*=.81; Cohen *d*=0.09; 95% CI –0.62 to 0.80). Participants most often strongly agreed with the item “I find the system useful in my job” (51.6%), followed by “Using the AI scribe in my job enables me to accomplish tasks more quickly” (45.2%; Table 4). PU was positively correlated with. No significant correlation was found between PU and age or profession (Table 3).

XSL•FO
RenderX

Table 4. Descriptive statistics for perceived usefulness, perceived ease-of-use, job relevance, intention to use, and quality output[a].

TAM[b] statements	Rating Scale							
	Strongly disagree	Disagree	Slightly disagree	Neither	Slightly agree	Agree	Strongly agree	Mean (SD)
PU[c] Items								
Using the AI scribe in my job enables me to accomplish tasks more quickly	1 (3.2)	0 (0.0)	1 (3.2)	2 (6.5)	4 (12.9)	9 (29.0)	*14 (45.2)* [d]	5.94 (1.41)
Using the AI scribe improves my job performance	1 (3.2)	1 (3.2)	0 (0.0)	3 (9.7)	*11 (35.5)*	8 (25.8)	7 (22.6)	5.39 (1.41)
Using the AI scribe in my job increases my productivity	1 (3.2)	1 (3.2)	2 (6.5)	3 (9.7)	5 (16.1)	*12 (38.7)*	7 (22.6)	5.39 (1.54)
Using the system enhances my effectiveness on the job	1 (3.2)	1 (3.2)	2 (6.5)	2 (6.5)	4 (12.9)	*13 (41.9)*	8 (25.8)	5.52 (1.55)
Using the system makes it easier to do my job	0 (0.0)	1 (3.2)	0 (0.0)	5 (16.1)	5 (16.1)	*12 (38.7)*	8 (25.8)	5.65 (1.23)
I find the system useful in my job	0 (0.0)	1 (3.2)	0 (0.0)	4 (12.9)	4 (12.9)	6 (19.4)	*16 (51.6)*	6.00 (1.32)
PEOU[e] Items								
Learning to operate the AI scribe was easy for me	0 (0.0)	1 (3.2)	1 (3.2)	2 (6.5)	4 (12.9)	6 (19.4)	*17 (54.8)*	6.06 (1.34)
I found it easy to get the AI scribe to do what I want it to do	1 (3.2)	1 (3.2)	4 (12.9)	4 (12.9)	*9 (29.0)*	*9 (29.0)*	3 (9.7)	4.87 (1.48)
My interaction with the AI scribe was clear and understandable	0 (0.0)	1 (3.2)	2 (6.5)	3 (9.7)	4 (12.9)	7 (22.6)	*14 (45.2)*	5.81 (1.45)
I find the AI scribe to be flexible to interact with	0 (0.0)	0 (0.0)	4 (12.9)	5 (16.1)	6 (19.4)	*8 (25.8)*	*8 (25.8)*	5.35 (1.38)
It was easy for me to become skillful at using the AI scribe	1 (3.2)	1 (3.2)	1 (3.2)	4 (12.9)	*10 (32.3)*	9 (29.0)	5 (16.1)	5.19 (1.42)
I find the AI scribe easy to use	0 (0.0)	0 (0.0)	2 (6.5)	0 (0.0)	5 (16.1)	9 (29.0)	*15 (48.4)*	6.13 (1.12)
Behavioral intention/intention to use								
Assuming I have access to the AI scribe, I intend to use it	0 (0.0)	0 (0.0)	3 (9.7)	0 (0.0)	1 (3.2)	5 (16.1)	*22 (71.0)*	6.39 (1.23)
Given that I have access to the AI scribe, I predict that I would use it.	0 (0.0)	1 (3.2)	1 (3.2)	1 (3.2)	1 (3.2)	4 (12.9)	*23 (74.2)*	6.42 (1.26)
Job relevance								
In my job, usage of the AI scribe is important	1 (3.2)	0 (0.0)	2 (6.5)	6 (19.4)	9 (29.0)	3 (9.7)	*10 (32.3)*	5.29 (1.53)
In my job, usage of the AI scribe is relevant	1 (3.2)	0 (0.0)	0 (0.0)	1 (3.2)	10 (32.3)	5 (16.1)	*14 (45.2)*	5.90 (1.33)
Output quality								
The quality of the output I get from the AI scribe is high	0 (0.0)	2 (6.5)	3 (9.7)	4 (12.9)	8 (25.8)	*8 (25.8)*	6 (19.4)	5.13 (1.48)
I have no problem with the quality of the AI scribe`s output	2 (6.7)	2 (6.7)	3 (10.0)	5 (16.7)	7 (23.3)	*8 (26.7)*	3 (10.0)	4.63 (1.69)

[a]Absolute number and percentage of participants in agreement with the statements.

[b]TAM: technology acceptance model

[c]PU: perceived usefulness.

[d]Italic text indicates the category with the highest response proportion.

[e]PEOU: perceived ease of use.

XSL•FO
RenderX

PEOU

Twenty-four of 31 participants (77.4%) expressed agreement with the ease of use of the AI scribe, selecting "agree" or "strongly agree" on the Likert scale. The mean PEOU score was 5.56 (SD 1.17) across all participants, with physicians scoring 5.54 (SD 1.16) and psychologists 5.58 (SD 1.22). There was no statistically significant difference between groups (P=.92; Cohen d=–0.04, 95% CI –0.74 to 0.67). Participants most often (strongly) agreed with the item "Learning to operate the AI scribe was easy for me" (54.8%), followed by "I find the AI scribe easy to use" (54.8%; Table 4). There was a significant and positive correlation between PEOU and SUS, QO, PIS, and AIAS. No significant correlation was found between PEOU and age or profession (Table 3).

Behavioral Intention

Twenty-seven of 31 participants (87.1%) agreed with the intention to use the AI scribe in the future, selecting either "agree" or "strongly agree" on the Likert scale. The mean BI score was 6.40 (SD 1.23) across all participants, with physicians scoring 6.41 (SD 0.96) and psychologists 6.39 (SD 1.55). There was no statistically significant difference between groups (P=.97; Cohen d=0.02, 95% CI –0.69 to 0.72). There was a significant positive correlation between BI and PEOU, SUS, and PU. No significant correlation was found between BI and age or profession (Table 3).

JR

Across all participants, the mean JR score was 5.63 (SD 1.35). Physicians had a mean score of 5.76 (SD 1.06), while psychologists had a mean score of 5.46 (SD 1.67). There was no statistically significant difference between groups (P=.55; Cohen d=0.22, 95% CI –0.49 to 0.93).

No significant correlation was found between JR and age or profession (Table 3).

QO

The mean QO score was 4.89 (SD 1.48) across all participants, with physicians scoring 5.03 (SD 1.44) and psychologists 4.71 (SD 1.57). There was no statistically significant difference between groups (P=.57; Cohen d=0.21, 95% CI –0.50 to 0.92). No significant correlation was found between QO and age or profession (Table 3).

System Usability

In total, 26 of 30 participants (86.7%) considered the usability of the AI scribe to be acceptable (SUS score >70). The mean SUS score was 78.3 (SD 10.4) across all participants, with physicians scoring 77.6 (SD 10.3) and psychologists 79.0 (SD 10.9). According to the adjective rating scale by Bangor et al [3] and van der Nat et al [29], this corresponds to the "Excellent" category, indicating a strong perceived usability of the AI scribe. There was no statistically significant difference between groups (P=.72; Cohen d=–0.13, 95% CI –0.85 to 0.59). Participants most often strongly agreed with the item "I think that I would like to use the AI scribe frequently" (80.6%, 25/31), followed by "I would imagine that most people would learn to use the AI scribe very quickly" (70.0%, 21/30). There was a significant positive correlation among SUS, QO, and PIS. No significant correlation was found between SUS and AIAS, age, or profession (Table 3).

General Attitude Toward AI Technology

Across all participants, the average AIAS score was 7.84 (SD 1.60). Physicians had a mean score of 7.78 (SD 1.49), while psychologists had a mean score of 7.91 (SD 1.78). There was no statistically significant difference between groups (P=.83; Cohen d=–0.08, 95% CI –0.79 to 0.63). There was a significant and positive correlation between AIAS and PU, PIS, BI, and JR. No significant correlation was found between AIAS and age or profession (Table 3).

Perception of Documentation Time and Documentation Fatigue

Most participants (80.6%, 25/31) expressed some agreement that the AI scribe could help them save time on documentation. About two-thirds of participants (66.6%, 20/30) expressed some level of agreement that the AI scribe had reduced fatigue due to a lighter documentation burden (Figure 1).

PIS

Most participants agreed that the AI scribe provided positive support for patient interaction (Figure 1). Across multiple items, responses clustered in the upper range of the Likert scale, indicating that the AI scribe facilitates attentiveness, active listening, and patient-centered interaction. Specifically, participants reported that the AI scribe helps them maintain focus on patients' verbal and nonverbal cues (67.7%, 21/31), reduces distractions from notetaking (70.9%, 22/31), and supports clearer, more engaged dialogue (64.5%, n0/31); Figure 1).

Qualitative Comments

In their open-ended feedback, participants described the AI tool as time-saving and reported that it reduced stress and fatigue in daily practice (Table 2). They highlighted that the generated summaries provided a starting point that could be finalized more quickly. Furthermore, participants reported that the AI scribe reduced the documentation backlog, allowed more time between consultations, and decreased the time and energy spent on notetaking.

Psychometric Properties of the PIS Scale

The items of PIS showed excellent internal consistency (Cronbach α=0.983), with corrected item-total correlations ranging from 0.872 to 0.976 and interitem correlations between 0.84 and 0.96. An EFA supported a unidimensional structure, with one factor explaining 92.5% of the total variance (KMO=0.924; Bartlett test P<.001), indicating that the items measured a coherent construct. The internal consistency remained high when any single item was removed, indicating that all items contributed meaningfully to the composite score. There was a significant and positive correlation between PIS and PU, PEOU, SUS, and AIAS. No significant correlation was found between PIS and age or profession (Table 3).

Quality of Notes

The mean modified PDQI-8 score across all participants was 4.3 (SD 0.3). Physicians had a mean score of 4.3 (SD 0.6), and

XSL•FO
RenderX

psychologists had a mean score of 4.3 (SD 0.6). There was no difference between the groups on the mean modified PDQI-8 score (P=.81). Overall ratings across documentation domains are summarized in Table 4. The thoroughness domain had the lowest mean score (3.9), followed by accuracy (4.1) and synthesis (4.1). The highest mean scores across all participants were observed for comprehensibility (4.6), internal consistency (4.5), and usefulness (4.4). Psychologists reported lower scores on the domain succinct than physicians (3.7 vs 4.7, P=.01; Table 5). One participant had a missing response for the modified PDQI-8 item "Internally consistent"; therefore, analyses for this item were based on 30 responses.

Table 5. Ratings of documentation quality in the modified PDQI-8[a] by domain.

Domain	All (n=30), mean (SD)	Physicians (n=18), mean (SD)	Psychologists (n=12), mean (SD)	*P* value	Cohen *d* (95% CI)
Accurate	4.1 (0.8)	4.1 (0.7)	4.0 (1.0)	.71	0.15 (–0.62 to 0.92)
Thorough	3.9 (1.0)	3.7 (1.3)	4.1 (0.5)	.31	–0.41 (–1.18 to 0.37)
Useful	4.4 (0.6)	4.4 (0.8)	4.5 (0.5)	.51	–0.26 (–1.03 to 0.51)
Organized	4.3 (0.7)	4.4 (0.7)	4.2 (0.8)	.38	0.35 (–0.43 to 1.12)
Comprehensible	4.6 (0.6)	4.7 (0.5)	4.5 (0.8)	.36	0.37 (–0.41 to 1.14)
Succinct	4.3 (1.0)	4.7 (0.5)	3.7 (1.3)	.01[b]	1.08 (0.25 to 1.90)
Synthesized	4.1 (0.7)	4.2 (0.8)	4.0 (0.6)	.50	0.27 (–0.51 to 1.03)
Internally consistent	4.5 (0.6)	4.5 (0.7)	4.5 (0.7)	.96	0.02 (–0.76 to 0.80)
Mean total (SD)	4.3 (0.5)	4.3 (0.5)	4.2 (0.6)	.46	0.29 (–0.49 to 1.06)

[a]PDQI-8: 8-item Physician Documentation Quality Instrument.

[b]Significance at P<.05.

In the open-ended questions, clinicians were divided in their experiences regarding the quality of the generated summaries. Some reported that the summaries were well-structured and occasionally captured more than they themselves had noted during the consultation, such as "it captured things I did not manage to note during the consultation" and "reduced the risk of missing out something important during the conversation." Some participants described the AI scribe as selective in what it included. While some clinicians appreciated that "small talk" was excluded, others considered it a drawback that important information was occasionally omitted. Some noted that the summaries lacked coherence and did not sufficiently capture biopsychosocial aspects, such as upbringing and conflicts, describing them as "lacking a holistic perspective." Participants reported that the summaries contained enough inaccuracies to require editing and proofreading (Table 2).

In open-ended feedback, clinicians reported that the AI scribe allowed them to draw more attention to the patient during consultations. They described being able to speak directly with the patient without taking handwritten notes or typing on a computer as an advantage. Another advantage mentioned was the ability to share the generated summary with the patient immediately after the consultation, enabling the patient to review and provide corrections or additional input. As a potential drawback, some clinicians noted that the recording could make some patients uncomfortable and "might limit the ability to talk about difficult topics" (Table 2).

Discussion

Principal Findings

To our knowledge, this is the first study to explore the PU and usability of an AI-based ambient scribe among physicians and psychologists working in specialized multidisciplinary rehabilitation. The results show that the AI scribe was perceived as both valuable and easy to use by most participants, with no significant differences between physicians and psychologists. Usability ratings were high, with over 86% of participants scoring above the accepted threshold on the SUS [3], indicating a positive usability experience. Although the study did not include a comparison with conventional documentation workflows, the mean SUS score of 78.3 exceeded the widely accepted benchmark score of 68 for average usability. Recent evidence has also suggested that this benchmark is applicable to digital health applications, supporting its use for contextualizing the present findings [36]. Comparisons across studies should nevertheless be interpreted cautiously because of differences in technologies, settings, and study populations. While overall perceptions were positive, some participants noted challenges, including occasional inaccuracies, selective omission of relevant information, and insufficient coverage of biopsychosocial aspects. In sum, these findings suggest that the AI scribe was generally well-received and perceived as useful for supporting clinical documentation in specialized rehabilitation settings.

Most participants considered the AI scribe important and relevant to their clinical work. JR showed a strong positive correlation with PU (r=0.731), indicating that clinicians who viewed the AI scribe as applicable to their professional responsibilities also tended to perceive it as beneficial. About 87% of participants reported an intention to use (behavioral intention [BI]) the AI scribe in the future, provided they were given the opportunity. These findings indicate a generally high level of acceptance and willingness to adopt the AI scribe among participants. Moreover, intention to use the AI scribe was positively correlated with PU, PEOU, and SUS. The observed

XSL•FO
RenderX

correlations indicate that PU and PEOU were associated with intention to use the AI scribe, consistent with the relationships proposed in the TAM model [24]. It should be noted that, in contrast to previous studies, PU has shown a clearly more decisive influence than PEOU on AI adoption in health care [37]. Our findings indicate a more balanced contribution of both constructs to BI (PU: r=0.585; PEOU: r=0.505). This may reflect the practical importance of ease of integration into clinical workflows, where time constraints, cognitive load, and the need for seamless documentation make usability a key consideration. However, it cannot be ruled out that the importance of PEOU might have diminished over time had clinicians gained more extensive experience with the AI scribe beyond the 4 weeks. This assumption is based on previous research suggesting that once the PU of a technology is firmly established, ease of use may become a less decisive factor in adoption decisions [38,39]. Nonetheless, in clinical practice, where time constraints, cognitive load, and the need for seamless documentation present ongoing challenges, continued attention to both PU and usability remains essential to support the long-term adoption of ambient AI scribe tools.

In addition to these system-specific factors, participants' general attitudes toward AI were assessed using the AIAS [33]. The mean score of 7.84 out of 10 indicates a broadly positive attitude toward AI. Although no significant differences were found between professions or age groups, AIAS scores correlated positively with PU, PIS, BI, and JR, suggesting that clinicians with more positive attitudes toward AI also tended to evaluate the AI scribe more favorably. It should be noted, however, that AIAS was administered after participants had used the AI scribe. While designed to measure general attitudes, responses may have been influenced by the specific experience, potentially affecting correlation coefficients. This type of response bias is a known concern in postintervention measurement of attitudes, particularly in technology acceptance research [40]. Future studies should consider measuring general AI attitudes prior to exposure to better distinguish pre-existing dispositions from experience-based evaluations.

Beyond general measures of technology acceptance and usability, a majority of participants reported that the AI scribe contributed positively to their clinical workflow. The results of this study suggest that ambient AI-scribing was perceived as contributing to greater efficiency, reduced documentation burden, and more attentive patient interactions during consultations. Specifically, 81% of participants responded with slight to strong agreement (Likert scale: 5–7) that the AI scribe reduced the time spent on documentation for the patient in question. In comparison, 66% reported reduced fatigue, addressing 2 commonly cited challenges in clinical contexts with high documentation demands [41]. While the current study did not directly measure time spent on clinical notetaking, the high level of agreement suggests that many participants perceived a meaningful reduction in documentation burden. This perception aligns with previous research suggesting that AI scribe technology may reduce perceived documentation burden and documentation-related workload while supporting clinician-patient interaction. However, reported effects on objective measures such as documentation time, EHR use, and productivity have varied across studies and clinical settings [42-47]. In their free-text answers, participants perceived the main benefit of the AI scribe as its ability to produce a well-structured initial draft that served as a foundation for documentation. Participants described the AI scribe as helping streamline note creation, reduce typing effort, and decrease time spent on postconsultation tasks. Several participants also perceived benefits related to reduced backlog, stress, and fatigue. On the other hand, some participants mentioned the omission of relevant information, inaccuracies, and increased editing time as disadvantages.

In addition to supporting documentation and workflow efficiency, participants also highlighted relational benefits of the AI scribe. Importantly, many perceived it as a support for patient interaction during consultations. Participants frequently reported that the AI scribe enhanced attentiveness, reduced distractions associated with notetaking, and facilitated more engaged patient interactions. These qualities are highly relevant in specialized rehabilitation, where therapeutic alliance and interaction are central to the quality of care. In fact, providing care in this context is both cognitively and relationally demanding [48,49], requiring clinicians to maintain focused attention to support complex decision-making, deliver individualized education, document clinical encounters, and foster a trusting relationship. Digital tools perceived as reducing cognitive load and administrative distractions, such as ambient AI scribes, may therefore be perceived as supportive of the therapeutic process.

In free-text responses in our study, participants reported that the AI scribe reduced reliance on memory for notetaking, as the generated draft captured key elements of the dialogue and minimized computer interaction during consultations. Participants reported that this supported more face-to-face engagement and attentiveness to patient needs. The positive impressions reported by many participants in our study are not isolated. Similar experiences have been documented in other clinical environments, reinforcing the notion that ambient AI scribe technology may support both interaction and workflow. For instance, studies have shown that such tools are associated with more personal and effective clinician–patient interactions, reduced clerical workload, and improved attentiveness [30,34,43]. In addition, a recently published study from Australia found that the AI scribe had a positive impact on factors considered important to the therapeutic alliance, such as listening, eye contact, and engagement, both among patients and allied health professionals [50]. These observations align with the results of our study, in which many participants reported increased engagement and attention during consultations.

In addition to the feedback provided in free-text fields by participants, this study introduced a composite variable, PIS, to quantitatively assess how use of the AI scribe influenced clinician-patient interaction and to explore its potential impact on technology acceptance constructs. Interestingly, PIS showed a strong correlation with PU (r=0.778), a moderate correlation with PEOU (r=0.519), and a weaker correlation with SUS (r=0.413). These results suggest that PIS was associated with clinicians' perceptions of usefulness and usability. However, the observed correlations reflect associations between variables

XSL•FO
RenderX

and do not permit conclusions about causality or the direction of these relationships. However, it should be noted that while the majority of participants in the present study reported supportive effects on patient interaction, some did not. This variability in PU may reflect contextual factors such as professional role, familiarity with digital tools, or integration into existing workflows. For example, some participants noted that the AI scribe did not align well with their documentation routines or clinical setting, which may have limited its perceived value. Thus, the results presented in this study offer a foundation for further research into how contextual factors may influence clinicians' acceptance of AI scribe technology in health care.

The AI-generated clinical notes were generally perceived as clear, well-organized, and clinically useful, according to assessments using the modified PDQI-8 instrument. Domains such as comprehensibility, internal consistency, and usefulness received particularly positive ratings, while thoroughness and accuracy were rated somewhat lower, indicating potential for improvement. Overall quality ratings were similar between physicians and psychologists. The only significant professional difference was related to succinctness, with psychologists rating the generated notes lower than physicians. This may reflect differing documentation preferences, as psychologists may place greater emphasis on preserving psychosocial context and narrative detail that can be reduced in more concise summaries. However, these findings should be interpreted cautiously given the small sample size and exploratory nature of the analyses. The generally positive PDQI ratings observed in this study are consistent with a recent study evaluating AI-generated clinical documentation, in which ambient AI-generated notes achieved quality ratings broadly comparable to physician-authored notes when assessed using modified versions of the PDQI framework [51]. However, methodological differences, including note types, evaluators, and scoring procedures, limit direct comparison with the present findings.

Assessments of output quality, as conceptualized in the TAM model, indicated that most participants rated the AI scribe's documentation as high quality. It should, however, be noted that a few participants in their free-text responses reported the clinical notes as incomplete and insufficiently addressing biopsychosocial dimensions, including social history, psychological factors, and family and work life. Furthermore, a few participants reported that the notes contained inaccuracies, which increased the time spent on proofreading and editing. One participant reported that the tool occasionally produced erroneous conclusions or misrepresented information from the consultation, including information about medication use. Although only a few participants raised these concerns, they highlight the need for a more comprehensive inclusion of biopsychosocial information to enhance clinical relevance and minimize the need for later editing. As the assessment of documentation quality was based on notes from initial patient consultations, future studies should examine how the AI scribe performs across a broader range of clinical contexts, where documentation needs and expectations may differ.

Although our study, along with previous studies, has found that clinicians generally rate AI-generated clinical notes as good to excellent in quality [52], several categories of errors warrant attention. Hallucinations, or fabricated content not present in the conversation, occurred more frequently in AI notes (31%) than in physician-authored notes (20%) [51]. Omission errors were also common, often involving details such as the history of present illness or social factors [52]. Some systems failed to differentiate and emphasize critical problems over less urgent ones, which may affect clinical prioritization [45]. Structural issues included misclassification in key sections of the clinical note and generalization errors, in which specific information was oversimplified (eg, "Colles fracture" reduced to "fracture") [52]. Finally, contextual factors such as off-topic dialogue and multi-speaker interactions affected accuracy, with some systems struggling to distinguish between 3 or more voices [52]. These findings underscore the need for ongoing refinement of AI scribe algorithms to minimize errors and ensure clinical reliability.

Relationship Between System Usability Score and TAM Constructs

To explore the role of usability in technology acceptance, the study examined how SUS scores related to TAM constructs, including BI. The observed correlation between SUS scores and BI (r=0.601) suggests that participants who perceived the AI scribe as more usable also tended to report a stronger intention to use it, complementing the TAM constructs of PEOU and PU. The stronger correlation between SUS and PEOU (r=0.794) than between SUS and PU (r=0.433) indicates that ease of use and usability may capture similar aspects of user experience. This is particularly relevant given that SUS is a standardized usability measure, whereas PEOU is a theoretical construct within TAM. This observation aligns with previous research suggesting that PEOU can reflect perceptions of usability in technology acceptance contexts [53]. Thus, future research could use the TAM framework to investigate the usability aspects of AI scribes further.

Limitations

The interpretation of this study's results must be tempered with an understanding of its limitations. The participants were from a similar rehabilitation organization, which may limit the generalizability of the findings. There was also a potential sampling bias among physicians and psychologists who volunteered to participate. To better understand the use of AI scribes in rehabilitation practices, future studies should consider clinicians who have concerns about incorporating an AI scribe in their practice. This explorative study also had a modest sample size and a relatively short intervention period, which together may limit the representativeness of the results. Unfortunately, due to technical errors, usage data, including the number of consultations and time spent on documentation, were not available. Consequently, individual exposure to the AI scribe could not be quantified, preventing analysis of actual usage patterns or potential differences in experiences according to frequency use. In addition, the study did not include baseline measurements or a control group, limiting the ability to determine whether the observed perceptions can be attributed specifically to AI scribe use.

Furthermore, the study did not include objective measurements of documentation time, independent assessments of note accuracy, or patient-reported outcomes. As a result, conclusions

XSL•FO
RenderX

are limited to clinicians' self-reported perceptions and cannot establish effects on documentation efficiency, note quality, accuracy, or patient experience. Finally, participants may have been more motivated to engage with new technology than non-participants, and the introduction of a novel tool may have contributed to more favorable responses, resulting in potential selection and novelty bias.

Moreover, self-reported data may introduce recall or reporting bias, limiting the accuracy of the findings. The modified PDQI-8 was administered following a single consultation, limiting the ability to assess the consistency and reliability of note-quality ratings over time. Although the rating of note quality was relatively consistent in this study, repeated assessments across multiple consultations could provide a more comprehensive understanding of the accuracy and quality of clinical notes in routine practice. It should be noted that the free-text answers may have been influenced by the preceding structured questionnaires (SUS, TAM, modified PDQI-8), as participants answered the open-ended questions after completing these scales. Consequently, the identified themes, such as time use, documentation burden, note quality, and patient interaction, reflect categories emphasized in the questionnaires, which may have shaped participants' focus and wording. In addition, questions related to the interaction aspect were phrased as positive statements, consistent with the TAM methodology. This may represent a limitation, as positive wording may lead to a tendency to agree, potentially reducing the ability to capture critical or negative perceptions. A limitation of this study is that participants were not invited to review or comment on the findings following the analysis.

While the composite measure of PIS showed high internal consistency, the strong interitem correlations suggest it may capture a narrow aspect of interaction support. This may imply that the measure primarily captures a specific dimension of clinician-patient interaction, such as attentional presence or reduced distraction, rather than the broader complexity of clinical interaction. Thus, future research should refine item wording to better capture the multifaceted nature of interaction in clinical practice and use qualitative methods to obtain more in-depth data on the communicative aspects of AI scribes.

Conclusion

This study provides initial evidence of acceptance of an AI-based ambient scribe among physicians and psychologists in specialized rehabilitation, with generally positive perceptions of its usefulness, usability, and documentation quality, and future use intentions. Most participants reported a perceived reduction in documentation burden and experienced the AI scribe as supportive of attentiveness during consultations. However, the findings reflect clinicians' subjective experiences and do not establish objective effects on documentation time, documentation accuracy, clinical outcomes, patient safety, or clinician well-being. Participants also reported omissions, occasional inaccuracies, and insufficient coverage of biopsychosocial aspects in generated notes, highlighting the continued need for clinician review and suggesting areas for further development. General attitudes toward AI were associated with key acceptance variables, indicating that broader views on AI may influence how such technologies are received in practice. Overall, the findings suggest that ambient AI-scribing is acceptable to many rehabilitation clinicians and is perceived as supportive of documentation processes and clinician-patient interactions. Future studies should investigate its implementation in routine care using objective measures of documentation time, note accuracy, workflow integration, quality of care, patient outcomes, and safety. Research should also explore long-term effects across diverse clinical contexts and professional roles and improve evaluation instruments to better capture the communicative and clinical impact of AI-assisted documentation.

Acknowledgments

The authors extend special thanks to candidate psychologist Kjetil Fjell, Associate Professor Simone Grassini, and senior researcher Filip Drozd, who provided valuable advice during the initial planning of the study, and to the psychologists and physicians at Unicare Rehabilitation, who volunteered for the study.

Funding

No external funding was obtained.

Data Availability

The dataset generated and/or analyzed in this study are available from the corresponding author upon reasonable request.

Generative AI Use

Microsoft Copilot was used to support language editing. All AI-assisted text was reviewed and revised by the authors, who take full responsibility for the final content.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Study-specific items included in the perceived interaction support (PIS) composite score.
[DOCX File , 31 KB - rehab_v13i1e90442_app1.docx]

References

1. Kuhn T, Basch P, Barr M, Yackel T, Medical Informatics Committee of the American College of Physicians. Clinical documentation in the 21st century: executive summary of a policy position paper from the American College of Physicians. Ann Intern Med 2015 Feb 17;162(4):301-303 [FREE Full text] [doi: 10.7326/M14-2128] [Medline: 25581028]
2. Vos JFJ, Boonstra A, Kooistra A, Seelen M, van Offenbeek M. The influence of electronic health record use on collaboration among medical specialties. BMC Health Serv Res 2020 Jul 22;20(1):676 [FREE Full text] [doi: 10.1186/s12913-020-05542-6] [Medline: 32698807]
3. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123. [doi: 10.5555/2835587.2835589]
4. Kjeken I, Bjor T, Westerlund D. Interdisciplinary documentation of the course of rehabilitation. Tidsskr Nor Laegeforen 2008;128(4):468-470 [FREE Full text] [Medline: 18274584]
5. Schwartz-Dillard J, Ng T, Villegas J, Johnson D, Murray-Weir M. Electronic documentation burden among outpatient rehabilitation therapists: a qualitative descriptive study and quality improvement initiative. J Am Med Inform Assoc 2024 Oct 01;31(10):2347-2355. [doi: 10.1093/jamia/ocae192] [Medline: 39042519]
6. Asan O, D Smith P, Montague E. More screen time, less face time - implications for EHR design. J Eval Clin Pract 2014 Dec;20(6):896-901 [FREE Full text] [doi: 10.1111/jep.12182] [Medline: 24835678]
7. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med 2016 Dec 06;165(11):753-760. [doi: 10.7326/M16-0961] [Medline: 27595430]
8. Tai-Seale M, Olson CW, Li J, Chan AS, Morikawa C, Durbin M, et al. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. Health Aff (Millwood) 2017 Apr 01;36(4):655-662 [FREE Full text] [doi: 10.1377/hlthaff.2016.0811] [Medline: 28373331]
9. Holmgren AJ, Downing NL, Bates DW, Shanafelt TD, Milstein A, Sharp CD, et al. Assessment of electronic health record use between US and non-US Health systems. JAMA Intern Med 2021 Feb 01;181(2):251-259 [FREE Full text] [doi: 10.1001/jamainternmed.2020.7071] [Medline: 33315048]
10. Gesner E, Dykes PC, Zhang L, Gazarian P. Documentation burden in nursing and its role in clinician burnout syndrome. Appl Clin Inform 2022 Oct;13(5):983-990 [FREE Full text] [doi: 10.1055/s-0042-1757157] [Medline: 36261113]
11. Mazur LM, Adapa K, Meltzer-Brody S, Karwowski W. Towards better understanding of workplace factors contributing to hospitalist burden and burnout prior to COVID-19 pandemic. Appl Ergon 2023 Jan;106:103884 [FREE Full text] [doi: 10.1016/j.apergo.2022.103884] [Medline: 36030734]
12. Alobayli F, O'Connor S, Holloway A, Cresswell K. Electronic health record stress and burnout among clinicians in hospital settings: a systematic review. Digit Health 2023;9:20552076231220241 [FREE Full text] [doi: 10.1177/20552076231220241] [Medline: 38130797]
13. Dugdale DC, Epstein R, Pantilat SZ. Time and the patient-physician relationship. J Gen Intern Med 1999 Jan;14 Suppl 1(Suppl 1):S34-S40 [FREE Full text] [doi: 10.1046/j.1525-1497.1999.00263.x] [Medline: 9933493]
14. van Buchem MM, Boosman H, Bauer MP, Kant IMJ, Cammel SA, Steyerberg EW. The digital scribe in clinical practice: a scoping review and research agenda. NPJ Digit Med 2021 Mar 26;4(1):57 [FREE Full text] [doi: 10.1038/s41746-021-00432-5] [Medline: 33772070]
15. Mess SA, Mackey AJ, Yarowsky DE. Artificial intelligence scribe and large language model technology in healthcare documentation: advantages, limitations, and recommendations. Plast Reconstr Surg Glob Open 2025 Jan;13(1):e6450. [doi: 10.1097/GOX.0000000000006450] [Medline: 39823022]
16. Bracken A, Reilly C, Feeley A, Sheehan E, Merghani K, Feeley I. Artificial intelligence (AI) - powered documentation systems in healthcare: a systematic review. J Med Syst 2025 Feb 18;49(1):28. [doi: 10.1007/s10916-025-02157-4] [Medline: 39966286]
17. Lee C, Britto S, Diwan K. Evaluating the impact of artificial intelligence (AI) on clinical documentation efficiency and accuracy across clinical settings: a scoping review. Cureus 2024 Nov;16(11):e73994. [doi: 10.7759/cureus.73994] [Medline: 39703286]
18. Shah SJ, Devon-Sand A, Ma SP, Jeong Y, Crowell T, Smith M, et al. Ambient artificial intelligence scribes: physician burnout and perspectives on usability and documentation burden. J Am Med Inform Assoc 2025 Feb 01;32(2):375-380. [doi: 10.1093/jamia/ocae295] [Medline: 39657021]
19. Ghatnekar S, Faletsky A, Nambudiri VE. Digital scribe utility and barriers to implementation in clinical practice: a scoping review. Health Technol (Berl) 2021;11(4):803-809 [FREE Full text] [doi: 10.1007/s12553-021-00568-0] [Medline: 34094806]
20. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational

XSL•FO
RenderX

studies. Ann Intern Med 2007 Oct 16;147(8):573-577 [FREE Full text] [doi: 10.7326/0003-4819-147-8-200710160-00010] [Medline: 17938396]
21. Boscardin CK, Gin B, Golde PB, Hauer KE. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. Acad Med 2024 Jan 01;99(1):22-27 [FREE Full text] [doi: 10.1097/ACM.0000000000005439] [Medline: 37651677]
22. Persico D, Manca S, Pozzi F. Adapting the technology acceptance model to evaluate the innovative potential of e-learning systems. Computers in Human Behavior 2014 Jan;30:614-622 [FREE Full text] [doi: 10.1016/j.chb.2013.07.045]
23. Hornbæk K, Hertzum M. Technology acceptance and user experience. ACM Trans Comput-Hum Interact 2017 Oct 06;24(5):1-30. [doi: 10.1145/3127358]
24. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340. [doi: 10.2307/249008]
25. Nguyen M, Fujioka J, Wentlandt K, Onabajo N, Wong I, Bhatia RS, et al. Using the technology acceptance model to explore health provider and administrator perceptions of the usefulness and ease of using technology in palliative care. BMC Palliat Care 2020 Sep 07;19(1):138 [FREE Full text] [doi: 10.1186/s12904-020-00644-8] [Medline: 32895060]
26. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
27. Lewis JR. The System Usability Scale: past, present, and future. International Journal of Human–Computer Interaction 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
28. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. Burlington, Massachusetts: Morgan Kaufmann; 2016.
29. van der Nat DJ, Huiskes VJB, Taks M, Pouls BPH, van den Bemt BJF, van Onzenoort HAW. Usability and perceived usefulness of patient-centered medication reconciliation using a personalized health record: a multicenter cross-sectional study. BMC Health Serv Res 2022 Jun 13;22(1):776 [FREE Full text] [doi: 10.1186/s12913-022-07967-7] [Medline: 35698220]
30. Duggan MJ, Gervase J, Schoenbaum A, Hanson W, Howell JT, Sheinberg M, et al. Clinician experiences with ambient scribe technology to assist with documentation burden and efficiency. JAMA Netw Open 2025 Feb 03;8(2):e2460637 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.60637] [Medline: 39969880]
31. Eo S, Park H, Kim YR. Predictive and representative composite score. Qual Quant 2022 Feb 24;57(1):1-14. [doi: 10.1007/s11135-022-01345-5]
32. Joshi A, Kale S, Chandel S, Pal D. Likert scale: explored and explained. Br J Appl Sci Technol 2015;7(4):396-403. [doi: 10.9734/BJAST/2015/14975]
33. Grassini S. Development and validation of the AI attitude scale (AIAS-4): a brief measure of general attitude toward artificial intelligence. Front Psychol 2023;14:1191628 [FREE Full text] [doi: 10.3389/fpsyg.2023.1191628] [Medline: 37554139]
34. Tierney A, Gayre G, Hoberman B, Mattern B, Ballesca M, Kipnis P. Ambient artificial intelligence scribes to alleviate the burden of clinical documentation. NEJM Catal Innov Care Deliv 2024;5(3):CAT.23.0404. [doi: 10.1056/CAT.23.0404]
35. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
36. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
37. Chaibi A, Brini R, Flah IB, Sobaih AE. Understanding AI acceptance in healthcare: a mixed-method study on trialability and doctor-patient relationships. J Posthumanism 2025;5(1):89-118. [doi: 10.63332/joph.v5i1.486]
38. Gefen D, Karahanna E, Straub DW. Trust and TAM in online shopping: an integrated model. MIS Q 2003;27(1):51-90. [doi: 10.2307/30036519]
39. Sobaih AEE, Elshaer IA, Hasanein AM. Examining students' acceptance and use of ChatGPT in Saudi Arabian higher education. Eur J Investig Health Psychol Educ 2024 Mar 17;14(3):709-721 [FREE Full text] [doi: 10.3390/ejihpe14030047] [Medline: 38534908]
40. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
41. Budd J. Burnout related to electronic health record use in primary care. J Prim Care Community Health 2023;14:21501319231166921 [FREE Full text] [doi: 10.1177/21501319231166921] [Medline: 37073905]
42. Albrecht M, Shanks D, Shah T, Hudson T, Thompson J, Filardi T, et al. Enhancing clinical documentation with ambient artificial intelligence: a quality improvement survey assessing clinician perspectives on work burden, burnout, and job satisfaction. JAMIA Open 2025 Feb;8(1):ooaf013 [FREE Full text] [doi: 10.1093/jamiaopen/ooaf013] [Medline: 39991073]
43. Olson KD, Meeker D, Troup M, Barker TD, Nguyen VH, Manders JB, et al. Use of ambient AI scribes to reduce administrative burden and professional burnout. JAMA Netw Open 2025 Oct 01;8(10):e2534976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2025.34976] [Medline: 41037268]

XSL•FO
RenderX

44. Haberle T, Cleveland C, Snow GL, Barber C, Stookey N, Thornock C, et al. The impact of nuance DAX ambient listening AI documentation: a cohort study. J Am Med Inform Assoc 2024 Apr 03;31(4):975-979. [doi: 10.1093/jamia/ocae022] [Medline: 38345343]
45. Alpert JM, Saper R, Boose E, Ruff J, Hopkins K, Gaskins D, et al. Evaluating an artificial intelligence scribe for clinical documentation. Digit Health 2025;11:20552076251395588 [FREE Full text] [doi: 10.1177/20552076251395588] [Medline: 41278378]
46. Sasseville M, Yousefi F, Ouellet S, Naye F, Stefan T, Carnovale V, et al. The impact of AI scribes on streamlining clinical documentation: a systematic review. Healthcare (Basel) 2025 Jun 16;13(12):20552076251395588 [FREE Full text] [doi: 10.3390/healthcare13121447] [Medline: 40565474]
47. Husa RA, Haggerty J, Nute AW, Levine J, Love K, Martinez X, et al. Ambient artificial intelligence use and clinician documentation burden, productivity, and efficiency. JAMA Netw Open 2026 May 01;9(5):e2615762 [FREE Full text] [doi: 10.1001/jamanetworkopen.2026.15762] [Medline: 42213439]
48. Pinto RZ, Ferreira ML, Oliveira VC, Franco MR, Adams R, Maher CG, et al. Patient-centred communication is associated with positive therapeutic alliance: a systematic review. J Physiother 2012 Jun;58(2):77-87 [FREE Full text] [doi: 10.1016/s1836-9553(12)70087-5]
49. Brondfield S, Lee K, O'Sullivan PS. The cognitive load of inpatient consults: a convergent parallel mixed methods study using the consult cognitive load instrument. Acad Med 2021 Nov 01;96(11S):S119-S125 [FREE Full text] [doi: 10.1097/ACM.0000000000004265] [Medline: 34348389]
50. Evans K, Papinniemi A, Ploderer B, Nicholson V, Hindhaugh T, Vuvan V, et al. Impact of using an AI scribe on clinical documentation and clinician-patient interactions in allied health private practice: perspectives of clinicians and patients. Musculoskelet Sci Pract 2025 Aug;78:103333 [FREE Full text] [doi: 10.1016/j.msksp.2025.103333] [Medline: 40328004]
51. Palm E, Manikantan A, Mahal H, Belwadi SS, Pepin ME. Assessing the quality of AI-generated clinical notes: validated evaluation of a large language model ambient scribe. Front Artif Intell 2025;8:1691499 [FREE Full text] [doi: 10.3389/frai.2025.1691499] [Medline: 41199808]
52. Ha E, Choon-Kon-Yune I, Murray L, Luan S, Montague E, Bhattacharyya O, et al. Evaluating the usability, technical performance, and accuracy of artificial intelligence scribes for primary care: competitive analysis. JMIR Hum Factors 2025 Jul 23;12:e71434 [FREE Full text] [doi: 10.2196/71434] [Medline: 40700466]
53. Lah U, Lewis JR, Šumak B. Perceived usability and the modified technology acceptance model. Int J Hum Comput Interact 2020 Feb 17;36(13):1216-1230. [doi: 10.1080/10447318.2020.1727262]

Abbreviations

AIAS: AI Attitude Scale
BI: behavioral intention
CGS: Curved Grading Scale
EFA: exploratory factor analysis
EHR: electronic health record
GDPR: General Data Protection Regulation
IU: intention to use
JR: job relevance
KMO: Kaiser-Meyer-Olkin
mPDQI-8: modified 8-item Physician Documentation Quality Instrument
NLP: natural language processing
PDQI: Physician Documentation Quality Instrument
PDQI-8: 8-item Physician Documentation Quality Instrument
PDQI-9: 9-item Physician Documentation Quality Instrument
PEOU: perceived ease of use
PIS: perceived interaction support
PU: perceived usefulness
QO: quality output
SUS: System Usability Scale
TAM: technology acceptance model
UMUX: usability metric for user experience

XSL•FO
RenderX

Edited by S Munce; submitted 28.Dec.2025; peer-reviewed by D Wright, S Das; comments to author 15.Aug.2026; revised version received 29.Sep.2026; accepted 05.Oct.2026; published 06.Nov.2026.

Please cite as:
Engan HK, Landstad I, Sand K, Evenshaug S, Henjum Halsnes E
AI Scribe in Specialized Rehabilitation: Usability Study of Perceived Usability and Usefulness Among Physicians and Psychologists
JMIR Rehabil Assist Technol 2026;13:e90442
URL: https://rehab.jmir.org/2026/1/e90442
doi:10.2196/90442
PMID:

©Harald Kåre Engan, Ingeborg Landstad, Kari Sand, Ståle Evenshaug, Einar Henjum Halsnes. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 06.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Robotic Rollator Assistance for Sit-to-Stand Transfers in Older Adults with Mobility Challenges: Within-Participant Feasibility Study

Anas Mahdi[1], MASc; Theresa Buchner[2], MASc; Christian Werner[2], PHD; Yue Hu[1], PHD; Katja Mombaur[3,4], PHD

[1]University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada

[2]Geriatric Center, Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany

[3]Karlsruhe Institute of Technology, Karlsruhe, Germany

[4]Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Anas Mahdi, MASc
University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada

Abstract

Background: Robotic mobility aids with sit-to-stand (STS) support may help older adults reduce physical effort during a demanding, fall-prone daily task.

Objective: This paper reports a feasibility evaluation of the "SkyWalker" robotic walker, which provides powered STS assistance via actuated lower handles. The 2 trajectories (linear and S-shape) were selected to examine how assistance profile shape influences feasibility, biomechanics, and user experience.

Methods: Seven older adult patients with mobility challenges were enrolled from a rehabilitation hospital. One withdrew during the robot familiarization (unable to stand up) without attempting any study trial, and 6 completers (4 women, 2 men; aged 70 - 84 years) form the analyzed sample. Participants completed 3 within-subject conditions: SkyWalker with the linear trajectory, SkyWalker with the S-shape trajectory, and a conventional rollator. We quantified STS task performance, perceived exertion Borg Category-Ratio 10 scale (Borg CR10), kinematics (center-of-gravity [CoG] height), kinetics (vertical ground-reaction force [Fz], normalized to body weight), and technology acceptance using an adapted Tele-healthcare Satisfaction Questionnaire–Wearable Technology (TSQ-WT). Participant-level data are the primary presentation. Condition effects were tested with exact nonparametric tests.

Results: Under linear assistance, 27 of 30 repetitions (90%) were successful (mean 4.50, SD 0.84; 4 of 6 participants at 5/5) and perceived exertion was the lowest of the 3 conditions (Borg CR10 mean 1.50, SD 1.38; lower than the rollator in 3 of 6 participants, unchanged in 2, higher in 1). Under S-shape assistance, 21 of 30 repetitions (70%) were successful (mean 3.50, SD 1.97; 3 of 6 at 5/5), with exertion comparable to the conventional rollator (mean 2.50, SD 2.43 vs mean 2.33, SD 1.63). All 6 completers achieved 5 of 5 repetitions with the conventional rollator, so robotic assistance did not increase the number of successful transfers in this cohort. Exact tests showed no significant condition effect on exertion (Friedman χ^2_2=2.1; exact *P*=.40; Kendall W=0.18) or success (Friedman χ^2_2=5.0, exact *P*=.11; Kendall W=0.42). Group CoG curves showed earlier, more monotonic elevation for linear, and per-participant Fz showed greater between-participant variability in load-transfer timing under S-shape. Technology acceptance was heterogeneous (TSQ-WT totals 22 - 77 of 80; 4 of 6 participants at 64 or higher).

Conclusions: Powered STS assistance was feasible for most assisted transfers: 27 of 30 repetitions (90%) succeeded under the linear trajectory, which also produced the lowest mean perceived effort of the 3 conditions. The S-shape trajectory was more demanding, producing more partial completions and no effort reduction relative to the conventional rollator. In this first evaluation of SkyWalker with its intended users, the assistance profile, not powered assistance itself, is the variable to prioritize in further testing, and the results, which motivate trajectory personalization and brief familiarization, are informing the design of the next SkyWalker generation.

(JMIR Rehabil Assist Technol 2026;13:e101303) doi:10.2196/101303

KEYWORDS

sit-to-stand; assistive robotics; smart walker; older adults; biomechanics; user experience

XSL•FO
RenderX

Introduction

Motivation: Sit-to-Stand as a High-Value Target for Mobility Support

Population aging is accompanied by a growing prevalence of mobility limitations that restrict independence and increase care needs. Mobility limitations are strongly associated with adverse health outcomes (eg, higher fall risk, health care usage, and reduced quality of life [QoL]), and they frequently co-occur with conditions that reduce strength, balance, and confidence in everyday movement [1]. Within this broader landscape, the sit-to-stand (STS) transition, rising from a chair to upright stance, is a particularly consequential activity of daily living. STS is performed repeatedly throughout the day (eg, from chairs, toilets, and during bedside transfers), and difficulty completing STS can become an early and persistent bottleneck for independent living: when standing up becomes unreliable, walking, toileting, and basic household participation are immediately threatened.

From a biomechanical perspective, STS is not simply "standing up"; it is a coordinated whole-body maneuver that requires (1) generating forward and upward momentum, (2) transferring support from a large base (chair + feet) to a smaller base (feet only), and (3) stabilizing the body in upright stance [2,3]. Classic analyses of chair-rise biomechanics describe characteristic phases (eg, flexion momentum, momentum transfer, extension, and stabilization), with the momentum-transfer portion often being the most balance-critical due to rapid changes in support and center-of-mass (CoM) motion [2]. Age and impairment-related reductions in lower-limb strength and postural control can therefore force compensatory strategies (greater trunk flexion, reliance on upper-limb push, slower or segmented transitions) that may maintain task completion yet increase fatigue, instability, and fear of falling [4,5].

Clinically, STS performance is widely treated as an indicator of lower-extremity function and balance capacity. Biomechanics and gerontology studies have shown that chair-rise performance differs systematically with age and functional status [4], and that functionally impaired older adults often adopt distinct strategies that reflect limited reserves and altered coordination [5]. These observations motivate STS not only as a measurement target but also as a design target: if an assistive device can make STS more feasible and less effortful while preserving stable and repeatable weight transfer, it could deliver immediate functional benefit (more successful transfers and less exertion) and potentially support safer daily routines. Accordingly, this paper focuses on walker-based assistance during STS, with the rationale for this choice discussed next.

Rationale for Rollator-Based Active STS Assistance

Walking aids, especially rollators, are among the most common mobility supports for older adults because they can provide balance assistance, allow weight-bearing through the upper limbs, and fit naturally into daily environments. In parallel, recent "smart" or instrumented rollators have expanded the role of walking aids from purely mechanical supports to systems that can sense user state, infer intent, and offer context-dependent assistance (eg, navigation support, fall-risk monitoring, task-specific stabilization) [6]. For STS specifically, many older adults already use a walker to pull or push themselves upward, even though conventional rollator use during STS can be mechanically suboptimal (eg, relying on arm strength and handle forces that may destabilize the device if brakes are not properly engaged). This makes STS a prime candidate for task-adaptive assistance within a walker form factor: the device is already present in the environment and physically coupled to the user at the moment STS assistance is needed.

Active (powered) STS assistance embedded in a walker can, in principle, address two intertwined demands of STS: (1) raising the CoM (reducing required lower-limb extensor effort) and (2) managing balance during momentum transfer (shaping the trajectory of support and encouraging a stable forward progression). However, achieving these benefits depends critically on how assistance is delivered. Assistance that is too abrupt, poorly timed, or misaligned with the user's intended strategy can increase coordination demands and reduce perceived safety, particularly for users with limited strength reserves or heightened concern about falling. This is where trajectory design and control policy matter: even with the same hardware, different assistance profiles may produce meaningfully different biomechanical and experiential outcomes.

Related Work: STS-Assist Rollators, Mechanisms, and Control Approaches

Research on STS supporting rollators spans a range of mechanical designs and control philosophies. Early work demonstrated feasibility of embedding a standing-assistance mechanism into a walker frame and coordinating force/position control to support a natural rise. For example, Chugo et al [7] proposed a rehabilitation walker with a compact parallel-link mechanism and a controller that switches assistance modes based on the user's posture, aiming to provide "natural assistance" informed by nursing practice. Related follow-up work expanded the system-level concept toward standing and walking assistance with a single device [8]. Building on these ideas, Chugo et al [9] also proposed motion control for continuous assistance across standing, walking, and seating operations, emphasizing mode switching and practical usability in daily-life sequences.

Beyond posture-based mode switching, other control approaches explicitly estimate user state/load to adapt assistance online. Morita et al proposed standing motion assistance on a robotic walker based on real-time estimation of the patient's load, and used that estimate to select force/position control modes intended to preserve the user's remaining physical capability [10]. More recent design efforts have also targeted domestic practicality and cost, for example, by combining minimal actuation (eg, gas-spring assisted lifting) with stabilization strategies and validating the concept in users with low levels of care [11]. Recent reviews and systems work continue to expand this design space, including intelligent walking-support robots that explicitly target the STS transition [12], proxy-based control of an assistive walker for intentional STS transfer [13],

XSL•FO
RenderX

lightweight assistive rollators evaluated with older users [14], and dedicated STS-assist devices for older adults [15].

In parallel, several studies have explored alternative assistive kinematics and walking-aid form factors for STS. Bulea and Triolo [16] presented a vertical lift walker designed to reduce upper- and lower-extremity force demands during sit-to-stand while remaining usable as a mobility aid, demonstrated in a prototype evaluation with an individual with spinal cord injury. Earlier walking-support concepts also addressed STS transfer assistance by guiding the motion through the walking-aid interface (eg, robotic walking support systems for sit-to-stand assistance) [17].

Alongside mechanism and control innovation, several studies explicitly ground assistance in human movement principles. Geravand et al [18] modeled the STS transfer to enable biologically inspired robot assistance, with the goal of coordinating robot behavior with human motion in an intuitive and safe way. Complementary biomechanical evidence highlights the balance demands around seat-off (momentum transfer), reporting similar seat-off balance measures in older versus younger adults despite slower execution in the older group [19]. Optimization-based work has further examined how external support forces should be applied to best assist geriatric STS transfers [20], and has been extended to optimize the physical design of assistive devices for STS [21].

Despite this progress, the STS-assist walker literature remains comparatively small relative to other assistive domains (eg, gait exoskeletons) and is methodologically heterogeneous. Importantly, evaluation in the intended population has begun to emerge: Werner et al [22] evaluated an STS assistance system integrated into a smart walker in older adults with motor impairments, reporting improved STS success rates and high user satisfaction. More broadly, systematic reviews of robotic rollators emphasize that evaluations vary widely in target users, outcome measures, and study settings, which makes it difficult to generalize device benefits or identify which design features drive user-relevant improvements [23,24].

Remaining Gaps Toward Real-World Adoption of STS-Assist Rollators

For STS-assist rollators aimed at older adults with mobility challenges, 3 gaps repeatedly limit translation from prototype promise to practical adoption.

Gap 1: Evaluation with the Intended Users in Realistic Contexts

Many assistive robotics studies demonstrate proof-of-concept performance in controlled settings, often using convenience samples. However, older adults with mobility challenges are a heterogeneous population with varied strength levels, balance strategies, and prior experience using mobility aids. Evidence gathered with healthy participants may therefore miss coordination difficulties, safety concerns, or training needs that emerge in clinical users [25,26]. Real-world deployment contexts such as geriatric hospitals, where users may already be fatigued, cautious, or recently injured, shape feasibility in ways that cannot be inferred from laboratory-only evaluations.

Gap 2: Linking Biomechanics to User Experience

Even when studies report performance improvements, they often do not explain *how* assistance reshapes movement strategy and loading, or how those biomechanical changes relate to perceived exertion and acceptance. Yet STS is biomechanically rich: assistance can alter CoM timing, foot loading progression, and the repeatability of weight transfer [2,3]. These characteristics are likely to influence perceived stability and effort, especially during momentum transfer. Without concurrent biomechanical and subjective data, it is difficult to determine whether a device reduces physical demand at the cost of increased coordination demand, or whether certain assistance profiles better support both mechanics and experience.

Gap 3: Acceptance Needs Explicit Measurement

Adoption of assistive technology depends on perceived usefulness, ease of use, trust/safety, and identity-related concerns (eg, stigma). Technology acceptance models emphasize that perceived usefulness and perceived ease of use drive intention to use, and that these determinants can be moderated by experience and context [27]. In mobility-impaired older adults, acceptance is additionally shaped by perceived safety, confidence, and fit with daily routines [28]. In social robotics and assistive robotics more broadly, scoping reviews similarly highlight that acceptability is multi-factorial and cannot be inferred from performance alone [29,30]. For STS-assist rollators, acceptance therefore needs to be measured explicitly and interpreted alongside task feasibility and mechanical effects.

Aim and Research Questions

The aim of this study was to evaluate the feasibility of powered STS assistance delivered by the SkyWalker robotic rollator in older adult patients with mobility challenges in a geriatric-hospital setting, comparing 2 assistance trajectories (linear and S-shape) against a conventional rollator. This is the first evaluation of the SkyWalker with its intended end users. Our earlier work on the platform addressed device development and control rather than user outcomes [31]. We therefore combine in a single protocol the outcomes that govern adoption: task performance, perceived exertion, instrumented kinematics and kinetics of the transfer, and technology acceptance, each read alongside the participant's baseline physical function and fall-related concern. We formulate the evaluation around 4 research questions (RQs) that are directly tied to measurable outcomes in our protocol:

1. RQ1 (Feasibility): Is active STS assistance with SkyWalker feasible for older adults with mobility challenges, as reflected by the ability to complete repeated STS trials safely?
2. RQ2 (Effort): Does SkyWalker STS assistance reduce perceived exertion during repeated STS compared with a conventional rollator condition?
3. RQ3 (Biomechanics): How do different assist trajectories (linear vs S-shape) influence STS biomechanics, specifically CoM elevation timing and vertical load transfer patterns, and do these profiles suggest differences in movement repeatability or coordination demand?

XSL•FO
RenderX

4. RQ4 (User experience): How do older adult users perceive the STS assistance in terms of benefit, usability, stigma/self-concept, and quality-of-life relevance, and how do these perceptions relate qualitatively to observed feasibility and biomechanics?

This study is a single-site within-participant feasibility study. Its contributions are (1) an evaluation with geriatric rehabilitation patients in a real clinical setting, (2) a within-subject comparison of assistance profiles (linear vs S-shape) against a conventional rollator, and (3) an integrated outcome set combining feasibility, perceived exertion, instrumented biomechanics, and user acceptance, addressing common evaluation fragmentation in robotic rollator studies [23].

Methods

SkyWalker Robot Platform

SkyWalker is a bilevel-handle powered rollator designed to support walking and to provide powered assistance during the STS transition. In this paper, we focus on the STS assistance function; walking support is not evaluated as an outcome. The device provides 2 sets of handles (Figure 1): fixed upper handles for post-STS standing/ambulation, and actuated lower handles that deliver the assist during STS. The lower handles execute a coordinated vertical lift with a gentle forward translation to support forward lean and safe CoG progression. For this study, 2 parameterized assist trajectories were used: a linear profile and an S-shaped profile.

Figure 1. Experimental setup and assistive conditions during sit-to-stand trials. Left: participant performing a sit-to-stand trial using the SkyWalker assistive device under the supervised study protocol. Right: participant performing the corresponding trial using a conventional rollator. In the SkyWalker conditions, 2 force plates recorded ground-reaction forces (1 under the chair and 1 under the participant's feet). The conventional-rollator condition was not instrumented with force plates. A Theia3D markerless motion-capture system recorded full-body kinematics in all conditions. Both panels illustrate the overall layout of the devices.



The actuated lower handles are position-controlled: each axis tracks a motion-profiled reference with continuous S-curve velocity shaping (bounded velocity, acceleration, and jerk) executed at the motor-controller level, and the coordinated 2-axis handle path during STS is generated by a kinematic trajectory generator. Assistance is therefore trajectory-based rather than force- or impedance-based. Handle start and end positions were set for each participant based on measured body height, while the trajectory execution time was fixed at 3 seconds for all participants and was identical for the linear and S-shape profiles, which differ in path shape rather than duration. Safety provisions during clinical testing included mechanical end stops on the actuated axes, software motion limits, overcurrent fuses that cut motor power on excessive current draw, and an operator emergency-stop button. A member of the study team (TB) supervised every repetition.

Participants

Seven patients (age 70 - 84 years; 5 women, 2 men) with mobility challenges were enrolled. One participant (woman, age 70; Clinical Frailty Scale [CFS] score 5, Short Physical Performance Battery [SPPB] score 4/12, Short Falls Efficacy Scale–International [Short FES-I] score 11/28, Mini-Mental State Examination [MMSE] score 28/30) withdrew during the initial "get to know the robot" familiarization (the study record notes that the patient was not able to stand up) and attempted no study trials, contributing no condition-level outcome data; thus, the analyzed sample comprised 6 participants (4 women and 2 men). The study took place at the rehabilitation wards of a geriatric hospital in Germany (Agaplesion Bethanien Heidelberg).

Inclusion criteria were age 65 years or older, ability to walk 4 meters with or without an assistive device, sufficient cognitive ability to follow study instructions, and self-reported difficulty performing the STS transition without external assistance. Examples of difficulty included requiring the use of the chair armrests or another support surface, needing physical assistance or close supervision, repeated attempts or prolonged time to stand, or feeling unstable during the rising phase. Exclusion criteria included cognitive impairment (MMSE score below 24 of 30) [32], as sufficient cognitive function was required to understand and follow the study instructions, interact safely with the device, and provide reliable self-reports on the questionnaires. All participants regularly used a walker or rollator for mobility. Notably, 3 of the 6 completers (and the

XSL•FO
RenderX

participant who did not complete testing) could not perform a single chair rise without arm use in the SPPB chair-stand test, confirming that the sample represents the intended user group with genuine STS difficulty.

Cognitive status was screened using the MMSE (0 - 30; higher =better cognition). Physical function was evaluated using the SPPB (0 - 12; higher =better function) under standardized instructions, with established interpretive bands in older adults: approximately 10 - 12 (near-normal), 6 - 9 (moderate limitation), and less than 6 (marked limitation) [33]. Concern about falling was assessed using the 7-item Short FES-I (range 7 - 28; higher =greater concern about falling) [34]. Notably, no participant reported being "very concerned" about falling in all situations; even those with higher FES-I scores identified only specific activities such as dressing or reaching that provoked concern rather than a generalized fear of movement. Fall history (number of falls in the previous year) was captured by patient self-report.

Study Design and Procedure

Participants completed three within-subject conditions in randomized order (a script generated a random condition sequence for each participant): conventional rollator (no powered assistance), SkyWalker linear trajectory, and SkyWalker S-shape trajectory. For each condition, participants attempted up to 5 STS repetitions, with rest as needed between conditions; if a participant could not complete an STS attempt, the trial was stopped and counted as a failure for that repetition. A repetition was classified as a user-initiated abort when the participant released the handles, verbally requested to stop, or sat back down before reaching upright stance. All STS trials were performed from the same height-adjustable chair, with seat height adjusted to each participant's body height. During the SkyWalker conditions, one force plate was positioned under the chair and one under the participant's feet, whereas the conventional-rollator condition was performed without force plates. Rollator handle height was likewise adjusted per participant (86 - 97 cm across the sample). Before recorded trials, the study team demonstrated hand placement and the handle motion. Participants then practiced assisted repetitions (approximately 3 - 5) until they reported feeling comfortable and ready to begin the experiment.

1. No powered assistance: Participants used a standard wheeled walker (conventional rollator) for balance/support. The participant was seated in a stable chair and instructed to stand using the rollator as needed, without any powered assistance. This condition reflects a typical mobility aid used in daily life.
2. SkyWalker linear trajectory: Participants used the SkyWalker with STS assistance activated. During the rising phase, the SkyWalker's lower handles moved in the horizontal and vertical directions simultaneously to support the lift, emphasizing a direct, coordinated upward assistance profile.
3. SkyWalker S-shape trajectory: Participants used the SkyWalker with STS assistance activated, but with an S-shape lower-handle trajectory. In this profile, the handles moved upward with a forward/curved component, resulting in a more complex handle motion pattern.

Outcome Measures and Instruments

Task Performance

During each condition, we recorded the number of successful STS repetitions out of five. A repetition was counted as successful if the participant managed to stand up fully (upright posture) without physical assistance from another person and, in the case of the SkyWalker conditions, without aborting the device mid-lift. Time to complete 5 repetitions was recorded where possible but is not analyzed here because several participants could not complete all 5 in some conditions; we therefore focus on the count of successful stands (success rate) as a more universally applicable metric.

Perceived Exertion

After each condition, participants immediately rated their perceived exertion using the Borg Category-Ratio 10 scale (Borg CR10 scale; 0 - 10 rating, where 0="No effort at all," 5="Hard," and 10="Maximal effort"). This provided a subjective measure of how physically demanding the STS task felt under each condition.

Technology Acceptance

After completing the three conditions, participants completed a German-language adaptation of the Tele-healthcare Satisfaction Questionnaire–Wearable Technology (TSQ-WT) [35], adapted for the SkyWalker STS assistance system. The original TSQ-WT comprises 6 domains (benefit, usability, self-concept, privacy and loss of control, QoL, and wearing comfort) scored on a 5-point Likert scale from 0 ("not at all / completely disagree") to 4 ("fully agree"), with higher values reflecting more positive evaluations. The instrument is designed to be adaptable to different systems; for this study, it was tailored to the STS assistance context by removing the domains that were not appropriate: "wearing comfort," which targets body-worn devices, and "privacy and loss of control," which focuses on long-term use. The resulting 20-item version evaluates user satisfaction and acceptance across 4 domains: benefit, usability, self-concept/stigma, and QoL. The same German-language adaptation, removing these 2 domains, has been used previously by our group to evaluate the STS assistance of a smart walker in older adults with motor impairments [22]. Negatively phrased items (items 2, 7, 8, 11, and 15) were reverse-coded so that, for all items, higher scores indicate a more positive attitude. Items are distributed unevenly across domains in the adapted instrument: benefit (7 items, maximum 28), usability (6 items, maximum 24), self-concept/stigma (4 items, maximum 16; higher =less stigma), and QoL (3 items, maximum 12). Domain scores were obtained by summing the items within each domain, and the total TSQ-WT score (range 0 - 80) was obtained by summing all 20 items. The questionnaire was administered once, after exposure to both robotic trajectories. It therefore reflects acceptance of the SkyWalker system as a whole rather than of an individual trajectory, and it was not administered for the conventional rollator. The questionnaire was administered in German, using the German-language version of the TSQ-WT adopted from the

XSL•FO
RenderX

HeiAge research project. No translation was performed within this study, and the English item wordings reported in Multimedia Appendix 1 are translations of the administered German items. To our knowledge, a formal comprehensive psychometric validation of the TSQ-WT, including validation of the German translation, has not been published. Its scores are therefore reported as descriptive indicators of acceptance rather than as validated acceptance measurements. The full list of items and their domain assignments is provided in Multimedia Appendix 1. The TSQ-WT has previously been used to assess wearable technology and robotic devices [22,36,37].

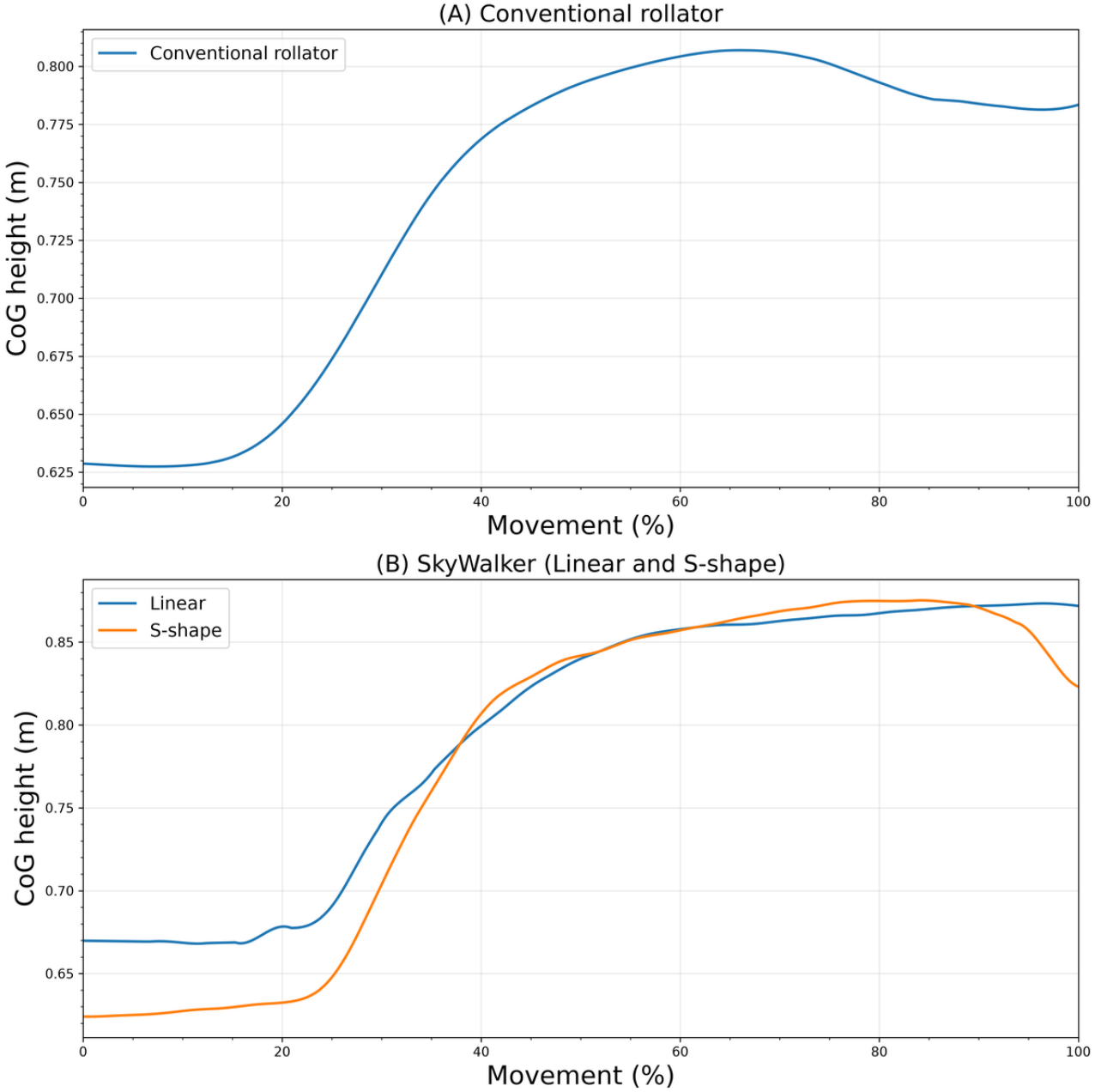
Biomechanical Data Acquisition and Processing

Kinematics

Full-body kinematics were recorded with a markerless multicamera motion-capture system and processed with the Theia3D full-body model. The whole-body center of gravity (CoG) was taken from the model output (100 Hz), interpolated, and smoothed (Savitzky–Golay filter). Movement onset was detected with a sliding-window change-point rule on CoG height (rise of at least 2.5 cm between adjacent 0.5-s windows with confirming vertical velocity). For each participant and condition, repetitions were aligned at onset and averaged. Group curves are means of the per-participant mean curves resampled to a fixed per-condition analysis window (SkyWalker conditions: 1 s before onset to 4.5 s after onset; conventional rollator: onset to 2.0 s after onset, reflecting the shorter unassisted transfer). Each curve is displayed against movement percentage, defined as the percentage of the fixed analysis window rather than of each repetition's own duration (Figure 2). Group curves include successful repetitions only and only participants with at least one successful repetition in the respective condition. Accordingly, the S-shape group curve comprises 5 participants (P3 completed no successful S-shape repetition).

XSL•FO
RenderX

Figure 2. Motion-capture center of gravity height during sit-to-stand: group-mean curves (means of per-participant mean curves over successful repetitions), aligned at movement onset. (A) Conventional rollator (n=6 participants; analysis window: onset to 2 seconds). (B) SkyWalker linear (n=6) and S-shape (n=5; analysis window: 1 second before onset to 4.5 seconds). P3 is not included in the S-shape curve because the patient completed no successful S-shape repetition. The horizontal axis is the percentage of the fixed analysis window, not of each repetition's own duration. CoG: center of gravity.

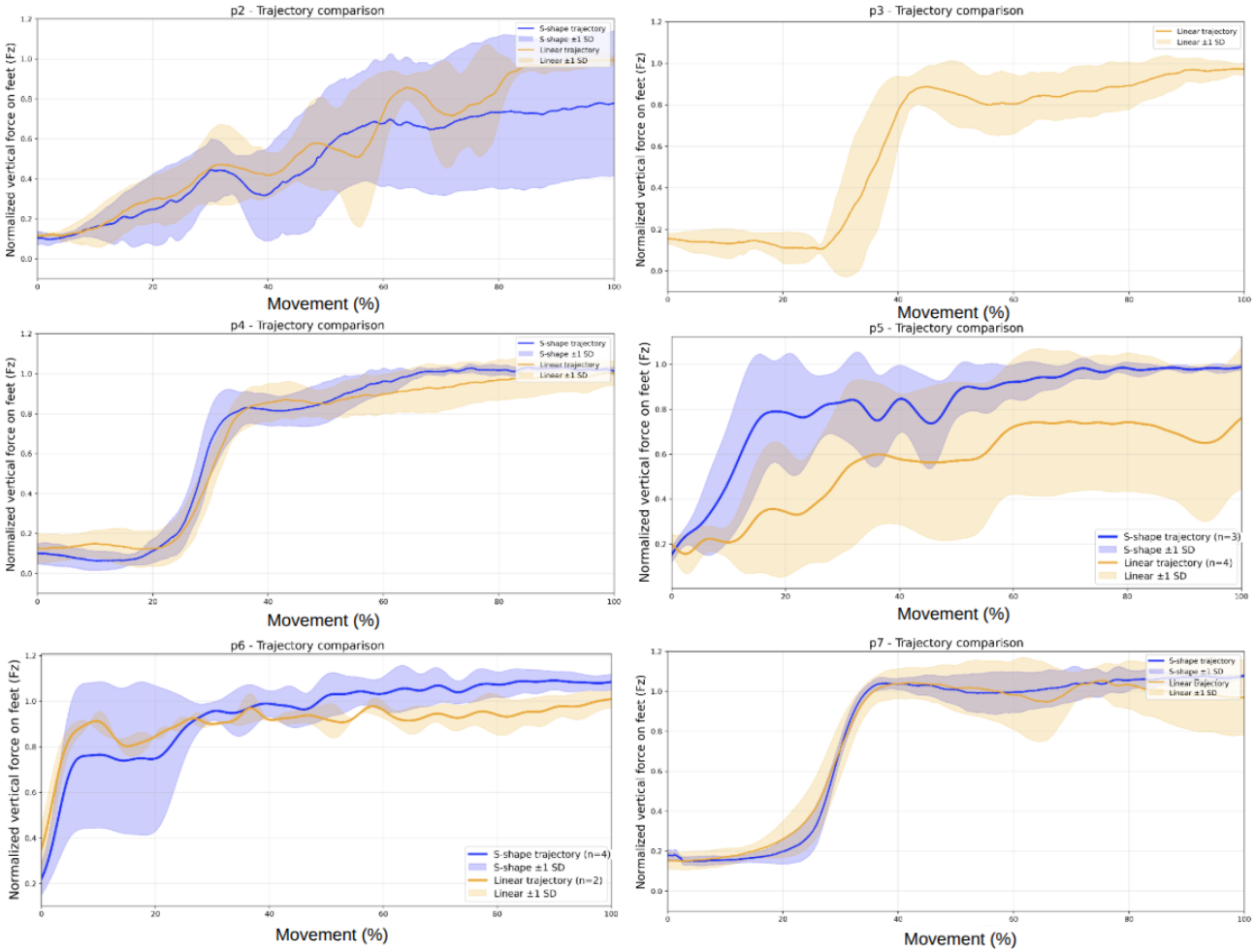


Kinetics

Ground-reaction forces were recorded during the SkyWalker conditions with two force plates, one under the chair and one under the participant's feet, at 1000 Hz. The conventional-rollator condition was not instrumented with force plates. Force signals were low-pass filtered (fifth-order Butterworth) and smoothed with a 10-sample moving average. Recordings were aligned at movement onset detected from the force traces, truncated to a common analysis window, and the vertical component measured under the feet (Fz) was normalized to body weight. Per-participant condition curves are means across the repetitions retained after excluding aborted repetitions and technically unusable recordings (exact counts in the caption of Figure 3).

XSL•FO
RenderX

Figure 3. Force-plate vertical ground-reaction force measured under the feet (Fz), normalized to body weight, across the analysis window (0% - 100%) under SkyWalker assistance. Mean curves for linear and S-shape profiles are shown with shading of 1 SD for each participant (panels: P2, P3 top row; P4, P5 middle row; P6, P7 bottom row). Curves average successful repetitions with usable recordings (Linear: 4, 5, 5, 4, 2, and 4 repetitions for P2-P7, respectively; S-shape: 4, none, 4, 3, 4, and 3). No S-shape curve is shown for P3 because the patient completed no successful S-shape repetition.



Statistical Analysis

Given the sample size (n=6) and the ordinal nature of several measures, we used exact nonparametric statistical tests, whose *P* values do not rely on large-sample approximations, and we present participant-level data and descriptive statistics as the primary results. For outcomes that involved repeated measures across the three conditions, we used the Friedman test with tie correction, with exact *P* values obtained by exhaustively permuting condition labels within participants, and tie-adjusted Kendall W reported as the effect size. Because zero (tied) paired differences are frequent at ceiling in these data and carry no directional information, pairwise condition contrasts are summarized by direction counts and analyzed with exact 2-sided sign tests (zeros reported separately), with Bonferroni correction across the 3 pairwise contrasts per outcome. Because the omnibus tests were not significant, all pairwise results are post hoc analyses and are reported as such. Paired effect estimates are given as Hodges–Lehmann pseudomedians with conservative 96.9% CIs under the symmetric-difference model, and matched rank-biserial correlations (computed after excluding zero differences) are reported descriptively. Descriptive statistics (mean, SD, median, and range, as appropriate) were calculated for baseline measures (SPPB, FES-I) and for outcome measures under each condition. Test selection was based a priori on the measurement scales and the paired design (ordinal exertion ratings; bounded success counts with substantial ceiling ties), not on pretesting for normality or homogeneity, which is uninformative at this sample size. Analyses were performed with custom enumeration scripts in Python 3.12 (NumPy/SciPy), with exact null distributions computed by exhaustive enumeration.

Ethical Considerations

The study was approved by the Ethics Committee of the Medical Faculty Heidelberg (approval no. S-654/2019) and conducted in accordance with the Declaration of Helsinki and its later amendments. All participants provided informed consent prior to participation; participation was voluntary, and participants could withdraw at any time. Personal identifiers in figures depicting participants have been anonymized, and participants consented to the use of study photographs for publication.

Results

Overview

Results are reported for the 6 participants who completed the protocol (P2–P7; Table 1). Each participant attempted up to 5 STS repetitions in each within-subject condition (SkyWalker

XSL•FO
RenderX

linear, SkyWalker S-shape, and conventional rollator). Because responses differed between participants, we present both participant-level outcomes (Table 2, Figures 3 and 4) and group-level summaries (Table 3).

Table . Participant demographics and baseline measures (N=6).

Participant	Age (years)	Sex	Frailty status[a]	Falls[b]	SPPB[c]	FES-I[d]	MMSE[e]
P2	70	Female	5 (Mild)	2	4	15	30
P3	84	Female	5 (Mild)	2	9	14	30
P4	82	Female	4 (Vulnerable)	2	9	8	28
P5	75	Male	4 (Vulnerable)	1	5	8	29
P6	73	Male	4 (Vulnerable)	5	6	8	28
P7	75	Female	5 (Mild)	1	3	18	27

[a]Frailty status: Clinical Frailty Scale (CFS) score (5=mildly frail, 4=vulnerable).

[b]Falls count in past year.

[c]SPPB: Short Physical Performance Battery (0 - 12).

[d]Short FES-I: 7-item Short Falls Efficacy Scale–International (7-28; higher=greater concern about falling).

[e]MMSE: Mini-Mental State Examination (0 - 30).

Table . Participant-level summary of baseline function, STS[a] performance, perceived exertion, and acceptance (N=6).

Participant	SPPB[b]	FES-I[c]	Linear[d] (out of 5)	S-shape[d] (out of 5)	Rollator[d] (out of 5)	Borg[e]	TSQ-WT[f]
P2	4	15	5	5	5	3.7	64
P3	9	14	5	0	5	3.7	67
P4	9	8	5	5	5	0.7	77
P5	5	8	4	3	5	3.3	43
P6	6	8	3	3	5	0.0	22
P7	3	18	5	5	5	1.3	69

[a]STS: sit-to-stand.

[b]SPPB: Short Physical Performance Battery.

[c]FES-I: Falls Efficacy Scale–International.

[d]Successful STS repetitions out of 5 per condition.

[e]Mean Borg Category-Ratio 10 scale rating across the 3 conditions.

[f]Telehealthcare Satisfaction Questionnaire–Wearable Technology (TSQ-WT) total score (0 - 80). Domain scores are shown in Figure 4.

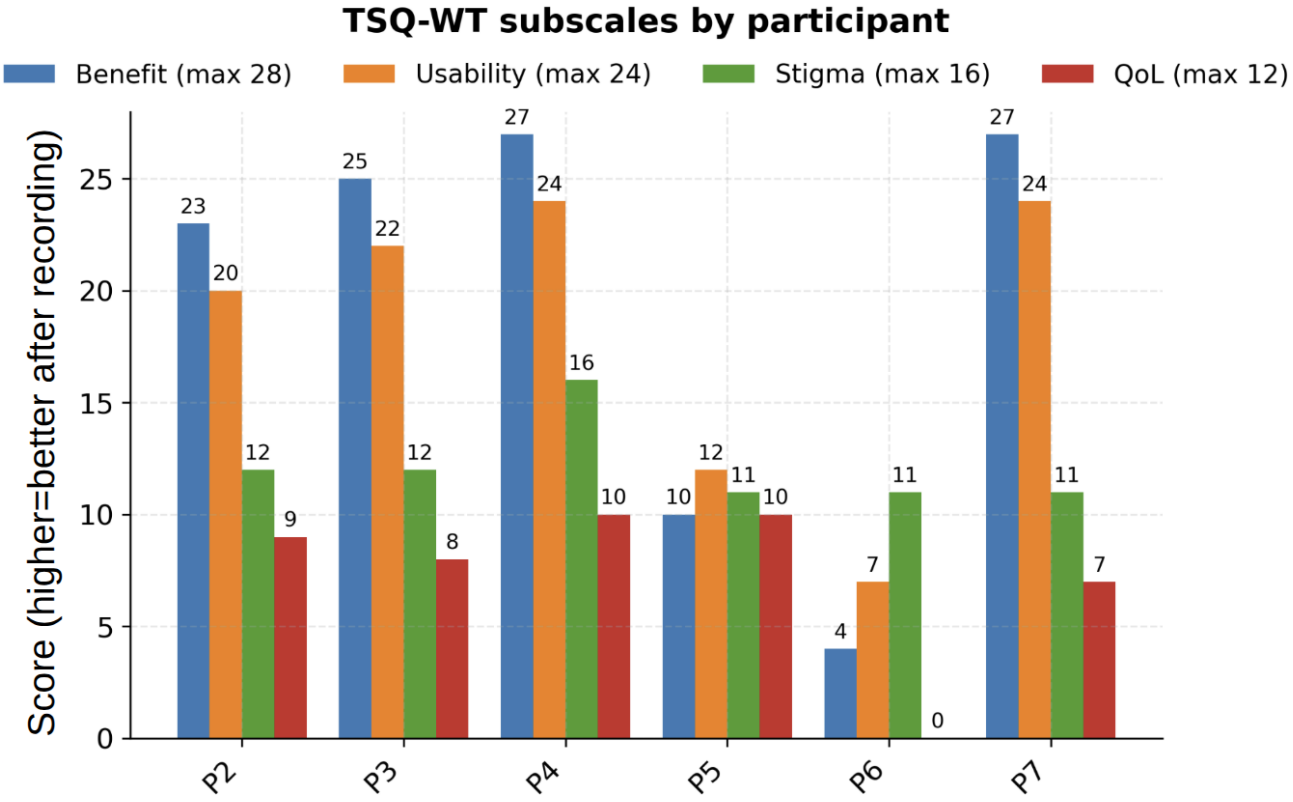
Table . Task performance and perceived exertion across conditions (N=6).

Condition	Success (out of 5), mean (SD)	Completed (all 5 repetitions), n (%)	Borg CR10[a], mean (SD)
Robot			
Linear trajectory	4.50 (0.84)	4/6 (67)	1.50 (1.38)
S-shape trajectory	3.50 (1.97)	3/6 (50)	2.50 (2.43)
Conventional rollator	5.00 (0)	6/6 (100)	2.33 (1.63)

[a]Borg CR10: Borg Category-Ratio 10 scale.

XSL•FO
RenderX

Figure 4. Technology acceptance (Tele-healthcare Satisfaction Questionnaire–Wearable Technology) subscales by participant (P2-P7). Bars show postrecoding subscale totals: benefit (maximum 28), usability (maximum 24), stigma (maximum 16; higher=less stigma), and quality of life (maximum 12). Items were rated 0 - 4 (0=completely disagree to 4=completely agree). Negatively worded items (2, 7, 8, 11, and 15) were reverse-coded (recoded as 4 minus the raw score) so that higher scores consistently indicate more favorable responses on every item and subscale. max: maximum; QoL: quality of life; TSQ-WT: Tele-healthcare Satisfaction Questionnaire–Wearable Technology.



Inferential Statistics

Neither perceived exertion (Friedman χ^2_2=2.1, exact *P*=.40; tie-adjusted Kendall W=0.18) nor STS success (Friedman χ^2_2=5.0, exact *P*=.11; W=0.42) differed significantly across the 3 conditions, and no pairwise comparison was significant after Bonferroni correction (Table 4; pairwise detail in Multimedia Appendix 2). Of note, with the observed tie structure the success statistic of 5.0 is the maximum attainable value, so its exact *P*n value cannot fall below .11 in this design regardless of effect size. The participant-level data below are the primary findings.

Table . Exact nonparametric analyses of the condition comparisons (N=6)[a].

Outcome	Friedman chi-square (*df*)	Exact *P* value	Kendall W[b]
Borg CR10[c]	2.1 (2)	.40	0.18
Success (out of 5)	5 (2)	.11	0.42

[a]Tie-corrected Friedman statistic; the exact *P* value was obtained by exhaustive within-participant permutation of condition labels (*df* shown conventionally). With the observed tie structure, the success statistic of 5 is the maximum attainable value, so its exact *P* value cannot fall below .11 in this design.

[b]Tie-adjusted Kendall W. Pairwise contrasts (exact sign tests, direction counts, Hodges-Lehmann estimates with exact confidence intervals, and rank-biserial correlations) are post hoc analyses and are reported in Multimedia Appendix 2.

[c]Borg CR10: Borg Category-Ratio 10 scale.

Feasibility and Task Performance

Completion and Safety Observations

Six of seven enrolled participants completed the full protocol (linear, S-shape, and rollator conditions). Across completers, no adverse events or device malfunctions were observed during STS attempts (as recorded by the study team). During P3's 2 unsuccessful S-shape attempts, the patient was unable to complete the rise and returned to the chair under supervision without contacting the floor or sustaining injury. These events were classified as incomplete repetitions rather than falls or adverse events. Partial STS failures occurred only in SkyWalker conditions and reflected user-initiated aborts or inability to complete the full standing transition on a given repetition.

STS Success Across Conditions

All 6 completers achieved ceiling performance with the conventional rollator (5/5 repetitions for every participant; Table 3). With SkyWalker assistance, success was near-ceiling for most participants but showed trajectory- and participant-specific reductions. At the group level, mean success was 4.50 (SD 0.84)

XSL•FO
RenderX

for linear and 3.50 (SD 1.97) for S-shape out of 5 (Table 3). At the participant level (Table 2), 3 individuals showed partial completion with SkyWalker: P3 completed no S-shape repetition (two attempts, both incomplete, after which the patient declined further S-shape attempts; 5/5 in linear), P5 completed 4/5 (linear) and 3/5 (S-shape), and P6 completed 3/5 in both robot modes. The remaining participants (P2, P4, and P7) completed 5/5 repetitions in both robot conditions. At the repetition level, 27 of 30 planned repetitions (90%) were successful under linear, 21 of 30 (70%) under S-shape, and 30 of 30 (100%) with the rollator (noting that repetitions within a participant are not independent observations). Robotic assistance therefore did not improve task success relative to the conventional rollator in this cohort. Its potential value lies in reshaping effort and movement quality rather than enabling otherwise impossible transfers.

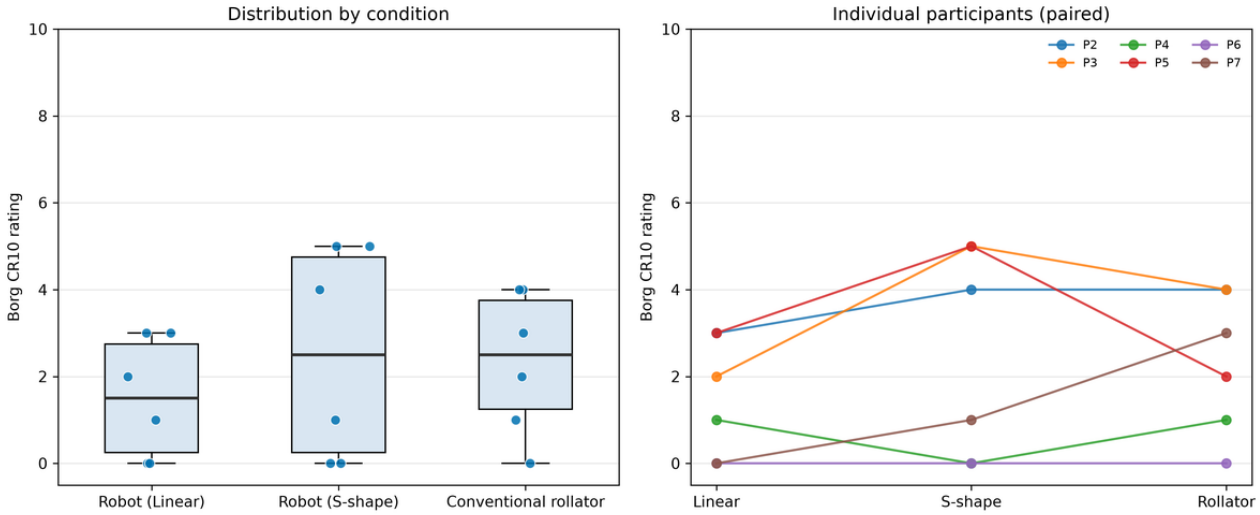
Baseline Function Versus Success

Neither baseline physical function (SPPB) nor fall concern (Short FES-I) showed a monotonic descriptive relationship with the corresponding participant-level outcomes (scatter plots in Multimedia Appendix 3). Baseline physical function did not relate strictly monotonically to mean success across conditions. For example, P3 and P4 had similarly high SPPB scores (9/12) but differed in SkyWalker S-shape feasibility (P3: 0/5; P4: 5/5), while P7 had the lowest SPPB (3/12) yet completed 5/5 repetitions in all conditions (Table 2). This pattern motivates explicit case-based reporting of feasibility rather than relying on group averages alone.

Perceived Exertion

Group-level perceived exertion differed descriptively by condition (Table 3, Figure 5). Mean Borg CR10 was lowest for SkyWalker linear (mean 1.50, SD 1.38; median 1.5, IQR 2.50), whereas SkyWalker S-shape (mean 2.50, SD 2.43; median 2.5, IQR 4.50) was comparable to the conventional rollator (mean 2.33, SD 1.63; median 2.5, IQR 2.50). At the participant level, linear was rated lower than the rollator by 3 of 6 participants, equal by 2, and higher by 1 (Figure 5, right panel). The lower mean under linear is therefore a participant-dependent pattern rather than a uniform effect, and it was specific to the linear trajectory rather than a general effect of powered support. Because exertion was rated once per condition after unequal numbers of completed repetitions, condition ratings are not perfectly workload-matched. The distribution in Figure 5 shows that low exertion values (0 - 1) occurred in all conditions, but higher exertion ratings were concentrated in the S-shape condition (two participants, P3 and P5, rated it 5, “Hard”) and rollator (values up to 4).

Figure 5. Perceived exertion (Borg CR10) by condition. Left: boxplots of postcondition exertion ratings for robot (linear), robot (S-shape), and conventional rollator, with individual participants overlaid. Right: paired participant-level ratings connected across conditions. Lower scores indicate less perceived effort.



Inter-individual variability was substantial (participant-level scatter in Multimedia Appendix 3). The relationship between fall concern (Short FES-I) and mean Borg was nonmonotonic: higher fall concern did not necessarily correspond to higher perceived exertion (eg, P7 had the highest FES-I but comparatively low mean Borg), whereas P2 and P3 reported the highest mean Borg despite only moderate-to-high FES-I scores (Table 2). The highest single-condition exertion ratings in the dataset were reported for the S-shape condition (Borg 5 by both P3, who could not complete it, and P5, who completed 3 of 5 repetitions and commented that the S-shape assistance “felt unnatural”), consistent with the larger spread seen for S-shape in Figure 5.

Biomechanics of Assistance Profiles (CoG Kinematics and Fz Kinetics)

Group CoG Kinematics

Figure 2 shows the vertical CoG group-mean curves (means of the per-participant mean curves over successful repetitions), aligned at movement onset and displayed against movement percentage of each condition’s fixed analysis window. Because the rollator window is shorter than the SkyWalker window, timing should be compared within a panel rather than across panels. With the conventional rollator (Figure 2A), the CoG rose to a peak and settled slightly toward terminal stance. With SkyWalker (Figure 2B), the linear curve rose earlier within the window and more monotonically and plateaued stably, whereas

XSL•FO
RenderX

the S-shape curve began its rise later within the window. Among participants with successful S-shape repetitions (n=5), the S-shape curve then overlapped the linear plateau through the mid-to-late window, with a small downward adjustment over the final portion of the window consistent with the added fore–aft component.

These kinematics align with behavioral outcomes (Table 3 and Figure 5): smoother, earlier CoG elevation in linear coincides with the highest STS success and the lowest exertion. In contrast, the S-shape's delayed vertical commitment and its feasibility failures (which, by construction, do not appear in the successful-repetition curves) co-occur with greater variability in success and occasionally higher perceived effort for a subset. Participants with higher SPPB and smooth S-shape CoG traces generally achieved 5/5 with low exertion; those with interrupted S-shape elevation were more likely to show partial successes and higher exertion. Thus, while S-shape can be effective, its coupled forward component appears to increase coordination/balance demands during momentum transfer, which can amplify perceived effort in users with lower reserves or higher fall concern.

Per-Participant Kinetics

Figure 3 depicts per-participant vertical ground-reaction force measured under the feet (Fz), normalized to body weight, across the analysis window for linear and S-shape assistance (mean curves with shading of 1 SD). Linear typically produces an earlier, more progressive load transfer to the feet that stabilizes near body weight as standing is completed, with narrower SD envelopes, ie, more repeatable weight-transfer strategies across repetitions. S-shape often shows a slightly delayed Fz rise and broader variability during midcycle in several participants (eg, P5 and P6), consistent with, though not direct evidence of, an additional demand of coordinating forward translation with vertical lift. Because upper-limb and handle forces were not measured, lower foot loading cannot be uniquely attributed to device support. It may also reflect altered balance strategy or incomplete weight transfer (see Limitations section).

These kinetic signatures are descriptively consistent with the subjective and task outcomes: S-shape's delayed and more variable Fz progression co-occurs with reduced or absent completion (0/5–3/5) and higher exertion ratings in some users, whereas participants whose linear and S-shape Fz curves overlap with narrow variability achieved 5/5 with low exertion. Rollator-condition ground-reaction forces were not recorded, so no claim is made about foot loading under assistance relative to the rollator. Likewise, because the SD envelopes summarize only the retained repetitions per curve (counts in the caption of Figure 3), they should be read as descriptive variability rather than a validated measure of movement stability or coordination.

TSQ-WT Responses

Figure 4 summarizes TSQ-WT domain scores by participant after reverse-coding all negatively worded items so that higher scores consistently indicate more favorable evaluations. Acceptance was high for most participants: 4 of 6 scored 64 or higher of a maximum 80 (P4 77/80; P2, P3, and P7 64–69/80), with the remaining two at 43/80 and 22/80 (Table 2).

Benefit and usability were high for most participants (eg, P4 and P7 near ceiling in usability), and QoL scores were generally positive (eg, P5 QoL high within its domain range), as shown in Figure 4. One participant (P6) reported very low benefit, usability, and QoL despite a mid-range Stigma score comparable to others. These domain-level patterns support a case-based interpretation alongside feasibility and biomechanics.

Participant-Level Patterns

Given the between-participant heterogeneity across feasibility, biomechanics, and TSQ outcomes, participant-level results are central rather than incidental. Detailed case descriptions for each participant (P2–P7), integrating baseline function, task success, exertion, Fz transfer signatures, and acceptance profiles, are provided in Multimedia Appendix 3. The recurring patterns across cases are synthesized in the Discussion. Three qualitative observations recorded by the study team merit note: P5 described the linear assistance as "unusual" and the S-shape assistance as feeling "unnatural"; P7 was observed to synchronize correctly with the S-shape assistance only from the third–fourth repetition onward, suggesting a short familiarization effect; and P3 declined to continue the S-shape condition after unsuccessful attempts.

Discussion

Principal Findings

This is the first study in which older adults with mobility challenges used the SkyWalker, rated it, and were instrumented while doing so. This single-site within-participant feasibility study evaluated SkyWalker with six older adult patients with mobility challenges, comparing two assist trajectories (linear and S-shape) against a conventional rollator. The principal findings are: (1) STS with the conventional rollator was at ceiling for all completers, so powered assistance did not improve task success in this cohort, and it introduced partial failures for three participants (particularly under S-shape); (2) perceived exertion was lowest under linear assistance, whereas S-shape exertion was comparable to the conventional rollator and included the two highest individual ratings; (3) biomechanically, linear produced earlier and more monotonic CoG elevation and generally more repeatable load transfer, while S-shape increased between-participant variability; and (4) responses on the unvalidated adapted TSQ-WT were heterogeneous (totals 22 - 77 of 80), with 2 participants reporting low benefit/usability. One of seven enrolled participants withdrew during the robot familiarization, before attempting any study trial. In this sample, linear had a lower mean Borg rating and more completed repetitions than S-shape. These observations identify the assistance profile as a variable to prioritize in further testing, not as an established determinant of user response. Below we discuss the results explicitly RQ-by-RQ and then synthesize individual-case profiles that carry direct design and clinical implications.

RQ1: Feasibility and Safety

RQ1 asks whether active STS assistance is feasible for older adults with mobility challenges. Feasibility of powered STS assistance was demonstrated for most participants: 6 of 7

XSL•FO
RenderX

enrolled participants completed the full protocol, 27 of 30 linear repetitions (90%) and 21 of 30 S-shape repetitions (70%) were successful, and no adverse events or device malfunctions were recorded by the study team. One enrollee withdrew during the robot familiarization (unable to stand up) before attempting any study trial, and one completer completed no S-shape repetition. Because no feasibility thresholds had been prespecified, these metrics are reported descriptively rather than as a pass/fail verdict. However, feasibility was not uniform across users or trajectories. Even among completers, SkyWalker success was not ceiling in all cases (Table 2), and S-shape showed the greatest vulnerability to partial completion (eg, P3: 0/5).

It is important to note that "feasibility" here is not equivalent to "superiority to rollator." The rollator condition was universally successful for completers (5/5), meaning that in this cohort the value proposition of the robot is not primarily about enabling STS that is otherwise impossible; rather, it is about reshaping the perceived and biomechanical demands of the transfer. The noncompletion during early exposure is equally important evidence for feasibility: it indicates that, in a rehabilitation hospital where fatigue and confidence vary substantially, a subset of intended users may require additional familiarization, confidence-building, or trajectory personalization before they can safely and comfortably engage with powered STS support.

RQ2: Effort Reduction and Perceived Support

RQ2 concerns perceived exertion during repeated STS. At the group level, mean Borg CR10 ratings were lowest with SkyWalker linear (lower than the rollator for 3 of 6 participants, equal for 2, higher for 1). S-shape ratings were comparable to the conventional rollator on average and included the two highest individual ratings in the dataset (Tables 3 and 4). Any effort benefit of powered assistance was therefore trajectory-specific and participant-dependent: it was observed under linear for half the participants, and not under S-shape. Because STS success with the rollator was ceiling for all completers (5/5), perceived exertion becomes a primary indicator of perceived support: the task was feasible across conditions, but it did not feel equally demanding. This is particularly relevant for adoption, as older adults may continue to complete STS with conventional aids while still experiencing high cumulative effort and fatigue across daily repetitions.

Importantly, the exertion data are not only separated by mean values but also by their dispersion. Figure 5 shows a narrower distribution for linear and the broadest spread for S-shape, whose ratings ranged from 0 to 5. This pattern supports a trajectory-specific interpretation: linear appears to provide more consistent perceived assistance across users, whereas S-shape yields more heterogeneous experiences: for some participants it remains low-effort, but for others it can feel demanding despite being "assisted." In evaluations with the intended users, this difference in variance is often as informative as the difference in central tendency because it signals robustness (or lack thereof) across users.

We also examined whether exertion aligned with baseline fall-related concern. The relationship between Short FES-I and mean Borg was weak and non-monotonic (Multimedia Appendix 3), indicating that generalized concern about falling was not sufficient to explain who experienced the task as effortful. For example, P7 exhibited the highest FES-I score yet relatively low mean Borg, whereas P2 and P3 reported high exertion with only moderate-to-high FES-I. One plausible interpretation is that the Borg scale here captures physical effort during execution more than anticipatory concern, and that exertion rises primarily when the user must invest additional muscular or coordination effort to maintain a stable transfer under a given assistance profile. Practically, this implies that screening users only by fear of falling measures would not reliably predict perceived effort with powered assistance; evaluation and personalization should consider how the user couples their motor strategy to the device trajectory.

Finally, the effort findings are mechanically interpretable in light of the kinematics and kinetics. Linear assistance produced earlier and more monotonic CoG elevation (Figure 2) and often narrower Fz variability envelopes (Figure 3), consistent with a more repeatable weight-transfer strategy across repetitions. Such repeatability plausibly reduces the need for compensatory upper-limb pushing and stabilizing corrections, which can lower perceived exertion even when task success is unchanged. In contrast, S-shape frequently exhibited delayed/variable load transfer timing and broader mid-cycle Fz dispersion in some participants (Figure 3), consistent with a higher coordination burden during momentum transfer. This provides a coherent explanation for why S-shape can remain low-effort for some users yet produce higher exertion (and occasionally aborted repetitions) for others.

Taken together, RQ2 provides participant-dependent evidence that SkyWalker with the linear trajectory can reduce perceived effort during repeated STS, with linear appearing more robust across participants and S-shape more sensitive to individual coordination demands. From a design and deployment perspective, these findings motivate (1) treating the linear profile as the candidate default for further study, (2) brief familiarization to reduce user–device synchronization demands (directly supported by the observation that P7 synchronized with the S-shape assistance only from the third–fourth repetition), and (3) evaluation of adaptive assistance strategies in the intended population. In our prior development work [31], we implemented a vision-based adaptive trajectory controller and validated it with healthy participants; a key next step is to test such adaptive STS control with older adults with mobility challenges to determine whether real-time trajectory adaptation improves consistency of weight transfer, perceived effort, and feasibility in clinically relevant users.

RQ3: Biomechanical Mechanisms and Trajectory-Dependent Coordination Demands

RQ3 asks how trajectory design influences STS biomechanics and whether the resulting profiles imply differences in repeatability or coordination demand. The kinematic summary in Figure 2 indicates that linear assistance is associated with earlier and more monotonic CoG elevation, whereas the S-shape profile delays vertical commitment. Among participants who completed S-shape repetitions successfully, the S-shape curve reached the same late-window plateau as linear, with a small

XSL•FO
RenderX

adjustment at the end of the window. In the context of STS phase structure, these differences matter because small shifts in the timing of forward/vertical commitment can change how the user manages momentum transfer and stabilization.

The participant-level Fz results in Figure 3 add specificity: several participants exhibited narrow SD envelopes and trajectories approaching body weight late in the cycle (eg, P4 and P7), while others displayed broad SD regions and/or sustained under-loading (eg, P5; and P2 in S-shape). We interpret a broader SD envelope cautiously, as descriptive repetition-to-repetition variability over the retained repetitions rather than a validated stability metric, but in a task such as STS, where stability depends on coordinating trunk motion, foot loading, and upper-limb support, it is at least suggestive that the user is not repeating a single consistent motor strategy in that condition.

A central mechanistic implication is that S-shape likely changes the coupling between forward progression and vertical lift relative to linear. When users can accommodate this coupling, Fz rises smoothly and repeatably (eg, P4 and P7), and feasibility remains high. When users cannot, we observe either (1) noncompletion (P3 in S-shape) or (2) high variability/under-loading patterns (eg, P5), which may reflect a more cautious or inconsistent strategy under that profile. Because the current dataset does not include direct measures of upper-limb forces, attention, or balance margins, we avoid stronger causal claims; in particular, lower foot loading under assistance is compatible with several explanations (greater device support, reduced engagement, altered balance strategy, or incomplete transfer) that cannot be separated without handle-force sensing. Nevertheless, the combined CoG timing differences and the participant-level Fz variability are descriptively consistent with linear being the better-tolerated trajectory in this sample, while S-shape can be effective for some users but appears more sensitive to individual capacity; coordination demand itself was not measured.

RQ4: Acceptance, Stigma, and Quality-of-Life Relevance

RQ4 concerns user experience as measured by TSQ-WT domains (benefit, usability, self-concept/stigma, and quality of lQoL). Overall acceptance was high for most participants (4 of 6 at 64 or higher of a maximum 80), with domain-level differences across the cohort (Figure 4). Several participants reported high benefit and usability with low stigma after reverse-coding (eg, P4 and P7), suggesting that powered STS assistance can be perceived as both useful and acceptable from a self-concept perspective in this cohort. Because stigma and self-concept concerns can be important barriers to adoption of assistive technologies in older adults, it is notable that stigma was not a dominant negative factor for most participants in this study.

The key nuance is dissociation across domains. P6 reported very low benefit and usability while having a mid-range Stigma score comparable to others (Figure 4), implying that low acceptance can stem from perceived lack of utility/usability even when stigma is not high. Conversely, P3 reported high Benefit and Usability despite reduced S-shape feasibility (Table 2), suggesting that some users may recognize the potential value of the system even if one control profile does not yet fit their motor strategy. For design and deployment, this means that acceptance should not be treated as a single scalar; rather, domain-level patterns can help distinguish “the device feels stigmatizing” from “the device does not help me” from “the device might help me but needs tuning/training.” Because the TSQ-WT was administered once after exposure to both robotic trajectories, these acceptance ratings reflect the system as a whole and cannot be attributed to a specific trajectory. Trajectory-specific acceptance measurement is an important refinement for future studies.

Individual Differences and Case-Based Interpretation

Overview

The cohort was heterogeneous enough that individual differences are themselves a result. We therefore use a case-based lens and summarize recurring case patterns across feasibility, biomechanics, and acceptance, with direct design and deployment implications (full case narratives in Multimedia Appendix 3):

Case Pattern 1: Robust Users with Consistent Biomechanics and High Acceptance (P4, P7)

These participants showed ceiling feasibility across conditions, smooth and repeatable Fz transfers (narrow SD envelopes and loading near body weight late in the cycle), and high TSQ benefit/usability with low Stigma (Figures 3 and 4). Notably, this pattern includes both high-function/low-concern (P4) and low-function/high-concern (P7) participants (Table 2), indicating that baseline function alone does not determine successful adoption; stable coupling between user strategy and assistance profile appears critical. An unmeasured but plausible mediator is upper-limb capacity: 3 of 6 completers (P2, P5, and P7) could not rise once without arm use in the SPPB chair-stand test, yet completed assisted repetitions by loading the handles. This observation shows that arm support was necessary for these users. Whether upper-limb strength, joint range of motion, or the ability to time arm loading with the moving handles mediates success with handle-based assistance is a hypothesis this dataset cannot test, and direct upper-limb measurements are needed.

Case Pattern 2: Feasible but Biomechanically Trajectory-Sensitive Users (P2)

P2 achieved ceiling success in all conditions yet displayed a marked trajectory separation in Fz (S-shape remaining under body weight late in the movement with larger variability) and high mean Borg across conditions (Figure 3). This suggests that feasibility alone can mask important differences in how assistance is being used: a user may complete every repetition while still experiencing higher effort or relying on a less repeatable strategy under one profile. These cases motivate monitoring transfer quality rather than success/failure alone.

Case Pattern 3: Users with Reduced Feasibility or Low Perceived Benefit/Usability (P5, P6, and the Noncompletion Case)

P5 and P6 showed partial completion with SkyWalker, broad Fz variability, and comparatively lower acceptance (especially

XSL•FO
RenderX

P6 with very low benefit/usability) (Table 2, Figures 3 and 4). P5's remarks that the assistance felt "unusual"/"unnatural" put words to this pattern. The noncompletion during early exposure underscores that a subset of intended users may be unable to use the system at all; the documented circumstances are reported in the participants section. Together, these cases suggest that a one-size-fits-all trajectory may be insufficient for clinical deployment.

Across patterns, acceptance can remain high even when one trajectory is not feasible (P3), supporting the interpretation that trajectory adaptation and/or additional familiarization may improve usability for some users without reducing perceived benefit.

Limitations and Future Work

This study evaluated six completers in a single session, which restricts statistical power and the ability to characterize learning effects. The sample size is largely structural: recruiting and testing patients in a rehabilitation hospital requires aligning with clinical schedules, medical status, fatigue tolerance, and safety considerations, and it is difficult to access larger cohorts with uniform diagnoses or impairment profiles within the same time window. As a result, the cohort is clinically heterogeneous, which reflects the intended user group in a clinical setting but constrains between-participant comparability. We note that this remains a supervised, instrumented, single-session experiment. It does not establish real-world or unsupervised use, and the comparator rollator differs from the SkyWalker in geometry, handle configuration, and permitted support strategy, so condition differences cannot be attributed to actuation alone.

The rollator condition exhibited ceiling success, which limits the interpretability of success-rate comparisons and shifts attention to effort and biomechanical quality. Additional limitations concern the measurements themselves. First, upper-limb and handle forces were not recorded, so reduced foot loading under assistance cannot be uniquely attributed to device support. Second, although the TSQ-WT has previously been used to evaluate STS assistance provided by a smart walker in older adults [22], a formal comprehensive evaluation of its psychometric properties is still lacking, which limits the interpretability of the questionnaire-derived findings. In addition, it was administered once for the robotic system as a whole, so acceptance cannot be linked to a specific trajectory. Third, biomechanical group curves average repetitions aligned at movement onset over a fixed time window rather than normalizing each repetition to its own movement duration. Per-curve repetition counts are disclosed in the figure captions. Fourth, seat height was adjusted to each participant's body height rather than fixed across participants, which affects raw between-participant comparisons of vertical CoG height. Fifth, although condition order was randomized, a short familiarization/learning component was observed for one participant, so within-session learning effects cannot be excluded. Future work should therefore include larger samples, multi-session exposure to capture learning/adaptation, direct comparative satisfaction measures across conditions, trajectory-specific acceptance measurement, and additional sensing (eg, upper-limb load sharing) to more directly explain why some participants exhibit sustained under-loading or high variability in Fz under specific trajectories.

Conclusions

This study is the first evaluation of SkyWalker with its intended users. It evaluated the feasibility, perceived effort, biomechanics, and user experience of SkyWalker, a rollator providing powered STS assistance, in older adults with mobility challenges, using a within-subject comparison of two assistance trajectories (linear and S-shape) against a conventional rollator. Addressing RQ1, SkyWalker STS assistance was feasible for most enrolled users who completed testing (6/7), with no observed adverse events among completers; however, feasibility was not uniform, as one participant was unable to complete testing and several participants showed partial completion in at least one powered condition, particularly with the S-shape trajectory. Robotic assistance did not improve task success relative to the conventional rollator. Addressing RQ2, perceived exertion (Borg CR10) was lowest with the linear trajectory, whereas the S-shape trajectory offered no exertion advantage over the conventional rollator. The effort benefit of powered assistance was therefore trajectory-specific rather than general. Addressing RQ3, the assistance profiles produced distinct biomechanical signatures: linear was associated with earlier and more monotonic CoG elevation, while participant-level Fz patterns revealed substantial inter-individual variability in load transfer timing and repeatability, with several users exhibiting broader variability or sustained under-loading in specific trajectories. These findings support the interpretation that trajectory design influences coordination constraints and that robust assistance should prioritize repeatable, stable weight transfer rather than success alone. Addressing RQ4, responses on the unvalidated adapted TSQ-WT were heterogeneous (22 - 77 of 80): benefit, usability, and self-concept ratings were high for 4 participants but low for 2, and these descriptive totals do not establish validated technology acceptance, including one markedly low-acceptance profile.

Overall, the results suggest that SkyWalker STS assistance with a simple, monotonic trajectory can reduce perceived effort for a subset of users and was rated positively by most, while trajectory-specific and participant-specific feasibility constraints point to personalization, training, and adaptive trajectory policies. In particular, the linear trajectory is the profile to prioritize in further testing in this population, whereas S-shape may be effective for some users but appears more sensitive to individual capacity. These observations come from a within-participant feasibility study with six completers and are reported as feasibility evidence, not as efficacy claims. The results are already being used to design the next generation of the SkyWalker, which addresses the challenges identified here. Future work should evaluate multi-session learning effects, include direct comparative satisfaction/preference measures against conventional aids, and incorporate larger samples with stratification by functional status to determine which users benefit most and how assistance trajectories should be tuned for safe, reliable clinical and home deployment.

XSL•FO
RenderX

Acknowledgments

The authors thank the clinical staff at Agaplesion Bethanien Krankenhaus Heidelberg for their support during participant recruitment and testing, and the participants who volunteered their time.

Funding

This research received funding from the Canada Excellence Research Chair program at the University of Waterloo, the Hector Foundation at Karlsruhe Institute of Technology, and the Carl Zeiss Foundation within the HeiAge Project (grant #: P2017-01-002) at Heidelberg University

Data Availability

The individual participant data underlying this study (raw recordings and questionnaire booklets) cannot be shared publicly owing to patient-privacy and data-protection restrictions at the participating rehabilitation hospital. The participant-level outcome data supporting the findings are reported in the manuscript and its multimedia appendices.

Authors' Contributions

Methodology: AM
Software: AM
Formal analysis: AM
Writing of the original draft: AM
Investigation (participant recruitment, patient interaction and feedback recording during testing, and questionnaire administration and collection): AM, TB
Data curation: TB
Resources: CW
Visualization: AM
Conceptualization: AM, CW, KM
Supervision: CW, YH, KM.
All authors reviewed and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tele-healthcare Satisfaction Questionnaire.
[DOCX File, 11 KB - rehab_v13i1e101303_app1.docx]

Multimedia Appendix 2
Pairwise analyses.
[DOCX File, 11 KB - rehab_v13i1e101303_app2.docx]

Multimedia Appendix 3
Case descriptions.
[DOCX File, 183 KB - rehab_v13i1e101303_app3.docx]

References

1. Musich S, Wang SS, Ruiz J, Hawkins K, Wicker E. The impact of mobility limitations on health outcomes among older adults. Geriatr Nurs 2018;39(2):162-169. [doi: 10.1016/j.gerinurse.2017.08.002] [Medline: 28866316]
2. Schenkman M, Berger RA, Riley PO, Mann RW, Hodge WA. Whole-body movements during rising to standing from sitting. Phys Ther 1990 Oct;70(10):638-648. [doi: 10.1093/ptj/70.10.638] [Medline: 2217543]
3. Nuzik S, Lamb R, VanSant A, Hirt S. Sit-to-stand movement pattern. A kinematic study. Phys Ther 1986 Nov;66(11):1708-1713. [doi: 10.1093/ptj/66.11.1708] [Medline: 3774881]
4. Alexander NB, Schultz AB, Warwick DN. Rising from a chair: effects of age and functional ability on performance biomechanics. J Gerontol 1991 May;46(3):M91-M98. [doi: 10.1093/geronj/46.3.m91] [Medline: 2030272]
5. Hughes MA, Schenkman ML. Chair rise strategy in the functionally impaired elderly. J Rehabil Res Dev 1996 Oct;33(4):409-412. [Medline: 8895136]
6. Resch S, Zirari A, Tran TDQ, Bauer LM, Sanchez-Morillo D. Smart walking aids with sensor technology for gait support and health monitoring: a scoping review. Technologies 2025;13(8):346. [doi: 10.3390/technologies13080346]

XSL•FO
RenderX

7. Chugo D, Matsuoka W, Jia S, Takase K. Rehabilitation walker with standing-assistance device. J Robot Mechatron 2007 Dec 20;19(6):604-611. [doi: 10.20965/jrm.2007.p0604]
8. Chugo D, Kitamura T, Takase K, Asawa T, Jia S. A rehabilitation walker with standing and walking assistance. Presented at: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Sep 22-26, 2008; Nice, France p. 260-265. [doi: 10.1109/IROS.2008.4650845]
9. Chugo D, Asawa T, Kitamura T, Songmin J, Takase K. A motion control of a robotic walker for continuous assistance during standing, walking and seating operation. Presented at: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Oct 10-15, 2009; St Louis, MO p. 4487-4492. [doi: 10.1109/IROS.2009.5354523]
10. Morita Y, Chugo D, Sakaida Y, Yokota S, Takase K. Standing motion assistance on a robotic walker based on the estimated patient's load. Presented at: IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob); Jun 24-27, 2012; Rome, Italy p. 1721-1726. [doi: 10.1109/BIOROB.2012.6290737]
11. Chugo D, Kawazoe S, Yokota S, et al. Development of a standing assistance walker for a patient with low level of care. Presented at: 14th International Conference on Informatics in Control, Automation and Robotics; Jul 26-28, 2017; Madrid, Spain p. 142-153. [doi: 10.5220/0006469501420153]
12. Sun Y, Xiao C, Chen L, et al. A review of intelligent walking support robots: aiding sit-to-stand transition and walking. IEEE Trans Neural Syst Rehabil Eng 2024;32:1355-1369. [doi: 10.1109/TNSRE.2024.3379453] [Medline: 38502616]
13. Huang J, Yan S, Yang D, et al. Proxy-based control of intelligent assistive walker for intentional sit-to-stand transfer. IEEE/ASME Trans Mechatron 2022;27(2):904-915. [doi: 10.1109/TMECH.2021.3074800]
14. Moustris G, Kardaris N, Tsiami A, et al. The i-walk lightweight assistive rollator: first evaluation study. Front Robot AI 2021;8:677542. [doi: 10.3389/frobt.2021.677542] [Medline: 34604315]
15. Xu Y, Zhang B, Li Y, et al. Design and evaluation of the sit-to-stand movement assistive device for elderly. Actuators 2024;13(3):102. [doi: 10.3390/act13030102]
16. Bulea TC, Triolo RJ. Design and experimental evaluation of a vertical lift walker for sit-to-stand transition assistance. J Med Device 2012 Mar;6(1):14504-14NaN. [doi: 10.1115/1.4005786] [Medline: 23904904]
17. Chuy O, Hirata Y, Wang Z, Kosuge K. Approach in assisting a sit-to-stand movement using robotic walking support system. Presented at: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems; Oct 9-15, 2006. [doi: 10.1109/IROS.2006.282007]
18. Geravand M, Korondi PZ, Werner C, Hauer K, Peer A. Human sit-to-stand transfer modeling towards intuitive and biologically-inspired robot assistance. Auton Robot 2017 Mar;41(3):575-592. [doi: 10.1007/s10514-016-9553-5]
19. Sloot LH, Millard M, Werner C, Mombaur K. Slow but steady: similar sit-to-stand balance at seat-off in older vs. younger adults. Front Sports Act Living 2020;2:548174. [doi: 10.3389/fspor.2020.548174] [Medline: 33345111]
20. Mombaur K, Ho Hoang KL. How to best support sit to stand transfers of geriatric patients: motion optimization under external forces for the design of physical assistive devices. J Biomech 2017 Jun 14;58:131-138. [doi: 10.1016/j.jbiomech.2017.04.037] [Medline: 28551097]
21. Ho Hoang KL, Mombaur KD. Optimal design of a physical assistive device to support sit-to-stand motions. Presented at: 2015 IEEE International Conference on Robotics and Automation (ICRA); May 26-30, 2015; Seattle, WA p. 5891-5897. [doi: 10.1109/ICRA.2015.7140024]
22. Werner C, Geravand M, Korondi PZ, Peer A, Bauer JM, Hauer K. Evaluating the sit‐to‐stand transfer assistance from a smart walker in older adults with motor impairments. Geriatr Gerontol Int 2020 Apr;20(4):312-316. [doi: 10.1111/ggi.13874] [Medline: 32006458]
23. Werner C, Ullrich P, Geravand M, Peer A, Hauer K. Evaluation studies of robotic rollators by the user perspective: a systematic review. Gerontology 2016;62(6):644-653. [doi: 10.1159/000444878] [Medline: 27008089]
24. Werner C, Ullrich P, Geravand M, Peer A, Bauer JM, Hauer K. A systematic review of study results reported for the evaluation of robotic rollators from the perspective of users. Disabil Rehabil Assist Technol 2018 Jan;13(1):31-39. [doi: 10.1080/17483107.2016.1278470] [Medline: 28125298]
25. Reicherzer L, Scheermesser M, Kläy A, Duarte JE, Graf ES. Barriers and facilitators to the use of wearable robots as assistive devices: qualitative study with older adults and physiotherapists. JMIR Rehabil Assist Technol 2024 Aug 9;11:e52676. [doi: 10.2196/52676] [Medline: 39132691]
26. Potter S, Hawley M, Higgins A, et al. Assistive robotics for healthy aging: a foundational phenomenological co-design exercise. J Med Internet Res 2026 Jan 28;28:e77179. [doi: 10.2196/77179] [Medline: 41605501]
27. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
28. Mitzner TL, Stuck R, Hartley JQ, Beer JM, Rogers WA. Acceptance of televideo technology by adults aging with a mobility impairment for health and wellness interventions. J Rehabil Assist Technol Eng 2017;4:2055668317692755. [doi: 10.1177/2055668317692755] [Medline: 31186923]
29. David D, Thérouanne P, Milhabet I. The acceptability of social robots: a scoping review of the recent literature. Comput Human Behav 2022 Dec;137:107419. [doi: 10.1016/j.chb.2022.107419]
30. Getson C, Nejat G. Socially assistive robots helping older adults through the pandemic and life after COVID-19. Robotics 2021;10(3):106. [doi: 10.3390/robotics10030106]

XSL•FO
RenderX

31. Mahdi A, Dong Z, Lin JFS, Hu Y, Hirata Y, Mombaur K. Enabling safe sit-to-stand support with mobile robots: 3D visual pose estimation in close quarters and support adaptation. In: Converging Clinical and Engineering Research on Neurorehabilitation V: Springer Nature Switzerland; 2025:524-528. [doi: 10.1007/978-3-031-77588-8_103]
32. Folstein MF, Robins LN, Helzer JE. The Mini-Mental State Examination. Arch Gen Psychiatry 1983 Jul;40(7). [doi: 10.1001/archpsyc.1983.01790060110016] [Medline: 6860082]
33. Pavasini R, Guralnik J, Brown JC, et al. Short physical performance battery and all-cause mortality: systematic review and meta-analysis. BMC Med 2016 Dec 22;14(1):215. [doi: 10.1186/s12916-016-0763-7] [Medline: 28003033]
34. Yardley L, Beyer N, Hauer K, Kempen G, Piot-Ziegler C, Todd C. Development and initial validation of the Falls Efficacy Scale-International (FES-I). Age Ageing 2005 Nov;34(6):614-619. [doi: 10.1093/ageing/afi196] [Medline: 16267188]
35. Zijlstra W, Becker C, Pfeiffer K. Wearable systems for monitoring mobility related activities. In: Röcker C, Ziefle M, editors. E-Health, Assistive Technologies and Applications for Assisted Living: IGI Global; 2011:244-267. [doi: 10.4018/978-1-60960-469-1.ch011]
36. Moulaei K, Moulaei R, Bahaadinbeigy K. The most used questionnaires for evaluating the usability of robots and smart wearables: a scoping review. Digit Health 2024;10:20552076241237384. [doi: 10.1177/20552076241237384] [Medline: 38601185]
37. Schmidt LI, Jansen CP, Depenbusch J, Gabrian M, Sieverding M, Wahl HW. Using wearables to promote physical activity in old age: feasibility, benefits, and user friendliness. Z Gerontol Geriatr 2022 Aug;55(5):388-393. [doi: 10.1007/s00391-022-02083-x] [Medline: 35849158]

Abbreviations

Borg CR10: Borg Category-Ratio 10 scale
CFS: Clinical Frailty Scale
CoG: center of gravity
CoM: center of mass
FES-I: Falls Efficacy Scale–International
Fz: vertical ground-reaction force
MMSE: Mini-Mental State Examination
QoL: quality of life
RQ: research question
SPPB: Short Physical Performance Battery
STS: sit-to-stand
TSQ-WT: Tele-healthcare Satisfaction Questionnaire–Wearable Technology

Edited by A Scano; submitted 13.May.2026; peer-reviewed by D Trabassi, L Herdiman; revised version received 03.Sep.2026; accepted 01.Oct.2026; published 11.Nov.2026.

Please cite as:
Mahdi A, Buchner T, Werner C, Hu Y, Mombaur K
Robotic Rollator Assistance for Sit-to-Stand Transfers in Older Adults with Mobility Challenges: Within-Participant Feasibility Study
JMIR Rehabil Assist Technol 2026;13:e101303
URL: https://rehab.jmir.org/2026/1/e101303
doi:10.2196/101303

© Anas Mahdi, Theresa Buchner, Christian Werner, yue Hu, Katja Mombaur. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 11.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and Acceptance of a Social Robot in People With Multiple Sclerosis: Comparative Cross-Sectional Study

Jessica Podda[1], PhD; Alice Bollini[1], PhD; Alessia Susini[1], MSc; Domenico Presenza[2], MSc; Keven T Kearney[2], MSc; Ester Cecchella[1], MD; Erica Grange[1,3], MSc; Mario Alberto Battaglia[4,5], MD; Andrea Tacchino[1*], PhD; Giampaolo Brichetto[1,6*], MD, PhD

[1]neuroBRITE Research Center, Italian Multiple Sclerosis Foundation, Via Operai 40, Genoa, Italy

[2]R&I Lab, Engineering Ingegneria Informatica S.p.A, Rome, Lazio, Italy

[3]Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genoa, Genoa, Italy

[4]Department of Life Science, University of Siena, Siena, Italy

[5]Italian Multiple Sclerosis Foundation, Genoa, Italy

[6]AISM Rehabilitation Service, Italian Multiple Sclerosis Society, Genoa, Italy

[*]these authors contributed equally

Corresponding Author:
Jessica Podda, PhD
neuroBRITE Research Center, Italian Multiple Sclerosis Foundation, Via Operai 40, Genoa, Italy

Abstract

Background: Socially assistive robotics (SAR) is an emerging field integrating robotics, medicine, cognitive and social sciences, and neuroscience. Its potential in multiple sclerosis (MS) remains largely unexplored, although MS affects sensory-motor, cognitive, and psychological domains where SAR could offer substantial benefits.

Objective: This study aimed to evaluate the usability of a humanoid robot in people with MS, compare it with age-matched healthy individuals (HIs), and explore usability predictors. Automated written-language analysis was tested as a complementary approach.

Methods: In this 2025 cross-sectional study at the Italian MS Foundation laboratory, 20 people with MS (mean age 51.4, SD 11.0 years; n=13, 65% female; mean education 13.0, SD 3.4 years) and 20 age-matched HIs (mean age 47.3, SD 10.1 years; n=12, 60% female; mean education 14.8, SD 2.2 years) completed 4 predefined tasks with the robot Pepper. Usability was assessed using the System Usability Scale (SUS). Attitudes toward robots, robot-related factors, and task-related effort were measured with the General Attitudes Towards Robots Scale (GAToRS), Godspeed questionnaire, and NASA-TLX, respectively. Psychological variables (eg, personality traits from the Italian Ten-Item Personality Inventory [I-TIPI]) and clinical characteristics for people with MS (eg, cognitive functions via the app DIGICOG-MS, upper-limb difficulties with the Arm Function in Multiple Sclerosis [AMSQ]) were collected. Written responses to open-ended questions were analyzed using natural language processing (NLP).

Results: SUS scores exceeded the standard cut-off in both groups, with no between-group differences (people with MS: mean 81.00, SD 12.31; HIs: mean 77.75, SD 11.24; r=0.14, 95% CI 0.01-0.42; P=.39), indicating high usability. People with MS reported more negative attitudes toward robots at the societal level (GAToRS S–: r=0.36, 95% CI 0.08-0.63; P=.02) and judged Pepper as more anthropomorphic (r=0.42, 95% CI 0.15-0.67; P=.007) and intelligent (r=0.37, 95% CI 0.09-0.61; P=.02) than HIs. Anthropomorphism was strongly associated with visuospatial memory (DIGICOG-MS: ρ=–0.632, 95% CI –0.84 to –0.26; P=.003) and relatively strong with both verbal memory (DIGICOG-MS: ρ=–0.504, 95% CI –0.77 to –0.08]; P=.02) and agreeableness (I-TIPI: ρ=0.491, 95% CI 0.06-0.77; P=.03). Exploratory forward stepwise regression showed that conscientiousness (I-TIPI: B=2.064, 95% CI 0.679-3.449), negative personal attitudes toward robots (GAToRS P–: B=0.169, 95% CI 0.033-0.306), and perceived upper-limb difficulties (AMSQ: B=0.169, 95% CI 0.033-0.306) explained 64% of SUS variance (adjusted R^2=0.568); task-related frustration (NASA-TLX: B=–1.497, 95% CI –2.276 to –0.719) explained 48% (adjusted R^2=0.451). NLP analyses revealed no between-group differences in lexical productivity or sentiment-lexicon coverage.

Conclusions: This study provides first preliminary evidence that people with MS perceive a social robot as highly usable, comparably to HIs, suggesting MS-related challenges may not hinder human-robot interaction. Psychological factors, prior attitudes toward robots, and clinical characteristics were associated with user experience in the MS group. Future research should explore whether integrating these measures with language-based approaches informs tailored interventions in MS.

(JMIR Rehabil Assist Technol 2026;13:e94068) doi:10.2196/94068

XSL•FO
RenderX

KEYWORDS

socially assistive robotics; multiple sclerosis; usability; technology acceptance; anthropomorphism; human-robot interaction; natural language processing; Pepper

Introduction

Socially assistive robotics (SAR) is a growing interdisciplinary research field aiming at supporting people in different social situations through robot-based solutions [1]. This area integrates robotics, medicine, social and cognitive sciences, and neuroscience. Social robots are used in various domains, such as health care, education, business, industry, and culture [2], where interaction between individuals is essential. One of the main challenges of human-robot interaction applied to different research areas, such as physical and cognitive rehabilitation or emotional support, lies in ensuring a fluid and effective interplay so that users have the subjective impression that they are interacting with an intelligent agent. In this context, achieving a positive user experience (UX) is crucial, as it directly influences engagement, acceptance, and the overall impact of SAR on human life [2]. SAR has demonstrated its value to promote people's well-being by creating new avenues to access services and resources for mental health in real-world contexts [3], offering numerous possibilities for the innovation and implementation of robots in education, medicine, rehabilitation, and psychotherapy, improving treatment compliance in children with specific clinical conditions such as autism spectrum disorder [4-8] and in reducing anxiety among older adults, with subsequent improvements in their quality of life [9]. As indicated in a recent review [10], SAR has been used with promising success specifically in supporting everyday life, providing interaction, and facilitating cognitive and physical training in the older population. Despite these advances, the use of SAR in other neurological populations [11-13], such as people with multiple sclerosis (MS), remains largely unexplored. This is surprising given that MS is characterized by a heterogeneous combination of sensory-motor impairments, cognitive deficits, and psychological symptoms, all domains in which SAR could potentially offer substantial benefits [14,15]. The impact of MS extends into work roles, economic status, and social relationships. It is reasonable to surmise, therefore, that MS and its manifestations make those living with this disease vulnerable to the experience of loneliness and social isolation [16,17]. In this context, SAR may offer a promising avenue not only for physical and cognitive rehabilitation but also for addressing the social and emotional needs of people with MS—a gap that has received little attention in the existing literature. However, before implementing such interventions, it is crucial to understand how people with MS perceive and interact with social robots. To date, there is no empirical evidence on how this clinical population perceives key social dimensions of human-robot interaction, such as robot anthropomorphism. Without this foundational knowledge, deploying SAR in MS rehabilitation risks being ineffective or misaligned with people's needs.

Based on these premises, the aims of this study were to (1) evaluate the usability of humanoid robot Pepper (SoftBank Robotics) in people with MS; (2) investigate any differences with an age-matched group of healthy individuals (HIs) in both perceived usability and task-related effort; and (3) exploratorily examine potential predictors of usability, including psychological variables, prior attitudes toward robots, and clinical characteristics. Furthermore, as a complementary component of the study, we applied natural language processing (NLP) techniques to participants' written responses to open-ended questions with the aim of examining whether automated language analysis could serve as a proxy for evaluating user appreciation of the social robot alongside the primary quantitative measures. This analysis was not intended as a substitute for qualitative inquiry, but rather as a preliminary, hypothesis-generating exploration of the linguistic correlates of UX in this population.

Methods

Study Population

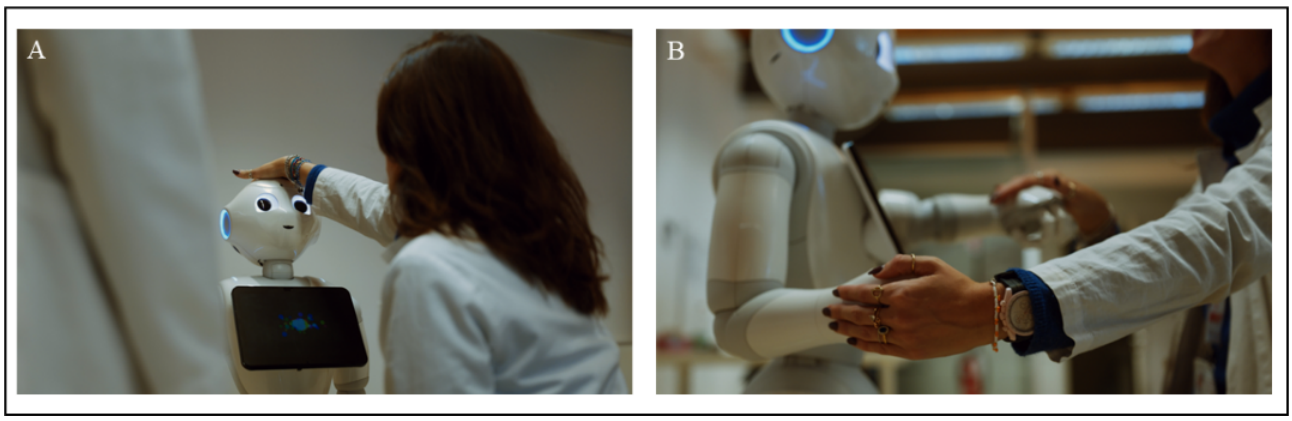
This cross-sectional study was conducted in November 2025 at the laboratory of the Fondazione Italiana Sclerosi Multipla (FISM) in Genoa (Italy). The study sample consisted of people with MS, followed as outpatients at the Italian Multiple Sclerosis Society (AISM) Rehabilitation Service of Liguria (Genoa, Italy). Potential participants were identified by the AISM clinical staff and approached directly during routine clinical appointments or contacted by telephone by the FISM Secretariat. All individuals contacted agreed to participate, except for 1 individual who expressed willingness but was unable to attend due to work commitments. Eligibility criteria included an age of 18 years and older; a confirmed MS diagnosis [18]; any disease course (relapsing-remitting, secondary progressive, primary progressive MS); the absence of relapses in the last 3 months; an Expanded Disability Status Scale (EDSS) [19] of 6.5 or below; and adequate visual, hearing, and motor capabilities to complete the experimental tasks. Exclusion criteria were a Montreal Cognitive Assessment score of 18 or below [20], neurological and major psychiatric illness, past serious head trauma, and alcohol or drug abuse. Inclusion criteria were intentionally broad to reflect the heterogeneity of the real-world MS population; eligible individuals were first identified by the clinical staff in collaboration with the research team, ensuring that only those meeting all eligibility criteria and able to safely complete the experimental tasks were approached. A healthy control group was included to provide a reference for comparison, allowing us to determine whether people with MS respond to the robot differently from HIs, and to identify usability patterns and attitudes that may be specific to the MS population rather than common to all users. HIs were recruited on a voluntary basis among AISM employers through institutional email distribution, matched by age, and were required to have no history of neurological or psychiatric diseases. All participants provided written informed consent prior to enrollment.

XSL•FO
RenderX

The Social Robot and Software Used in the Experiment

The device used in this study was the humanoid social robot Pepper. Pepper is designed to interact with people and is capable of autonomous navigation, face recognition, scripted dialogue, and adopting life-like human poses, such as performing arm and hand gestures. In this study, Pepper was programmed with a small and simple set of behaviors: greeting participants, providing instructions for each experimental task, and guiding participants through the tasks themselves, as detailed in the next section (Figure 1). The software controlling the robot's behaviors in the study was developed by Engineering Ingegneria Informatica S.p.A. (EII) as part of their long-term Socially Assistive Robotic Solution for Ambient (SARA)–assisted living, a complex distributed home automation and robotics system providing health monitoring and (socially interactive) assistance in daily living tasks to older adults and their caregivers at home, with the aim of prolonging their autonomy and delaying their institutionalization. Following [21], the SARA has the ambitious goal to develop robots that provide physical assistance, for example, rehabilitation robots, wheelchair robots, mobility aids, manipulator arms), and are able to engage humans in interaction, for example, from [22] engaging in high-level dialog, using and understanding natural cues (gaze, gesture, etc), expressing/perceiving emotions, and establishing and maintaining social relationships (among other competencies).

Figure 1. Participant-robot interaction with the humanoid social robot Pepper (SoftBank Robotics, Tokyo, Japan). (A) All interactions started with a head-touch gesture to activate the robot's social engagement mode. (B) Keyframe from 1 experimental task, showing the reciprocal movement dynamics between the participant and the robot.



Experimental Procedures: Human-Robot Interaction

Prior to the experimental sessions, a preliminary focus group was conducted with a multidisciplinary team, including clinicians with a high experience in MS (eg, psychologists, physiotherapists, occupational therapists) and engineers, to identify the most appropriate and meaningful tasks for the study. Insights from this focus group were used to select and refine task scenarios that were engaging and feasible for interaction with the robot. This collaborative approach ensured that the activities included in the experiment were ecologically valid and aligned with both abilities and practical considerations of people with MS for the robot's implementation.

Each participant, in a structured setting, was invited to engage in 4 predefined interaction sequences with the robot (inspired by Tleubayev et al [23]), served to document engagement levels and usability perceptions of the robotic system, under the supervision of a trained health care professional (ie, a psychologist). The first 2 tasks were "static," with the participant sitting in front of the robot, which remained docked (in a standing pose) on its recharging station throughout. Both tasks relied principally on the robot's touch sensors and its upper body motion capabilities. The tasks were as follows:

- Task 1—Make me pose: The goal for the participant, seated in front of Pepper, was to make the robot adopt a target pose (demonstrated by the robot at the beginning of the interaction). Starting from a random initial pose, the robot moved into a new pose each time the participant touched the robot's hands. The task stopped when the participant recognized that Pepper reached the target pose, indicated by the participant touching the robot's head. The aim of this exercise was to stimulate the participant's memory ability.
- Task 2—Keep in touch: The objective for the participant, again seated in front of Pepper, was to maintain physical contact between her/his hands and the robot's hands while the robot executed a slow preplanned sequence of arm movements. The task stopped at the end of the sequence. The aim of this exercise was to stimulate participants' focused attention while engaging in controlled upper body movements.

The remaining tasks were more dynamic, with both Pepper and the participant free to move around. The robot's behavior in these tasks relied on its autonomous navigation capabilities and used proximity sensors (infrared, sonars, and lasers) embedded within its mobile base to detect the relative location of the participant. The final task also depended on the native visual localization functions of the robot. The 2 dynamic tasks were as follows:

- Task 3—Dance with me: The participant stood at 1 m or so in front of Pepper. The goal was to guide the robot through a short sequence of locomotive movements (forward, backward, and turning). The robot moved in response to changes in the relative positions of the participant and the robot—for example, if the participant stepped backward, Pepper moved forward—to maintain a

XSL•FO
RenderX

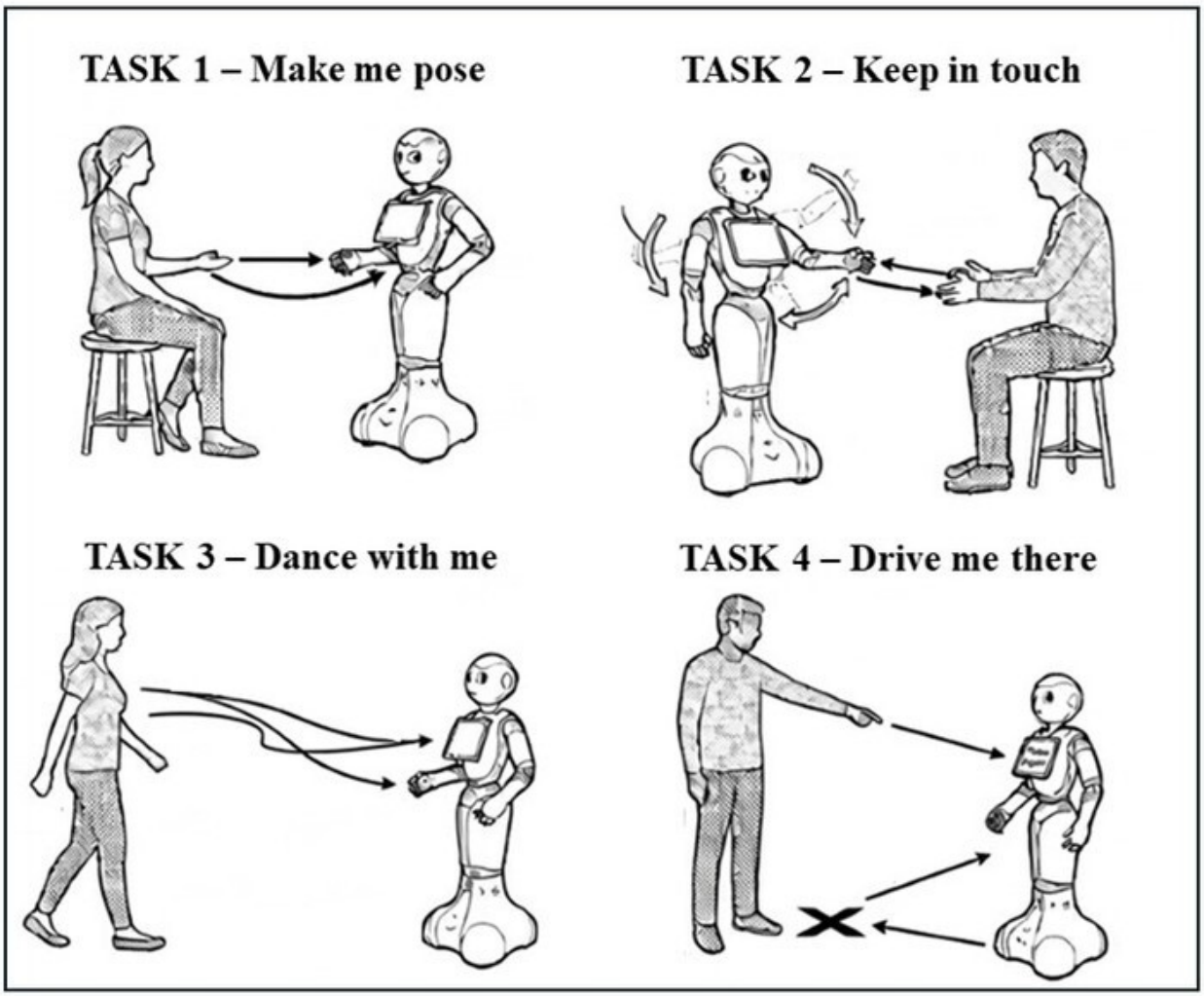
constant relative distance and orientation. The aim of this exercise was to stimulate working memory and spatio-temporal orientation with the participant engaging in simple locomotion.

- Task 4—Drive me there: Once again the participant was standing, but now the objective was to guide the robot to a nearby location (a few meters away). The robot moved in response to the participant's movements in the same way as in the previous Task 3. The aim of this exercise was to stimulate spatio-temporal orientation and planning abilities with the participant engaging in simple locomotion.

Each task was introduced and explained to the participant by the robot itself. During the task, the robot also offered occasional encouraging remarks and provided feedback on the participant's progress, which were all accompanied with appropriate hand/arm and head gestures to favor communication and avoid ambiguity.

The experimental session lasted 45 minutes, comprising interaction with Pepper (20 min) and test and questionnaire administration (25 min). Figure 2 shows an overview of the 4 experimental tasks illustrating human-robot interaction.

Figure 2. Overview of the 4 experimental tasks illustrating human-robot interaction. Participants performed 4 tasks with Pepper. In Task 1 (Make me pose), a seated participant guided the robot to a target pose via touch. In Task 2 (Keep in touch), a seated participant maintained continuous contact with the robot's hands while it executed a sequence of movements. In Task 3 (Dance with me), a standing participant led the robot through locomotor movements using a "follow-the-leader" approach. In Task 4 (Drive me there), a standing participant guided the robot to a target location.



Demographic and Clinical Outcomes

Sociodemographic data (ie, age, sex, educational level) were collected for all participants, while disease-related information (ie, MS course, disease duration, and disability level as indicated by EDSS) was gathered only for people with MS. Furthermore, MS cognitive functioning was tested using DIGICOG-MS, a mHealth app including digital tests that measure visuo-spatial and verbal memory, information processing speed, and semantic fluency [24]. Patient-reported outcomes (PROs), routinely used in MS clinical practice and well-established in this population, were administered to people with MS to assess upper-limb functionality (Arm Function in Multiple Sclerosis [AMSQ] [25]), perceived fatigue (Modified Fatigue Impact Scale [MFIS] [26,27]), perceived cognitive deficit (Multiple Sclerosis Neuropsychological Questionnaire [MSNQ] [28]), and impact of MS (29-item Multiple Sclerosis Impact Scale [MSIS 29] [29]). State and trait anxiety (State-Trait Anxiety Inventory [STAI] [30]), risk of depression (9-item Patient Health Questionnaire [PHQ-9] [31,32]), and personality traits (Italian Ten-Item Personality Inventory [I-TIPI] [33]) were collected

XSL•FO
RenderX

for both people with MS and HIs. All questionnaires were administered in paper-and-pencil format, except for DIGICOG-MS, which was completed digitally. Participants completed all measures independently, with a psychologist available to clarify any questions. All instruments have been validated in Italian, with the exception of the MSIS-29, for which no formal Italian-language validation study has been published to date; however, its Italian translation is widely used in MS research internationally.

Human-Robot Interaction Outcomes

Before starting the experiment, all participants completed the General Attitudes Towards Robots Scale (GAToRS) [34] to explore positive and negative attitudes toward robots at a personal and societal level. After the task, the following questionnaires were administered: System Usability Scale (SUS) [35] for overall usability (primary outcome); Godspeed questionnaire [36,37] for robot-related factors such as Anthropomorphism (ANTR), Animacy (ANIM), Likeability (LIKE), Perceived Intelligence (INT), or Perceived Safety (SAF) of robots; the National Aeronautics and Space Administration Task Load Index (NASA-TLX) [38] for mental and physical workload.

After the interaction with the robot, participants responded to three open-ended written questions: (1) “How would you describe your experience with the social robot during the activity?” (2) “What kind of feelings or thoughts did the interaction with the robot evoke?” (3) “Was there anything that struck you or surprised you about the way the robot interacted with you?”

Statistical Analysis

Given the absence of prior SAR studies in people with MS, no formal a priori power analysis was feasible. The sample size of 20 per group was chosen in line with established practice in pilot SAR research, where small samples are commonly adopted when exploring novel populations or contexts [39,40]. Descriptive statistics including means, SDs, SEs, and medians with IQRs were calculated to summarize the data. Depending on the distribution of the data, parametric tests, nonparametric tests, or chi-square tests were used to compare demographic, psychological, and technological/robotic factors between people with MS and HIs, as appropriate. To explore factors associated with the group differences, exploratory bivariate correlations with clinical, psychological, and functional variables were run. Pearson or Spearman correlation analyses, as appropriate, were performed to examine associations between variables, with correlation coefficients interpreted as follows: 0.20-0.39 were interpreted as moderate, 0.40-0.59 as relatively strong, 0.60-0.79 as strong, and ≥0.80 as very strong correlation [41]. Effect sizes for Mann-Whitney *U* tests were reported as *r* ($Z/\sqrt{N}$) and interpreted as small (<0.30), medium (0.30 - 0.50), or large (>0.50). Effect sizes for χ^2 tests were reported as Cramér *V*. Given the lack of evidence on factors influencing the usability of social robots in people with MS and the small sample size (n=20), 2 exploratory stepwise linear regressions were conducted—one examining psychological factors, the other technological/robotic factors—to identify possible predictors, both using a forward selection procedure, entering all candidate predictors simultaneously and retaining only those meeting the p-to-enter threshold of .05 and p-to-remove threshold of .10. Particularly, the first stepwise multiple linear regression analysis was conducted to explore the extent to which personality traits (I-TIPI), attitudes toward robots (GAToRS), and other clinical variables such as perceived upper-limb difficulty (AMSQ), depression (PHQ-9), perceived cognitive abilities (MSNQ), impact of MS (MSIS), and perceived fatigue (MFIS) predicted SUS scores. The second analysis was run with workload as measured by NASA-TLX subdomains and Godspeed. Both R^2 and adjusted R^2 were reported to account for the number of predictors. Influential-case diagnostics were conducted for both regression models using Cook distance, with a threshold of 4/n=0.20. Prior to interpretation, model assumptions were verified through visual inspection. Normality of residuals was assessed using P-P plots, and homoscedasticity was examined through residual scatterplots, which revealed no systematic patterns, supporting the assumption of homoscedasticity in both models. Multicollinearity was evaluated using tolerance statistics and variance inflation factor (VIF) values. Given the exploratory nature of this study, no formal corrections for multiple comparisons were applied. While this increases the risk of type I error, applying corrections such as Bonferroni or FDR in exploratory analyses raises the risk of type II error and may obscure potentially meaningful patterns worthy of further investigation in confirmatory studies. Our analytical strategy aimed to capture the widest possible range of behavioral and clinical predictors, allowing for a more inclusive exploration of the factors shaping robot-assisted interaction in this clinical population.

In addition, participants provided short written responses to 3 open-ended questions about their experience interacting with the social robot. These responses were analyzed using a lightweight NLP pipeline aimed at extracting basic lexical and evaluative features. This analysis was conceived as exploratory and complementary to the primary quantitative outcomes, rather than as a stand-alone qualitative investigation.

Text preprocessing and lemmatization were conducted using the Italian language model in the spaCy NLP library [42]. Punctuation, numbers, and Italian stopwords were removed before analysis. Stopword lists were based on standard Italian Snowball resources [43] and were supplemented with additional context-specific terms related to the experimental setting to avoid artificially inflating lexical counts. The additional terms excluded were the following Italian words: *robot* (“robot”), *pepper* (“Pepper”), *esperienza* (“experience”), *interazione* (“interaction”), *attività* (“activity”), *esercizio* (“exercise”), *studio* (“study”), *compito* (“task”), *task* (“task”), *persona* (“person”), and *paziente* (“patient”). Lemmatized tokens were then used to compute basic lexical features, including total word count (NumWords) and the number of unique lemmas. To capture evaluative language, tokens were matched against the Sentix Italian sentiment lexicon [44,45]. For each participant, we calculated the number of tokens appearing in the sentiment lexicon (SentN). Because longer responses naturally increase the likelihood of lexicon matches, we derived an additional measure to estimate sentiment-lexicon coverage independently of response length. Specifically, SentN was regressed on total

XSL•FO
RenderX

word count, and the residuals from this model were extracted. This residualized measure reflects whether a response contains more or fewer sentiment-lexicon matches than would be expected given its length. Responses to the 3 questions were concatenated to obtain a single text sample for each participant. The corpus therefore comprised 120 written responses from 40 participants (3 responses per participant). After preprocessing, 648 lemmatized tokens were retained for analysis, corresponding to a mean of 16.2 tokens per participant (SD 11.9; median 11.5; range 2‐50). Given the exploratory nature of the NLP component, no between-group test was formally prespecified. Both Welch 2-tailed *t* tests and Mann-Whitney *U* tests were conducted as complementary robustness checks. Within the MS group, exploratory associations between linguistic variables and UX measures (SUS and GAToRS) were examined using Kendall tau (τ), which is suitable for small samples and variables with tied ranks. To assess the sensitivity of the association between residualized sentiment-lexicon coverage and SUS to individual observations in the small MS sample, leave-one-participant-out and leave-two-participants-out analyses were additionally performed, reestimating the residualization model at each iteration. No formal multiplicity correction was prespecified for the exploratory NLP analyses. For the 4 reported NLP associations, Benjamini-Hochberg FDR-adjusted *P* values were calculated post hoc as a sensitivity analysis.

Ethical Considerations

The study has been approved by the Regional Ethics Committee (229/2022 - id 12330), and participants provided their written consent to participate after reading information about the study in accordance with the revised Declaration of Helsinki. Data collected were stored in an anonymized format to properly protect the privacy and confidentiality of participants, ensuring that no individual can be identified from the data provided. Participants have been informed that data collection could be used only for research purposes.

Results

Sample Characteristics

Twenty people with MS and 20 HIs were enrolled and completed the study. As indicated by Mann-Whitney *U* tests, the groups did not differ in terms of demographic variables (all *P* values >.05), including sex distribution (χ^2_1 =0.107, *P*=.74, Cramér *V*=0.052; Table 1). Considering psychological variables, MS groups showed higher depressive symptoms compared to HIs (*P*<.001, *r*=0.70). This finding is consistent with evidence showing a higher prevalence of mood disturbances in people with MS compared to the healthy population [46]. No other differences have been observed between groups (all *P* values >.05). See Table 1 for a detailed summary of descriptive analysis of both groups considering demographic and psychological variables and Table 2 for clinical and disease-related information from people with MS.

XSL•FO
RenderX

Table . Demographic and psychological characteristics of people with multiple sclerosis (MS; n=20) and age-matched healthy individuals (HIs; n=20), with between-group comparisons[a].

Variable and group	Values	Median (IQR)	*R* (95% CI)	*P* value
Age (y), mean (SD)				
MS	51.40 (11.02)	52.50 (45.0 - 57.0)	0.20 (0.01-0.49)	.21
HI	47.30 (10.13)	49.00 (40.8 - 54.2)	—[b]	—
Educational level (y), mean (SD)				
MS	13.00 (3.36)	13.00 (11.5 - 15.2)	0.24 (0.02-0.51)	.12
HI	14.75 (2.20)	15.00 (13.0 - 16.2)	—	—
Sex (female), n (%)				
MS	13 (65)	—	0.05[c]	.74
HI	12 (60)	—	—	—
STAI-State[d], mean (SD)				
MS	47.35 (4.97)	47.50 (43.8 - 50.0)	0.02 (0.00-0.34)	.92
HI	47.05 (4.19)	48.00 (44.0 - 50.2)	—	—
STAI-Trait[e], mean (SD)				
MS	45.50 (5.92)	47.00 (43.2 - 50.0)	0.12 (0.01-0.41)	.46
HI	44.70 (4.49)	45.00 (41.0 - 47.2)	—	—
I-TIPI[f] Extraversion, mean (SD)				
MS	9.15 (2.46)	9.50 (7.0 - 11.0)	0.13 (0.01-0.44)	.43
HI	8.65 (1.60)	9.00 (8.0 - 10.0)	—	—
I-TIPI Agreeableness, mean (SD)				
MS	11.30 (1.98)	11.00 (10.0 - 13.0)	0.17 (0.01-0.45)	.29
HI	10.65 (1.81)	10.50 (9.0 - 12.0)	—	—
I-TIPI Conscientiousness, mean (SD)				
MS	10.30 (2.87)	10.00 (8.0 - 13.0)	0.24 (0.02-0.54)	.13
HI	11.75 (1.48)	12.00 (11.0 - 13.0)	—	—
I-TIPI Emotional Stability, mean (SD)				
MS	8.05 (3.80)	8.50 (5.0 - 11.0)	0.31 (0.04-0.58)	.051
HI	10.40 (1.88)	11.00 (9.0 - 12.0)	—	—
I-TIPI Openness, mean (SD)				
MS	9.30 (2.32)	9.00 (8.0 - 11.0)	0.02 (0.00-0.36)	.89
HI	9.45 (2.09)	10.00 (8.0 - 10.2)	—	—
PHQ-9[g], mean (SD)				
MS	9.60 (5.78)	9.00 (5.8 - 12.5)	0.70 (0.53-0.82)	<.001
HI	1.85 (2.23)	1.00 (0.0 - 3.0)	—	—

[a] Between-group differences were assessed using Mann-Whitney *U* tests for continuous variables and X^2 test for sex distribution. Effect sizes are reported as *r* ($Z/\sqrt{N}$) with 95% bootstrap CIs for continuous variables, and as Cramér *V* for sex. *P* value is reported for each comparison.

[b] Values refer to the MS vs HI comparison.

[c] This value indicates Cramér *V*.

[d] STAI-State: State-Trait Anxiety Inventory (State subscale).

[e] STAI-Trait: State-Trait Anxiety Inventory (Trait subscale).

[f] I-TIPI: Italian Ten-Item Personality Inventory.

[g] PHQ-9: 9-item Patient Health Questionnaire.

XSL•FO
RenderX

Table . Disease related and clinical information of people with multiple sclerosis (n=20).

Variable	Mean (SD; SE)
Disease duration (y)	17.65 (11.69; 2.61)
EDSS[a]	3.15 (1.92; 0.43)
DIGICOG-MS (remember and place)[b]	20.70 (4.93; 1.10)
DIGICOG-MS (listen and repeat)[c]	54.65 (9.08; 2.03)
DIGICOG-MS (generate words)[d]	26.05 (6.64; 1.49)
DIGICOG-MS (associate numbers)[e]	31.60 (10.05; 2.25)
AMSQ[f]	53.75 (28.93; 6.47)
MSIS-29[g]	69.10 (28.79; 6.44)
MFIS[h]	36.90 (24.51; 5.48)
MSNQ[i]	25.55 (13.75; 3.08)

[a]EDSS: Expanded Disability Status Scale.
[b]DIGICOG-MS (visuospatial memory test).
[c]DIGICOG-MS (verbal memory test).
[d]IGICOG-MS (semantic fluency test).
[e]DIGICOG-MS (information processing speed test).
[f]AMSQ: Arm Function in Multiple Sclerosis.
[g]MSIS: Multiple Sclerosis Impact Scale.
[h]MFIS: Modified Fatigue Impact Scale.
[i]MSNQ: Multiple Sclerosis Neuropsychological Questionnaire.

Results on UX, Usability, and Human-Robot Interaction

Table 3 shows a detailed summary of results from Mann-Whitney *U* tests. The average SUS score exceeded the standard cut-off of 68 and was comparable in both groups (MS: mean 81.00, SD 12.31; HI: mean 77.75, SE 2.51, r=0.14; P=.39), indicating that participants reported a highly positive experience with Pepper.

XSL•FO
RenderX

Table . Results of statistical analyses comparing people with multiple sclerosis (MS; n=20) and age-matched healthy individuals (HIs; n=20) on technological and human-robot interaction outcomes at baseline[a].

Variable and group	Mean (SD)	Median (IQR)	*R* (95% CI)	*P* value
SUS (%)[b]			0.14 (0.01-0.42)	.39
MS	81.00 (12.31)	81.25 (77.5 - 90.0)		
HI	77.75 (11.24)	80.00 (70.0 - 85.0)		
GAToRS *P*+[c]			0.17 (0.01-0.45)	.29
MS	26.55 (5.36)	27.00 (21.8 - 32.0)		
HI	24.70 (4.77)	25.00 (22.0 - 29.0)		
GAToRS P–[d]			0.09 (0.01-0.42)	.56
MS	11.35 (5.25)	10.50 (7.8 - 13.0)		
HI	10.25 (3.78)	9.00 (8.0 - 12.2)		
GAToRS S+[e]			0.05 (0.00-0.37)	.75
MS	28.25 (5.17)	28.50 (24.8 - 32.5)		
HI	27.50 (4.92)	28.00 (25.0 - 30.0)		
GAToRS S–[f]			0.36 (0.08-0.63)	.02
MS	25.90 (6.95)	26.50 (20.0 - 32.2)		
HI	21.20 (5.55)	20.00 (16.8 - 24.8)		
NASA-TLX[g] Mental Demand			0.17 (0.01-0.47)	.29
MS	5.35 (3.74)	4.12 (2.9 - 5.9)		
HI	5.86 (3.24)	5.88 (4.2 - 6.6)		
NASA Physical Demand			0.03 (0.01-0.35)	.88
MS	3.17 (2.95)	2.25 (1.0 - 3.8)		
HI	2.48 (1.32)	2.12 (1.5 - 3.3)		
NASA Temporal Demand			0.09 (0.01-0.39)	.59
MS	3.30 (3.15)	2.50 (1.4 - 3.9)		
HI	3.16 (1.81)	3.25 (1.4 - 4.6)		
NASA Effort			0.14 (0.01-0.43)	.40
MS	3.83 (3.07)	2.38 (1.6 - 5.2)		
HI	4.17 (2.47)	3.88 (2.4 - 5.6)		
NASA Performance			0.10 (0.01-0.42)	.52
MS	16.31 (4.09)	17.12 (14.9 - 19.4)		
HI	15.70 (3.79)	15.88 (13.0 - 19.0)		
NASA Frustration			0.01 (0.01-0.37)	.96
MS	5.01 (5.70)	2.50 (1.0 - 5.8)		
HI	3.06 (1.80)	3.00 (1.0 - 4.6)		
Godspeed ANTR[h]			0.42 (0.15-0.67)	.007
MS	3.10 (0.79)	3.20 (2.6 - 3.6)		
HI	2.38 (0.79)	2.40 (1.9 - 2.8)		
Godspeed ANIM[i]			0.28 (0.03-0.57)	.08
MS	3.33 (0.57)	3.40 (3.0 - 3.6)		
HI	2.98 (0.82)	2.80 (2.4 - 3.5)		

XSL•FO
RenderX

Variable and group	Mean (SD)	Median (IQR)	*R* (95% CI)	*P* value
Godspeed LIKE[j]			0.28 (0.03-0.55)	.07
MS	4.49 (0.62)	4.80 (4.2 - 5.0)		
HI	4.15 (0.65)	4.10 (4.0 - 4.6)		
Godspeed INT[k]			0.37 (0.09-0.61)	.02
MS	4.16 (0.69)	4.00 (3.6 - 5.0)		
HI	3.49 (0.93)	3.40 (3.0 - 4.2)		
Godspeed SAF[l]			0.05 (0.01-0.36)	.77
MS	2.68 (0.58)	2.50 (2.3 - 3.0)		
HI	2.68 (0.69)	2.50 (2.3 - 3.0)		

[a]Between-group differences were assessed using Mann-Whitney *U* tests. Effect sizes are reported as *r* (Z/√N) with 95% bootstrap CIs. *P* value is reported for each comparison.

[b]SUS: System Usability Scale.

[c]GAToRS *P*+: General Attitudes Towards Robots Scale—Personal Positive subscale.

[d]GAToRS P–: General Attitudes Towards Robots Scale—Personal Negative subscale.

[e]GAToRS S+: General Attitudes Towards Robots Scale—Social Positive subscale.

[f]GAToRS S–: General Attitudes Towards Robots Scale—Social Negative subscale.

[g]NASA-TLX: National Aeronautics and Space Administration Task Load Index.

[h]Godspeed (Anthropomorphism).

[i]Godspeed (Animacy).

[j] Godspeed (Likeability).

[k]Godspeed (Perceived Intelligence).

[l]Godspeed (Perceived Safety).

As shown by the score obtained at GAToRS S– (General Attitudes Towards Robots Scale—Social Negative subscale), people with MS perceived the introduction of robots into daily life at a community level with greater apprehension than HIs, with more focusing on the potential risks that SAR could pose at a societal level (MS: mean 25.90, SD 6.95; HI: mean 21.20, SD 5.5, r=0.36; P=.02).

Considering the mental and physical workload measured by the NASA-TLX, no differences were found between the groups (all P values >.05), suggesting that the proposed tasks were feasible and not excessively demanding. On the Godspeed scale, differences in ANTR and INT subscales were found. People with MS tended to perceive the robot as more anthropomorphic (natural, human-like, realistic, and fluid) than HIs (MS: mean 3.10, SD 0.79; HI: mean 2.38, SD 0.79, r=0.42; P=.007). Additionally, the INT subdomain differed between the groups (MS: mean 4.16, SD 0.69; HI: mean 3.49, SD 0.93, r=0.37; P=.02), suggesting that people with MS considered Pepper to be more competent, well-educated, responsible, and judicious. The results of Spearman correlation analyses between ANTR, INT, and psychological and clinical variables showed that ANTR was strongly associated with visuospatial memory (DIGICOG-MS: ρ=–0.632, 95% CI –0.84 to –0.26; P=.003), relatively strong with both verbal memory (DIGICOG-MS: ρ=–0.504, 95% CI [–0.77 to –0.08]; P=.02), and moderate with Agreeableness (I-TIPI: ρ=0.491, 95% CI 0.06-0.77; P=.03). Besides the correlation between ANTR and INT (ρ=0.645, P=.002, 95% CI 0.29-0.85), no other meaningful associations were observed (Ps>.05). See Table S1 in Multimedia Appendix 1 for the full results of the correlation analyses.

The first exploratory stepwise multiple linear regression on psychological factors was conducted to predict SUS scores. Conscientiousness (I-TIPI), negative personal attitudes toward robots (General Attitudes Towards Robots Scale—Personal Negative subscale [GAToRS P–]), and perceived upper-limb difficulty (AMSQ) explained 64% of the variance in usability (R^2=0.637, adjusted R^2=0.568; $F_{3,16}$=9.342; P=.001). As shown in Table 4, higher conscientiousness was associated with greater perceived usability, whereas lower negative personal attitudes toward robots correlated with higher perceived usability—a finding consistent with prior literature suggesting that people with more open and positive dispositions toward technology tend to rate it as more usable and accessible. In addition, people with MS with higher perceived difficulties in upper-limb function reported greater usability. The second exploratory stepwise multiple linear regression on technological/robotic factors revealed that the final model with task-related frustration from NASA-TLX was retained as the sole predictor. Higher frustration was associated with lower perceived usability (B=−1.497, 95% CI −2.276 to −0.719, SE 0.367, β=−.693; t_{1} =−4.077; P=.001). The model was significant ($F_{1,1}$ =16.620; P=.001) and explained 48% of the variance in SUS scores (R^2=0.480, adjusted R^2=0.451), suggesting that, among people with MS, lower levels of task-related frustration were associated with higher perceived usability of the robot. Visual inspection of P-P plots and residual scatterplots supported the assumptions of normality and homoscedasticity in both models.

XSL•FO
RenderX

Influential-case diagnostics were conducted for both regression models using Cook distance, with a threshold of 4/N=0.20. In the first model, 9 cases exceeded this threshold (Cook *D* range: 0.00 - 0.72, mean 0.126, SD 0.196), and in the second model, 2 cases exceeded the threshold (Cook *D* range: 0.00 - 0.365), confirming that both models should be interpreted with caution given the sensitivity to individual observations inherent in small samples. No multicollinearity was detected (model 1: VIF range 1.005 - 1.022, tolerance range 0.979 - 0.995; Model 2: VIF=1.000, tolerance=1.000). See Figure 3 for the heatmap of Spearman coefficients from regression analyses.

Table . Results of the first exploratory forward stepwise linear regression model predicting System Usability Scale (SUS) in people with multiple sclerosis (MS)[a].

Predictors	B[b] (95% CI)	SE	ß	*t* test (*df*)	*P* value
I-TIPI[c]-Conscientiousness	2.064 (0.679 to 3.449)	0.653	0.481	3.159 (16)	.006
AMSQ[d]	0.169 (0.033 to 0.306)	0.064	0.398	2.631 (16)	.02
GAToRS P–[e]	–0.937 (–1.694 to –0.180)	0.357	–0.400	–2.625 (16)	.02

[a]The final model included Conscientiousness (I-TIPI), negative personal attitudes toward robots (GAToRS P–), and perceived upper-limb difficulty (AMSQ) as predictors. *P* values and effect sizes with 95% CIs are reported for each comparison. R^2=0.637, adjusted R^2=0.568, $F_{3,16}$=9.342; *P*=.001.

[b]Nonstandardized coefficient.

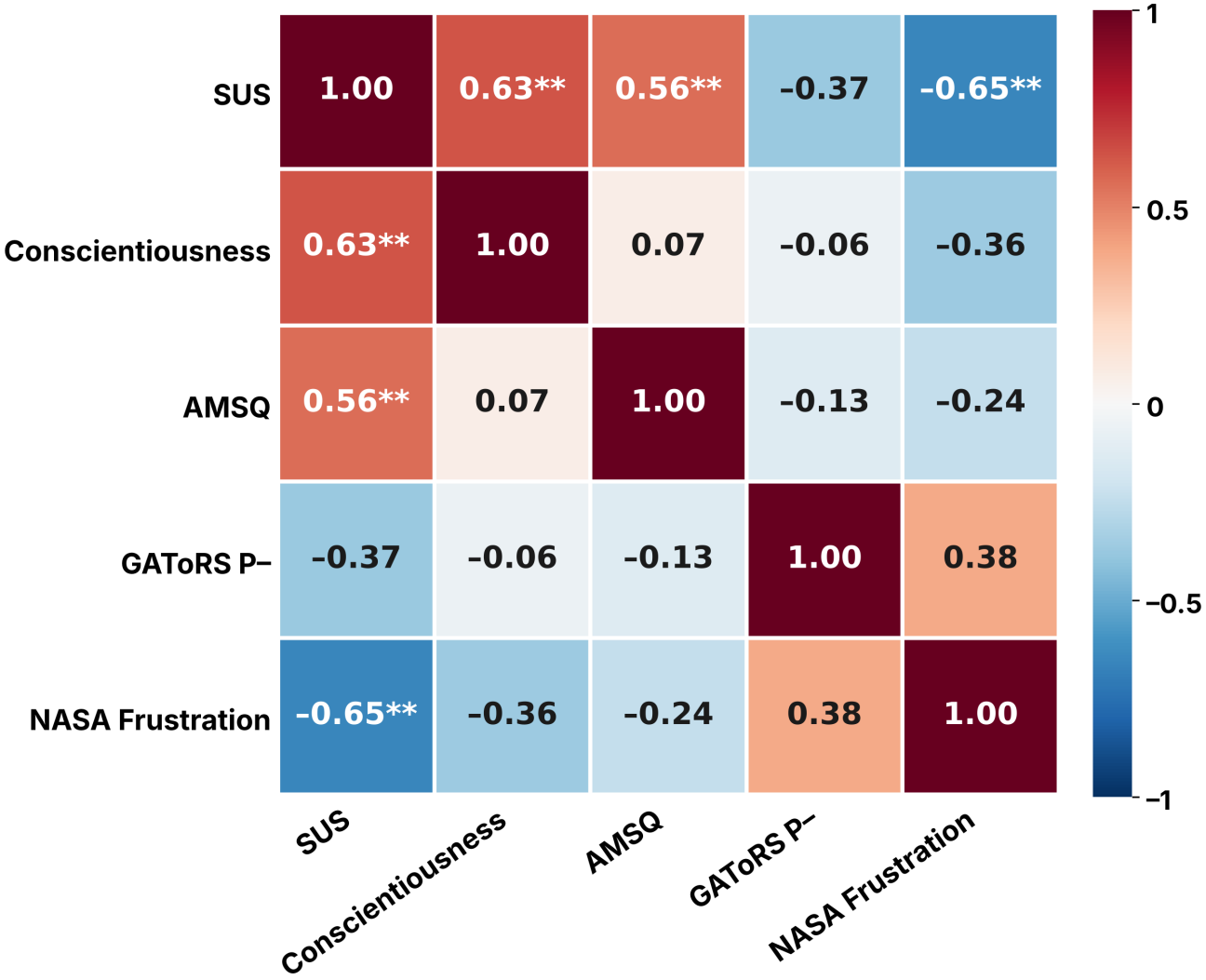
[c]I-TIPI: Italian Ten-Item Personality Inventory.

[d]AMSQ: Arm Function in Multiple Sclerosis.

[e]GAToRS P–: General Attitudes Towards Robots Scale—Personal Negative subscale.

XSL•FO
RenderX

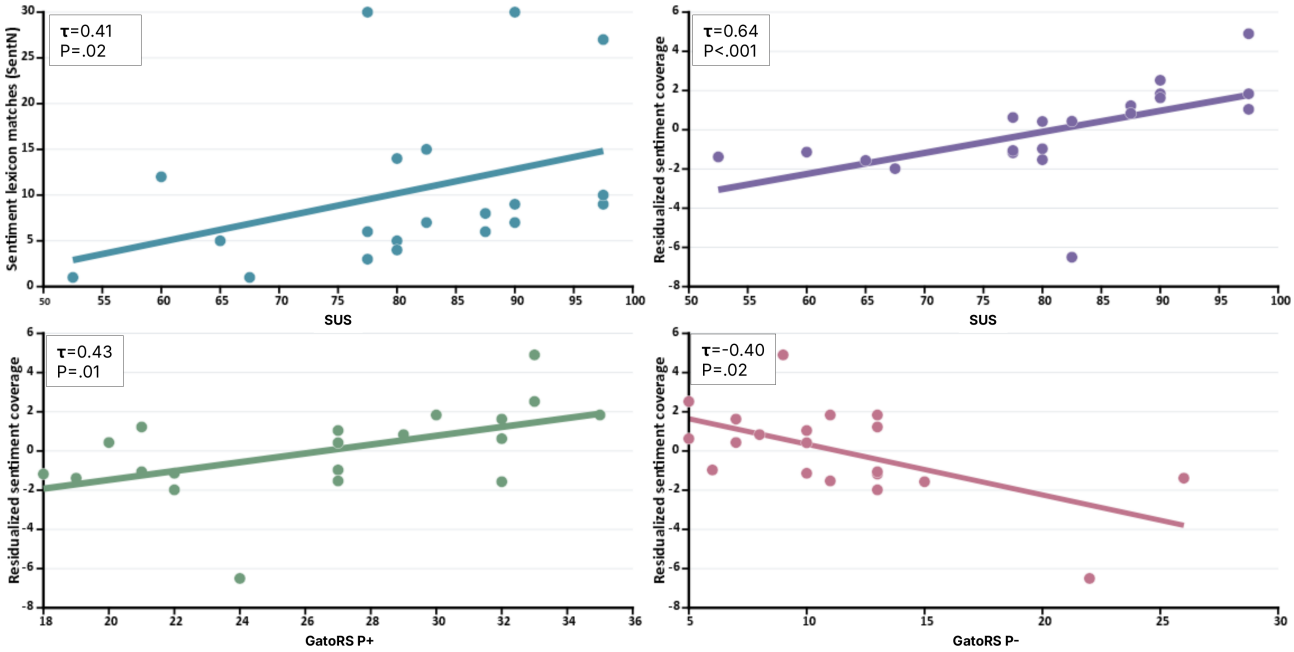
Figure 3. Heatmap showing Spearman correlation coefficients between meaningful predictors and System Usability Scale (SUS) scores from exploratory forward regression analyses. Color intensity represents the strength of the correlation, with darker colors indicating stronger correlations. *P* value is reported for each comparison. Positive correlations (ρ>0) are shown in red, and negative correlations (ρ<0) are shown in blue. **P<.01. AMSQ: Arm Function in Multiple Sclerosis Questionnaire; GAToRS P–: General Attitudes Towards Robots Scale—Personal Negative subscale; NASA: National Aeronautics and Space Administration.



From the exploratory NLP analysis, no meaningful differences emerged between people with MS and HIs in the extracted linguistic features (all *P* values >.05), including lexical productivity and sentiment-lexicon coverage. Within the MS group, exploratory analyses revealed meaningful associations between evaluative lexical content and psychological and user-experience measures. The number of sentiment-lexicon matches (SentN) was positively associated with SUS (τ=0.411; *P*=.02). After controlling for response length, residualized sentiment-lexicon coverage was positively associated with usability (τ=0.642; *P*<.001). This association was stable to case deletion: leave-one-participant-out Kendall's τ values ranged from 0.592 to 0.669, while leave-two-participants-out values ranged from 0.515 to 0.715, indicating that the association was not attributable to any single participant or pair of participants. Residualized sentiment-lexicon coverage was also positively associated with GAToRS P+ (General Attitudes Towards Robots Scale—Personal Positive subscale) (τ=0.431; *P*=.01) and negatively associated with GAToRS P– (τ=–0.403; *P*=.02; Figure 4).

XSL•FO
RenderX

Figure 4. Natural language processing (NLP) associations with attitudes toward robots and perceived usability in people with multiple sclerosis. Sentiment lexicon matches and residualized sentiment coverage are reported on the *y*-axes, whereas System Usability Scale (SUS), GAToRS P+, and GAToRS P– are reported on the *x*-axes. GAToRS P–: General Attitudes Towards Robots Scale—Personal Negative subscale; GAToRS S–: General Attitudes Towards Robots Scale—Societal Negative subscale.



Discussion

Although SAR has shown promising results in supporting everyday life, providing interaction, and facilitating cognitive and physical training in neurological and older populations, its application in MS remains largely unexplored. Our findings bridge this gap, demonstrating that Pepper was perceived as highly usable not only by HIs but also, and notably, by people with MS, a population often characterized by significant motor, cognitive, and psychological challenges. Critically, SUS scores in both groups consistently exceeded the standard cut-off of 68, and the absence of between-group differences indicates that, at least in this pilot study, MS-related challenges may not substantially hinder human-robot interaction. These preliminary findings indicate that Pepper was perceived as usable, even in the presence of MS-related challenges, and support the rationale for further investigation of SAR in MS rehabilitation contexts. These results align with Technology Acceptance Model (TAM) [47] and Unified Theory of Acceptance and Use of Technology (UTAUT) [48], 2 of the most widely used frameworks to understand why people adopt or reject new technologies, in which perceived ease of use and low effort are considered key drivers of acceptance. The high SUS scores and the absence of workload differences observed in both groups suggest that Pepper met these basic requirements, laying a favorable ground for its integration in clinical settings. Despite this high technical evaluation, meaningful differences emerged regarding social integration. People with MS expressed more cautious attitudes toward the introduction of robots in daily life, as measured by GAToRS S–, suggesting that while the robot's functional usability is well accepted, people with MS could have greater concerns about the broader social and practical implications of SAR. This finding is consistent with studies reporting that negative attitudes and avoidance tendencies toward robots may hinder their acceptance across various societal contexts [49] and resonates with the Social Influence construct in UTAUT, which highlights how perceived community norms and collective attitudes can affect individual acceptance of new technologies. Addressing these broader social narratives around robots in health care will therefore be essential to support long-term adoption [44]. To favor a smooth introduction of SAR into human lives, it is essential to explore effective strategies targeting the reduction in negative attitudes toward robots or the fear of potential harm from their extensive use, highlighting the importance of addressing trust, transparency, and perceived risks when introducing such technologies in clinical settings.

The absence of between-groups differences in mental and physical workload, as measured by NASA-TLX, indicates that the tasks were generally manageable for both people with MS and healthy controls. This aligns with TAM and UTAUT constructs, especially with the principle that good usability relies on systems that are intuitive and efficient, minimizing the learning curve and reducing potential user frustration [50]. Surprisingly, people with MS tended to perceive the robot as more human-like (anthropomorphic) and intelligent compared to HIs, as reflected in the ANTR and INT subscales of the Godspeed questionnaire, respectively. ANTR was further associated with cognitive abilities and personality traits. Participants with lower visuospatial and verbal memory scores perceived the robot as more human-like. This pattern could be interpreted in 2 ways. First, as posited by Epley [51], anthropomorphism, as the process of attributing human-like qualities and characteristics to nonhuman agents and objects [52], may function as a cognitive compensatory strategy for individuals facing reduced processing resources [51]. Under this framework, individuals with reduced cognitive capacities may rely more heavily on heuristic social attributions, perceiving

XSL•FO
RenderX

the robot as more human-like to simplify the understanding of its behavior. Second, the positive correlation between ANTR and Agreeableness from I-TIPI suggests a dispositional influence. Consistent with Epley et al [53], individual differences in anthropomorphizing nonhuman agents are often rooted in stable prosocial traits and sociality motivation—the fundamental human drive for interpersonal connection. Consequently, those who self-identify as kind, sympathetic, and cooperative may be more naturally inclined to extend these social categories to robotic partners. The finding that people with MS attributed higher levels of anthropomorphism and perceived intelligence to Pepper compared to HIs may be interpreted considering the heightened social needs commonly reported in this population. People with MS experience higher rates of social isolation and loneliness than healthy adults [16]. Building on social identity theory and research on human-robot interaction, it may be speculated that for individuals who perceive themselves as belonging to a socially marginalized group, such as people living with a chronic condition like MS, anthropomorphism may not merely humanize the robot, but symbolically reposition it as an ingroup partner, potentially attenuating intergroup boundaries between human and artificial agents [49,50]. While speculative, this interpretation is consistent with evidence that social needs and group membership influence the attribution of human-like qualities to nonhuman agents [53]. Understanding these psychological factors is crucial for tailoring interventions, improving user engagement, and optimizing the effectiveness of robot-assisted in MS.

Exploratory multiple linear regression analyses revealed that Conscientiousness (I-TIPI), personal negative attitudes toward robots (GAToRS P–), and upper-limb difficulties (AMSQ) were associated with perceived usability. Specifically, higher levels of Conscientiousness and less negative personal predispositions toward robots were associated with enhanced usability ratings. These findings also resonate with TAM and UTAUT constructs. Conscientiousness, which reflects a tendency toward goal-directed and organized behavior, may foster a more positive performance expectancy, another key driver of technology acceptance according to UTAUT. Similarly, negative personal attitudes toward robots closely mirror the concept of Behavioral Intention in TAM—that is, the degree to which an individual is predisposed toward a technology influences how they ultimately experience it. Furthermore, increased perceived difficulty in upper-limb function positively predicted usability. This suggests that individuals facing greater motor limitations may perceive assistive technologies as more indispensable, leading to a more favorable evaluation of the social robot's use. This finding aligns with prior research indicating that participants with significant challenges in activities of daily living express a stronger preference for robotic support [54]. Finally, lower task-related frustration, as measured by the NASA-TLX, was associated with higher usability, suggesting that people with MS who experienced the interaction as less frustrating tended to evaluate the robot more favorably.

Overall, these findings highlight that human-robot interaction in MS is not uniform; psychological, prior attitudes, and functional abilities which jointly influence how people with MS perceive and interact with social robots should be considered when developing personalized rehabilitation approaches. Understanding these factors is crucial for tailoring interventions, improving user engagement, and optimizing the effectiveness of robot-assisted rehabilitation in MS. Exploratory NLP analyses revealed no between-group differences in the lexical measures examined. Within the MS group, however, associations emerged between evaluative lexical content and UX measures: participants whose written responses contained a higher density of sentiment-lexicon matches relative to response length tended to report higher usability scores and more favorable attitudes toward robots. These findings suggest that free-text language may provide information complementary to structured UX measures, although this interpretation should remain cautious given the exploratory nature of the analysis, the small sample, and the brevity of the written responses. In particular, sentiment-lexicon coverage reflects evaluative lexical density rather than positive emotional valence and should not be interpreted as evidence of a positive emotional tone. Future work using longer spoken or written responses may help determine whether linguistic signals can provide reliable complementary indicators of user experience in clinical populations. Such studies could also investigate whether richer language measures may contribute to the development of more responsive and user-centered human-robot interactions and may help clarify the role of linguistic signals as complementary indicators of UX in clinical populations. The results from NLP are further in line with key recommendations in the literature about the importance of implementing empathy-driven design in SAR, particularly ensuring that social robots can recognize and respond to human emotions through language and behavior, as discussed by Park and Whang [55]. SAR must move beyond basic functionality to provide meaningful and personalized interactions. This is confirmed by studies in autism spectrum disorders, dementia, and psychiatric disorders that emphasized that the effectiveness of SAR is linked to its ability to adapt to users' specific needs and behavioral cues.

To our knowledge, this study represents the first preliminary exploration into the usability of social robots specifically within the MS population. By identifying the multifaceted predictors of perceived usability, spanning cognitive, personality, and clinical domains, this work marks a foundational step in understanding how these individuals perceive and evaluate SAR. These insights are instrumental for the design of tailored, ethically grounded, and user-accepted interventions, ensuring that robot-assisted rehabilitation is optimized for the unique needs of those living with MS. SAR can offer real-time feedback, maintain patient engagement, and reduce the need for continuous human supervision, thereby improving the scalability and efficiency of rehabilitation programs [13,56]. In principle, given the observed high appreciation toward robots, SAR could be strategically used in MS care as motivational agents, real-time feedback providers, or digital companions that help sustain adherence and engagement. From a rehabilitation perspective, these preliminary findings provide an initial indication that social robots may warrant further exploration as complementary tools in MS rehabilitation, though firm conclusions about their role and implementation require replication in larger and more representative samples.

XSL•FO
RenderX

Addressing concerns identified by GAToRS, including perceived social risks and apprehension toward robotic integration, will be essential to foster acceptance and long-term adoption of social robots in rehabilitation contexts.

Several limitations should be acknowledged. In the absence of prior studies about the use of SAR in the MS population, no formal sample size calculation was available to conduct a formal a priori power analysis. The limited sample size, while consistent with pilot SAR studies in neurological and older populations [39,40], may reduce the statistical power to detect small effects, limiting the generalizability of the findings and precluding causal interpretations. Future studies should replicate these findings in larger samples to confirm and extend the present results. Additionally, the study focused on a limited set of tasks, and the robot did not capture other potentially informative measures of human interaction, such as gaze, number of touches, or facial expressions, due to privacy considerations. This issue is in line with Song and Luximon [56] that discussed that ethical considerations around privacy and consent in the context of human-robot interaction should be carefully managed to ensure these technologies are used responsibly and effectively. As SAR becomes more advanced, understanding how they can be seamlessly integrated into clinical workflows is crucial.

The NLP approach was selected to provide a transparent and replicable method for extracting evaluative features from a small corpus, with the explicit aim of capturing dimensions of UX not fully covered by structured questionnaires. Given the limited number of participants and the brevity of the written answers, a full qualitative content analysis was beyond the scope of the present work. Moreover, NLP results may have been influenced by response modalities: recording oral responses could have facilitated participants' performance while reducing fatigue, particularly among people with MS. The absence of between-group differences may also reflect the limited sensitivity of lexical metrics when applied to very short texts. Future studies with larger samples and richer verbal or written data should consider complementing NLP approaches with systematic qualitative or content analysis methods to provide a more comprehensive picture of participants' subjective experience with social robots. In this study, verbal interaction between Pepper and participants was limited and restricted to instructions. However, as indicated by [57], even minimal speech is generally well received, suggesting that the mere presence of verbal communication can enhance engagement.

Acknowledgments

The authors would like to thank Maria Madera for her help in participants' enrollment and Simone Boccardo for his support in data acquisition. Finally, special thanks go to all the people with multiple sclerosis that took part in the study.

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) and Claude (Anthropic) to ensure linguistic accuracy and readability. Figure 2 was generated with the assistance of Gemini (Google). After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the manuscript.

Funding

This study was supported by Fondazione Italiana Sclerosi Multipla (FISM) (project name: DIGICOG-PLUS, number: FISM 2025/S/5) and by the European Union – NextGenerationEU, the Ministry of University and Research (MUR), and National Recovery and Resilience Plan (NRRP) (project name: Robotics and AI for Socio-economic Empowerment - RAISE; number: ECS00000035).

Data Availability

All data produced in this study are available upon reasonable request to the corresponding author.

Authors' Contributions

Conceptualization: JP, AT
Data curation: AS, EC
Formal analysis: JP, AB
Funding acquisition: MAB, GB
Investigation: AS, EC
Methodology: JP, EG, DP, KTK, AT
Project administration: JP
Software: DP, KTK
Supervision: JP, GB
Visualization: JP, AB
Writing – original draft: JP, AB
Writing – review and editing: JP

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Spearman correlation between Anthropomorphism and Perceived Intelligence (Godspeed subscales) and psychological variables in the multiple sclerosis group.
[DOCX File, 20 KB - rehab_v13i1e94068_app1.docx]

References

1. Feil-Seifer D, Mataric M. Socially assistive robotics. IEEE Robot Automat Mag 2011;18(1):24-31. [doi: 10.1109/MRA.2010.940150]
2. Shourmasti ES, Colomo-Palacios R, Holone H, Demi S. User experience in social robots. Sensors (Basel) 2021 Jul 26;21(15):5052. [doi: 10.3390/s21155052] [Medline: 34372289]
3. Jeong S, Aymerich-Franch L, Alghowinem S, Picard RW, Breazeal CL, Park HW. A robotic companion for psychological well-being: a long-term investigation of companionship and therapeutic alliance. ACM IEEE Int Conf Hum Robot Interact 2023 Mar;2023:484-495. [doi: 10.1145/3568162.3578625] [Medline: 38751573]
4. Belpaeme T, Kennedy J, Ramachandran A, Scassellati B, Tanaka F. Social robots for education: a review. Sci Robot 2018 Aug 15;3(21):eaat5954. [doi: 10.1126/scirobotics.aat5954] [Medline: 33141719]
5. Chevalier P, Isableu B, Martin JC, Tapus A. Individuals with autism: analysis of the first interaction with NAO robot based on their proprioceptive and kinematic profiles. In: Borangiu T, editor. Advances in Robot Design and Intelligent Control Proceedings of the 24th International Conference on Robotics in Alpe-Adria-Danube Region (RAAD): Springer; 2016:225-233. [doi: 10.1007/978-3-319-21290-6_23]
6. Diehl JJ, Schmitt LM, Villano M, Crowell CR. The clinical use of robots for individuals with autism spectrum disorders: a critical review. Res Autism Spectr Disord 2012 Jan;6(1):249-262. [doi: 10.1016/j.rasd.2011.05.006] [Medline: 22125579]
7. Alhmiedat T, Alotaibi M. Employing social robots for managing diabetes among children: SARA. Wireless Pers Commun 2023 May;130(1):449-468. [doi: 10.1007/s11277-023-10293-8]
8. Alhmiedat T, AlBishi LA, Alnajjar F, Alotaibi M, Marei AM, Shalayl R. Improving diabetes education and metabolic control in children using social robots: a randomized trial. Technologies 2024;12(11):209. [doi: 10.3390/technologies12110209]
9. Pu L, Moyle W, Jones C, Todorovic M. The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies. Gerontologist 2019 Jan 9;59(1):e37-e51. [doi: 10.1093/geront/gny046] [Medline: 29897445]
10. Andtfolk M, Nyholm L, Eide H, Fagerström L. Humanoid robots in the care of older persons: a scoping review. Assist Technol 2022 Sep 3;34(5):518-526. [doi: 10.1080/10400435.2021.1880493] [Medline: 33481675]
11. Yuan F, Anderson JG, Wyatt TH, et al. Assessing the acceptability of a humanoid robot for Alzheimer's disease and related dementia care using an online survey. Int J of Soc Robotics 2022 Jul;14(5):1223-1237. [doi: 10.1007/s12369-021-00862-x]
12. Khan A, Wahid SR, Sattar J, Khashayar P. Reactions and engagement of individuals with dementia toward humanoid assistive robots: a study using the Pepper robot. Stud Health Technol Inform 2025 May 12;326:65-69. [doi: 10.3233/SHTI250239] [Medline: 40357603]
13. Carnevale A, Raso A, Antonacci C, et al. Exploring the impact of socially assistive robots in rehabilitation scenarios. Bioengineering (Basel) 2025 Feb 19;12(2):204. [doi: 10.3390/bioengineering12020204] [Medline: 40001723]
14. Yuan F, Klavon E, Liu Z, Lopez RP, Zhao X. A systematic review of robotic rehabilitation for cognitive training. Front Robot AI 2021;8(May):605715. [doi: 10.3389/frobt.2021.605715] [Medline: 34046433]
15. Li J. The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum Comput Stud 2015 May;77:23-37. [doi: 10.1016/j.ijhcs.2015.01.001]
16. Balto JM, Pilutti LA, Motl RW. Loneliness in multiple sclerosis: possible antecedents and correlates. Rehabil Nurs 2019;44(1):52-59. [doi: 10.1097/rnj.0000000000000128] [Medline: 30601432]
17. Khatooni M, Dehghankar L, Bahrami M, Panahi R, Hajnasiri H. Perceived social isolation among patients with multiple sclerosis suffering from disease-induced pain and disability. J Nurs Res 2024 Aug 1;32(4):e341. [doi: 10.1097/jnr.0000000000000624] [Medline: 39037382]
18. Montalban X, Lebrun-Frénay C, Oh J, et al. Diagnosis of multiple sclerosis: 2024 revisions of the McDonald criteria. Lancet Neurol 2025 Oct;24(10):850-865. [doi: 10.1016/S1474-4422(25)00270-4] [Medline: 40975101]
19. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an Expanded Disability Status Scale (EDSS). Neurology 1983 Nov;33(11):1444-1452. [doi: 10.1212/wnl.33.11.1444] [Medline: 6685237]
20. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]

XSL•FO
RenderX

21. Kearney KT, Presenza D, Sacca F, Wright P. Key challenges for developing a socially assistive robotic (SAR) solution for the health sector. IEEE Int Workshop Comput Aided Model Des Commun Links Netw 2018:1-7. [doi: 10.1109/CAMAD.2018.8515005]
22. Feil-Seifer D, Mataric MJ. Defining socially assistive robotics. In: 9th International Conference on Rehabilitation Robotics (ICORR 2005): IEEE; 2005:465-468. [doi: 10.1109/ICORR.2005.1501143]
23. Tleubayev B, Zhexenova Z, Zhakenova A, Sandygulova A. Robot-assisted therapy for children with ADHD and ASD. Proc 2019 2nd Int Conf Serv Robot Technol 2019 Mar 22:58-62. [doi: 10.1145/3325693.3325703]
24. Podda J, Tacchino A, Ponzio M, et al. Mobile health app (DIGICOG-MS) for self-assessment of cognitive impairment in people with multiple sclerosis: instrument validation and usability study. JMIR Form Res 2024 Jun 20;8:e56074. [doi: 10.2196/56074] [Medline: 38900535]
25. Tacchino A, Ponzio M, Pedullà L, et al. Italian validation of the Arm Function in Multiple Sclerosis Questionnaire (AMSQ). Neurol Sci 2020 Nov;41(11):3273-3281. [doi: 10.1007/s10072-020-04363-z] [Medline: 32394274]
26. Kos D, Kerckhofs E, Carrea I, Verza R, Ramos M, Jansa J. Evaluation of the modified fatigue impact scale in four different European countries. Mult Scler 2005 Feb;11(1):76-80. [doi: 10.1191/1352458505ms1117oa] [Medline: 15732270]
27. Piscitelli D, Brichetto G, Geri T, et al. Italian adaptation and psychometric validation of the Fatigue Impact Scale (FIS) and its modified versions in adults with multiple sclerosis: a Rasch analysis study. Disabil Rehabil 2024 Nov;46(22):5366-5379. [doi: 10.1080/09638288.2024.2302878] [Medline: 38236054]
28. Benedict RHB, Munschauer F, Linn R, et al. Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. Mult Scler 2003 Feb;9(1):95-101. [doi: 10.1191/1352458503ms861oa] [Medline: 12617275]
29. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain 2001 May;124(Pt 5):962-973. [doi: 10.1093/brain/124.5.962] [Medline: 11335698]
30. Santangelo G, Sacco R, Siciliano M, et al. Anxiety in multiple sclerosis: psychometric properties of the State-Trait Anxiety Inventory. Acta Neurol Scand 2016 Dec;134(6):458-466. [doi: 10.1111/ane.12564] [Medline: 27219913]
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
32. Mazzotti E, Fassone G, Picardi A, et al. Il Patient Health Questionnaire (PHQ) per lo screening dei disturbi psichiatrici: uno studio di validazione nei confronti della Intervista Clinica Strutturata per il DSM-IV asse I (SCID-I) [Article in Italian]. G Ital Psicopatol 2003;9(3):235-242 [FREE Full text]
33. Chiorri C, Bracco F, Piccinno T, Modafferi C, Battini V. Psychometric properties of a revised version of the Ten Item Personality Inventory. Eur J Psychol Assess 2015 May;31(2):109-119. [doi: 10.1027/1015-5759/a000215]
34. Carradore M, Artioli G, Sarli A. The general attitudes towards robots scale (GAToRS): a preliminary validation of the Italian version. Int J of Soc Robotics 2024 Oct;16(9-10):2001-2018. [doi: 10.1007/s12369-024-01170-w]
35. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194. [doi: 10.1201/9781498710411-35]
36. Bartneck C. Godspeed questionnaire series: translations and usage. In: Krägeloh CU, Alyami M, Medvedev ON, editors. International Handbook of Behavioral Health Assessment: Springer; 2023:1-35. [doi: 10.1007/978-3-030-89738-3_24-1]
37. Stefania O. HRI: l’interazione tra esseri umani e macchine. Dall’interazione sociale all’interazione sociotecnica [Master’s thesis in Italian]. : University of Genoa; 2021 URL: https://hdl.handle.net/20.500.14242/70811 [accessed 2026-10-03]
38. Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Human Mental Workload: North-Holland; 1988:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
39. Raz D, Barkan-Slater S, Baum-Cohen I, et al. A novel socially assistive robotic platform for cognitive-motor exercises for individuals with Parkinson’s disease: a participatory-design study from conception to feasibility testing with end users. Front Robot AI 2023;10:1267458. [doi: 10.3389/frobt.2023.1267458] [Medline: 37868274]
40. Feingold-Polak R, Elishay A, Shahar Y, Stein M, Edan Y, Levy-Tzedek S. Differences between young and old users when interacting with a humanoid robot: a qualitative usability study. Paladyn 2018 Jul 1;9(1):183-192. [doi: 10.1515/pjbr-2018-0013]
41. Rea LM, Parker RA. Designing and Conducting Survey Research: A Comprehensive Guide, 4th edition: John Wiley & Sons; 2014.
42. Honnibal M, Montani I. SpaCy 2: natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. Open Research Knowledge Graph (ORKG). 2017. URL: https://orkg.org/resources/R11007?noRedirect= [accessed 2026-09-28]
43. Viriyayudhakorn K, Kunifuji S. Automatic topic identification for idea summarization in idea visualization programs. IEICE Trans Inf Syst 2013;E96-D(1):64-72. [doi: 10.1587/transinf.E96.D.64]
44. Baccianella S, Esuli A, Sebastiani F. SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. Proc Seventh Int Conf Lang Resour Eval 2010. [doi: 10.63317/2jktnrykui9s]
45. Basile V, Nissim M. Sentiment analysis on Italian tweets. In: Balahur A, van der Goot E, Montoyo A, editors. Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis: Association for Computational Linguistics; 2013:100-107 URL: https://aclanthology.org/W13-1614/ [accessed 2026-09-28]

XSL•FO
RenderX

46. Sparaco M, Lavorgna L, Bonavita S. Psychiatric disorders in multiple sclerosis. J Neurol 2021 Jan;268(1):45-60. [doi: 10.1007/s00415-019-09426-6] [Medline: 31197511]
47. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
48. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012 Mar 1;36(1):157-178. [doi: 10.2307/41410412]
49. Bartneck C, Nomura T, Kanda T, Suzuki T, Kato K. Cultural differences in attitudes towards robots. In: Proceedings of the AISB Symposium on Robot Companions: Hard Problems And Open Challenges In Human-Robot Interaction: Society for the Study of Artificial Intelligence and the Simulation of Behaviour (SSAISB); 2005:1-4 URL: https://www.bartneck.de/publications/2005/cultureNars/bartneckAISB2005.pdf [accessed 2026-09-28]
50. Schraick LM, Ehrlich-Sommer F, Stampfer K, Meixner O, Holzinger A. Usability in human-robot collaborative workspaces. Univ Access Inf Soc 2025 Jun;24(2):1609-1622. [doi: 10.1007/s10209-024-01163-6]
51. Epley N. Mindwise: How We Understand What Others Think, Believe, Feel, and Want: Penguin Books Limited; 2014.
52. Złotowski J, Strasser E, Bartneck C. Dimensions of anthropomorphism: from humanness to humanlikeness. In: HRI '14: Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction: Association for Computing Machinery (ACM); 2014:66-73. [doi: 10.1145/2559636.2559679]
53. Epley N, Waytz A, Akalis S, Cacioppo JT. When we need a human: motivational determinants of anthropomorphism. Soc Cogn 2008 Apr;26(2):143-155. [doi: 10.1521/soco.2008.26.2.143]
54. Sørensen L, Johannesen DT, Melkas H, Johnsen HM. Care-receivers with physical disabilities' perceptions on having humanoid assistive robots as assistants: a qualitative study. BMC Health Serv Res 2024 Apr 25;24(1):523. [doi: 10.1186/s12913-024-10857-9] [Medline: 38664810]
55. Park S, Whang M. Empathy in human-robot interaction: designing for social robots. Int J Environ Res Public Health 2022 Feb 8;19(3):1889. [doi: 10.3390/ijerph19031889] [Medline: 35162909]
56. Song Y, Luximon Y. Trust in AI agent: a systematic review of facial anthropomorphic trustworthiness for social robot design. Sensors (Basel) 2020 Sep 7;20(18):5087. [doi: 10.3390/s20185087] [Medline: 32906760]
57. Bradwell HL, Winnington R, Thill S, Jones RB. Longitudinal diary data: six months real-world implementation of affordable companion robots for older people in supported living. In: HRI '20: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction: Association for Computing Machinery (ACM); 2020:148-150. [doi: 10.1145/3371382.3378256]

Abbreviations

AISM: Italian Multiple Sclerosis Society
AMSQ: Arm Function in Multiple Sclerosis Questionnaire
ANTR: Anthropomorphism (Godspeed subscale)
EDSS: Expanded Disability Status Scale
GAToRS: General Attitudes Towards Robots Scale
GAToRS P+: General Attitudes Towards Robots Scale—Personal Positive subscale
GAToRS P–: General Attitudes Towards Robots Scale—Personal Negative subscale
GAToRS S–: General Attitudes Towards Robots Scale—Societal Negative subscale
HI: healthy individual
I-TIPI: Italian Ten Item Personality Inventory
INT: Perceived Intelligence (Godspeed subscale)
MFIS: Modified Fatigue Impact Scale
MS: multiple sclerosis
MSIS-29: 29-item Multiple Sclerosis Impact Scale
MSNQ: Multiple Sclerosis Neuropsychological Questionnaire
NASA-TLX: National Aeronautics and Space Administration Task Load Index
NLP: natural language processing
PHQ-9: 9-item Patient Health Questionnaire
PRO: patient-reported outcome
SAR: socially assistive robotics
SARA: Socially Assistive Robotic Solution for Ambient
STAI: State-Trait Anxiety Inventory
SUS: System Usability Scale
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology
UX: user experience
VIF: variance inflation factor

XSL•FO
RenderX

Edited by S Munce; submitted 23.Mar.2026; peer-reviewed by CY Yang, DS Mcconnell, T Alhmiedat; revised version received 28.Aug.2026; accepted 14.Sep.2026; published 21.Oct.2026.

Please cite as:
Podda J, Bollini A, Susini A, Presenza D, Kearney KT, Cecchella E, Grange E, Battaglia MA, Tacchino A, Brichetto G
Usability and Acceptance of a Social Robot in People With Multiple Sclerosis: Comparative Cross-Sectional Study
JMIR Rehabil Assist Technol 2026;13:e94068
URL: https://rehab.jmir.org/2026/1/e94068
doi:10.2196/94068

© Jessica Podda, Alice Bollini, Alessia Susini, Domenico Presenza, Keven T Kearney, Ester Cecchella, Erica Grange, Mario Alberto Battaglia, Andrea Tacchino, Giampaolo Brichetto. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 21.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Using a Lumbar-Type Hybrid Assistive Limb by Older Adults With Frailty on Exercise Self-Efficacy: Retrospective Analysis Based on the Kanagawa Prospective ME-BYO Cohort Study

Makoto Nagasawa[1,2,3], PhD; Sho Nakamura[1,2], MD, PhD; Yoshinobu Saito[4,5,6], PhD; Yuko Oguma[7,8], MD, PhD; Takashi Kasukawa[9]; Kaname Watanabe[1,3], MD, PhD; Hiroto Narimatsu[1,2,3,6], PhD, MD

[1]Cancer Prevention and Control Division, Kanagawa Cancer Center Research Institute, Yokohama, Kanagawa, Japan

[2]Graduate School of Health Innovation, Kanagawa University of Human Services, Research Gate Building Tonomachi 2-A, 3-25-10 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, Japan

[3]Department of Genetic Medicine, Kanagawa Cancer Center, Yokohama, Kanagawa, Japan

[4]Faculty of Sport Management, Nippon Sport Science University, Kanagawa, Japan

[5]Graduate School of Physical Education, Health and Sport Studies, Nippon Sport Science University, Tokyo, Japan

[6]Center for Innovation Policy, Kanagawa University of Human Services, Kanagawa, Japan

[7]Sports Medicine Research Center, Keio University, Kanagawa, Japan

[8]Graduate School of Health Management, Keio University, Kanagawa, Japan

[9]Shonan Robocare Center, Kanagawa, Japan

Corresponding Author:
Hiroto Narimatsu, PhD, MD
Cancer Prevention and Control Division, Kanagawa Cancer Center Research Institute, Yokohama, Kanagawa, Japan

Abstract

Background: Improving exercise self-efficacy (ESE) is a key strategy for promoting sustained physical activity among older adults with frailty. The lumbar-type Hybrid Assistive Limb (HAL), a wearable device that supports voluntary movement through bioelectric signals, may facilitate mastery experiences in individuals with low baseline self-efficacy via adjustable assistance. However, its effects on ESE remain unclear.

Objective: This study evaluated the effects of lumbar-type HAL exercise on ESE in older adults with frailty and examined whether responsiveness differed according to baseline readiness for exercise.

Methods: A retrospective pooled analysis was conducted using data from 2 prospective studies of a lumbar-type HAL–based frailty prevention program: a feasibility study (n=20) and a randomized controlled trial (RCT; n=76; lumbar-type HAL group: n=38; waitlist control: n=38). Participants aged 50 to 84 years performed sit-to-stand and squat exercises using the lumbar-type HAL twice weekly. ESE was assessed using a validated 4-item questionnaire (range 4‐20). The prespecified primary analysis examined within-person changes in ESE in the lumbar-type HAL group (2-tailed paired *t* test; n=96). A secondary analysis compared changes in ESE between the lumbar-type HAL and waitlist control groups in the RCT using the Welch *t* test and baseline-adjusted analysis of covariance. A prespecified subgroup analysis compared participants in stages 1 to 4 of the transtheoretical model of behavior change (exercise habit not established) with those in stage 5 (maintenance). All statistical tests were 2 sided, with significance at $P<.05$.

Results: The primary pooled paired analysis demonstrated a significant increase in ESE of 1.05 points (95% CI 0.41-1.69; $P=.002$). The subgroup analysis was also significant: those in stages 1 to 4 (n=56) improved by 1.73 points vs 0.05 points for those in stage 5 (n=38), a difference of 1.68 points (95% CI 0.41-2.95; $P=.01$). The largest within-group improvement was observed in stage 1 (precontemplation: +3.50 points, 95% CI 0.27-6.73; $P=.04$; Cohen $d=1.14$); 4 of 6 participants improved and none deteriorated. In the secondary RCT between-group comparison, the lumbar-type HAL group showed a directionally favorable improvement (between-group difference=1.21 points, 95% CI −0.33 to 2.75; $P=.12$; Cohen $d=0.36$). The subgroup differences were 2.10 points in stages 1 to 4 ($P=.07$; Cohen $d=0.58$) and 0.02 points in stage 5 ($P=.99$). Sensitivity analyses confirmed robustness to source-related heterogeneity (heterogeneity test: $P=.74$; source-adjusted regression: $P=.76$).

Conclusions: The pooled analysis demonstrated significant within-person improvement in ESE among older adults with frailty following lumbar-type HAL training, with the greatest improvement in individuals without established exercise habits. The RCT indicated a favorable but inconclusive between-group effect, and adequately powered confirmatory RCTs remain necessary. The

XSL•FO
RenderX

exploratory response among precontemplators—typically difficult to engage in interventions—suggests that HAL-supported mastery experiences may support disengaged older adults at high risk of frailty in starting physical exercise.

Trial Registration: University Hospital Medical Information Network Clinical Trials Registry UMIN000037911; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000043213 and UMIN000042352; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000048345

(JMIR Rehabil Assist Technol 2026;13:e79358) doi:10.2196/79358

KEYWORDS

exercise self-efficacy; frailty; Hybrid Assistive Limb; lumbar type; physical activity; transtheoretical model

Introduction

Frailty and Exercise Self-Efficacy in Older Adults

Frailty is a common clinical syndrome in older adults and is associated with increased risk of hospitalization, falls, and disability, highlighting the need for targeted interventions [1,2]. Exercise self-efficacy (ESE), a psychological construct linked to frailty [3], reflects confidence in maintaining exercise despite barriers [4]. Lower ESE is associated with reduced physical activity (PA) [3], supporting its relevance as an outcome in exercise interventions. According to the social cognitive theory by Bandura [5], self-efficacy is shaped by mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states, with mastery experiences being most influential [5,6]. Mastery experiences involve successful performance in challenging tasks. In rehabilitation, task difficulty is adjusted to enable success, supporting motor learning and improving efficiency [7]. Thus, appropriate task adjustment may facilitate mastery experiences and enhance ESE.

Challenge of Reaching Disengaged Older Adults

Behavioral readiness for exercise, as defined by the transtheoretical model (TTM) of behavior change, is a key determinant of intervention responsiveness [8]. Older adults in precontemplation (stage 1)—who have no intention of initiating exercise—are typically difficult to engage in conventional exercise programs and rarely accumulate mastery experiences. Although exercise interventions can improve physical function and frailty status in older adults [9], systematic reviews of TTM-based interventions report limited and inconsistent evidence for promoting PA in this population [10]. Therefore, how intervention responsiveness varies by baseline behavioral readiness remains poorly understood.

Lumbar-Type Hybrid Assistive Limb and Mastery Experience

The Hybrid Assistive Limb (HAL; Cyberdyne Inc) is a wearable cyborg that supports movement in response to voluntary intent [11,12]. It uses skin-mounted sensors to detect bioelectric signals associated with movement intention and converts these signals into motor torque assistance. Unlike conventional exoskeletons that rely primarily on preprogrammed motion patterns, the HAL generates assistance directly in response to the wearer's voluntary activity [13]. Clinical studies have shown that a lower limb–type HAL improves walking ability in patients with stroke and spinal cord injury [13-17]. The lumbar-type HAL consists of an exoskeleton frame and a power unit comprising angle sensors and actuators for both hip joints. It assists hip extension during transitions from flexion to extension, supporting standing and lifting movements. Compared with the lower limb type, it is lighter and more compact, enabling its use in settings such as home-based training and group exercise programs in long-term care. Previous studies in frail populations have reported improvements in walking speed and stride length [18], although evidence in older adults with frailty remains limited. By providing consistent, scaffolded assistance, the lumbar-type HAL may facilitate mastery experiences—particularly in individuals with low baseline self-efficacy—thereby enhancing ESE. However, its effect on ESE has not yet been examined.

Study Aim and Hypotheses

This study aimed to evaluate the effects of a lumbar-type HAL exercise program on ESE in older adults with frailty: (1) lumbar-type HAL training was hypothesized to improve ESE in this population, and (2) the magnitude of improvement was hypothesized to vary by baseline TTM stage, with the greatest gains expected among participants who had not yet established a regular exercise habit (stages 1 - 4) compared with those in the maintenance stage (stage 5).

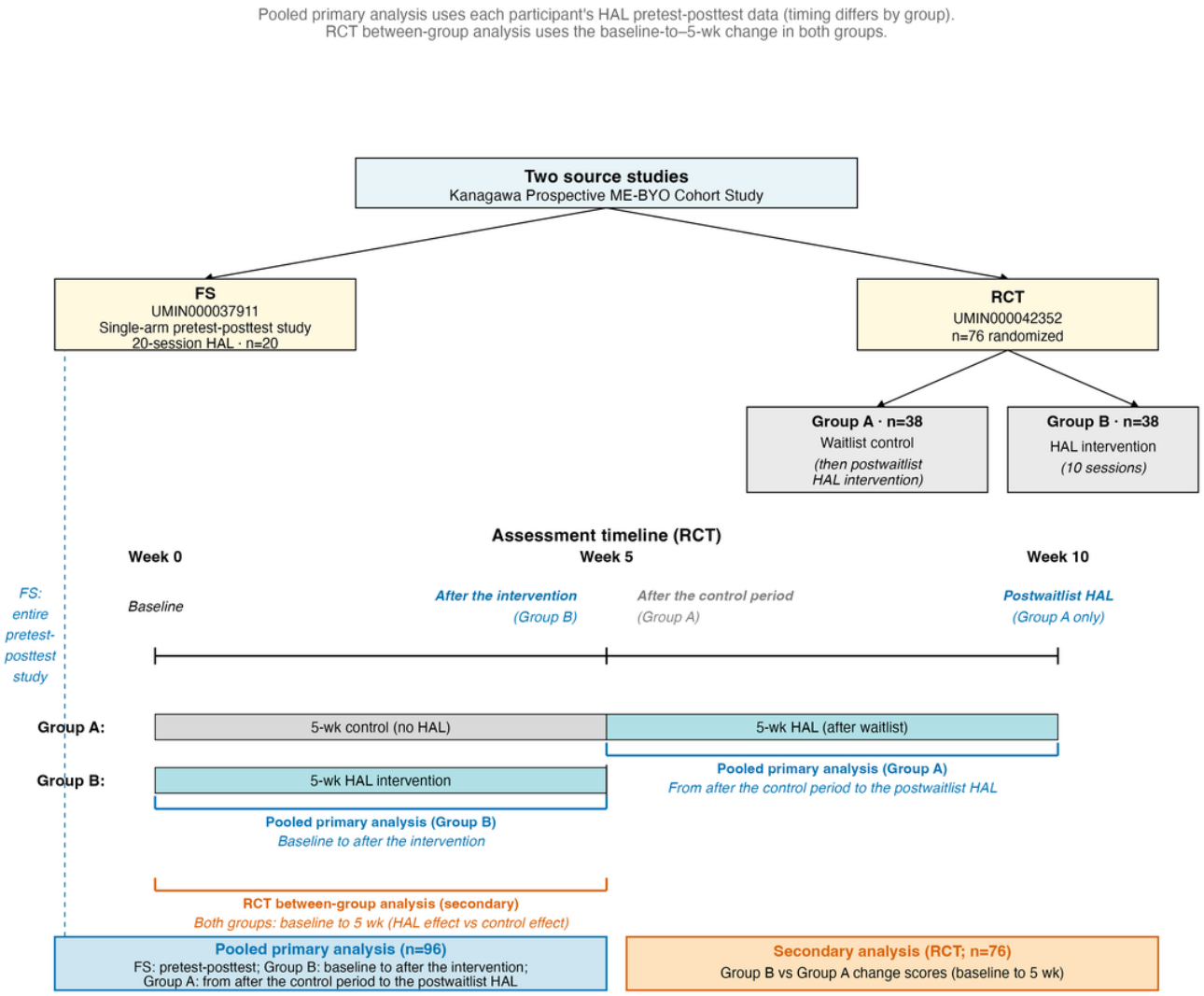
Methods

Study Design and Data Source

This study used a retrospective secondary analysis of pooled outcome data derived from 2 prospective studies of a lumbar-type HAL–based care prevention program (see Figure 1 for the participant flow). The source studies were a feasibility study (FS; UMIN000037911) and a randomized controlled trial (RCT; UMIN000042352) based on the Kanagawa Prospective ME-BYO Cohort Study entitled "Research on the development of a nursing care prevention program using healthcare robots" [19]. Data from both the FS and the RCT were pooled to evaluate the effect of a lumbar-type HAL exercise program on ESE.

XSL•FO
RenderX

Figure 1. Participant flow and analytic use of assessment time points. FS: feasibility study; HAL: Hybrid Assistive Limb; RCT: randomized controlled trial.



The FS aimed to evaluate the feasibility and effectiveness of a care prevention program using the lumbar-type HAL. The primary end point comprised program completion rate. Participants aged below 85 years who met at least one of the frailty criteria based on the Japanese version of the Cardiovascular Health Study, including weight loss, fatigue, reduced PA level, decreased grip strength, and slow walking speed, were included in the analysis. The intervention consisted of lumbar-type HAL training delivered twice weekly for 90 minutes per session over 20 sessions. Ultimately, 20 participants completed the program.

The RCT evaluated the effectiveness of lumbar-type HAL training for care prevention, with normal walking speed as the primary end point. A randomized, single-blind, parallel-group design was used, with participants randomly assigned to either the intervention group or a waitlist control group. Individuals aged 65 to 85 years with a score of 6 or higher on the 5-question Geriatric Locomotive Function Scale, a normal walking speed of less than 1.0 m/s (eligibility based on either the Geriatric Locomotive Function Scale or walking speed criteria), and no certification for long-term care were analyzed. The intervention group received lumbar-type HAL training twice weekly for 90 minutes per session over 10 sessions, whereas the waitlist control group received no intervention during the 5-week period [19]. Ultimately, 77 participants completed the study.

Intervention: Exercise Program Using Lumbar-Type HAL

In both the FS and the RCT, lumbar-type HAL training was conducted primarily in cybernic voluntary control mode, in which assistive torque supporting hip extension during standing and lifting tasks was generated in response to bioelectric signals detected by skin surface sensors placed over the erector spinae muscles. Assistance levels were individually adjusted according to each participant's functional capability to ensure appropriate support during training. The program included repetitive exercises such as sit-to-stand movements and squats.

Assessment Measures

Assessment measures included ESE, the 10-m walk test (10MWT), the timed up and go (TUG) test, amount of PA, TTM stage of behavior change [8], and a self-administered questionnaire. ESE was assessed using a 4-item questionnaire developed by Oka [4] evaluating confidence in exercising when slightly tired, unmotivated, busy, or under unfavorable weather conditions. Responses were rated on a 5-point Likert scale (1=not at all confident; 5=very confident), with higher scores

XSL•FO
RenderX

indicating greater ESE. In the 10MWT, participants walked at a usual pace along a 10-m walkway, and walking speed (m/s) was calculated using stopwatch timing. The TUG test measured the time required to rise from a chair, walk 3 m, turn, return, and sit. PA was assessed using 2 items: weekly exercise frequency (5 - 7 days, 3 - 4 days, 1 - 2 days, or none) and daily exercise duration. Average daily duration was estimated by multiplying frequency (coded as 6, 3.5, 1.5, and 0 days per week) by duration per day and dividing by 7 [20]. TTM stage was rated on a 5-point scale: stage 1 (precontemplation: no intention to perform PA), stage 2 (contemplation: planning to start soon), stage 3 (preparation: performing activity once a week), stage 4 (action: regular activity for <6 months), and stage 5 (maintenance: regular activity for ≥6 months). Regular PA was defined as engaging in PA for 60 minutes per day or more (≥40 minutes for individuals aged ≥65 years) even at low intensity. The self-administered questionnaire was structured as an interview to capture participants' subjective perceptions of changes after program completion. Data from both studies were pooled for analysis, with ESE as the primary end point and the 10MWT, the TUG test, PA, and questionnaire responses as secondary end points.

Sample Size Calculation

Sample size calculation was based on ESE data from the FS (mean change 1.25, SD 2.88 points; paired Cohen *d*=0.43) [21]. For the primary pooled within-person analysis, 44 participants were required to achieve 80% power with a 2-sided α of .05; the achieved sample size (n=96) exceeded this requirement. For the secondary RCT between-group comparison, 172 participants (86 per group) were required; the achieved sample size (n=76) was below this threshold, resulting in limited statistical power (post hoc: 30% - 34% at a 2-sided α=.05).

Statistical Analysis

Overview

Descriptive data were expressed as means and SDs for continuous variables and numbers and percentages for categorical variables. The 95% CIs were reported for all primary effect estimates. All statistical tests were 2 sided, with a significance threshold of .05. Statistical analyses were performed using the R software (version 4.2.2; R Foundation for Statistical Computing). One participant with incomplete primary outcome data was excluded from the analysis.

Primary Analysis

The primary analysis comprised a within-person pretest-posttest comparison of ESE among all lumbar-type HAL recipients (n=96) pooled from the FS and the RCT (the RCT intervention group plus the waitlist control group after completion of the lumbar-type HAL exercise program) using a paired *t* test. In the RCT, ESE was assessed at baseline (week 0), after the intervention (week 5), and—for the waitlist control group only—after the subsequent 5-week lumbar-type HAL program (postwaitlist HAL; week 10). For the pooled primary analysis, each participant contributed a pretest-posttest comparison corresponding to lumbar-type HAL exposure: from baseline to week 5 for the intervention group (group B) and from after the control period to after the intervention for the waitlist group (group A; Figure 1). The waitlist group received lumbar-type HAL training after the control period for ethical reasons in accordance with the registered protocol (UMIN000042352) [19]. This pooled pretest-posttest comparison constituted the primary inferential analysis specified in the study design, which was powered to detect changes in ESE, an outcome not individually powered in either source study. Paired *t* tests were also applied to the 10MWT, TUG, and PA outcomes.

Subgroup Analyses

An a priori dichotomous subgroup comparison was specified prior to pooled data analysis based on TTM theory [8] and findings from the FS. The comparison contrasted stages 1 to 4 (precontemplation, contemplation, preparation, and action; no established exercise habit) vs stage 5 (maintenance; regular exercise for ≥6 months). This contrast was not included in the original University Hospital Medical Information Network Clinical Trials Registry protocols (UMIN000037911 and UMIN000042352) as those preceded the development of the pooled analysis plan; however, specification occurred prior to data integration from the 2 source studies. The empirical rationale was derived from the FS, in which participants in lower-readiness stages (n=10) demonstrated a mean ESE increase of 2.5 points compared with 0 points in higher-readiness participants (n=10; *P*=.052) [21]. The theoretical basis aligns with the social cognitive theory by Bandura [5], which predicts stronger mastery-related effects among individuals with lower baseline self-efficacy, and with the TTM framework, in which stage 5 reflects a maintenance phase associated with attenuated intervention effects due to ceiling effects. Within-stratum changes in the pooled dataset were analyzed using paired *t* tests, and between-stratum differences in change scores were assessed using the Welch *t* test. ESE changes were also described across all 5 TTM stages; given the small sample sizes within some stages, these analyses were considered exploratory.

Secondary Analysis (RCT; Between Groups)

The randomized design of the RCT (n=76) was used as a secondary analysis to compare ESE change over the same 5-week period between the lumbar-type HAL group (n=38) and the waitlist control group (n=38). Between-group differences in change scores were assessed using the Welch *t* test. Results were further confirmed using analysis of covariance, with postintervention ESE as the dependent variable, group as the fixed factor, and baseline ESE as a covariate. The Cohen *d* was calculated as a standardized effect size. The prespecified TTM dichotomy was also examined within this between-group framework.

Sensitivity Analyses for the Pooled Estimate

Robustness of the pooled estimate was evaluated against potential heterogeneity between the source studies (FS: 20 sessions, single arm; RCT: 10 sessions with a waitlist control group) using three prespecified sensitivity analyses within a fixed-effects framework: (1) source-stratified paired *t* tests performed separately within each study, (2) source-adjusted linear regression including a binary study indicator (FS=1; RCT=0), and (3) a between-study heterogeneity assessment

XSL•FO
RenderX

using the Welch *t* test on change scores comparing FS and RCT participants.

Ethical Considerations

Both source studies were approved by the Research Ethics Review Committee of the Graduate School of Health Innovation, Kanagawa University of Human Services (FS: approval number 20019-36-004; RCT: approval number Hodai 30-008). Written informed consent was obtained from all participants, including permission for secondary analyses of deidentified data. The present pooled analysis was reviewed by the same committee and deemed exempt from full ethics review (review notification number SHI 62) as it used existing deidentified data. This study was conducted in accordance with the Declaration of Helsinki.

Results

Participant Characteristics

ESE data were available for 99% (96/97) of the participants who completed either the FS or RCT programs. Participants' age ranged from 50 to 84 years. Baseline characteristics of these participants are shown in Table 1.

Table . Participant characteristics (n=96)[a].

Characteristics	Participants, n (%)
Age (y)	
50 - 59	4 (4.2)
60 - 69	10 (10.4)
70 - 79	66 (68.8)
80 - 84	16 (16.7)
Sex (n=95)	
Male	28 (29.5)
Female	67 (70.5)
Source study	
FS[b]	20 (20.8)
RCT[c]	76 (79.2)
Stage of TTM[d] (n=94)	
Stage 1 (precontemplation)	6 (6.4)
Stage 2 (contemplation)	28 (29.8)
Stage 3 (preparation)	13 (13.8)
Stage 4 (action)	9 (9.6)
Stage 5 (maintenance)	38 (40.4)

[a]Sex was not available for 1% (1/96) of the participants. Transtheoretical model stage data were missing for 2.1% (2/96) of the participants (1 in the postwaitlist group A and 1 in the randomized controlled trial lumbar-type Hybrid Assistive Limb group).

[b]FS: feasibility study.

[c]RCT: randomized controlled trial.

[d]TTM: transtheoretical model.

Pooled Pretest-Posttest Analysis (Primary Analysis)

In the primary pooled within-person analysis of all lumbar-type HAL recipients (n=96), ESE increased significantly by 1.05 points (95% CI 0.41-1.69; *P*=.002) from a baseline mean of 11.94 (SD 3.05) points to 12.99 (SD 3.21) points after the intervention (Table 2). Secondary outcomes also showed significant improvement: 10MWT walking speed increased by 0.37 m/s (95% CI 0.33-0.40; *P*<.001), TUG time decreased by 1.68 seconds (95% CI −2.01 to −1.35; *P*<.001), and PA increased by 18.28 minutes per day (95% CI 12.95-23.62; *P*<.001; Table 2).

XSL•FO
RenderX

Table . Effect of lumbar-type Hybrid Assistive Limb training on primary and secondary outcomes (pooled paired analysis; n=96)[a].

Outcome	Participants, n	Baseline, mean (SD)	After the intervention, mean (SD)	Change, mean (95% CI)	*P* value	Cohen *d*
ESE[b] total score (range 4 - 20)	96	11.94 (3.05)	12.99 (3.21)	+1.05 (+0.41 to +1.69)	.002	0.33
10MWT[c] speed (m/s)	96	0.99 (0.28)	1.36 (0.37)	+0.37 (+0.33 to +0.40)	<.001	—[d]
TUG[e] time (s)	96	10.12 (8.47)	8.45 (7.64)	–1.68 (–2.01 to –1.35)	<.001	—
PA[f] (min per d)	96	30.47 (29.27)	48.75 (33.99)	+18.28 (+12.95 to +23.62)	<.001	—

[a]Exercise self-efficacy is the primary outcome; 10-m walk test speed, timed up and go time, and physical activity are secondary outcomes. Pooled paired analysis (n=96): paired *t* test, 2-sided α=.05. The 10-m walk test speed (m/s) was calculated as 10 m divided by walking time.

[b]ESE: exercise self-efficacy.

[c]10MWT: 10-m walk test.

[d]Cohen *d* is reported only for the primary outcome (ESE).

[e]TUG: timed up and go test.

[f]PA: physical activity.

Sensitivity analyses supported the robustness of the pooled estimate to source-related heterogeneity (Table 3). Source-stratified analyses demonstrated similar improvements (FS: Δ=+1.25 points, *P*=.07, and n=20; RCT lumbar-type HAL group: Δ=+1.00 point, *P*=.009, and n=76). The between-source heterogeneity test was nonsignificant (*P*=.74), and source-adjusted regression produced a comparable estimate (+1.00 point) with a nonsignificant source effect (*P*=.76).

Table . Sensitivity analyses for the pooled estimate of exercise self-efficacy change[a].

Analysis	Participants, n	Point estimate (95% CI)	*P* value
Primary pooled paired analysis	96	+1.05 (+0.41 to +1.69)	.002
Source stratified			
FS[b] only	20	+1.25 (−0.10 to +2.60)	.07
RCT[c] lumbar-type HAL[d] recipients	76	+1.00 (+0.26 to +1.74)	.009
Between-source heterogeneity[e]			
FS − RCT difference	96	+0.25 (−1.26 to +1.76)	.74
Source-adjusted regression[f]			
Pooled estimate (RCT reference)	96	+1.00 (+0.28 to +1.72)	.007
Source coefficient (FS vs RCT)	—[g]	+0.25 (−1.33 to +1.83)	.76

[a]Prespecified sensitivity analyses for the pooled primary estimate. A fixed-effects framework was applied as the 2 source studies represent sequential phases of a single research program.

[b]FS: feasibility study.

[c]RCT: randomized controlled trial.

[d]HAL: Hybrid Assistive Limb.

[e]Between-source heterogeneity test: Welch *t* test comparing change scores between FS and RCT participants.

[f]Source-adjusted regression: linear regression of exercise self-efficacy change on a source-study indicator (FS=1; RCT=0). The intercept estimate represents the pooled mean change, with RCT as the reference. The source coefficient tests for between-source difference.

[g]Not applicable (the source coefficient tests the between-source difference within the regression model including all 96 participants and does not correspond to a separate participant subgroup).

Pooled TTM Subgroup Analysis

Subgroup analyses stratified by baseline TTM stage were conducted to characterize changes in ESE (Table 4). Participants in the precontemplation stage showed a mean increase of 3.50 (SD 3.08) points following lumbar-type HAL training. The mean ESE increases were 1.14 (SD 3.27), 1.46 (SD 2.79), and 2.78 (SD 3.56) points in the contemplation (n=28), preparation (n=13), and action (n=9) stages, respectively. In contrast, the maintenance stage (n=38) showed a minimal change of 0.05 (SD 2.86) points. The prespecified between-stratum comparison was significant. Stages 1 to 4 combined (n=56) demonstrated

XSL•FO
RenderX

a mean ESE increase of 1.73 points (95% CI 0.87-2.60; *P*<.001) compared with 0.05 points in stage 5 (n=38; 95% CI −0.90 to 1.01; *P*=.91), yielding a between-stratum difference of +1.68 points (95% CI 0.41-2.95; *P*=.01; Table 4).

Table . Change in exercise self-efficacy by baseline transtheoretical model stage (pooled analysis; n=94)[a].

	Participants, n	Baseline score (4-20), mean (SD)	Postintervention score (4-20), mean (SD)	Change (points), mean (95% CI)	*P* value
Stage 1 (precontemplation)	6	8.00 (3.16)	11.50 (4.18)	+3.50 (+0.27 to +6.73)	.04
Stage 2 (contemplation)	28	11.46 (3.17)	12.61 (3.65)	+1.14 (−0.13 to +2.41)	.08
Stage 3 (preparation)	13	12.00 (2.58)	13.46 (2.67)	+1.46 (−0.22 to +3.15)	.08
Stage 4 (action)	9	12.22 (2.28)	15.00 (3.32)	+2.78 (+0.04 to +5.52)	.047
Stage 5 (maintenance)	38	12.76 (2.89)	12.82 (2.81)	+0.05 (−0.90 to +1.01)	.91
Stage 1 - 4 (combined)	56	11.34 (3.08)	13.07 (3.51)	+1.73 (+0.87 to +2.60)	<.001
Difference (stages 1 - 4 vs 5)	—[b]	—	—	+1.68 (+0.41 to +2.95)	.01

[a]Paired *t* tests within each subgroup; between-stratum comparisons were performed using the Welch *t* test on change scores. Two-sided α=.05. Per-stage analyses are exploratory and hypothesis generating due to the small sample sizes within several stages.

[b]Baseline and postintervention values are not applicable for the between-stratum difference row as it reflects a contrast between groups rather than a within-group estimate.

RCT Between-Group Comparison (Secondary Analysis)

The secondary analysis used the randomized design of the RCT to support causal inference (Table 5 and Figure 2). The lumbar-type HAL group (n=38) showed a mean ESE change of +1.16 (SD 3.67) points compared with −0.05 (SD 3.05) points in the waitlist control group (n=38). The between-group difference in change was 1.21 points (95% CI −0.33 to 2.75; *P*=.12; Cohen *d*=0.36). Analysis of covariance adjusting for baseline ESE yielded a similar adjusted between-group difference of 1.08 points (95% CI −0.28 to 2.44; *P*=.12). Limited statistical power was present for this between-group comparison (post hoc: 30% - 34% at α=.05), and interpretation relied on consistency with the pooled primary estimate. Within the prespecified TTM dichotomy, the between-group difference was 2.10 points among participants in stages 1 to 4 (lumbar-type HAL: n=21; control: n=21; 95% CI −0.16 to 4.35; *P*=.07; Cohen *d*=0.58) and 0.02 points among participants in stage 5 (lumbar-type HAL: n=16; control: n=17; 95% CI −2.21 to 2.24; *P*=.99; Cohen *d*=0.01; Table 5 and Figure 2). Stage-specific descriptive results are provided in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Effect of lumbar-type Hybrid Assistive Limb (HAL) training on exercise self-efficacy in the randomized controlled trial: between-group comparison overall and by transtheoretical model stage dichotomy[a].

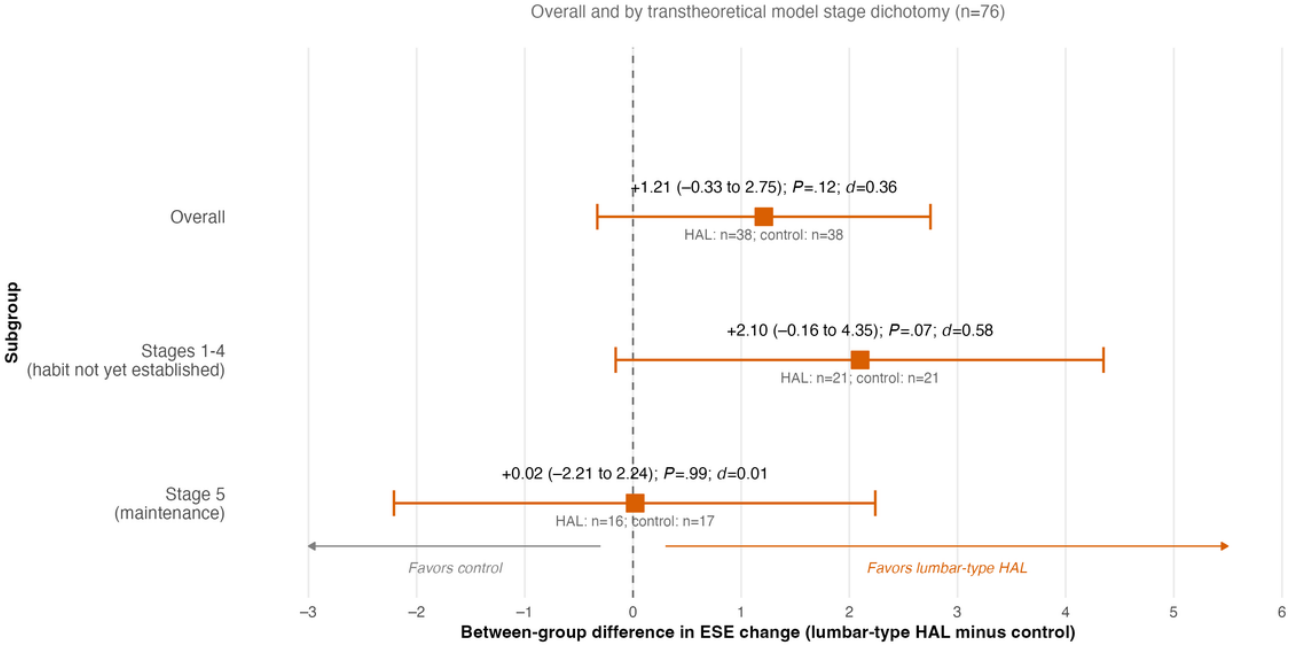
Group and subgroup	Participants, n	Baseline score (4-20), mean (SD)	Postintervention score (4-20), mean (SD)	Change (points), mean (SD)	Difference (95% CI)	*P* value	Cohen *d*
Overall					+1.21 (−0.33 to +2.75)	.12	0.36
Lumbar-type HAL	38	11.63 (2.84)	12.79 (3.31)	1.16 (3.67)			
Waitlist control	38	11.87 (3.19)	11.82 (3.21)	−0.05 (3.05)			
Stage 1-4[b]					+2.10 (−0.16 to +4.35)	.07	0.58
Lumbar-type HAL	21	10.57 (2.64)	12.33 (3.65)	1.76 (3.81)			
Waitlist control	21	11.67 (3.15)	11.33 (3.06)	−0.33 (3.41)			
Stage 5[c]					+0.02 (−2.21 to +2.24)	.99	0.01
Lumbar-type HAL	16	12.88 (2.66)	13.19 (2.86)	0.31 (3.53)			
Waitlist control	17	12.12 (3.31)	12.41 (3.39)	0.29 (2.59)			

[a]Between-group difference (lumbar-type HAL minus waitlist control) was calculated using the Welch *t* test on change scores. Analysis of covariance adjusting for baseline exercise self-efficacy yielded similar results: +1.08 points (95% CI −0.28 to +2.44; *P*=.12).

[b]Stages 1-4: precontemplation, contemplation, preparation, and action stages (regular exercise habit not yet established).

[c]Stage 5: maintenance stage (regular exercise habit established for ≥6 months).

Figure 2. Randomized controlled trial between-group comparison of exercise self-efficacy (ESE) change. HAL: Hybrid Assistive Limb.



Subjective Feedback

Responses to the self-administered questionnaire included individual statements such as the following:

My motivation to exercise has increased.

I want to continue exercising in the future.

It has become easier to move. Standing up takes less effort now.

I can stand up without having to hold on.

I felt that my back was more comfortable after using the HAL for the first time. My sitting posture has improved.

XSL•FO
RenderX

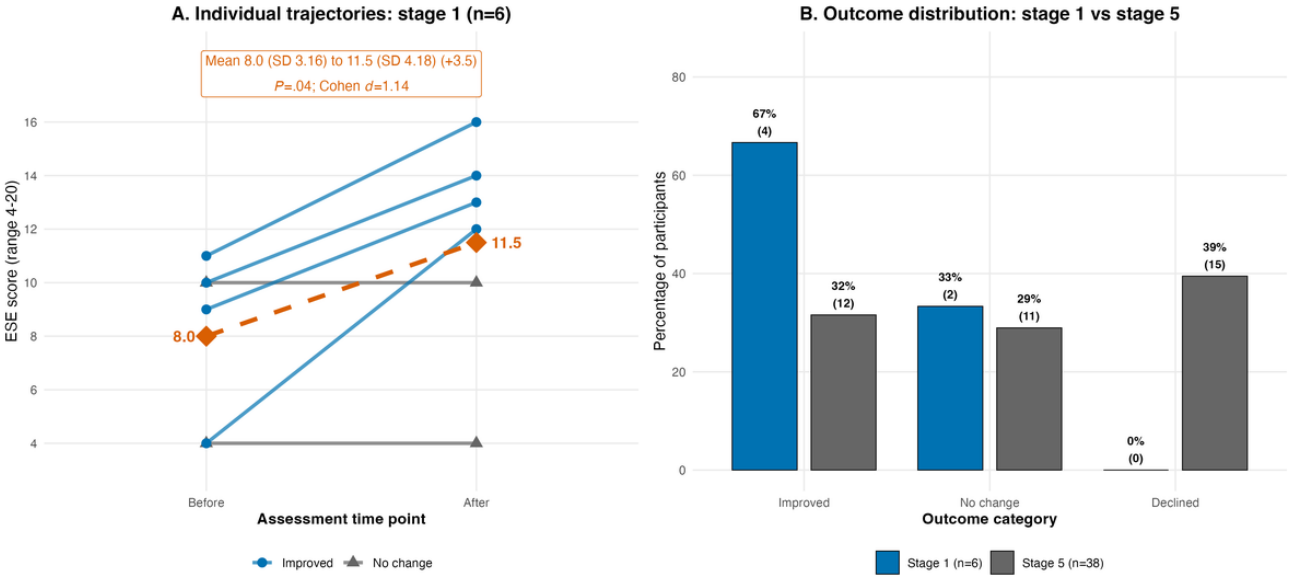
Discussion

Principal Findings

This retrospective pooled analysis combining an FS and an RCT of lumbar-type HAL training in older adults with frailty demonstrated a significant improvement in ESE of 1.05 points (95% CI 0.41-1.69; *P*=.002) among 96 recipients with lumbar-type HAL. Improvements in ESE were accompanied by significant gains in physical function (10MWT and TUG test) and PA, indicating concurrent enhancement of physical and psychological readiness for exercise. Improvements in physical function likely reinforced mastery experiences, thereby contributing to increased ESE [5-7]. By providing consistent, scaffolded assistance, the lumbar-type HAL enables successful completion of functional tasks such as sit-to-stand movements and squats that may otherwise exceed baseline capacity, thereby strengthening confidence in performing exercise under everyday conditions.

Subgroup analyses demonstrated differential effects across baseline TTM stages. The prespecified between-stratum contrast (stages 1‐4 vs stage 5) was significant (between-stratum difference of +1.68 points, 95% CI 0.41-2.95; *P*=.01), indicating that responsiveness varied by behavioral readiness. The largest effect was observed in precontemplation (stage 1), where 4 of 6 participants improved (none deteriorated; mean change +3.50, SD 3.08 points; Cohen *d*=1.14; Figure 3). This exploratory finding (n=6) is of interest given that precontemplators are typically difficult to engage in conventional exercise programs and have limited opportunities for mastery experiences [6,10]. In contrast, participants in stage 5 (maintenance) showed minimal change (+0.05 points; *P*=.91). Although baseline ESE in this subgroup did not approach the scale maximum, it was already relatively high for this population, leaving little room for further improvement; in addition, the intervention tasks (sit-to-stand exercises and squats) may have provided limited additional challenge for individuals with established exercise habits [22]. Sensitivity analyses supported the validity of pooling FS and RCT data (between-source heterogeneity: *P*=.74; source-adjusted regression: *P*=.76; RCT-only results were directionally consistent), suggesting minimal influence of methodological differences between studies. The secondary RCT between-group analysis was not significant (between-group difference of +1.21 points; *P*=.12; Cohen *d*=0.36), likely reflecting limited statistical power (post hoc: 30% ‐ 34%; required n=172 for 80% power). However, the effect estimate aligned closely with the pooled primary result (+1.21 vs +1.05 points), and the TTM pattern was consistent (Cohen *d*=0.58 for stages 1 ‐ 4 vs Cohen *d*=0.01 for stage 5); thus, the direction of the effect estimates was consistent across analyses rather than contradictory. However, confirmation through a sufficiently powered, dedicated RCT is required before causal inference can be established.

Figure 3. Stage 1 (precontemplation) detailed analysis. ESE: exercise self-efficacy.



Comparison With Previous Studies

These findings extend those of previous studies on the lumbar-type HAL by demonstrating effects on a psychological outcome (ESE) in addition to previously reported physical function [18]. To our knowledge, this study is the first to examine lumbar-type HAL responsiveness according to TTM stage and describe an exploratory finding among precontemplators. The observed effect size in this subgroup (Cohen *d*=1.14) aligns with the theory by Bandura [5] that mastery experience interventions yield the largest gains among individuals with low baseline self-efficacy. The minimal change observed in the maintenance stage is consistent with prior exercise intervention studies in older adults with established activity habits, in which ceiling effects on self-efficacy measures are commonly reported [22]. A recent systematic review reported limited and inconsistent evidence on TTM-based interventions for PA in older adults [10], highlighting the scarcity of evidence on stage-tailored exercise strategies. The present findings provide preliminary evidence that lumbar-type HAL training may serve as a stage-appropriate entry strategy for previously disengaged individuals. Notably, a recent study of a wearable walking assistance robot in community-dwelling older adults [23] reported feasibility and acceptability without

XSL•FO
RenderX

examining ESE; the present study extends this line of research by linking robot-assisted exercise to a psychological readiness outcome.

Limitations

This study has some limitations. First, the FS and RCT differed in the structure and delivery of the lumbar-type HAL–based exercise programs. The FS implemented an individualized program, whereas the RCT combined individual and group-based training. Group-based settings provide opportunities for social encouragement and observational learning (vicarious experiences), both of which enhance self-efficacy [5]. Accordingly, participants in the RCT may have had greater exposure to self-efficacy–enhancing mechanisms than those in the FS. Moreover, the number of lumbar-type HAL training sessions differed between studies (20 in the FS vs 10 in the RCT), which may have influenced the magnitude of ESE improvement. These heterogeneities were addressed through prespecified sensitivity analyses (source stratified, source-adjusted regression, and between-source heterogeneity testing), all of which supported the robustness of the pooled estimate. Second, the secondary RCT between-group analysis was underpowered: the achieved sample size (n=76; 38 per group) was below the required 172 to detect a medium effect (Cohen *d*=0.43) at a 2-sided α of .05 with 80% power, corresponding to a post hoc power of 30% to 34%. The directionally consistent effects observed in the pooled primary analysis (significant; +1.05 points; *P*=.002) and the secondary RCT analysis (nonsignificant; +1.21 points; *P*=.12) suggest a potentially meaningful intervention effect; however, a larger adequately powered RCT is required for definitive between-group inference. Third, subgroup analyses by TTM stage were limited by small sample sizes within some strata (eg, stage 1: n=6 in the pooled analysis; n=2 in the lumbar-type HAL group and n=4 in the control group of the RCT). Consequently, stage-specific findings should be considered exploratory and hypothesis generating. The prespecified dichotomous comparison between stages 1 to 4 vs stage 5 (*P*=.01) therefore served as the primary stratified analysis. Although no participant in stage 1 showed a decline in ESE, confirmation in studies adequately powered to assess effects within individual TTM stages remains necessary. Fourth, open recruitment may have preferentially attracted participants with greater health awareness, potentially limiting the generalizability of the findings.

Future Directions

Future research should recruit larger numbers of participants in the precontemplation stage (stage 1) to confirm the large effect estimate observed in this exploratory analysis (n=6; Cohen *d*=1.14). A purposive sampling strategy targeting disengaged older adults—who are often underrepresented in exercise research—would enable rigorous evaluation of whether the pronounced ESE gains observed in precontemplators can be replicated. Adequately powered studies stratified by baseline TTM stage are needed to determine whether lumbar-type HAL training reliably benefits older adults with low readiness for exercise. The differential effects observed across TTM stages also support a stratified intervention approach whereby lumbar-type HAL training may be appropriately targeted to lower-readiness individuals (stages 1 - 4), whereas alternative strategies may better serve those in the maintenance stage.

Conclusions

In this retrospective pooled analysis, a lumbar-type HAL exercise program demonstrated a significant within-person improvement in ESE in older adults with frailty, with the greatest improvement observed among individuals without established exercise habits. Evidence from the RCT indicated a favorable but inconclusive between-group effect, and adequately powered confirmatory RCTs remain necessary.

The exploratory response in precontemplators—a group traditionally difficult to engage in exercise interventions—suggests that HAL-supported mastery experiences may provide a novel entry point to PA for disengaged older adults at high risk of frailty. Future studies should specifically target participants in stage 1 using purposive sampling, confirm the differential effects in adequately powered stratified designs, and examine whether improvements in ESE translate into sustained PA over time.

Acknowledgments

The authors would like to thank all study participants for their contribution. Appreciation is extended to Editage for providing English-language editing support.

Funding

This study received no specific funding. The feasibility study and randomized controlled trial were supported by the Regional Innovation Ecosystem Formation Program of the Ministry of Education, Culture, Sports, Science, and Technology of Japan (https://www.mext.go.jp/a_menu/kagaku/chiiki/program/1367366.html).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy and ethical restrictions but are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

MN and HN conceptualized and designed the study. MN and HN analyzed and interpreted the data and drafted the manuscript. SN, YS, YO, KW, and TK critically revised the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest. MN has been an employee of CYBERDYNE, Inc. since October 2022. However, the company has no role in this study. None of the other authors have any conflicts of interest to declare.

Multimedia Appendix 1
Stage-by-stage descriptive results from the randomized controlled trial stratified by transtheoretical model stage (1-5) and intervention group.
[DOCX File, 18 KB - rehab_v13i1e79358_app1.docx]

References

1. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
2. Xue QL. The frailty syndrome: definition and natural history. Clin Geriatr Med 2011 Feb;27(1):1-15. [doi: 10.1016/j.cger.2010.08.009] [Medline: 21093718]
3. Neupert SD, Lachman ME, Whitbourne SB. Exercise self-efficacy and control beliefs: effects on exercise behavior after an exercise intervention for older adults. J Aging Phys Act 2009 Jan;17(1):1-16. [doi: 10.1123/japa.17.1.1] [Medline: 19299835]
4. Oka K. Stages of change for exercise behavior and self-efficacy for exercise among middle-aged adults [Article in Japanese]. Nihon Koshu Eisei Zasshi 2003 Mar;50(3):208-215. [Medline: 12704833]
5. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
6. Maeba K, Takenaka K. Associations between sources of exercise self-efficacy and stages of change for exercise among older adults. Jpn J Behav Med 2012;18(1):12-18. [doi: 10.11331/jjbm.18.12]
7. Zhao J, Zhang G, Xu D. The effect of reward on motor learning: different stage, different effect. Front Hum Neurosci 2024;18:1381935. [doi: 10.3389/fnhum.2024.1381935] [Medline: 38532789]
8. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
9. Prommaban A, Moonkayaow S, Phinyo P, Siviroj P, Sirikul W, Lerttrakarnnon P. The effect of exercise program interventions on frailty, clinical outcomes, and biomarkers in older adults: a systematic review. J Clin Med 2024 Nov 1;13(21):6570. [doi: 10.3390/jcm13216570] [Medline: 39518709]
10. Fawole HO, Itua SA, Idowu OA, et al. Effects of transtheoretical model of change-based interventions on physical activity among older adults: a systematic review of randomised controlled and non-randomised controlled trials. Eur Rev Aging Phys Act 2025 Dec 2;23(1):1. [doi: 10.1186/s11556-025-00391-w] [Medline: 41331559]
11. Maeshima S, Osawa A, Nishio D, et al. Efficacy of a Hybrid Assistive Limb in post-stroke hemiplegic patients: a preliminary report. BMC Neurol 2011 Sep 27;11:116. [doi: 10.1186/1471-2377-11-116] [Medline: 21943320]
12. Suzuki K, Mito G, Kawamoto H, Hasegawa Y, Sankai Y. Intention-based walking support for paraplegia patients with Robot Suit HAL. Adv Robot 2007;21(12):1441-1469. [doi: 10.1163/156855307781746061]
13. Kawamoto H, Kamibayashi K, Nakata Y, et al. Pilot study of locomotion improvement using Hybrid Assistive Limb in chronic stroke patients. BMC Neurol 2013 Oct 7;13:141. [doi: 10.1186/1471-2377-13-141] [Medline: 24099524]
14. Kubota S, Nakata Y, Eguchi K, et al. Feasibility of rehabilitation training with a newly developed wearable robot for patients with limited mobility. Arch Phys Med Rehabil 2013 Jun;94(6):1080-1087. [doi: 10.1016/j.apmr.2012.12.020] [Medline: 23313372]
15. Wall A, Borg J, Palmcrantz S. Clinical application of the Hybrid Assistive Limb (HAL) for gait training-a systematic review. Front Syst Neurosci 2015;9:48. [doi: 10.3389/fnsys.2015.00048] [Medline: 25859191]
16. Watanabe H, Tanaka N, Inuta T, Saitou H, Yanagi H. Locomotion improvement using a Hybrid Assistive Limb in recovery phase stroke patients: a randomized controlled pilot study. Arch Phys Med Rehabil 2014 Nov;95(11):2006-2012. [doi: 10.1016/j.apmr.2014.07.002] [Medline: 25010538]
17. Yoshimoto T, Shimizu I, Hiroi Y, Kawaki M, Sato D, Nagasawa M. Feasibility and efficacy of high-speed gait training with a voluntary driven exoskeleton robot for gait and balance dysfunction in patients with chronic stroke: nonrandomized pilot study with concurrent control. Int J Rehabil Res 2015 Dec;38(4):338-343. [doi: 10.1097/MRR.0000000000000132] [Medline: 26288120]
18. Kotani N, Morishita T, Yatsugi A, et al. Biofeedback core exercise using Hybrid Assistive Limb for physical frailty patients with or without Parkinson's disease. Front Neurol 2020;11:215. [doi: 10.3389/fneur.2020.00215] [Medline: 32328021]

XSL•FO
RenderX

19. Saito Y, Nakamura S, Kasukawa T, Nagasawa M, Oguma Y, Narimatsu H. Efficacy of exercise with the Hybrid Assistive Limb lumbar type on physical function in mobility-limited older adults: a 5-week randomized controlled trial. Exp Gerontol 2024 Oct 1;195:112536. [doi: 10.1016/j.exger.2024.112536] [Medline: 39098361]
20. Saito Y, Oguma Y, Tanaka A, et al. Community-wide physical activity intervention based on the Japanese physical activity guidelines for adults: a non-randomized controlled trial. Prev Med 2018 Feb;107:61-68. [doi: 10.1016/j.ypmed.2017.11.008] [Medline: 29126918]
21. Nagasawa M, Nakamura S, Oguma Y, Saito Y, Kasukawa T, Narimatsu H. Effects of an exercise program using the lumbar-type Hybrid Assistive Limb on exercise self-efficacy in frail older adults [Article in Japanese]. Rigaku Ryoho Gaku 2021;48(Suppl 1):B-202_1. [doi: 10.14900/cjpt.48S1.B-202_1]
22. Garin O. Ceiling effect. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research: Springer; 2014:631-633. [doi: 10.1007/978-94-007-0753-5_296]
23. Cho E, Hwang S, Heo SJ, Lim B, Lee J, Lee Y. Feasibility and effects of a gait assistance and gait resistance training program using a walking-assist wearable robot for community-dwelling older adults: single-group, pre-, and posttest study. JMIR Mhealth Uhealth 2025 May 26;13:e58142. [doi: 10.2196/58142] [Medline: 40418849]

Abbreviations

10MWT: 10-m walk test
ESE: exercise self-efficacy
FS: feasibility study
HAL: Hybrid Assistive Limb
PA: physical activity
RCT: randomized controlled trial
TTM: transtheoretical model
TUG: timed up and go

Edited by S Munce; submitted 30.Jun.2025; peer-reviewed by A Samuel, W Peersman; revised version received 19.Jul.2026; accepted 22.Jul.2026; published 24.Aug.2026.

Please cite as:
Nagasawa M, Nakamura S, Saito Y, Oguma Y, Kasukawa T, Watanabe K, Narimatsu H
Effects of Using a Lumbar-Type Hybrid Assistive Limb by Older Adults With Frailty on Exercise Self-Efficacy: Retrospective Analysis Based on the Kanagawa Prospective ME-BYO Cohort Study
JMIR Rehabil Assist Technol 2026;13:e79358
URL: https://rehab.jmir.org/2026/1/e79358
doi:10.2196/79358

© Makoto Nagasawa, Sho Nakamura, Yoshinobu Saito, Yuko Oguma, Takashi Kasukawa, Kaname Watanabe, Hiroto Narimatsu. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Concurrent Validity of Isokinetic Hip and Knee Strength Measurements Using the UGO Exoskeleton in Healthy Young Adults: Cross-Sectional Validation Study

Hao Xie[1*], BS; Ming-Hui Ding[1*], MS; Pei-Ming Chen[2], PhD; Shamay S M Ng[2], PhD; Chu-Huai Wang[1], MS; Jiang-Li Zhao[1,3], MS

[1]Department of Rehabilitation Medicine, the First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China

[2]Department of Rehabilitation Sciences, the Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[3]Department of Rehabilitation Medicine, Guizhou Hospital of the First Affiliated Hospital of Sun Yat-sen University, Guiyang, Guizhou, China

[*]these authors contributed equally

Corresponding Author:
Jiang-Li Zhao, MS
Department of Rehabilitation Medicine, the First Affiliated Hospital of Sun Yat-sen University
58 Zhongshan Er Road
Guangzhou, Guangdong, 510080
China
Phone: 86 13430204648
Email: zhaojl6@mail.sysu.edu.cn

Abstract

Background: The UGO lower limb exoskeleton rehabilitation robot is designed for patients with lower limb motor dysfunction, and it can be used to assess the peak torque of hip and knee flexors and extensors.

Objective: The primary aim of this study was to establish the concurrent validity of peak isokinetic torque using the UGO exoskeleton and Humac Norm dynamometers in healthy young adults.

Methods: Twenty healthy young adults were enrolled in this study. All participants underwent isokinetic muscle strength tests using both the UGO exoskeleton and the Humac Norm dynamometers. Peak torque (Nm) of the hip and knee flexors and extensors was measured. The Pearson correlation coefficient was used to determine concurrent validity. Bland-Altman analysis was used to assess the agreement between the 2 devices.

Results: At the knee, peak torque of the extensors measured by the UGO exoskeleton correlated strongly with that measured by the Humac Norm dynamometer (right: r=0.662; P=.001; left: r=0.768; P<.001), with 95% limits of agreement (LoA) ranging from −22.89 to 82.21 Nm and from −24.60 to 75.65 Nm, respectively; the right knee flexors showed a moderate correlation (r=0.547; P=.01), with 95% LoA ranging from −26.27 to 36.01 Nm, whereas the left knee flexors showed no significant association (r=0.347; P=.12), with 95% LoA ranging from −2.40 to 75.02 Nm. At the hip, peak torque of the flexors correlated strongly on the right (r=0.615; P=.003) and moderately on the left (r=0.443; P=.04), with 95% LoA ranging from −15.52 to 86.57 Nm and from −12.53 to 98.50 Nm, respectively, whereas the bilateral hip extensors showed no significant association (right: r=0.337; P=.14; left: r=0.391; P=.20), with 95% LoA ranging from −51.57 to 76.80 Nm and from −28.19 to 79.59 Nm, respectively.

Conclusions: In healthy young adults, the UGO exoskeleton demonstrated moderate to strong concurrent validity for peak torque of the bilateral knee extensors and hip flexors, but not for the nondominant knee flexors or bilateral hip extensors. The device may serve as a convenient torque assessment tool for selected muscle actions, but systematic overestimation and posture-related differences must be considered, and validation in the intended clinical population of patients with neurological conditions is required in further studies.

(JMIR Rehabil Assist Technol 2026;13:e99564) doi:10.2196/99564

KEYWORDS

exoskeleton rehabilitation robot; isokinetic strength testing; peak torque; validity; lower extremity

XSL•FO
RenderX

Introduction

Lower limb motor dysfunction is a significant complication associated with normal aging or nervous system diseases that may result in a considerable decline in quality of life [1-3], impose a heavy health burden, and lead to substantial economic losses to society [4]. Lower limb muscle strength is an important factor in maintaining balance and gait [5,6]. Among lower limb muscles, knee extensors and flexors play a critical role in weight bearing, propulsion [7], and stability during ambulation [8], making their accurate assessment essential for rehabilitation planning and outcome evaluation. Traditional muscle strength testing methods, such as manual muscle testing, are widely used in clinical settings due to their simplicity and low cost. However, these methods suffer from well-documented limitations, including subjective interpretation, assessor strength bias, and poor reliability [9], particularly when testing stronger muscle groups such as the knee extensors in individuals who are healthy or moderately impaired. Therefore, the ability to assess muscle strength efficiently and with high precision has become a clinical imperative. Isokinetic dynamometers are known as the gold standard for muscle strength testing because of their high accuracy. However, their clinical utility is limited by their cumbersome design, stationary nature, and high cost [10].

Recently, exoskeleton-based rehabilitation robotics has emerged as a cutting-edge intervention for restoring function in patients with neurological sequelae [11-13]. A hallmark of these systems is the capacity to autonomously deliver highly repeatable, standard rehabilitation exercises, dramatically reducing the physical load on therapists [14]. Exoskeleton rehabilitation robots have emerged as a promising intervention for neurological recovery. Current hot topics include upper limb goal-directed training, lower limb motion control strategies, and exoskeleton rehabilitation robot power [15]. A pilot randomized controlled trial of the REX robotic exoskeleton (REX Bionics PLC), a self-stabilizing lower limb device that supports patients in a hands-free upright standing position without crutches or walking frames, included 3 supplementary training programs—standing balance activities (pushing a Bobath ball and multidirectional reaching), thera-band elastic band resistance training following an upper limb PNF diagonal spiral pattern, and lower limb function training (single-leg weight bearing, lateral and repeated strides, squatting, and alternating strides)—to promote balance and lower limb motor recovery in patients with subacute stroke [16]. The integration of series elastic actuators in lower limb exoskeletons has significantly advanced robotic rehabilitation, particularly for enhancing human-robot interaction safety and motion control [17]. Prasad et al [11] proposed 4 conceptual configurations for cable-driven exoskeletons, which address key power-related challenges inherent in traditional actuation systems.

Exoskeleton rehabilitation robots overcome some limitations of traditional clinical assessment by providing more objective, sensitive, reliable, and efficient measurements [18]. Several groups have proposed using robotic devices to assess proprioception and motor function [18-20], such as gait and joint position sense; the Lokomat, for example, has been used to assess lower limb joint position sense [19], and the bilateral Harmony upper limb exoskeleton can measure joint angles across the shoulder, forearm, and wrist [21]. An intelligent assessment study using a lower limb exoskeleton for patients with stroke reported predictive accuracies of 85% for the 6-minute walk distance and 93% for the Fugl-Meyer assessment lower extremity score, outperforming traditional manual measurement in efficiency and objectivity [22]. Although these studies establish that exoskeleton and robotic platforms can quantify kinematic and functional variables, the agreement of an exoskeleton's isokinetic torque output with a criterion standard dynamometer has, to our knowledge, not been reported. The UGO rehabilitation exoskeleton is seeing growing adoption in clinical practice for treating patients with lower limb motor dysfunction resulting from stroke, spinal cord injury, or peripheral nerve injury [23]. Designed to provide high-frequency, repetitive, standardized gait training, the device additionally features isokinetic muscle strength testing as a key diagnostic function. Unlike conventional isokinetic dynamometers, it is equipped with force sensors at both the hip and knee joints on each side, so that bilateral testing does not require positional changes—a design that theoretically makes testing more convenient and efficient. Although the UGO exoskeleton is widely used and has recognized theoretical benefits, the measurement properties of its strength assessment capability remain to be established. These properties are conventionally established in 2 distinct steps. In a previous study, we examined the intratester and intertester reliability of the device (an internal, device-only property—whether it produces reproducible values) [24]. The present study addresses the separate and complementary question of concurrent validity (an external property—whether the device agrees with an accepted reference standard). Reliability is a prerequisite for, but is conceptually and statistically independent of, validity: a device may be reliable yet invalid. The 2 studies drew on an overlapping healthy participant pool but used entirely different analyses and report no duplicate results; the present work provides the first agreement-based, reference standard comparison of the UGO's isokinetic torque function.

The aim of this study was to determine the concurrent validity of isokinetic hip and knee peak torque measurements obtained with the UGO rehabilitation exoskeleton against those obtained with the Humac Norm dynamometer in healthy young adults, using both agreement and association analyses. We hypothesized that peak torque measured by the 2 devices would show at least moderate, statistically significant correlations and clinically moderate to strong agreement for the major muscle actions.

Methods

Study Setting

This study was carried out in the Biomechanical Rehabilitation Engineering Laboratory of the Department of Rehabilitation Medicine at the First Affiliated Hospital of Sun Yat-sen University. Participants were randomly recruited from interns within the department. Measurements were taken by the same physical therapist, who had >5 years of clinical experience and was proficient in operating both pieces of equipment, using the 2 isokinetic dynamometers. This tester had previously undergone

XSL•FO
RenderX

training and passed evaluations related to the operation of isokinetic muscle strength testing equipment for both devices.

Sample Size Calculation

According to a previous study [25] that investigated the correlation between a Biodex isokinetic dynamometer (Evome Medical Technologies Inc) and a hand-held dynamometer in male football players (r=0.322-0.617), a sample of 17 participants provides approximately 80% power to detect a correlation of r=0.60 at α=.05 (2-tailed) in this study, as calculated using G*Power. To draw a stronger conclusion, the sample size was finally determined to be 20.

Recruitment

This study recruited participants via 2 approaches. First, a recruitment announcement was posted in the WeChat (Tencent Holdings) group for interns within the department, followed by on-site screening to identify eligible candidates. Baseline information such as age, sex, height, weight, BMI, and the dominant lower limb were collected for each participant. Researchers contacted potential participants who satisfied the testing equipment requirements (height between 150 cm and 190 cm, weight not exceeding 100 kg) to confirm their willingness to participate [23].

Inclusion Criteria

The inclusion criteria were as follows: (1) aged 20 to 30 years, (2) absence of any history of cardiovascular or respiratory diseases, (3) lack of a history of lower limb musculoskeletal conditions (eg, hip or knee joint surgery, ligament injury, or muscle strain or sprain) that would affect the test, and (4) absence of high-intensity physical activities during the past week.

Ethical Considerations

The study was approved by the medical ethics committee of the First Affiliated Hospital of Sun Yat-sen University in Guangzhou, China ({2020}430) and conducted according to the Declaration of Helsinki. This was a safe and noninvasive experiment. Prior to the commencement of the experiment, a member of the research team elucidated the experimental procedures to all participants and obtained their informed consent. Furthermore, throughout the experiment, the participants had the option to withdraw from the experiment at any time without providing any reason.

Instrument

Two isokinetic dynamometers were used in this study: the UGO rehabilitation exoskeleton (Chengtian) and the Humac Norm (CSMi Solutions) (Figure 1). The UGO lower limb exoskeleton robot is widely used as a research and clinical platform for validating gait planning algorithms. Notably, its capacity for isokinetic concentric torque testing at the hip and knee joints constitutes one of its most important evaluation functions, providing reliable and functionally relevant measurements in a natural upright posture. Many studies have confirmed that the Humac Norm is a highly reliable and accurate isokinetic dynamometer [10,26-28]. Identical angular velocity (60°/s) and joint range-of-motion settings were configured on both dynamometers; body posture, however, differed between the devices by design (UGO: upright; Humac Norm: seated for the knee, supine for the hip).

Figure 1. Isokinetic muscle strength testing of participants using the UGO rehabilitation exoskeleton (Chengtian) and Humac Norm isokinetic dynamometer (CSMi Solutions).



XSL•FO
RenderX

Parameter

This study primarily focused on collecting peak torque data for the knee and hip flexor and extensor muscle groups during concentric contractions using the 2 isokinetic dynamometers. The isokinetic dynamometers measure muscle strength by collecting mechanical signals within the range of motion through sensors and calculating the measurements using their respective analysis software. The parameter that was recorded was absolute peak torque (in Newton-meters). Peak torque has been defined as the maximum force produced by a muscle during voluntary contraction within the range of motion under a given set of conditions [29,30].

Procedure

To minimize the potential influence of biological rhythms on the results, the 2 tests were conducted at the same time on 2 consecutive days. Participants were instructed to abstain from strenuous physical activity and refrain from consuming any nutritional supplements or functional beverages that could enhance physical performance, starting 48 hours before the tests and continuing until their completion. They were permitted to maintain their usual daily activities and diet. Prior to the formal test, participants performed 3 minutes of lower limb muscle stretching, followed by 10 knee and hip flexion and extension repetitions at a submaximal voluntary contraction level to familiarize themselves with the testing procedure. After the warm-up, each participant rested for 5 minutes to prevent fatigue. The testing sequence involved assessing the dominant side first, followed by the nondominant side after a 2-minute rest, and the knee and hip tests were performed 5 minutes apart.

To ensure consistency and comparability, in the knee test, both devices used the lateral femoral condyle as the axis of motion, with a movement speed of 60°/s and a range of 0° to 70° (0°: the fully extended position of the knee joint) for knee testing. For the UGO device, thigh and calf lengths of participants were adjusted so that the axis of movement was aligned with the lateral femoral condyle in a seated posture, and their trunk, distal femur, proximal tibia, and foot were securely fixed to avoid compensatory forces. Testing was performed in an upright position, with the external frame fixed first, and the muscle strength test of the flexion muscle group was started from the extension position of the knee joint. Each movement was followed by a cue tone, followed by movement in the opposite direction. Verbal instruction for peak torque was provided as “flexion and extension as hard and as fast as you can” to ensure maximal effort during each action.

The greater trochanter of the femur was used as the axis of motion, with a movement speed of 60°/s and a range of −10° to 60° (0° being the neutral position of the hip) in hip testing. The position was the same as that for knee testing and started from the extension position. During the UGO isokinetic muscle strength test, repositioning was not required when switching between test sides or joint angles, which saved therapists both time and operational steps. The bilateral test took 5 minutes to complete for each joint, during which 5 consecutive flexion and extension movements were executed, and peak torque values were recorded after each movement.

In accordance with the standardized testing protocol of Humac Norm, participants were seated on the isokinetic dynamometer with their hip joints flexed at 85°. The tension sensor’s rotation angle was set at 40°, and its height was adjusted to 8 cm. The rotational axis of the knee was placed in line with the dynamometer axis of rotation, and 0° was determined as 0° knee extension. The resistance arm pad was secured just proximally to the medial malleolus, so that movement of the ankle was not constricted. To minimize compensatory trunk movements, straps were used to secure the participant’s trunk, distal femur, and distal tibia. During the test, participants were required to cross their arms in front of their chest to prevent compensatory force generation by gripping the seat. The hip test was performed in the supine position; the tension sensor’s rotation angle was set at 0° and its height was set at 9 cm. The resistance arm pad was secured just distal to the femur, and a fixation band was used to fix the pelvis. Upon program initiation, 5 consecutive flexion and extension movements were completed.

Statistical Analysis

Descriptive statistics were used to analyze the demographic characteristics of the participants in this study. A 1-sample Kolmogorov-Smirnov test was used to test whether the data were normally distributed. The Pearson correlation coefficient (r) or Spearman rank correlation coefficient (ρ) was used to determine concurrent validity. Here, r or ρ values between 0.00 and 0.19, between 0.20 and 0.39, between 0.40 and 0.59, between 0.60 and 0.79, and ≥0.80 represented very weak, weak, moderate, strong, and very strong correlations, respectively [31]. A scatter diagram was constructed to evaluate agreement between the UGO and HUMAC measurements and to identify any systematic bias [10]. A Bland-Altman plot was used to compare the mean difference and the 95% limits of agreement (LoA) between the 2 devices. The upper and lower 95% LoA were calculated as the mean difference±1.96×SD. All statistical analyses were performed using SPSS (version 20.0; IBM Corp), with a significance level set at $P<.05$.

Results

Overview

Twenty healthy young adults were recruited for this study. All participants cooperated well during the tests, and each participant exerted maximal effort. The demographic characteristics of the participants are provided in Table 1. Of the 20 participants, 9 (45%) were male, 11 (55%) were female, and all 20 (100%) were right-lower-limb dominant.

XSL•FO
RenderX

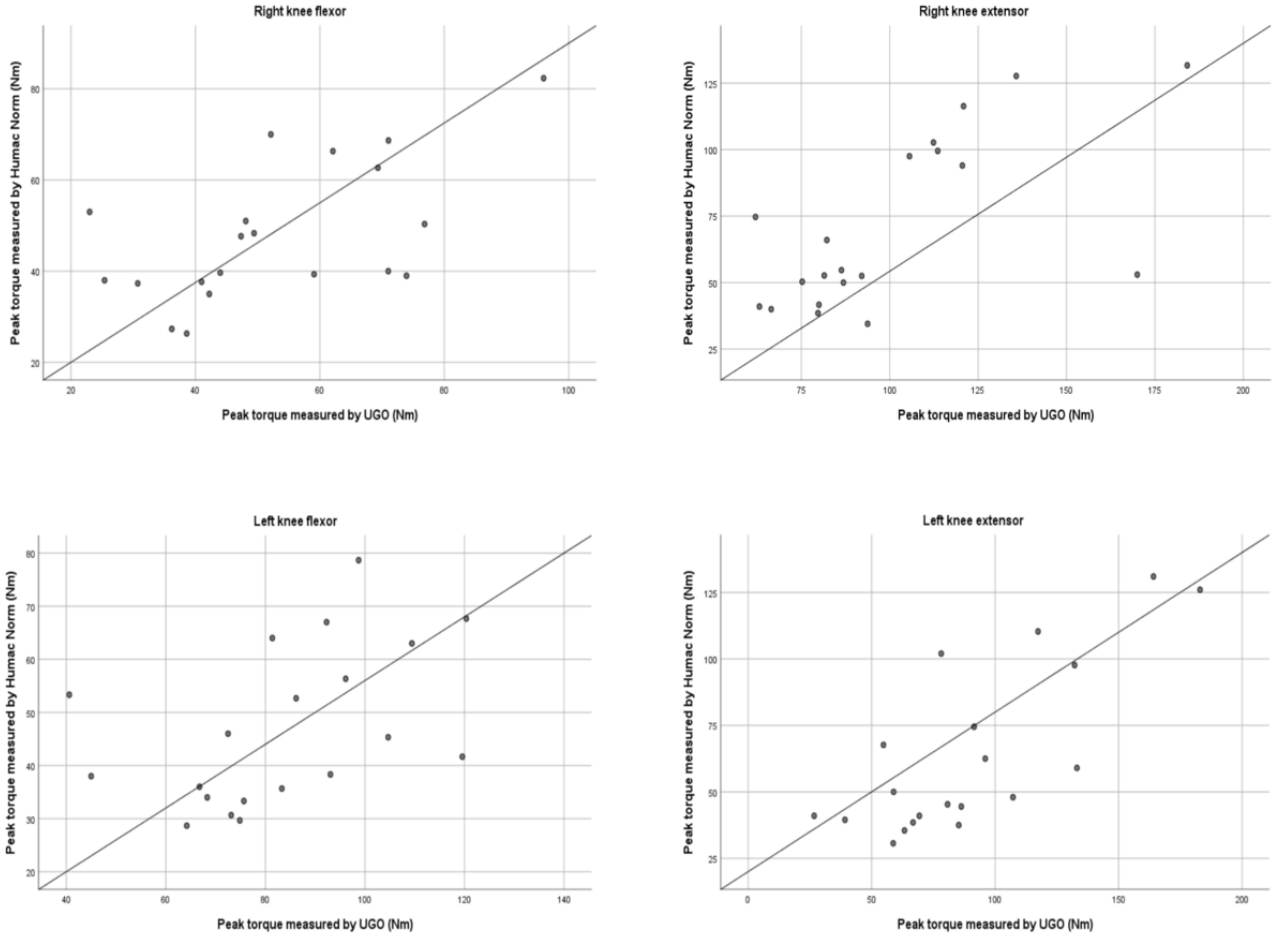
Table 1. Demographic characteristics of the study participants (N=20).

Variables	Values
Age (years), mean (SD; range)	22.70 (0.98; 21.0-25.0)
Height (cm), mean (SD; range)	167.65 (7.87; 156.0-180.0)
Weight (kg), mean (SD; range)	62.08 (12.16; 47.0-87.0)
BMI (kg/m^2), mean (SD; range)	21.90 (2.75; 18.6-28.4)
Sex, n (%)	
Male	9 (45)
Female	11 (55)
Dominant side, n (%)	
Left	0 (0)
Right	20 (100)

Concurrent Validity of Isokinetic Knee Strength Measurements

On the dominant (right) side, the correlation coefficients for peak torque of knee extension and flexion were 0.662 (P=.001) and 0.547 (P=.01), respectively, with 95% LoA ranging from −22.89 to 82.21 Nm and from −26.27 to 36.01 Nm, respectively. Specifically, there was a strong correlation for peak torque during knee extension and a moderate correlation for peak torque during knee flexion. On the nondominant (left) side, the correlation coefficients for peak torque of knee extension and flexion were 0.768 (P<.001) and 0.347 (P=.12), respectively, with 95% LoA ranging from −24.60 to 75.65 Nm and from −2.40 to 75.02 Nm, respectively. The peak torque of knee extension exhibited a strong correlation, while that of knee flexion showed a weak correlation. These results are demonstrated in the scatter plot (Figure 2) and the Bland-Altman plot assessing agreement between the 2 devices (Figure 3). Detailed results of the concurrent validity analysis of isokinetic knee strength measurements are shown in Table 2.

Figure 2. Scatter plot of knee isokinetic peak torque measured by the UGO exoskeleton and Humac Norm isokinetic dynamometer.



XSL•FO
RenderX

Figure 3. Bland-Altman plot of knee isokinetic peak torque measured by the UGO exoskeleton and Humac Norm isokinetic dynamometer.

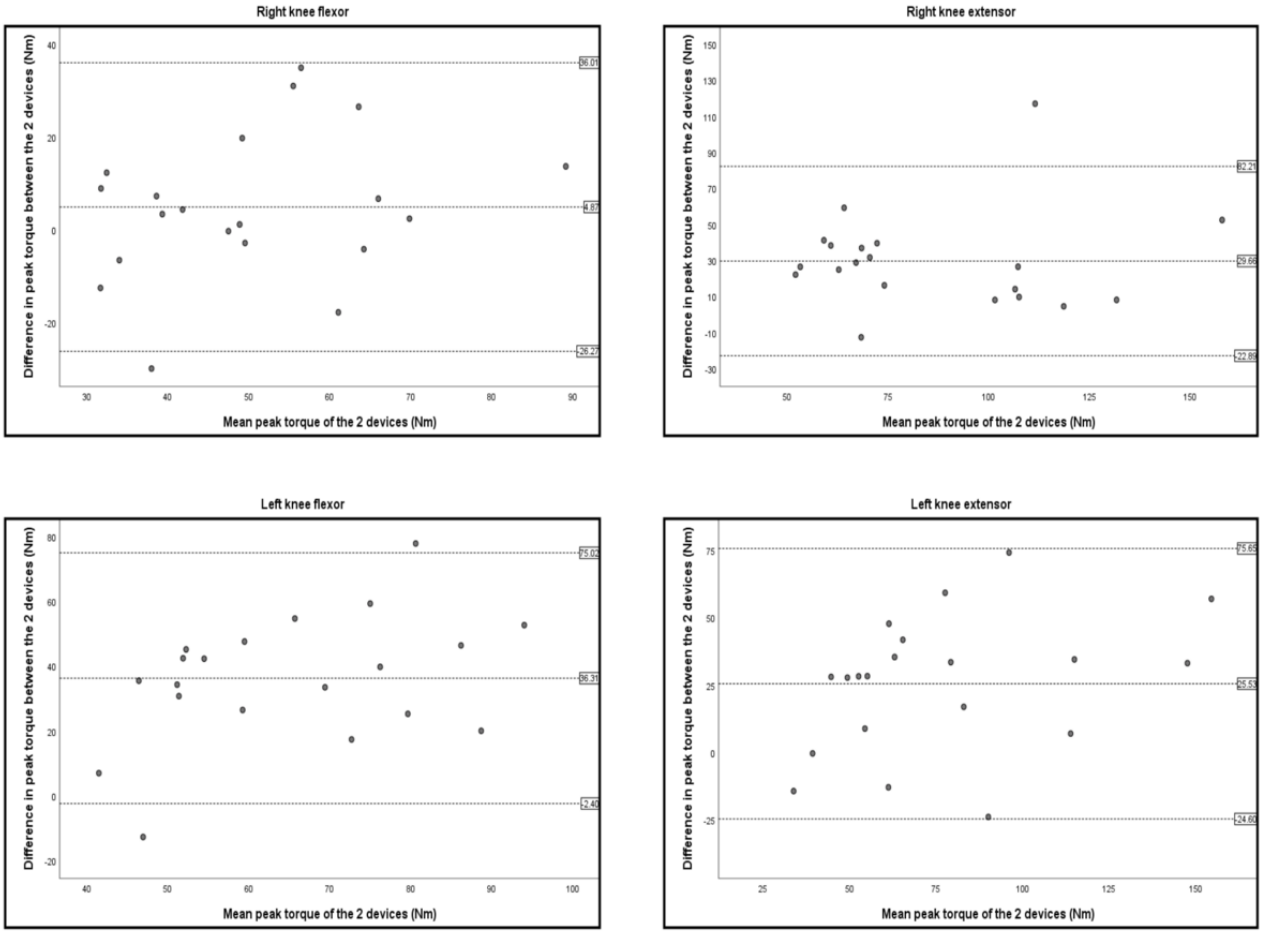


Table 2. Concurrent validity of 2 isokinetic dynamometers for measuring peak torque during isokinetic knee strength testing (N=20).

Variables	Peak torque (Nm), mean (SD)		Bias (Nm; 95% limits of agreement[a])	*r*	*P* value[b]
	UGO exoskeleton	Humac Norm			
Dominant side					
Extensor	101.14 (32.46)	71.31 (31.09)	4.87 (−22.89 to 82.21)	0.662	*.001*
Flexor	52.89 (18.64)	49.36 (16.03)	29.66 (−26.27 to 36.01)	0.547	*.01*
Nondominant side					
Extensor	89.39 (39.11)	64.44 (31.12)	36.31 (−24.60 to 75.65)	0.768	*<.001*
Flexor	83.02 (21.16)	49.21 (17.71)	25.53 (−2.40 to 75.02)	0.347	.12

[a]Bland-Altman means bias (UGO–Humac Norm) with 95% limits of agreement.

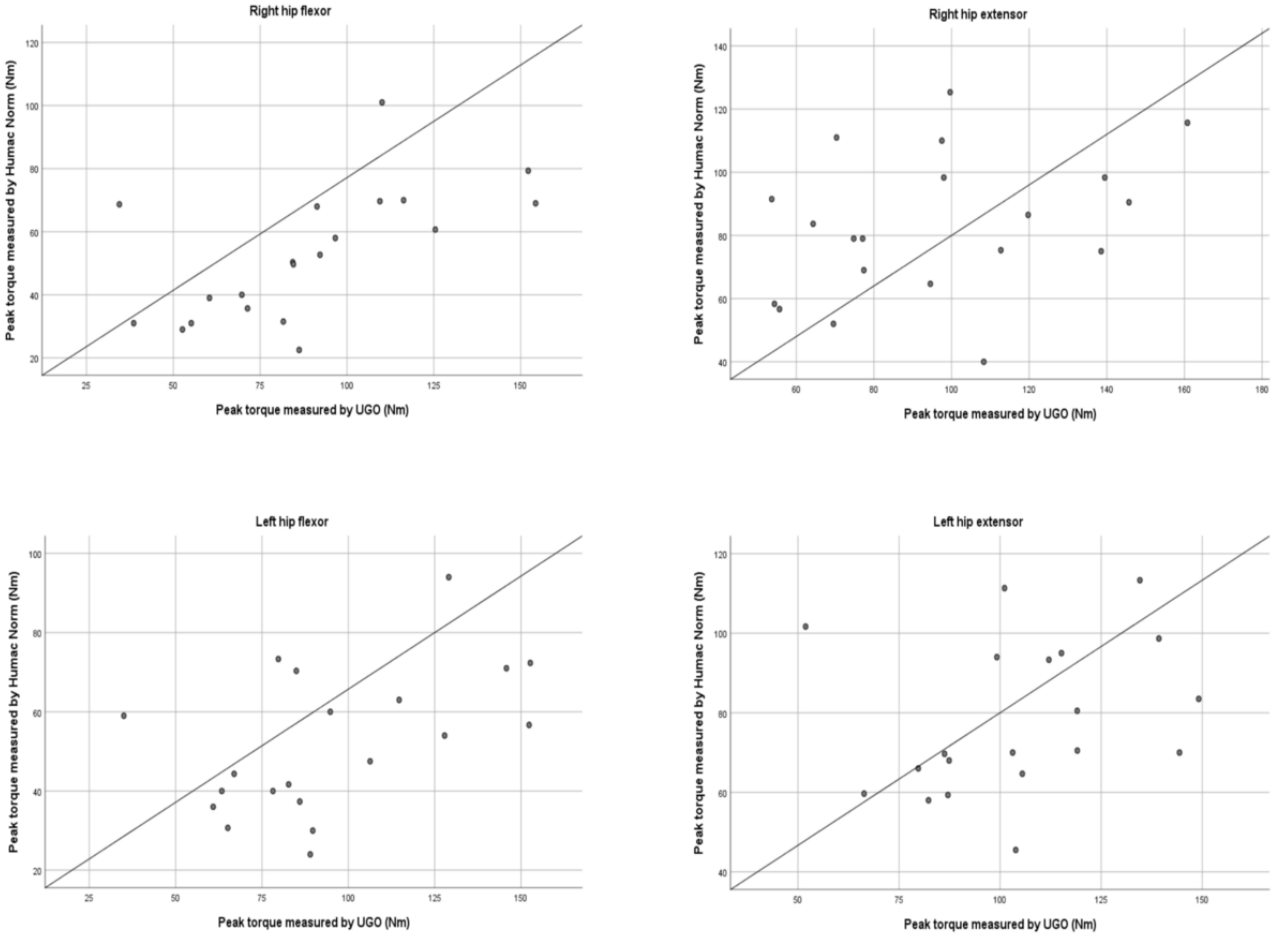
[b]Italicized values indicate significant correlations (*P*<.05).

Concurrent Validity of Isokinetic Hip Strength Measurements

The correlation coefficients for peak torque of hip extension were 0.337 (*P*=.14) and 0.391 (*P*=.20) on the right and left sides, respectively, with 95% LoA ranging from −51.57 to 76.80 Nm and from −28.19 to 79.59 Nm, respectively, and showed no significant correlation. The correlation coefficients for peak torque of hip flexion were 0.615 (*P*=.003) and 0.443 (*P*=.04), on the right and left sides, respectively, with 95% LoA ranging from −15.52 to 86.57 Nm and from −12.53 to 98.50 Nm, respectively. The peak torque of right hip flexion exhibited a strong correlation, while that of left hip flexion showed a moderate correlation. These results are demonstrated in the scatter plot (Figure 4) and the Bland-Altman plot assessing agreement between the 2 devices (Figure 5). Detailed results of the concurrent validity analysis of isokinetic hip strength measurements are shown in Table 3.

XSL•FO
RenderX

Figure 4. Scatter plot of hip isokinetic peak torque measured by the UGO exoskeleton and Humac Norm isokinetic dynamometer.



XSL•FO
RenderX

Figure 5. Bland-Altman plot of hip isokinetic peak torque measured by the UGO exoskeleton and Humac Norm isokinetic dynamometer.

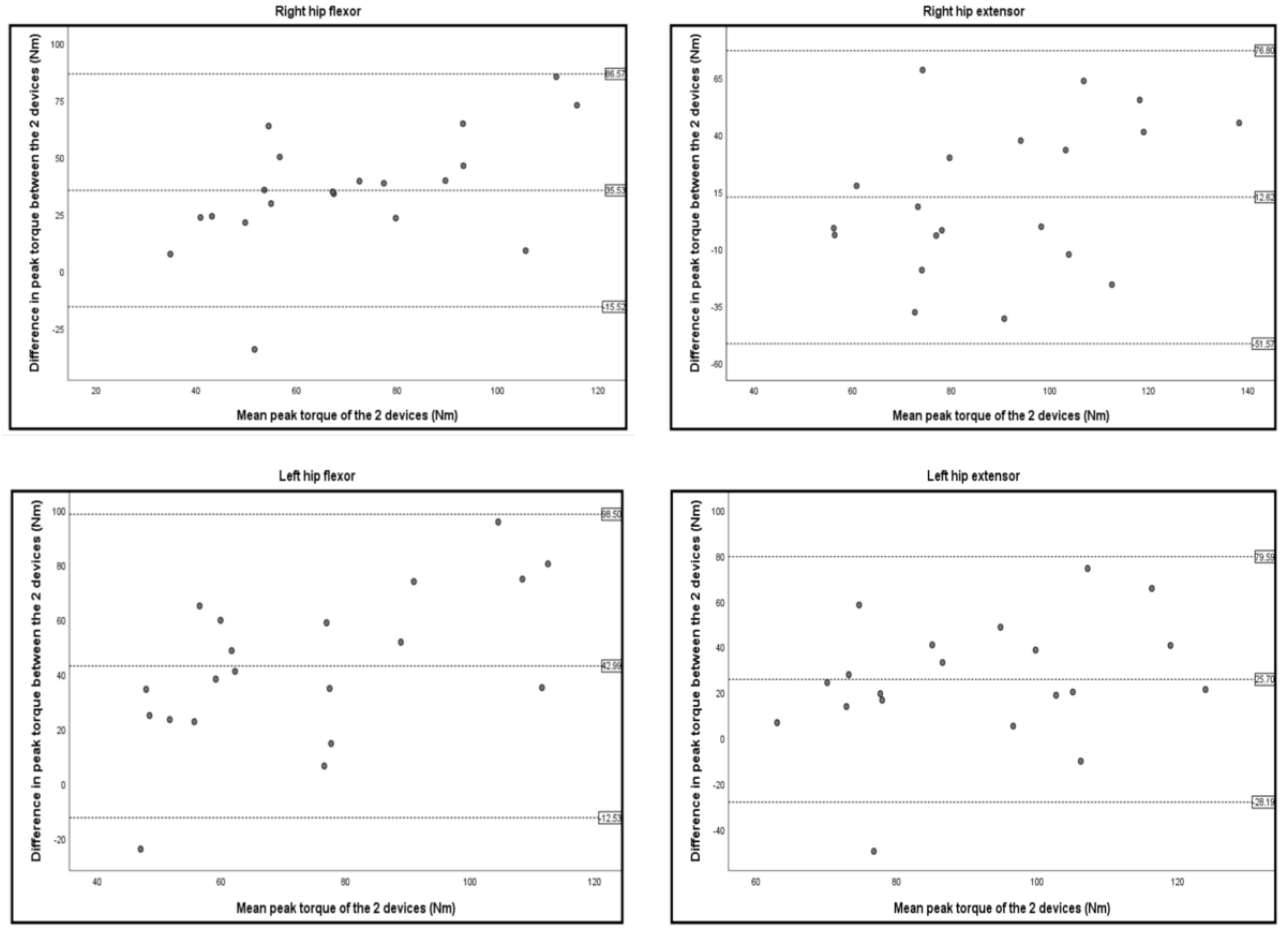


Table 3. Concurrent validity of 2 isokinetic dynamometers for measuring peak torque during isokinetic hip strength testing (N=20).

Variables	Peak torque (Nm), mean (SD)		Bias (Nm; 95% limits of agreement[a])	*r*	*P* value[b]
	UGO exoskeleton	Humac Norm			
Dominant side					
Extension	95.61 (31.65)	82.80 (22.13)	35.53 (−51.57 to 76.80)	0.337	.14
Flexion	88.17 (32.29)	53.06 (20.02)	12.62 (−15.52 to 86.57)	0.615	*.003*
Nondominant side					
Extension	103.70 (25.25)	80.62 (20.77)	42.99 (−28.19 to 79.59)	0.391	.20
Flexion	94.24 (32.20)	53.29 (18.31)	25.70 (−12.53 to 98.50)	0.443	*.04*

[a]Bland-Altman means bias (UGO–Humac Norm) with 95% limits of agreement.

[b]Italicized values indicate significant correlations (*P*<.05).

Discussion

Principal Findings

To the best of our knowledge, this study represents the first attempt to investigate the concurrent validity of isokinetic muscle strength testing in the context of lower limb exoskeleton rehabilitation robots. In this cross-sectional validation study, the UGO exoskeleton showed moderate to strong concurrent validity—with significant correlations of at least moderate strength and Bland-Altman agreement—for the bilateral knee extensors and the hip flexors, but not for the left knee flexors or bilateral hip extensors, which showed nonsignificant associations and greater systematic bias. Across all muscle actions, the UGO systematically overestimated peak torque relative to the Humac Norm.

Previous studies comparing torque measurements obtained with hand-held dynamometers and isokinetic dynamometers have found that tests of the knee extensor and flexor muscles show moderate to strong correlations [32], while tests of internal and external rotation of the shoulder joint show weak correlations

XSL•FO
RenderX

[33]. O'Neill et al [10] found that the C-station (Fysiometer) appeared to provide valid measurements compared with an isokinetic dynamometer for plantar flexor isometric strength. However, these studies only compared devices with an isokinetic dynamometer during isometric contraction tests. Our study compared 2 devices during isokinetic knee and hip strength testing. We found that the UGO showed strong correlations for both knee extensor torque measurements, a moderate correlation for right knee flexor torque, and a nonsignificant correlation for left knee flexor torque compared to the reference standard, the Humac Norm. In hip testing, the correlation of the flexor torque values was stronger than that of the extensor torque values on both sides. The underlying reasons for these differences included 4 factors. First, the test positions differed. During the UGO isokinetic strength test, the knee and hip flexion and extension muscles were tested in the upright position. In contrast, during the Humac Norm isokinetic strength test, the knee was positioned in the seated posture, while the hip was placed in the supine position. Diverse test positions can influence force production, thereby leading to varying torque values. During the upright position test, participants consistently reported that exerting force for hip flexion combined with knee extension required less effort than hip extension combined with knee flexion. Second, the torque signal acquisition and calculation formulas differed between the 2 devices. Despite the sampling rate of both torque sensors being 100 Hz [23,34], the torque sensor of the UGO exoskeleton has a measurement range of 150 Nm, a resolution of 0.05 Nm, and a force-sensing accuracy of 0.3% [23]. The resolution and accuracy of the Humac Norm isokinetic dynamometer force sensor have not been reported in the existing literature, but Habets et al [35] reported that the Humac Norm dynamometer showed good to excellent reliability in strength tests of the knee and shoulder joints. The disparities in the mechanical structure and software of the 2 devices may likewise affect the calculation of the lever arm and the torque value of the generated force [27]. As torque is obtained by multiplying the measured force by the perpendicular lever-arm length ($\tau=F\times d\perp$), such device-related discrepancies have been shown to produce systematic torque differences of around 8 Nm between hand-held and isokinetic dynamometers [36]. Third, the number of torque sensors differed between the 2 devices. The UGO dynamometer has 4 torque sensors, and the Humac Norm has only 1 sensor. We hypothesize that the presence of numerous force sensors may increase the potential for error associated with torque calculation [37]. Finally, factors related to the patient's maximum voluntary force contraction may have contributed to the differences. In knee testing, the correlation coefficient for peak hamstring torque appeared to be lower than that for peak quadriceps torque [25]. This finding may be partly explained by the characteristically lower test-retest reliability of hamstring strength measurements relative to those of the quadriceps [38].

In hip testing, the complexity and standardization of hip joint testing present additional challenges. Hip movements are more difficult to isolate and standardize than knee tasks because of greater degrees of freedom, pelvic tilt variability, and compensatory trunk or lumbar involvement. The UGO, as a functional exoskeleton, may capture more multijoint, closed-chain dynamics, while Humac Norm emphasizes open-chain, single-plane isokinetic testing, leading to divergent torque profiles. Hip extension involves strong gluteal or hamstring recruitment with potential antagonist cocontraction (eg, flexors and abdominals), which exoskeletons may detect differently from laboratory dynamometers. Participants may unconsciously alter their movement strategies (eg, trunk leaning) on the UGO, inflating or deflating measured torque relative to Humac Norm. Interestingly, we found that the measurement of peak torque exhibited significantly higher correlation coefficients for the anterior muscle group than the posterior muscle group.

For patients, a significant inconvenience arises from standard isokinetic muscle strength testing, which requires repositioning the body when assessing multiple joints in the lower limbs. Although the UGO's isokinetic muscle strength test results **e**xhibit certain discrepancies compared to those of the Humac Norm reference standard, it offers superior clinical convenience. In particular, the isokinetic muscle strength testing function of the UGO device enables therapists to quantitatively assess changes in patients' lower limb function during clinical treatment and across treatment courses (preintervention and postintervention). The exoskeleton robot establishes a closed-loop rehabilitation training protocol for patients, allowing the rehabilitation therapist to adjust the training program based on the results of muscle strength testing.

Overall, the UGO's isokinetic muscle strength test is more convenient and maneuverable than traditional isokinetic dynamometers. This device not only reduces the operational requirements of therapist-operated devices and saves time, but it also alleviates patient discomfort by minimizing repositioning—especially for those with mobility difficulties. Therefore, the isokinetic muscle strength assessment capability of the UGO device enables therapists to evaluate lower limb motor function and implement targeted treatment strategies more effectively in clinical rehabilitation settings.

Limitations

This study has several limitations. First, the inherent hardware differences between the devices inevitably lead to discrepancies in test posture and lever-arm length. Second, device-specific error sources for the UGO include reliance on estimated segment lengths for lever-arm or torque computation, the distributed 4-sensor architecture, and a proprietary, nontransparent torque calculation algorithm; for the reference device, fatigue status before Humac Norm testing was assessed only subjectively, which may partly explain the relatively lower recorded torques. Third, the relatively small sample size in this study may limit the statistical power of the analyses and preclude definitive conclusions. Fourth, only healthy young adults were studied; findings cannot be extrapolated to patients with neurological impairment, in whom spasticity, weakness, altered tone, impaired trunk control, and reduced range of motion may alter both torque magnitude and device-reference agreement. Finally, as the device testing order was fixed, potential order effects （learning or fatigue） cannot be fully excluded, although the 24-hour interdevice interval and activity restriction were intended to mitigate them.

XSL•FO
RenderX

Conclusions

In healthy young adults, the UGO exoskeleton demonstrated moderate to strong concurrent validity for peak torque of the bilateral knee extensors and the hip flexors, but not for the nondominant knee flexors and bilateral hip extensors, and it systematically overestimated peak torque relative to the Humac Norm. The device may serve as a convenient torque assessment tool for selected muscle actions; however, its use for the knee flexors and hip extensors is not yet supported, and validation in the intended clinical population is required in future studies.

Acknowledgments

The authors thank all the patients who participated in this study. The authors also thank all the colleagues in the Department of Rehabilitation Medicine at the First Affiliated Hospital of Sun Yat-sen University, China.

Funding

The authors declare that this research received financial support from the National Natural Science Foundation of China (grant 82102649).

Data Availability

The data are available from the corresponding author upon reasonable request.

Authors' Contributions

HX, JLZ, and CHW conceived the study; HX and JLZ performed data collection and statistical analysis; HX, MHD, and JLZ drafted the manuscript; JLZ, PMC, and SSMN revised the manuscript; and HX assisted in recruiting participants. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Grimmer M, Riener R, Walsh CJ, Seyfarth A. Mobility related physical and functional losses due to aging and disease - a motivation for lower limb exoskeletons. J Neuroeng Rehabil 2019 Jan 03;16(1):2 [FREE Full text] [doi: 10.1186/s12984-018-0458-8] [Medline: 30606194]
2. Wang D, Wu Y, Yu H. State of the art of brain function detection technologies in robot-assisted lower limb rehabilitation. Brain Connect 2024 Oct;14(8):401-417. [doi: 10.1089/brain.2024.0005] [Medline: 39001823]
3. Tang C, Zhang Y, Zhou T, Yin R, Zhang L, Ren C, et al. Effects of lower limb exoskeleton robot-assisted rehabilitation on motor function, activities of daily living, and electromyography in hemiplegic stroke: a randomized controlled trial. BMC Neurol 2026 May 18;26(1):440 [FREE Full text] [doi: 10.1186/s12883-026-04975-6] [Medline: 42144590]
4. Chen YN, Wu YN, Yang BS. The neuromuscular control for lower limb exoskeleton- a 50-year perspective. J Biomech 2023 Sep;158:111738. [doi: 10.1016/j.jbiomech.2023.111738] [Medline: 37562276]
5. Flansbjer UB, Downham D, Lexell J. Knee muscle strength, gait performance, and perceived participation after stroke. Arch Phys Med Rehabil 2006 Jul;87(7):974-980. [doi: 10.1016/j.apmr.2006.03.008] [Medline: 16813786]
6. Muehlbauer T, Gollhofer A, Granacher U. Associations between measures of balance and lower-extremity muscle strength/power in healthy individuals across the lifespan: a systematic review and meta-analysis. Sports Med 2015 Dec;45(12):1671-1692 [FREE Full text] [doi: 10.1007/s40279-015-0390-z] [Medline: 26412212]
7. Tayfur B, Johnson AK, Palmieri-Smith RM. Compromised quadriceps and hamstring force control, not maximal strength, is associated with gait biomechanics at 9 months following anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 2025 Dec;33(12):4466-4474. [doi: 10.1002/ksa.70129] [Medline: 41147842]
8. Hsu WH, Hsu WB, Lin ZR, Chang SH, Fan CH, Kuo LT, et al. Effects of 24 weeks of a supervised walk training on knee muscle strength and quality of life in older female total knee arthroplasty: a retrospective cohort study. Healthcare (Basel) 2023 Jan 26;11(3):356 [FREE Full text] [doi: 10.3390/healthcare11030356] [Medline: 36766931]
9. Bohannon RW. Considerations and practical options for measuring muscle strength: a narrative review. Biomed Res Int 2019;2019:8194537 [FREE Full text] [doi: 10.1155/2019/8194537] [Medline: 30792998]
10. O'Neill S, Weeks A, Nørgaard JE, Jorgensen MG. Validity and intrarater reliability of a novel device for assessing plantar flexor strength. PLoS One 2023 May 31;18(3):e0282395 [FREE Full text] [doi: 10.1371/journal.pone.0282395] [Medline: 37000780]
11. Prasad R, El-Rich M, Awad MI, Hussain I, Jelinek HF, Huzaifa U, et al. A framework for determining the performance and requirements of cable-driven mobile lower limb rehabilitation exoskeletons. Front Bioeng Biotechnol 2022 Jun 20;10:920462 [FREE Full text] [doi: 10.3389/fbioe.2022.920462] [Medline: 35795162]

XSL•FO
RenderX

12. Huo C, Shao G, Chen T, Li W, Wang J, Xie H, et al. Effectiveness of unilateral lower-limb exoskeleton robot on balance and gait recovery and neuroplasticity in patients with subacute stroke: a randomized controlled trial. J Neuroeng Rehabil 2024 Dec 05;21(1):213 [FREE Full text] [doi: 10.1186/s12984-024-01493-9] [Medline: 39639336]
13. Zhang Q, Chen Y, Lewek M, Si J, Huang H. Towards adaptive optimal personalization control of robotic hip exoskeleton assistance for individuals with paretic stroke. IEEE Trans Biomed Eng 2026 Apr 17;PP:41996429. [doi: 10.1109/TBME.2026.3685273] [Medline: 41996429]
14. Yang J, Gong Y, Yu L, Peng L, Cui Y, Huang H. Effect of exoskeleton robot-assisted training on gait function in chronic stroke survivors: a systematic review of randomised controlled trials. BMJ Open 2023 Sep 14;13(9):e074481 [FREE Full text] [doi: 10.1136/bmjopen-2023-074481] [Medline: 37709309]
15. Zhang Y, Liu X, Qiao X, Fan Y. Characteristics and emerging trends in research on rehabilitation robots from 2001 to 2020: bibliometric study. J Med Internet Res 2023 May 31;25:e42901 [FREE Full text] [doi: 10.2196/42901] [Medline: 37256670]
16. Zhang Y, Zhao W, Wan C, Wu X, Huang J, Wang X, et al. Exoskeleton rehabilitation robot training for balance and lower limb function in sub-acute stroke patients: a pilot, randomized controlled trial. J Neuroeng Rehabil 2024 Jun 08;21(1):98 [FREE Full text] [doi: 10.1186/s12984-024-01391-0] [Medline: 38851703]
17. Liu H, Zhu C, Zhou Z, Dong Y, Meng W, Liu Q. Synergetic gait prediction and compliant control of SEA-driven knee exoskeleton for gait rehabilitation. Front Bioeng Biotechnol 2024 Jan 26;12:1358022 [FREE Full text] [doi: 10.3389/fbioe.2024.1358022] [Medline: 38344287]
18. Riener R, Nef T, Colombo G. Robot-aided neurorehabilitation of the upper extremities. Med Biol Eng Comput 2005 Jan;43(1):2-10 [FREE Full text] [doi: 10.1007/BF02345116] [Medline: 15742713]
19. Domingo A, Lam T. Reliability and validity of using the Lokomat to assess lower limb joint position sense in people with incomplete spinal cord injury. J Neuroeng Rehabil 2014 Dec 16;11:167 [FREE Full text] [doi: 10.1186/1743-0003-11-167] [Medline: 25516305]
20. Xu D, Kang SH, Lee SJ, Oppizzi G, Zhang LQ. Multi-joint assessment of proprioception impairments poststroke. Arch Phys Med Rehabil 2024 Mar;105(3):480-486. [doi: 10.1016/j.apmr.2023.08.029] [Medline: 37714505]
21. De Oliveira AC, Sulzer JS, Deshpande AD. Assessment of upper-extremity joint angles using Harmony exoskeleton. IEEE Trans Neural Syst Rehabil Eng 2021 May 20;29:916-925. [doi: 10.1109/TNSRE.2021.3074101] [Medline: 33872155]
22. Zhang S, Fan L, Ye J, Chen G, Fu C, Leng Y. An intelligent rehabilitation assessment method for stroke patients based on lower limb exoskeleton robot. IEEE Trans Neural Syst Rehabil Eng 2023;31:3106-3117. [doi: 10.1109/TNSRE.2023.3298670] [Medline: 37490379]
23. Yang W, Zhang J, Zhang S, Yang C. Lower limb exoskeleton gait planning based on crutch and human-machine foot combined center of pressure. Sensors 2020 Dec 16;20(24):7216 [FREE Full text] [doi: 10.3390/s20247216] [Medline: 33339443]
24. Zhao JL, Xie H, Chen PM, Ng SS, Wang CH. Intra-tester and inter-tester reliability of maximum isokinetic hip and knee strength measurements with the exoskeleton in healthy young adults. Front Sports Act Living 2026 Jan 9;7:1696954 [FREE Full text] [doi: 10.3389/fspor.2025.1696954] [Medline: 41586009]
25. Whiteley R, Jacobsen P, Prior S, Skazalski C, Otten R, Johnson A. Correlation of isokinetic and novel hand-held dynamometry measures of knee flexion and extension strength testing. J Sci Med Sport 2012 Sep;15(5):444-450. [doi: 10.1016/j.jsams.2012.01.003] [Medline: 22424705]
26. de Araujo Ribeiro Alvares JB, Rodrigues R, de Azevedo Franke R, da Silva BG, Pinto RS, Vaz MA, et al. Inter-machine reliability of the Biodex and Cybex isokinetic dynamometers for knee flexor/extensor isometric, concentric and eccentric tests. Phys Ther Sport 2015 Feb;16(1):59-65. [doi: 10.1016/j.ptsp.2014.04.004] [Medline: 24913915]
27. Whinton AK, Thompson KM, Power GA, Burr JF. Testing a novel isokinetic dynamometer constructed using a 1080 Quantum. PLoS One 2018 Jul 20;13(7):e0201179 [FREE Full text] [doi: 10.1371/journal.pone.0201179] [Medline: 30028864]
28. Fouasson-Chailloux A, Daley P, Menu P, Gadbled G, Bouju Y, Gautier G, et al. Use of hand hydraulic dynamometers as an overall evaluation of the upper-limb weakness in patients with thoracic outlet syndrome. Front Neurol 2022 Aug 15;13:919312 [FREE Full text] [doi: 10.3389/fneur.2022.919312] [Medline: 36046630]
29. Feiring DC, Ellenbecker TS, Derscheid GL. Test-retest reliability of the biodex isokinetic dynamometer. J Orthop Sports Phys Ther 1990;11(7):298-300. [doi: 10.2519/jospt.1990.11.7.298] [Medline: 18796902]
30. Jaric S. Muscle strength testing: use of normalisation for body size. Sports Med 2002;32(10):615-631. [doi: 10.2165/00007256-200232100-00002] [Medline: 12141882]
31. Post MW. What to do with "moderate" reliability and validity coefficients? Arch Phys Med Rehabil 2016 Jul;97(7):1051-1052. [doi: 10.1016/j.apmr.2016.04.001] [Medline: 27095143]
32. Muff G, Dufour S, Meyer A, Severac F, Favret F, Geny B, et al. Comparative assessment of knee extensor and flexor muscle strength measured using a hand-held vs. isokinetic dynamometer. J Phys Ther Sci 2016 Sep;28(9):2445-2451 [FREE Full text] [doi: 10.1589/jpts.28.2445] [Medline: 27799667]
33. Loepfe JG, Nagelli CV, Larson DR, Tagliero AJ, Cherny CE, Johnson AJ. Reliability and validity of measuring shoulder internal and external rotation strength in volleyball specific positions. Int J Sports Phys Ther 2025 Apr 1;20(4):532-541. [doi: 10.26603/001c.132170] [Medline: 40182905]

XSL•FO
RenderX

34. Luo H, Xiong Y, Zhu M, Wei X, Tao X. Integrated wearable system for monitoring skeletal muscle force of lower extremities. Sensors (Basel) 2024 Jul 22;24(14):4753 [FREE Full text] [doi: 10.3390/s24144753] [Medline: 39066149]
35. Habets B, Staal JB, Tijssen M, van Cingel R. Intrarater reliability of the Humac NORM isokinetic dynamometer for strength measurements of the knee and shoulder muscles. BMC Res Notes 2018 Jan 10;11(1):15 [FREE Full text] [doi: 10.1186/s13104-018-3128-9] [Medline: 29321059]
36. Garcia MA, Fonseca DS, Souza VH. Handheld dynamometers for muscle strength assessment: pitfalls, misconceptions, and facts. Braz J Phys Ther 2021 Oct 3;25(3):231-232 [FREE Full text] [doi: 10.1016/j.bjpt.2020.09.003] [Medline: 33077344]
37. Edouard P, Codine P, Samozino P, Bernard PL, Hérisson C, Gremeaux V. Reliability of shoulder rotators isokinetic strength imbalance measured using the Biodex dynamometer. J Sci Med Sport 2013 Mar;16(2):162-165. [doi: 10.1016/j.jsams.2012.01.007] [Medline: 22749937]
38. Li RC, Wu Y, Maffulli N, Chan KM, Chan JL. Eccentric and concentric isokinetic knee flexion and extension: a reliability study using the Cybex 6000 dynamometer. Br J Sports Med 1996 Jun;30(2):156-160 [FREE Full text] [doi: 10.1136/bjsm.30.2.156] [Medline: 8799603]

Abbreviations

LoA: limits of agreement

Edited by S Munce; submitted 27.Apr.2026; peer-reviewed by B Ozgul, Y Jin; comments to author 03.Jul.2026; revised version received 29.Jul.2026; accepted 14.Sep.2026; published 05.Oct.2026.

Please cite as:
Xie H, Ding MH, Chen PM, Ng SSM, Wang CH, Zhao JL
Concurrent Validity of Isokinetic Hip and Knee Strength Measurements Using the UGO Exoskeleton in Healthy Young Adults: Cross-Sectional Validation Study
JMIR Rehabil Assist Technol 2026;13:e99564
URL: https://rehab.jmir.org/2026/1/e99564
doi:10.2196/99564
PMID:

©Hao Xie, Ming-Hui Ding, Pei-Ming Chen, Shamay S M Ng, Chu-Huai Wang, Jiang-Li Zhao. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 05.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Economic, Organizational, and Research Impact Evaluation of Introducing Robotic Rehabilitation in a Medium-Sized Hospital: Scenario-Based Comparative Analysis

Alessandro Scano[1], MSci, PhD

Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA), National Research Council of Italy (CNR), Milano, Lombardia, Italy

Corresponding Author:
Alessandro Scano, MSci, PhD
Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing (STIIMA)
National Research Council of Italy (CNR)
via Corti 12
Milano, Lombardia, 20133
Italy
Phone: 39 3804666915
Email: alessandro.scano@cnr.it

Abstract

Background: While the clinical efficacy of robotic rehabilitation is well documented, its widespread adoption is often hindered by uncertainty regarding its organizational and systemic impact. A comprehensive evaluation that integrates economic sustainability with research and innovation potential is currently lacking.

Objective: This study evaluates the economic, organizational, and scientific impact of introducing a robotic rehabilitation unit in a medium-sized hospital, comparing it with standard care through a health technology assessment–based simulation model.

Methods: A comparative scenario-based model was developed to simulate 2 equivalent facilities over a 5-year horizon: a "hybrid" clinic (robotics + conventional therapy) and a standard care clinic. The model uses a "workforce multiplier" framework and evaluates efficiency through the cost per point of clinical outcome (Fugl-Meyer Assessment).

Results: Under the simulation hypothesis, robotic integration allows for a 16.1% reduction in cost per patient and an increase in workforce efficiency. Our scenario-based analysis suggests that robotic-assisted rehabilitation units can achieve favorable cost-effectiveness profiles compared with standard manual care, provided that patient throughput remains high. However, robotic systems do not replace the necessity of careful capital budgeting and workflow reorganization.

Conclusions: Robotic rehabilitation should be appraised not merely as a technological cost, but as a multiplier of organizational capacity and a driver of digital innovation. Institutional decision-makers should view robotic adoption as a strategic operational shift rather than a stand-alone guaranteed cost-reduction tool.

(JMIR Rehabil Assist Technol 2026;13:e93572) doi:10.2196/93572

KEYWORDS

cost-effectiveness; health technology assessment; hospital management; robotic rehabilitation; scenario analysis; health systems engineering

Introduction

Neurological and orthopedic conditions represent leading causes of long-term disability worldwide and account for a large proportion of rehabilitation demand in hospital settings [1]. Poststroke rehabilitation plays a critical role in promoting functional recovery and reducing long-term dependency, with evidence supporting the importance of early, intensive, and task-specific therapy [2,3]. Conventional rehabilitation relies heavily on one-to-one therapist-patient interaction and manual guidance. This model presents intrinsic limitations in terms of achievable treatment intensity, reproducibility of exercises, and scalability with increasing patient volumes [4-6]. High-intensity and repetitive motor training is often constrained by therapist fatigue, time availability, and organizational limits [7,8]. From an organizational perspective, rehabilitation departments face growing pressure to increase performance while maintaining quality of care, in a context of constrained budgets and

XSL•FO
RenderX

workforce shortages [9]. Long waiting times, heterogeneous treatment protocols, and limited access to high-dose therapy for patients with severe impairments remain common challenges [2,6]. These factors have stimulated interest in technological solutions capable of supporting therapists and enhancing the efficiency of rehabilitation delivery.

Robotic rehabilitation systems have been developed to deliver high-intensity, repetitive, and quantifiable motor training. Clinical studies have reported improvements in motor function, particularly in patients who have had a stroke, when robotic therapy is used to support conventional treatment [5,8,10]. There is evidence that robotic therapy has a small, statistically significant positive effect [11] or no improvement compared with standard therapy [12]. Meta-analyses suggest that robotic-assisted therapy can increase the amount of practice delivered per session when appropriately integrated into multidisciplinary rehabilitation programs [8,13]. Robotic systems offer potential organizational advantages, including the possibility of automating treatments, reducing therapist physical load, and enabling more standardized and monitored therapy protocols [14,15]. The automatic recording of kinematic and kinetic data further opens opportunities for quantitative outcome assessment and longitudinal monitoring of patient recovery.

However, adoption in routine hospital practice remains limited. High acquisition costs, maintenance expenses, and the need for dedicated infrastructure represent major economic barriers [16-18]. Additional obstacles include the need for specialized staff training, integration into existing clinical workflows, and uncertainty regarding reimbursement mechanisms and long-term sustainability [19,20]. As a result, robotic rehabilitation is often confined to research-oriented centers or highly specialized units rather than being systematically implemented at the hospital level. The majority of published studies on robotic rehabilitation focus on clinical efficacy and motor outcomes, typically through randomized controlled trials or controlled cohort studies [5,8,10]. They provide limited insight into the broader consequences of technology adoption for hospital organization. Only a small number of investigations have addressed economic aspects, often using cost-effectiveness frameworks applied to specific devices or short time horizons [16,21]. Even fewer studies have examined the systemic impact of introducing robotic rehabilitation units in real hospital contexts, considering patient throughput, staff workload, and infrastructural requirements simultaneously [17,19]. Moreover, the potential research and innovation benefits of robotic rehabilitation—such as the generation of structured clinical datasets and the attraction of academic and industrial collaborations—are rarely included in formal evaluations. These dimensions are increasingly relevant for medium-sized hospitals seeking to strengthen their role within clinical research networks and innovation ecosystems [22]. Finally, few studies adopt a comparative scenario-based approach contrasting a hospital implementing robotic rehabilitation with an equivalent hospital providing only standard care. This limits the ability of decision-makers to assess the opportunity cost and evaluate alternative organizational strategies under realistic operational constraints.

This study addresses an evidence gap in the evaluation of rehabilitation technologies by focusing on organizational and economic consequences rather than clinical efficacy alone. The aim is to evaluate the organizational, clinical, economic, and scientific impact of introducing a robotic rehabilitation unit in a medium-sized hospital through a scenario-based comparative model. Two equivalent hospital scenarios are simulated over a multiyear time horizon: clinic A, implementing robotic-assisted rehabilitation alongside conventional therapy, and clinic B, relying exclusively on standard rehabilitation care. This approach aims to provide a multidimensional assessment of robotic rehabilitation adoption at the hospital level.

Methods

Study Design

This study adopted a scenario-based simulation design to evaluate the economic, organizational, and research impact of introducing a robotic rehabilitation unit in a medium-sized hospital. Two hypothetical but structurally equivalent hospitals were modeled and compared over a multiyear time horizon: clinic A, implementing a novel robotic rehabilitation unit integrated within standard rehabilitation services; and clinic B, providing conventional rehabilitation care only. Both scenarios were assumed to operate within the same regional health care system and under identical regulatory and reimbursement frameworks. The model was developed from the hospital perspective, focusing on direct capital and operating costs, organizational processes, and patient throughput. The time horizon of the analysis was set at 5 years to capture both initial investment costs and medium-term operational dynamics, including device depreciation and stabilization of clinical workflows. The model structure combined elements of budget impact analysis and organizational simulation, with baseline assumptions and sensitivity analyses to explore parameter uncertainty. No individual patient-level simulation was performed; instead, average patient profiles and treatment pathways were modeled based on typical caseloads in neurological and orthopedic rehabilitation units.

Description of the 2 Scenarios

Baseline Hospital Characteristics

The reference setting was a medium-sized general hospital with an inpatient and outpatient rehabilitation department. The modeled hospital was assumed to manage a mixed caseload of neurological and orthopedic patients, primarily including stroke, traumatic brain injury, spinal cord injury (in the subacute phase), and postsurgical orthopedic conditions, and others. Baseline characteristics included a stable annual volume of patients undergoing rehabilitation, including inpatients and a comparable number of external patients; standard multidisciplinary staff composed of physiotherapists, physicians, nurses, and occupational therapists; and conventional therapy pathways based on individualized sessions of manual therapy and task-oriented exercises. Average treatment duration, number of sessions per patient, and staffing ratios were derived from or inspired by published benchmarks and typical organizational patterns observed in European hospital rehabilitation units.

XSL•FO
RenderX

Clinic A Scenario: A Novel Robotic Rehabilitation Unit

In scenario A, the hospital was assumed to introduce a robotic rehabilitation unit integrated into the existing rehabilitation department. The unit included robotic devices for upper-limb rehabilitation, selected to represent commonly adopted clinical platforms for poststroke and orthopedic motor recovery. The robotic unit was characterized by: (1) a predefined number of robotic systems for the upper-limb extremities (in this model, 2 upper-limb robots), (2) dedicated therapy spaces adapted to accommodate robotic equipment and ensure patient safety, (3) a core team of trained physiotherapists responsible for delivering robotic-assisted sessions, and (4) support from a clinical engineer or technical specialist for maintenance and troubleshooting. Robotic therapy was modeled as being delivered in addition to conventional rehabilitation, partially replacing manual therapy time for selected patient groups. Patients eligible for robotic therapy were assumed to be those with moderate to severe motor impairment requiring high-intensity repetitive training or selected patients deemed fit for medical robotics by physicians. The robotic unit was also assumed to potentially generate structured clinical datasets that could be used for observational studies and applied research projects, thereby increasing the hospital's research capacity.

Clinic B Scenario: Standard Rehabilitation

Scenario B represented a hospital providing only conventional rehabilitation services, without robotic devices. Therapy consisted of manual and task-oriented exercises delivered by physiotherapists in individual or small-group sessions. The available therapy time was assumed to be identical to that in scenario A at baseline. However, treatment intensity and throughput were constrained by therapist availability and physical workload. No automated collection of kinematic data or technology-driven standardization of therapy protocols was assumed in this scenario. Research activity related to rehabilitation was limited to traditional clinical observation, administration of clinical scales, and sporadic data collection, without access to high-frequency quantitative motor measurements.

Model Hypothesis

Table 1 summarizes the quantitative hypotheses and assumptions made for the model. Model parameters were derived from the following main sources: published literature on robotic rehabilitation effectiveness, costs, and organizational use; publicly available data on patient volumes in comparable European rehabilitation hospitals; price ranges reported by manufacturers, brochures, and clinical engineering reports; and feedback from clinics operating in the field.

When direct data were unavailable, conservative assumptions were adopted and tested through sensitivity analysis. All assumptions were explicitly reported in the model description.

Robotic-assisted rehabilitation was assumed to increase the achievable intensity of therapy by enabling longer or more frequent sessions without a proportional increase in therapist physical workload. Functional effectiveness values (Fugl-Meyer Assessment [FMA] motor score improvement) were derived from published randomized trials and meta-analyses comparing robotic-assisted and conventional rehabilitation. A comparative therapy quality framework was derived from the literature to characterize differences between robotic-assisted and standard rehabilitation. Evidence indicates that robotic rehabilitation may provide superior intrinsic therapeutic mechanisms through ad hoc assist-as-needed mechanisms but allows the delivery of a higher and more standardized therapy dose under comparable staffing conditions [5,8,10,11]. These dimensions were therefore used to qualitatively compare the expected quality of therapy between scenarios. Quantitative indicators of rehabilitation effectiveness were derived from published randomized trials and meta-analyses comparing robotic-assisted and conventional therapy. Reported ranges for repetitions per session, therapy frequency, and functional outcome improvements (FMA and activities of daily life [ADL] scales) were extracted and used to estimate cost-normalized effectiveness metrics. These values were combined with model-based cost estimates to compute cost-normalized indicators. This approach aimed to link therapy effectiveness and economic outcomes without assuming superior intrinsic efficacy of robotic rehabilitation but rather by reflecting higher delivered therapy dose and access under comparable staff conditions.

XSL•FO
RenderX

Table 1. Hypotheses and assumptions for the model.

Parameter/assumption	Scenario A	Scenario B	Source/assumption
Number of robots available, n	2	0	Fundamental model assumption
Daily operating time per robot (hours)	8	N/A[a]	Standard hospital working hours per day
Average duration per robotic session (minutes)	30	30	Clinical literature/robotic physiotherapy protocol [23]
Robot use rate, %	100	N/A	Model assumption, subject to sensitivity analysis
Daily therapist working time (hours)	8	8	Standard hospital working hours
Duration per standard occupational therapy session (minutes)	30	30	Clinical literature/hospital protocols [23]
Hourly cost per therapist (€)[b]	30	30	Real cost/hospital contracts
Robot purchase cost (unit cost; €)	200,000	0	Manufacturer quote/literature
Robot installation cost (€)	100,000	0	Estimated from manufacturer quote, includes structural costs
Robot maintenance cost (€)	50,000	0	Estimated from literature [24]
Robot depreciation (years)	5	N/A	Standard accounting practice
Number of patients treated per month, n	80 (40 inpatients, 40 external patients)	60 (40 inpatients, 20 external patients)	Hospital data/realistic estimate
Number of sessions per patient (1 month), n	13	13	Clinical literature/rehabilitation protocols
Average clinical improvement per patient (FMA[c]), points	+5/+7	+3/+5	Clinical trial/reference [10,25,26]
Treatment cycle duration (months), n	1	1	Standard clinical protocol
Annual robot maintenance cost per robot (€)	5000	N/A	Manufacturer quote/literature
Consumables cost per session per patient (€)	5	5	Model estimate/clinical data [24]
Maximum robot capacity, %	100	N/A	Model assumption
Percentage of sessions lost (absence/failure), %	10	0	Hospital historical data/assumption [27]
Number of robot days available per month, n	20	N/A	Working days per months
Maximum number of patients administered with robot therapy per month, n	45	N/A	Derived from model
Corresponding full-time therapists needed (FTE[d])	2.34	2.44	Derived from model
Reimbursement per session (€)	33.75	33.75	Reasonable model assumption
External patients included	Yes	Yes	Model assumption
Robot therapist supervises multiple patients, n	2	N/A	Model assumption/clinical practice [10,11,28,29]
Maximum number of patients administered by 1 occupational therapy therapist during a session, n	1	1	Model assumption
Maximum robotic sessions per day	Limited by robot capacity and hours	Limited by robot capacity and hours	Derived from model
Break-even calculation	Based on cumulative direct costs only	Based on cumulative direct costs only	Model assumption

[a]N/A: not applicable.

[b]€1=US $1.12 as of October 26, 2026.

[c]FMA: Fugl-Meyer Assessment.

[d]FTE: full-time equivalent.

XSL•FO
RenderX

Clinical and Organizational Assumptions

Other hypotheses for the model include the following assumptions. Klamroth-Marganska et al [23] reported at least 45 minutes per session per week. Based on these data and considering the capabilities of low-functioning patients, we opted for a conservative duration of 30 minutes per session.

Previous studies have reported modest but clinically relevant improvements in upper-limb motor function following robot-assisted or electromechanical rehabilitation, with between-group differences in the FMA of approximately 2 to 3 points reported in controlled trials [10], while systematic reviews have confirmed overall beneficial effects of robot-assisted and electromechanical training on upper-limb function [25]. Moreover, real-world observational data have reported larger pre-post FMA improvements following robot-assisted rehabilitation, in the range of approximately 7 to 11 points depending on the robotic technology used [10,25,26], supporting the plausibility of a moderate improvement in motor outcomes under intensive robotic rehabilitation scenarios. Based on these data, we chose a realistic, cautious scenario (+5/+7 FMA points with robot, +3/+5 points with standard therapy).

Maximum robot capacity was set to 100%. Since this value might be unrealistic or at least difficult to achieve, sensitivity analysis will also consider lower robot use rates, as well as 10% of sessions being lost.

Auxiliary personnel expenditures—such as initial technology-specific training programs and variations in treatment efficiency linked to individual therapist expertise—were modeled as structural overheads rather than dynamic per-session variables. While onboarding new robotic devices requires initial staff training, these setup costs are typically amortized across institutional operational budgets and do not alter the macrolevel comparative cost-effectiveness between robotic and manual pathways.

Economic parameters related to hardware acquisition, installation, and maintenance were established using aggregated market benchmarks and standardized vendor quotations reflecting commercial upper-limb robotic technologies, ensuring generalizability beyond specific proprietary brand pricing.

Clinical outcomes are reported as aggregate average FMA score improvements across standardized virtual patient cohorts rather than stratified baseline severities, reflecting macrolevel therapeutic efficacy rather than individual patient-level recovery curves. Calculating cumulative annual clinical gains via weighted averages assumes population-level homogeneity, thereby omitting individual response variance and nonlinear recovery profiles typical of heterogeneous clinical cohorts.

Consumable costs per session were modeled using a standardized parity assumption across both pathways. While robotic systems may involve specific component replacements or specialized hygienic accessories, these marginal expenses account for a minimal share of total operational costs compared with labor and capital investments, preserving the structural validity of the cost comparison.

The 5-year analytical horizon focuses strictly on the initial capital amortization cycle. In real-world hospital operations, robotic rehabilitation devices frequently remain functional and heavily used well beyond this initial period, meaning that long-term operational phases often benefit from postamortization efficiency, where capital acquisition costs are fully absorbed. For this reason, no equipment disposal costs were considered. No discount rate was applied to costs or health outcomes over the 5-year time horizon.

We assumed institutional alignment and compliance with workflow reengineering. In clinical practice, successful technology integration depends on overcoming initial staff resistance through adequate training and change management. However, because robotic devices function as clinical support tools designed to reduce physical fatigue and augment supervision capacity rather than displace personnel, organizational integration is structured as a collaborative enhancement of the existing health care team.

Cost Model and Economic Evaluation Framework

Overview

The economic evaluation was conducted from the hospital's perspective over a 5-year time horizon, following the reporting guidance of the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement [30]. Multimedia Appendix 1 provides the CHEERS checklist. The model integrates capital, operational, and direct labor expenditures to assess the budget impact and cost-effectiveness of introducing a robotic rehabilitation unit (clinic A) vs standard manual care (clinic B).

Capital and Equipment Costs (Investment Phase)

For clinic A, capital expenditures include the acquisition of 2 upper-limb robotic systems (C_{acq}=€200,000 each; €1=US $1.12 as of October 26, 2026), site installation and structural safety adaptations (C_{inst}=€100,000 total), and long-term maintenance agreements (C_{maint}=€5000 per robot per year, totaling €50,000 over the 5-year horizon). Capital costs are annualized using a straight-line depreciation method over a 5-year useful life (L=5 years) with zero residual salvage value:

annual depreciation =  + annual maintenance

Clinic B incurs no capital expenditures for robotic technology.

Operating and Personnel Costs (Operational Phase)

Personnel expenditures are modeled dynamically based on the full-time equivalent (FTE) staff required to sustain the annual volume of therapy sessions. The hourly cost per physiotherapist (w) is fixed at €30 per hour based on standard European hospital employment contracts. Total annual personnel cost is computed as:

$$C_{Staff} = FTE_{needed} \cdot \text{annual working hours} \cdot w$$

In clinic A, the model explicitly accounts for the "workforce multiplier" effect, where a 1:2 therapist-to-patient supervision ratio reduces the relative labor input per session compared with the 1:1 manual ratio in clinic B. Supervision ratios are modeled

XSL•FO
RenderX

as macrolevel operational averages rather than static constraints. While early-stage patients with severe impairment require intensive 1:1 therapist-to-patient attention, the robotic model accommodates progressive skill acquisition, allowing patients with mild to moderate impairments to be managed under optimized multipatient supervision ratios (eg, up to 1:2 or 1:3 during autonomous execution phases). Consequently, the projected efficiency gains reflect an aggregate operational average across the continuum of the rehabilitation cycle rather than a uniform one-size-fits-all ratio.

Consumables and Direct Operational Logistics

Session-specific consumables (such as hygienic interfaces and accessories) are modeled at a flat rate of €5 per session for both scenarios. Administrative overhead and scheduling logistics are assumed to scale proportionally with patient throughput.

$C_{\text{Consumables}}$ = total annual sessions · €5

A uniform reimbursement rate of €33.75 per rehabilitation session was assumed for both clinics, irrespective of inpatient or outpatient status. Annual revenue was calculated as the number of patients treated per month × 12 months × 13 sessions per patient × €33.75 per session. This uniform rate was used as a simplifying assumption for the economic comparison and was not intended to represent setting-specific reimbursement tariffs.

Total Direct Costs and Margins

Total annual direct costs (DCs) for each clinic are calculated as the sum of personnel, technology/depreciation, and consumables (maintenance + consumables). Cost per patient and cost per FMA point gained are evaluated as core value-based metrics to determine economic efficiency.

$\text{DC} = C_{\text{Staff}} + C_{\text{Tech}} + C_{\text{Consumables}}$

Total annual revenue is derived from standard institutional reimbursement rates for outpatient and inpatient rehabilitation cycles. The gross operating margin (earnings before interest, taxes, depreciation, and amortization [EBITDA]) is calculated as:

EBITDA = total annual revenue – DC

Finally, cost per patient and cost per FMA point gained are evaluated as core value-based metrics to determine economic efficiency.

Outcome Model

The outcome model was developed to capture three complementary dimensions: clinical, organizational, and scientific/innovation-related outcomes.

Organizational Outcomes

Organizational outcomes focused on the operational performance of the rehabilitation service and included annual patient throughput, defined as the total number of patients who successfully complete their treatment program within a year. This metric reflects the facility's total capacity and is a driver for amortizing fixed technological costs. Total annual sessions refers to the aggregate number of individual therapy sessions delivered by the department annually. It is determined by the patient volume multiplied by the number of sessions prescribed per clinical protocol. Annual therapy hours represent the total volume of effective treatment time provided to the patient population. Given a standard session duration (eg, 30 minutes), this metric represents the cumulative "dose" of therapy delivered by the unit. Workforce needed (FTE) refers to the number of FTE staff members required to sustain the projected clinical workload. This measure highlights how robotic integration can decouple patient volume from human labor requirements. Productivity (sessions/FTE) is an efficiency ratio representing the average number of therapy sessions managed by 1 full-time therapist per year. It quantifies the "workforce multiplier" effect of technology, showing the increase in output per staff member enabled by robotic-assisted supervision models.

Regarding therapists, hours of therapy (annual therapy hours) represent the total cumulative volume of effective treatment time delivered to the entire patient population over 1 year. This metric is calculated based on the total number of sessions multiplied by the duration of each session. It serves as a measure of the total "therapeutic dose" the facility is capable of providing. FTE needed (workforce needed) indicates the number of FTE therapists required to cover the total annual therapy workload. In the model, it accounts for the different supervision ratios, showing how robotic integration (1:2 ratio) allows a facility to manage higher patient volumes with fewer staff members compared with the manual 1:1 model. Efficiency (therapy hours per FTE) measures the amount of therapy time successfully delivered by each full-time therapist unit per year. It quantifies the operational gain of the robotic unit, demonstrating how technology enables therapists to "produce" more treatment time by supervising multiple automated sessions simultaneously.

Economic Evaluation

Economic evaluation was performed using a combination of total annual revenue, which represents the total revenue generated from rehabilitation services in a year; DCs (staff + tech), which are the sum of personnel costs (physical therapists required to cover therapy hours) and technology costs (depreciation of machines, maintenance, and consumables); and operating margin (EBITDA), which indicates the gross operating margin (revenue – DCs).

Clinical Outcomes (Economically Weighted)

Clinical outcomes were modeled at an aggregate level and included weighted average FMA improvement, which represents the average change in the FMA score—the gold standard for evaluating motor recovery—achieved by patients after a full treatment cycle. It is "weighted" to reflect the specific clinical evidence found in the literature for both robotic-assisted (scenario A) and manual (scenario B) therapies. Annual clinical gain (total FMA points) is a systemic impact metric that quantifies the "total amount of recovery" produced by the department in 1 year. It is calculated by multiplying annual patient throughput by the weighted average FMA improvement. Therapy delivery effectiveness measures the clinical gain achieved relative to the resources invested (eg, FMA points per hour of therapy or per therapist FTE). It evaluates the "quality" of the organizational model, demonstrating how effectively the

XSL•FO
RenderX

clinical protocol translates staff effort and technological support into measurable patient recovery.

Health Technology Assessment

To provide a comprehensive evaluation of the transition from manual to robotic-assisted therapy, a multidimensional framework based on health technology assessment (HTA [31]) principles was adopted. Beyond clinical efficacy, specific metrics to quantify organizational scalability (eg, staff engagement ratio) and economic sustainability (eg, cost per FMA point gained) were introduced. These parameters allow for a value-based comparison, aligning clinical recovery with the financial viability of the health care provider.

DC per patient calculates the average direct expenditure incurred by the hospital to provide a complete treatment cycle to a single patient. It includes the cost of specialized staff (physiotherapists) and the proportional cost of technology (robot depreciation and maintenance). It is an indicator of operational efficiency, showing how the robotic model can lower the "unit price of care" by scaling the number of patients treated.

Cost per FMA point gained is a cost-effectiveness indicator. It is calculated by dividing the DC per patient by the weighted average FMA improvement. It represents the financial investment required to achieve a single point of motor recovery on the FMA scale.

Staff engagement ratio (or supervision ratio) defines the relationship between the number of therapists and the number of patients treated simultaneously. In clinic B, this is a 1:1 ratio (1 therapist per patient), whereas in clinic A, it shifts toward a 1:2 or hybrid model. It quantifies workforce leverage, allowing therapists to oversee more treatments without increasing their physical workload.

Marginal clinical gain per patient represents the additional clinical improvement achieved by a patient in the robotic scenario (clinic A) compared with standard manual care (clinic B). It is calculated as the difference between the average FMA score improvement in the 2 scenarios. It quantifies the "added value" of technology in terms of individual recovery.

Clinical reach potential measures the department's ability to extend specialized rehabilitation to a broader portion of the target population. It is linked to increased patient throughput and reduced waiting lists. It quantifies the public health impact by measuring how many more patients can access high-quality treatment within the same time frame and infrastructure.

Scientific and Innovation Outcomes

Scientific and innovation outcomes were modeled as proxy indicators of research capacity rather than as direct outputs. These included the potential number of observational or pilot studies enabled by structured data collection, the volume of performance datasets generated annually, the estimated number of scientific publications and conference contributions, and the potential to initiate funded research projects and industrial collaborations. These outcomes were estimated based on the volume of treated patients and the availability of quantitative motor data produced by robotic systems.

Sensitivity Analysis

To assess the robustness of the model to assumptions, one-way and multiparameter sensitivity analyses were conducted on key variables, considering the economic and clinical outcomes. Two variables were chosen: patient volume and the robot use rate, which indicates the extent to which the robot unit is used and encompasses, in aggregate, the effects of the number of patients treated, the time the robot is not in use, and the number of patients handled by the therapists; and the cost per FMA point gained, as a measure of the effectiveness of the robot unit in improving the clinical quality of treatments. Threshold analyses were performed to identify minimum use rates and patient volumes required for economic sustainability.

Ethical and Regulatory Context

The present study was based entirely on a hypothetical simulation model and did not involve direct patient recruitment, interventions, access to individual clinical records, or the collection or analysis of identifiable or pseudonymized patient-level data. The model incorporated assumptions informed by current clinical practice and relevant regulatory frameworks for medical devices, including the European Medical Device Regulation (MDR 2017/745) [32] and internationally recognized technical standards issued by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC), including ISO 13485 and the IEC 60601 series. Potential future clinical and kinematic data collection and secondary use were considered only as hypothetical scenarios, assuming appropriate anonymization and compliance with applicable data protection requirements, including the General Data Protection Regulation (GDPR) [33]. Any future observational studies conducted using the robotic unit would be subject to the applicable ethical review and regulatory requirements, including review by the relevant ethics committee where required, in accordance with the Declaration of Helsinki and applicable national regulations.

Results

Organizational Impact

The model suggests that clinic A achieved a higher annual patient throughput than clinic B (Table 2), with 960 vs 720 patients per year (+33%). This increase was associated with a higher total number of therapy sessions (12,480 vs 9360) and greater annual therapy hours (6240 vs 4680). The modeled transition from 720 to 960 annual patients reflects potential technical capacity rather than guaranteed market expansion. Actual patient volume intake is bounded by external constraints such as regional health care demand, referral networks, and institutional budgetary caps. When local demand limits volume growth, operational performance converges toward lower use bands, underscoring the importance of regional market alignment before capital adoption. Despite the higher activity volume, the required workforce was slightly lower in clinic A (2.34 vs 2.44 FTE), resulting in a substantial gain in productivity, expressed as sessions per FTE (+39%). Overall, the robotic-hybrid organizational model allowed higher service capacity with comparable or reduced staffing levels.

XSL•FO
RenderX

Table 2 also shows that clinic A could manage a higher monthly patient volume than clinic B (80 vs 60 patients, +33%), corresponding to a higher number of therapy hours (520 vs 390 hours per month). Despite this increased workload, the estimated staff requirement was slightly lower in clinic A (2.34 vs 2.44 FTE). Consequently, staff efficiency, expressed as therapy hours per FTE, was markedly higher in the robotic-hybrid model (223 vs 160 hours/FTE and +63 hours/FTE), indicating a more efficient allocation of human resources.

The introduction of robotic units allows for a decoupling of patient throughput and staff requirements. While clinic B requires 2.44 FTEs to treat 60 patients, clinic A handles a 33% higher volume (80 patients) with a lower staff demand (2.34 FTEs), effectively increasing the therapy-to-staff productivity ratio by 39%.

Table 2. Organizational outcomes and staff requirements: comparison table.

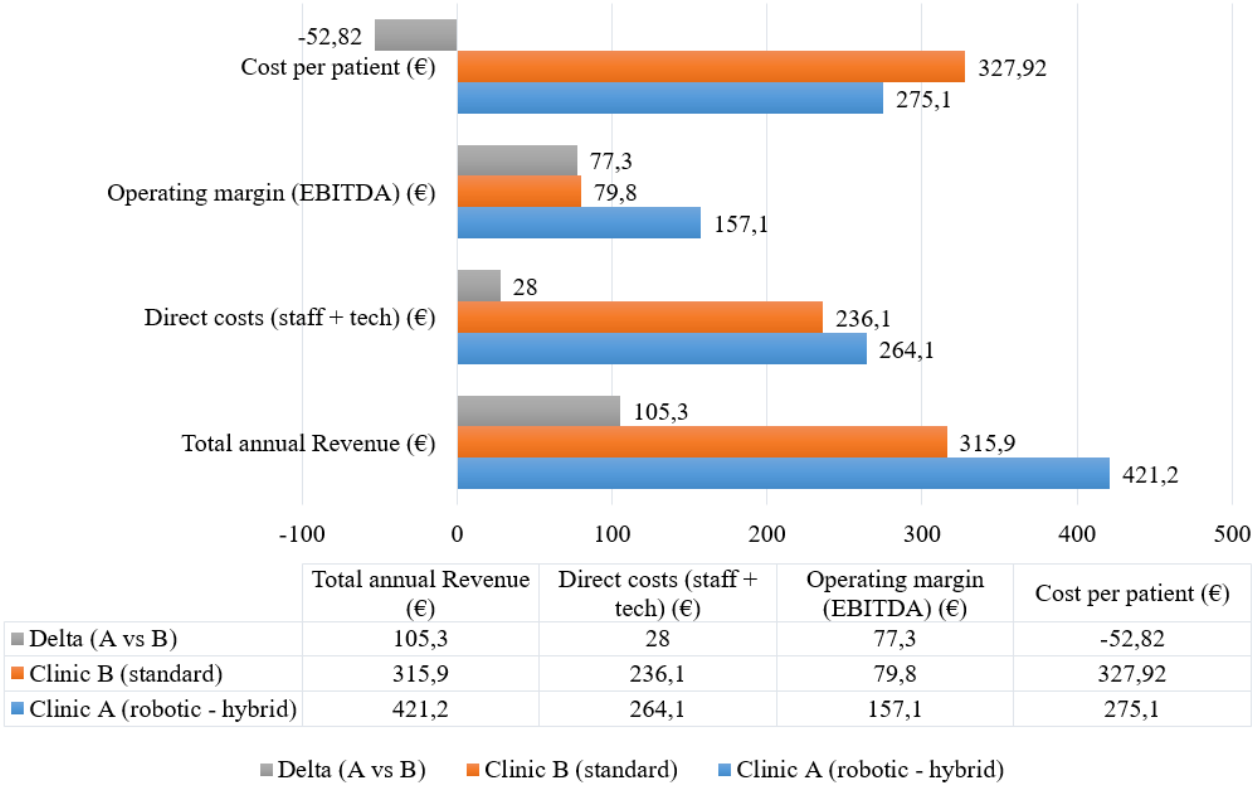
Parameter	Clinic A (robotic-hybrid model)	Clinic B (standard)	Difference (A vs B)
Patients per month, n	80	60	+20 (+33%)
Annual patient throughput, n	960	720	+240 (+33%)
Total annual sessions, n	12,480	9360	+3120
Annual therapy hours, n	6240	4680	+1560
Workforce needed, (FTE[a])	2.34	2.44	–0.10
Productivity (sessions/FTE)	5333	3836	+1497 (+39%)
Hours of therapy, n	520	390	+130 (+33%)
Efficiency (hours/FTE)	223	160	+63

[a]FTE: full-time equivalent.

Economic Impact

Under the assumptions, clinic A generated higher total annual revenue than clinic B (€421,200 vs €315,900), despite moderately higher DCs (€264,100 vs €236,100). This resulted in a substantially higher operating margin for clinic A (€157,100 vs €79,800). Furthermore, the robotic-hybrid model was associated with a lower cost per patient (€275.10 vs €327.92), corresponding to a reduction of €52.82 per treated patient (Figure 1).

Figure 1. Economic impact (per patient). Clinics A and B are compared with their difference. EBITDA: earnings before interest, taxes, depreciation, and amortization.



XSL•FO
RenderX

Expected Clinical and Organizational Outcomes

Therapy Quality

Therapy quality is based on literature and not on model assumptions. Compared with standard therapy, robotic-assisted rehabilitation showed slightly higher performance across most therapy quality dimensions, particularly in repetition rate, protocol reproducibility, and availability of objective feedback (Table 3). The 2 approaches were assumed to have comparable intrinsic effectiveness when therapy dose is matched, in line with current meta-analyses. The observed quality advantage of scenario A was therefore mainly attributable to the capacity to deliver more intensive and standardized training rather than to a distinct clinical mechanism.

When normalized by functional gains, robotic-assisted rehabilitation showed comparable or slightly lower cost per unit of improvement than standard therapy. This effect was primarily driven by a higher therapy dose (greater repetitions and session frequency), which translated into moderately higher average motor and ADL improvements reported in the literature [8,11]. These results suggest that robotic rehabilitation is not intrinsically more effective but may achieve similar or better cost-effectiveness by enabling higher intensity of practice without proportional increases in staffing costs.

Table 3. Comparative therapy quality indicators based on literature.

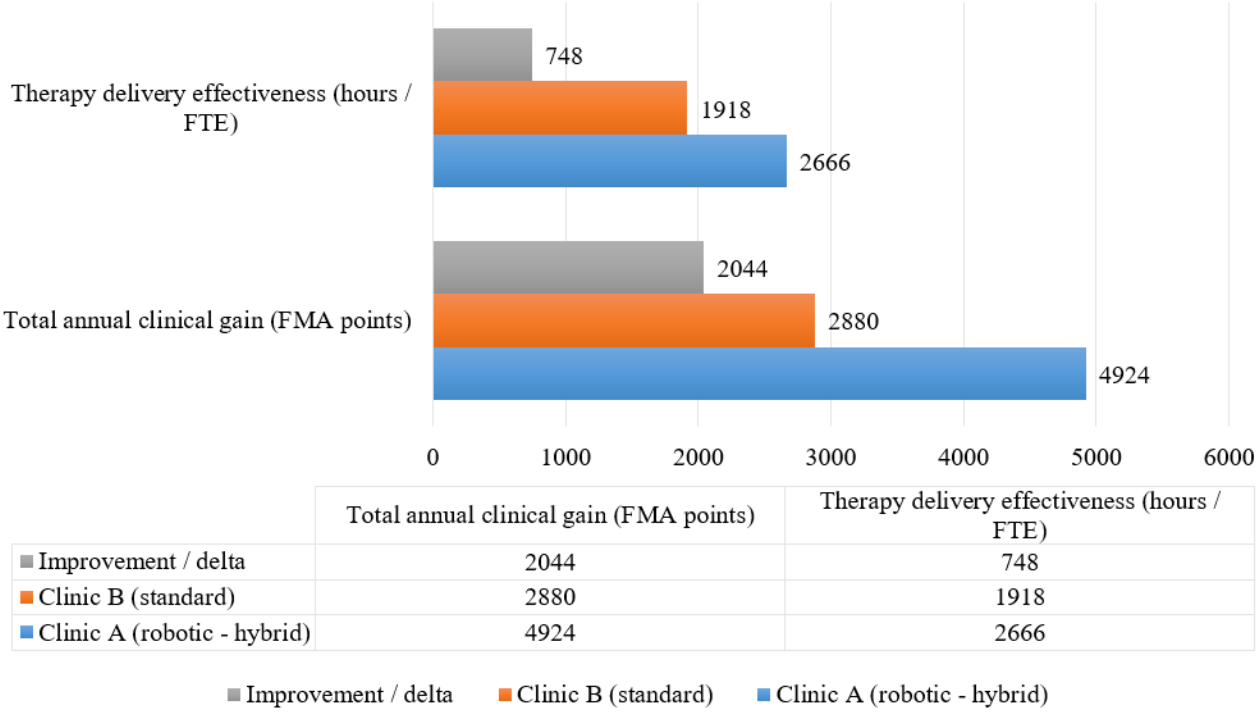
Dimension of therapy quality	Robotic-assisted rehabilitation	Standard rehabilitation	Evidence from literature
Repetitions per session	High (hundreds to >1000 movements per session)	Low to moderate (10 to approximately 100 repetitions per session)	Robotic devices enable massed practice and repetitive training [8,10,14]
Intensity control	Precisely programmable and reproducible	Therapist-dependent and variable	Robots allow controlled speed, force, and task difficulty [5,15]
Task specificity	High (goal-directed, trajectory-based, and task-oriented modes)	Variable (depends on therapist skill and fatigue)	Task-oriented robotic protocols support motor learning [4,14]
Intersession reproducibility	High (same protocol can be repeated identically)	Low to moderate (manual guidance variability)	Robotic training reduces intertherapist variability [15,17]
Objective performance feedback	Continuous (kinematics, forces, and errors)	Mainly qualitative and score-based	Robots provide high-frequency quantitative data [8]
Therapist physical workload	Reduced manual assistance	High physical effort for guidance and support	Robotic assistance reduces physical strain [14,17]
Access for patients with severe impairments	Higher (assistance-as-needed modes)	Limited by therapist burden	Robots allow early/intensive training in severe impairment [8,11]
Intrinsic therapeutic effectiveness	Comparable when dose is matched	Comparable when dose is matched	Still a matter of debate. No clear superiority when intensity is controlled [5,8]

Clinical Results (Economically Weighted)

The model suggests that clinic A has higher clinical performance than clinic B (Figure 2), with a higher weighted average improvement in FMA scores (5.13 vs 4.00 FMA points, +28.2%). This translated into a substantially greater total annual clinical gain (4924 vs 2880 FMA points, +71%), supported by higher therapy delivery effectiveness per FTE (+39%).

XSL•FO
RenderX

Figure 2. Quantitative comparative rehabilitation effectiveness and cost-related indicators. FMA: Fugl-Meyer Assessment; FTE: full-time equivalent.



	Total annual clinical gain (FMA points)	Therapy delivery effectiveness (hours / FTE)
Improvement / delta	2044	748
Clinic B (standard)	2880	1918
Clinic A (robotic - hybrid)	4924	2666

HTA

From an economic perspective (Table 4), clinic A was associated with lower DC per patient (−16.1%) and a markedly reduced cost per FMA point gained (€53.62 vs €81.98, −34.6%). In addition, the staff engagement ratio was lower in the robotic-hybrid model, indicating reduced relative personnel effort per unit of clinical output. Impact metrics further showed a marginal clinical gain per patient and an increased clinical reach potential (+240 patients per year, +33%), suggesting both clinical efficacy and expanded service capacity.

Table 4. Economic perspective.

Indicator	Clinic A (robotic-hybrid model)	Clinic B (standard)	Improvement/difference
Economic efficiency			
Direct cost per patient (€)[a]	275.10	327.92	–16.1%
Cost per FMA[b] point gained (€ per point)	53.62	81.98	–34.6%
Staff engagement ratio	0.72	1.00	–28% effort
Impact metrics			
Marginal clinical gain per patient (points)	+1.13	Baseline	Clinical advantage
Clinical reach potential (patients per year), n	+240	Baseline	+33% coverage

[a]€1=US $1.12 as of October 26, 2026.

[b]FMA: Fugl-Meyer Assessment.

Scientific and Innovation Perspectives (Qualitative Benefits)

While the simulation model focuses primarily on economic and organizational performance, the introduction of a robotic rehabilitation unit also creates qualitative opportunities for research and institutional development. Rather than acting as guaranteed measurable outputs, these dimensions represent potential structural benefits. The systematic logging of high-resolution kinematic and kinetic data provides a robust digital foundation for future observational studies and secondary data analyses, replacing sporadic traditional clinical scales with continuous multimodal datasets. Although absolute metrics such as publication counts or grant acquisitions cannot be reliably modeled, the presence of standardized technological infrastructure inherently enhances the facility's profile, facilitating academic networking, multicenter research alignment, and collaborative clinical innovation.

Sensitivity Analysis

Overview

To evaluate the robustness of the economic and organizational model, we performed a one-way sensitivity analysis on 2 critical parameters: robot use rate (reflecting clinical adherence) and total patient throughput (reflecting the hospital's ability to attract

XSL•FO
RenderX

volume), reported in Table 5. The baseline values for clinic A (45 robotic patients per month, 80 total patients per month) were tested against "low use" and "volume stagnation" scenarios to identify sustainability thresholds.

Table 5. Sensitivity analysis: economic and organizational thresholds.

Parameter scenario	Scenario A	Total patients, n	Workforce (FTE[a])	Cost per patient (€)[b]	Difference vs standard care (clinic B)
Clinic B (reference)	0	60	2.44	327.91	—[c]
A1. Low use	25	80	2.73	310.50	–5.3% (marginal saving)
A2. Baseline (optimal)	45	80	2.34	275.10	–16.1% (significant saving)
A3. Volume stagnation	45	60	2.14	365.40	+11.4% (cost increase)

[a]FTE: full-time equivalent.

[b]€1=US $1.12 as of October 26, 2026.

[c]Not applicable.

Impact of Robot Use Rate (Clinical Adherence)

We simulated a scenario in which the robotic unit is underused due to lower clinical adoption or technical constraints, serving only 25 patients per month (approximately 30% of total throughput) instead of the baseline 45. In this "low use" scenario, the organizational efficiency of scenario A decreases significantly. To manage the remaining 55 patients via standard manual therapy (1:1 ratio), the workforce requirement increases from 2.34 FTEs (baseline) to 2.73 FTEs. Consequently, the net cost per patient rises to €310.50. This indicates that the economic advantage of the robotic unit is strictly dependent on maximizing its use; if use drops below a critical threshold of 30 patients per month (approximately 60% of capacity), the efficiency gains in workforce productivity are insufficient to offset capital depreciation.

Impact of Patient Volume (Scalability)

We assessed the financial risk associated with a failure to increase patient volume. In the "stagnation" scenario, clinic A invests in robotics but fails to attract external patients, maintaining the same throughput as clinic B (60 patients per month). Under these conditions, the fixed costs of the technology weigh heavily on a smaller patient base. The cost per patient in clinic A rises to €365.40, which is 11% higher than the standard care cost in clinic B (€327.91). This confirms that robotic rehabilitation acts as a "scale-enabler": it becomes economically sustainable only if the facility leverages the technology to increase total throughput above a break-even volume of approximately 70 patients per month.

Sensitivity Thresholds

The analysis identifies 2 necessary conditions for the economic viability of the robotic unit. The operational threshold requires the device to be used at least 60% of its theoretical capacity to generate sufficient FTE savings. The volume threshold requires the introduction of the technology to be accompanied by a minimum 15% to 20% increase in total patient volume to amortize fixed capital costs effectively. Failure to meet either condition results in a scenario in which robotic rehabilitation becomes a net cost driver rather than an efficiency multiplier.

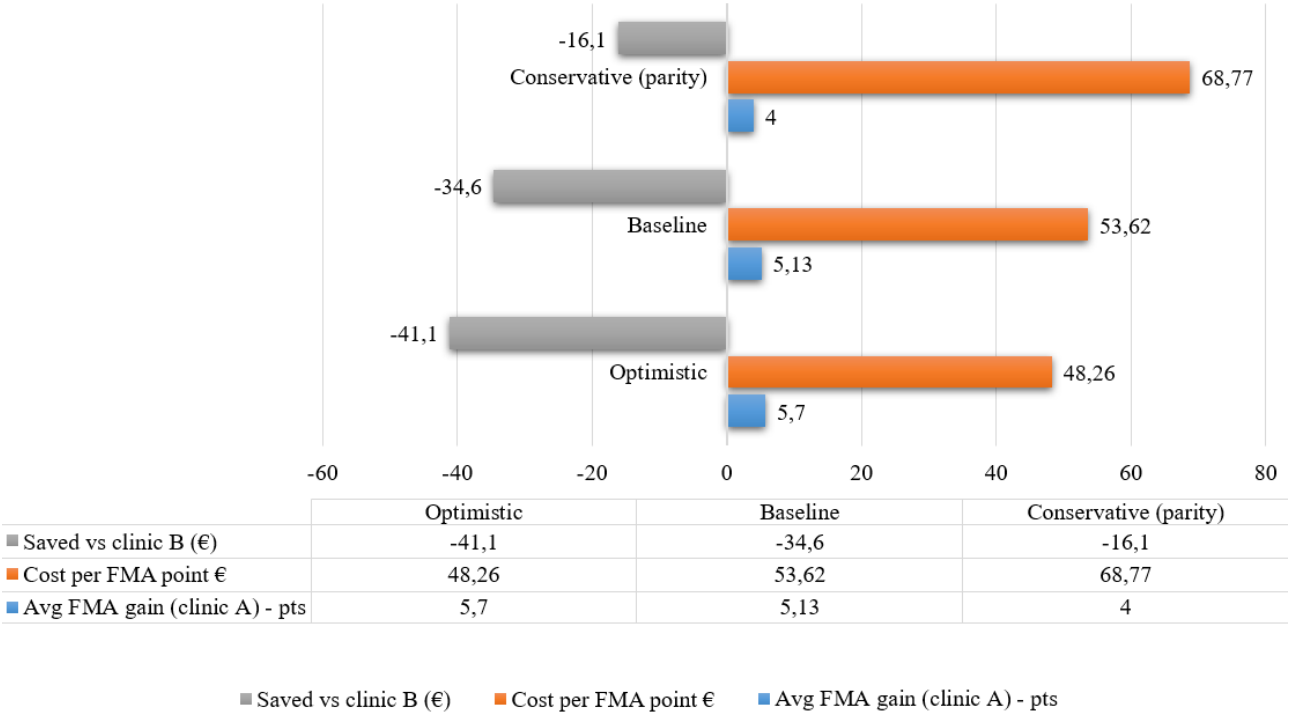
Sensitivity Analysis of Cost-Effectiveness: Cost Per FMA Point Gained

Sensitivity analysis was also conducted on the cost per FMA point gained, the primary indicator of value-based efficiency. The robustness of this metric was assessed by varying the clinical efficacy differential, defined as the additional FMA points gained through robotic therapy compared with standard care. In the baseline scenario, the robotic-hybrid model (clinic A) achieved an average improvement of 5.13 FMA points, compared with 4.00 points in clinic B, resulting in a 34.6% reduction in cost per point (€53.62 vs €81.98). Two alternative clinical outcomes were simulated. In the conservative scenario (clinical parity), both clinics achieved exactly 4.00 points, indicating that robotic therapy provided no additional clinical recovery over manual therapy. In the optimistic scenario (high intensity), the robotic group reached the upper bound of the expected improvement (7.0 points), increasing the weighted average improvement for clinic A to 5.70 points.

Assuming a fixed cost per patient of €275.10 (scenario A2 of the sensitivity analysis), even in the conservative scenario, where the robotic unit provides zero additional clinical benefit over manual therapy (clinical parity), clinic A remains 16.1% more cost-effective than clinic B (€68.77 vs €81.98). This suggests that the sustainability of robotic rehabilitation is "double-protected." The financial viability is guaranteed by the organizational efficiency (FTE optimization and higher throughput), even if the clinical superiority were to be marginal. Conversely, when clinical superiority is realized (baseline and optimistic scenarios), the cost-effectiveness scales exponentially, offering a massive advantage for health care providers focused on maximizing patient recovery within fixed budgets. Results are summarized in Figure 3.

XSL•FO
RenderX

Figure 3. Sensitivity analysis of cost per Fugl-Meyer Assessment (FMA) point gained. Avg: average; pts: points.



	Optimistic	Baseline	Conservative (parity)
Saved vs clinic B (€)	-41,1	-34,6	-16,1
Cost per FMA point €	48,26	53,62	68,77
Avg FMA gain (clinic A) - pts	5,7	5,13	4

Discussion

The Robotic Workforce Multiplier

The results of this scenario-based simulation suggest that the introduction of a robotic rehabilitation unit in a medium-sized hospital (clinic A) may contribute to improving the "production function" of rehabilitative care. Unlike the traditional manual model (clinic B), which relies on a rigid 1:1 therapist-to-patient ratio, the hybrid model in clinic A may enable a decoupling of patient throughput from human workload. By optimizing the robotic units to a saturation point of 45 patients per month (representing approximately 56% of the total monthly caseload of 80 patients), clinic A may achieve a 33% increase in total volume compared with clinic B. Crucially, this expansion could be achieved without additional staffing. In fact, our FTE-based analysis reveals that clinic A requires only 2.34 FTEs to manage 80 patients, whereas clinic B requires 2.44 FTEs for just 60 patients. This efficiency gain is not attributed to an inherent clinical superiority of the robot in terms of recovery mechanisms but rather to its role as a workforce multiplier. In clinic A, the productivity ratio reaches 2666 therapy hours per FTE annually, a 39% improvement over the 1918 hours per FTE in clinic B. This allows the facility to absorb previously unmet rehabilitation demand while reducing the net cost per patient from €328 to €275.

Clinical Efficiency

The integrated analysis of clinic A's hybrid model reveals the possibility of a dual advantage over traditional care. First, the introduction of robotic units may enhance the quality of clinical outcomes. By enabling higher dosage, precision, and repeatability, the robotic treatment could lead to a weighted average FMA improvement of 5.13 points, compared with 4.00 points in the manual-only setting of clinic B (+28.2%). When scaled to the annual hospital volume, the impact could be substantial: clinic A produces a total annual clinical gain of 4924 FMA points, which is 71% higher than the 2880 points generated by clinic B. This metric suggests that a robotic-equipped facility acts as a far more powerful "health generator" for the community, potentially providing not just more care, but more effective recovery per patient.

The "Cost Per FMA Point" Paradigm

One of the most compelling findings of this study is the hypothetical possibility of a reduction in the cost of clinical recovery. In clinic B, the lack of technological support results in a high human-labor cost per unit of recovery, with each FMA point gained costing the facility €81.98. Conversely, in clinic A, the combination of a lower net cost per patient (€275.10) and clinical outcomes leads to a cost of €53.62 per FMA point. This 34.6% reduction in the cost of recovery represents a fundamental shift in hospital management. It demonstrates that the initial capital investment in robotics is not merely a "tech-upgrade" but a strategic move toward value-based health care. By lowering the cost of "producing" recovery, the hospital can reinvest the saved resources (approximately €28 per FMA point) into further expanding services or research.

Organizational Synergy and Workforce Productivity

The hybrid configuration of clinic A—where 45 patients use robotic units while 35 remain in standard care—is potentially a compelling organizational model. This setup allows the facility to handle a 33% higher patient volume with a theoretical workforce requirement of 2.34 FTEs, compared with the 2.44 FTEs needed for a lower volume in clinic B. The resulting staff engagement ratio of 0.72 in clinic A (ie, only 0.72 therapist-hours are required to produce 1 hour of therapy)

XSL•FO
RenderX

highlights how technology mitigates the physical and temporal constraints of manual therapy. This "liberated time" is what enables the therapist to focus on high-level clinical decision-making and supervision, effectively increasing the therapy delivery effectiveness by 39% (2666 vs 1918 hours per FTE).

Organizational Implications: Scalability and Standardization

From an organizational perspective, the transition to an FTE-based resource management model highlights that technological investment is amortized through increased service capacity rather than staff reduction. The ability of a single therapist to supervise two robotic sessions simultaneously (1:2 ratio) effectively frees nearly 1 full therapist equivalent (0.91 FTE) compared with purely manual 1:1 delivery for the same volume. Furthermore, the robotic system introduces a level of operational standardization that is unattainable in manual therapy. Every session in clinic A follows precise, programmable protocols, reducing intertherapist variability and ensuring that patients with severe impairments receive a consistent, high-intensity dose of therapy (hundreds to thousands of repetitions per session) without increasing therapist physical fatigue.

Research and Innovation: Digital Transformation of Rehabilitation

Beyond economic and clinical metrics, the adoption of robotic technology offers qualitative advantages regarding data collection and research infrastructure. While our model does not formally quantify research output, the generation of continuous digital datasets ensures comprehensive traceability of training sessions. This technological environment serves as a qualitative enabler for observational research, supporting internal data consistency and fostering potential academic and industrial partnerships. Rather than positioning the clinical unit as an immediate quantitative "innovation hub," these capabilities should be viewed as long-term institutional opportunities that depend heavily on internal research governance and strategic academic engagement. The projected benefits in terms of clinical gains (FMA improvements) and cost efficiencies are heavily contingent upon institutional adoption rates, consistent device use, and local reimbursement structures. Future empirical implementations are necessary to validate these theoretical projections against real-world clinical variabilities.

Limitations

Despite the promising insights provided by this comparative analysis, some limitations should be acknowledged. First, this study relies on a scenario-based simulation model rather than longitudinal empirical data. While the model parameters—such as patient throughput, device costs, and staffing ratios—were derived from established clinical literature and realistic hospital benchmarks, they may not be applicable to every institutional setting. In fact, while these parameters were derived from robust clinical benchmarks and manufacturer specifications, the absence of real-world calibration means that the projected outcomes reflect idealized operational assumptions. Consequently, the quantitative outputs—such as exact cost reductions and throughput margins—should not be interpreted as absolute predictions but rather as comparative estimates illustrating structural organizational dynamics. Future research will prioritize multicenter prospective validation studies to calibrate these simulation parameters against longitudinal, real-world clinical and financial workflows.

We acknowledge that factors such as specific regional reimbursement policies, variations in labor costs, and the architectural constraints of different hospitals could influence the analysis. The clinical outcomes modeled in this study, specifically the weighted average FMA improvement, were estimated based on aggregated data from meta-analyses. While these values provide a robust baseline, they do not account for the high variability in individual patient recovery trajectories, which are influenced by the timing of intervention, lesion severity, and comorbidities. Third, the organizational "multiplier effect" assumes a consistent 1:2 therapist-to-patient ratio during robotic sessions. In practice, the initial learning curve for clinical staff and potential technical downtime may temporarily reduce this efficiency during the first year of implementation. Finally, the scientific and innovation outcomes are presented as proxy indicators of research potential. While the capacity for data generation is a clear technical advantage of robotic units, the transformation of these datasets into high-impact publications depends on the hospital's internal research infrastructure and academic partnerships, which were not modeled. Other factors, such as real-world operational complexities—including patient transfer delays, device donning and doffing procedures, hygiene and disinfection protocols, unexpected technical downtime, spatial limitations, and long-term mechanical wear and tear—were not modeled at a microanalytical level. While annualized maintenance costs partially offset hardware degradation, high-frequency operational frictions can temporarily reduce the effective daily use rate. Consequently, actual hospital implementations must account for these microenvironmental overheads when targeting the operational efficiency thresholds identified in our sensitivity analysis. Moreover, the baseline assumption of a 100% robot use rate represents a theoretical operational ceiling rather than a realistic daily average. In actual clinical practice, auxiliary activities such as patient positioning, device donning and doffing, and routine software calibration consume a portion of the operational window. Nevertheless, our sensitivity analysis explicitly models underuse thresholds (eg, lower operational saturation), demonstrating that while efficiency margins compress, the model retains economic viability provided that use remains above the critical break-even limit. The analysis was conducted strictly from the institutional hospital perspective, omitting a patient-level microsimulation or a broader societal perspective. While institutional decision-making is primarily governed by operational budgets, throughput capacities, and direct labor costs, future investigations should incorporate patient-level longitudinal tracking to capture individual variability in recovery trajectories and out-of-pocket expenses. Finally, while conventional therapy is subject to performance variability driven by physical therapist fatigue, robotic pathways rely on electromechanical consistency. Nonetheless, operational continuity in robotic settings can occasionally be modulated by technical factors such as routine software recalibrations or minor

XSL•FO
RenderX

maintenance interventions, elements that are mitigated by standardized service agreements and safety protocols. Future research should focus on multicenter prospective studies to validate these simulation-based findings with real-world longitudinal evidence. Results may not generalize to hospitals with different reimbursement schemes or staffing costs.

Addressing Secondary Economic Variables and Operational Downtime

Although the current sensitivity analysis focused on primary throughput and cost boundaries, we acknowledge that secondary operational factors—such as unexpected robotic downtime, fluctuations in specialized physiotherapist wages, regional variations in reimbursement tariffs, and escalating maintenance agreements—represent real-world variables that could influence economic margins. In practice, a severe increase in technical downtime (reducing device use below critical operational thresholds) or a downward revision of regional Diagnosis-Related Group tariffs would compress the gross operating margin (EBITDA) calculated in clinic A. Nevertheless, the structural robustness demonstrated across our high- and low-use sensitivity ranges suggests that the model maintains economic viability provided that patient volume remains stable. Future implementations should incorporate real-world stochastic tracking of these secondary parameters as longitudinal data become available.

Acknowledgments

The authors wish to acknowledge the Italian Ministry of University and Research under the complementary actions to the Plan of National Recovery and Resilience (PNRR), “Fit4MedRob—Fit for Medical Robotics” (grant number PNC0000007).

Funding

This work was funded by the Italian Ministry of University and Research under the complementary actions to the Plan of National Recovery and Resilience (PNRR), “Fit4MedRob—Fit for Medical Robotics” (grant number PNC0000007).

Generative AI Use

The authors declare the use of generative AI (GenAI) in the research and writing process. According to the GAIDeT (Generative AI Declaration Taxonomy; 2025), the following tasks were delegated to GenAI tools under full human supervision: reproducibility testing, proofreading and editing, translation, quality assessment, and identification of limitations. The GenAI tool used was Gemini 1.5 (Google). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Authors' Contributions

AS conceived the paper, analyzed the data, wrote the paper, provided funding, and administered the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CHEERS checklist.
[DOCX File , 20 KB - rehab_v13i1e93572_app1.docx]

References

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the global burden of disease study 2019. Lancet Neurol 2021;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
2. Langhorne P, Baylan S, Early Supported Discharge Trialists. Early supported discharge services for people with acute stroke. Cochrane Database Syst Rev 2017;7(7):CD000443 [FREE Full text] [doi: 10.1002/14651858.CD000443.pub4] [Medline: 28703869]
3. Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American heart association/American stroke association. Stroke 2016;47(6):e98-e169. [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
4. Kwakkel G, van Peppen R, Wagenaar RC, Wood Dauphinee S, Richards C, Ashburn A, et al. Effects of augmented exercise therapy time after stroke: a meta-analysis. Stroke 2004;35(11):2529-2539. [doi: 10.1161/01.STR.0000143153.76460.7d] [Medline: 15472114]
5. Veerbeek JM, van Wegen E, van Peppen R, van der Wees PJ, Hendriks E, Rietberg M, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. PLoS One 2014;9(2):e87987 [FREE Full text] [doi: 10.1371/journal.pone.0087987] [Medline: 24505342]

XSL•FO
RenderX

6. Bernhardt J, Hayward KS, Kwakkel G, Ward NS, Wolf SL, Borschmann K, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the stroke recovery and rehabilitation roundtable taskforce. Int J Stroke 2017;12(5):444-450 [FREE Full text] [doi: 10.1177/1747493017711816] [Medline: 28697708]
7. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 2008;51(1):S225-S239. [doi: 10.1044/1092-4388(2008/018)] [Medline: 18230848]
8. Mehrholz J, Kugler J, Pohl M, Elsner B. Electromechanical-assisted training for walking after stroke. Cochrane Database Syst Rev 2025;5(5):CD006185. [doi: 10.1002/14651858.CD006185.pub6] [Medline: 40365867]
9. Rehabilitation in Health Systems: Guide for Action. Geneva, Switzerland: World Health Organization; 2018.
10. Lo AC, Guarino PD, Richards LG, Haselkorn JK, Wittenberg GF, Federman DG, et al. Robot-assisted therapy for long-term upper-limb impairment after stroke. N Engl J Med 2010;362(19):1772-1783 [FREE Full text] [doi: 10.1056/NEJMoa0911341] [Medline: 20400552]
11. Boardsworth K, Rashid U, Olsen S, Rodriguez-Ramirez E, Browne W, Alder G, et al. Upper limb robotic rehabilitation following stroke: a systematic review and meta-analysis investigating efficacy and the influence of device features and program parameters. J Neuroeng Rehabil 2025;22(1):164 [FREE Full text] [doi: 10.1186/s12984-025-01662-4] [Medline: 40671078]
12. Rodgers H, Bosomworth H, Krebs HI, van Wijck F, Howel D, Wilson N, et al. Robot assisted training for the upper limb after stroke (RATULS): a multicentre randomised controlled trial. Lancet 2019;394(10192):51-62 [FREE Full text] [doi: 10.1016/S0140-6736(19)31055-4] [Medline: 31128926]
13. Bertani R, Melegari C, De Cola MC, Bramanti A, Bramanti P, Calabrò RS. Effects of robot-assisted upper limb rehabilitation in stroke patients: a systematic review with meta-analysis. Neurol Sci 2017;38(9):1561-1569. [doi: 10.1007/s10072-017-2995-5] [Medline: 28540536]
14. Hogan N, Krebs HI, Rohrer B, Palazzolo JJ, Dipietro L, Fasoli S, et al. Motions or muscles? Some behavioral factors underlying robotic assistance of motor recovery. J Rehabil Res Dev 2006;43(5):605-618. [doi: 10.1682/jrrd.2005.06.0103] [Medline: 17123202]
15. Krebs HI, Hogan N, Aisen ML, Volpe BT. Robot-aided neurorehabilitation. IEEE Trans Rehabil Eng 1998;6(1):75-87 [FREE Full text] [doi: 10.1109/86.662623] [Medline: 9535526]
16. Turchetti G, Vitiello N, Trieste L, Romiti S, Geisler E, Micera S. Why effectiveness of robot-mediated neurorehabilitation does not necessarily influence its adoption. IEEE Rev Biomed Eng 2014;7:143-153. [doi: 10.1109/RBME.2014.2300234] [Medline: 24803207]
17. Morone G, Paolucci S, Cherubini A, De Angelis D, Venturiero V, Coiro P, et al. Robot-assisted gait training for stroke patients: current state of the art and perspectives of robotics. Neuropsychiatr Dis Treat 2017;13:1303-1311 [FREE Full text] [doi: 10.2147/NDT.S114102] [Medline: 28553117]
18. de Witte L. Emerging technologies and the field of assistive technology: possibilities and challenges. In: Rehabilitation Technology in Transformation: A Human-Technology-Environment Perspective. Dortmund, Germany: Eldorado, TU Dortmund University; 2024:35-43.
19. Pignolo L. Robotics in neuro-rehabilitation. J Rehabil Med 2009;41(12):955-960. [doi: 10.2340/16501977-0434] [Medline: 19841823]
20. Hatem SM, Saussez G, Della Faille M, Prist V, Zhang X, Dispa D, et al. Rehabilitation of motor function after stroke: a multiple systematic review focused on techniques to stimulate upper extremity recovery. Front Hum Neurosci 2016;10:442 [FREE Full text] [doi: 10.3389/fnhum.2016.00442] [Medline: 27679565]
21. Norouzi-Gheidari N, Archambault PS, Fung J. Effects of robot-assisted therapy on stroke rehabilitation in upper limbs: systematic review and meta-analysis of the literature. J Rehabil Res Dev 2012;49(4):479-496. [doi: 10.1682/jrrd.2010.10.0210] [Medline: 22773253]
22. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
23. Klamroth-Marganska M, Blanco J, Campen K, Curt A, Dietz V, Ettlin T, et al. Three-dimensional, task-specific robot therapy of the arm after stroke: a multicentre, parallel-group randomised trial. Lancet Neurol 2014;13(2):159-166 [FREE Full text] [doi: 10.1016/S1474-4422(13)70305-3] [Medline: 24382580]
24. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford, United Kingdom: Oxford University Press; 2015.
25. Mehrholz J, Pohl M, Platz T, Kugler J, Elsner B. Electromechanical and robot-assisted arm training for improving activities of daily living, arm function, and arm muscle strength after stroke. Cochrane Database Syst Rev 2018;9(9):CD006876 [FREE Full text] [doi: 10.1002/14651858.CD006876.pub5] [Medline: 30175845]
26. Calabrò RS, Morone G, Naro A, Gandolfi M, Liotti V, D'aurizio C, et al. Robot-assisted training for upper limb in stroke (ROBOTAS): an observational, multicenter study to identify determinants of efficacy. J Clin Med 2021;10(22):5245 [FREE Full text] [doi: 10.3390/jcm10225245] [Medline: 34830527]
27. Dantas LF, Fleck JL, Cyrino Oliveira FL, Hamacher S. No-shows in appointment scheduling - a systematic literature review. Health Policy 2018;122(4):412-421. [doi: 10.1016/j.healthpol.2018.02.002] [Medline: 29482948]

XSL•FO
RenderX

28. Hesse S, Schulte-Tigges G, Konrad M, Bardeleben A, Werner C. Robot-assisted arm trainer for the passive and active practice of bilateral forearm and wrist movements in hemiparetic subjects. Arch Phys Med Rehabil 2003;84(6):915-920. [doi: 10.1016/s0003-9993(02)04954-7] [Medline: 12808550]
29. Li L, Tyson S, Weightman A. Professionals' views and experiences of using rehabilitation robotics with stroke survivors: a mixed methods survey. Front Med Technol 2021;3:780090 [FREE Full text] [doi: 10.3389/fmedt.2021.780090] [Medline: 35047969]
30. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs A, Carswell C, et al. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. BMJ 2022;376:e067975 [FREE Full text] [doi: 10.1136/bmj-2021-067975] [Medline: 35017145]
31. Banta D. The development of health technology assessment. Health Policy 2003;63(2):121-132. [doi: 10.1016/s0168-8510(02)00059-3] [Medline: 12543525]
32. Svempe L. Exploring impediments imposed by the Medical Device Regulation EU 2017/745 on software as a medical device. JMIR Med Inform 2024;12:e58080 [FREE Full text] [doi: 10.2196/58080] [Medline: 39235850]
33. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). Official Journal of the European Union. 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2026-10-22]

Abbreviations

ADL: activities of daily life
CHEERS: Consolidated Health Economic Evaluation Reporting Standards
DC: direct cost
EBITDA: earnings before interest, taxes, depreciation, and amortization
FMA: Fugl-Meyer Assessment
FTE: full-time equivalent
GDPR: General Data Protection Regulation
IEC: International Electrotechnical Commission
ISO: International Organization for Standardization
MDR: Medical Device Regulation

Edited by J Sarvestan; submitted 15.Feb.2026; peer-reviewed by K Khashu, Y Jin; comments to author 08.Jul.2026; revised version received 19.Oct.2026; accepted 20.Oct.2026; published 11.Nov.2026.

Please cite as:
Scano A
Economic, Organizational, and Research Impact Evaluation of Introducing Robotic Rehabilitation in a Medium-Sized Hospital: Scenario-Based Comparative Analysis
JMIR Rehabil Assist Technol 2026;13:e93572
URL: https://rehab.jmir.org/2026/1/e93572
doi:10.2196/93572
PMID:

©Alessandro Scano. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 11.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

From Metrics to Meaning in Neurological Rehabilitation: Clinicians' Perspectives on Digital Metrics of Upper Limb Functioning—A Focus Group Study

Johannes Pohl[1*], PhD; Laura Mayrhuber[2*], MSc; Olivier Lambercy[2], PhD; Chris Easthope Awai[1], PhD

[1]Data Analytics and Rehabilitation Technology (DART), Lake Lucerne Institute, Rubistrasse 9, Vitznau, Switzerland

[2]Rehabilitation Engineering Laboratory (RELab), Department of Health Sciences and Technology, ETH Zurich, Zürich, Switzerland

[*]these authors contributed equally

Corresponding Author:
Johannes Pohl, PhD
Data Analytics and Rehabilitation Technology (DART), Lake Lucerne Institute, Rubistrasse 9, Vitznau, Switzerland

Related Article:

This is a corrected version. See correction statement: https://rehab.jmir.org/2026/1/e106631

Abstract

Background: Digital assessment technologies, such as optical motion capture and inertial measurement units, enable detailed kinematic analysis and continuous monitoring of upper limb activity in persons with neurological conditions. While such *digital metrics of functioning* are increasingly recognized in research, their uptake in clinical neurorehabilitation is limited. It remains unclear which *digital metrics of functioning* clinicians perceive as most meaningful and how these are integrated into patient-centered care. Understanding clinicians' information needs and reasoning processes is a prerequisite for implementing digital assessment technology.

Objective: This study aims to characterize how rehabilitation professionals perceive, prioritize, and integrate *digital metrics of functioning* into clinical reasoning and to identify features that would support their routine use.

Methods: Three 90-minute focus groups were conducted in 3 Swiss neurorehabilitation centers, involving 11 clinicians with diverse professional backgrounds (5 physiotherapists, 4 occupational therapists, 1 movement scientist, and 1 medical practitioner). Participants discussed essential parameter domains and individually rated the relevance and meaningfulness of 17 kinematic metrics for the well-studied drinking task and 10 established arm use performance metrics. Verbatim transcripts were analyzed using reflexive thematic analysis, and rating data were summarized descriptively.

Results: Five main themes were identified. (1) *Functional requirements to interpret movement quality and performance* (active/passive range of motion, strength, selective muscle control, and grasp) form the basis for interpreting movement. (2) *Essential aspects of movement quality* (smoothness, efficiency, and compensatory movement) are valued when aligned with observable task execution. (3) *Added value of real-world performance* (hourly activity profiles, arm-use symmetry, and functional workspace) represents the reference for patient-centered reasoning. (4) *Individualizing what matters*, including diagnosis-specific preferences, shapes assessment selection. (5) *Blending clinical eye and reference data* reflects clinicians' reliance on visual judgment complemented by normative values. Intuitive metrics such as task duration, number of movement units, and range of motion were favored, whereas confidence was lower in more complex metrics (eg, jerk and interjoint coordination).

Conclusions: Clinicians value intuitive *digital metrics of functioning* when they are clearly linked to patient-centered outcomes and supported by normative references. The findings highlight the need for targeted educational strategies and digital competency training that help clinicians interpret digital metrics and integrate them with contextual information and clinical reasoning.

(JMIR Rehabil Assist Technol 2026;13:e87339) doi:10.2196/87339

KEYWORDS

digital metrics; upper limb; kinematic metrics; real-world performance metric; clinical reasoning; neurorehabilitation

XSL•FO
RenderX

Introduction

Upper limb dysfunction is among the most disabling consequences of neurological conditions such as stroke and spinal cord injury (SCI). Loss of dexterity and limited arm use reduce independence in activities of daily living (ADL) and restrict participation in social and professional roles [1,2]. The International Classification of Functioning, Disability and Health (ICF) captures these interactions under the umbrella term "functioning," which encompasses body functions, activities, and participation [3]. Assessment of upper limb functioning is therefore central to motor rehabilitation, informing both patient-centered goal setting [4] and comprehensive recovery monitoring [5,6].

Guidelines for clinical assessments after stroke emphasize the routine application of established core sets to ensure that clinically relevant and meaningful domains of functioning are systematically evaluated [7-9]. In conventional clinical assessments, clinicians evaluate the patient's maximal voluntary capacity, relying on observation and standardized rating scales to assess constructs such as active range of motion (ROM) and patient-reported ADL performance. These methods are often integrated into the decision-making process known as clinical reasoning, which clinicians use to interpret information and guide treatment. However, these assessments are limited in their ability to adequately and reliably capture movement quality and detect subtle changes [10,11]. Additionally, upper limb capacity assessments often fail to align with the actual performance in real-world settings. Multiple studies show a capacity-performance mismatch [12,13] and indicate that gains on standardized tests do not always translate into increased arm use [14,15].

To address these limitations, technology-based assessments have been developed to reliably quantify upper limb movements with high temporal resolution. In analogy to the ICF framework, we introduce a technology-based counterpart: *digital metrics of functioning*. These metrics are derived from technologies such as optical motion capture, computer vision, and wearable sensors, which enable the quantification of fine-grained movement execution and real-world activity. *Digital metrics of functioning* can be grouped into 2 complementary domains. In our framework, movement quality is the overarching construct that characterizes how a movement is executed. It can be evaluated through observational movement analysis (clinician-rated) or instrumented kinematic analysis, the latter yielding kinematic metrics that quantify aspects of movement quality such as smoothness, movement time, and compensatory movements during both standardized and habitual tasks. In standardized tasks, these metrics reflect what an individual can perform under controlled conditions, that is, upper limb capacity.

In contrast, *real-world performance metrics* reflect arm use behavior in daily life, for example, through average acceleration or arm use symmetry, thereby capturing arm use in people with neurological conditions in a habitual environment.

Although related, *kinematic* and *real-world performance metrics* represent distinct constructs. Improvements in task-based kinematic metrics, that is, movement quality or upper limb capacity, do not necessarily translate into arm use behavior in a habitual environment. Moreover, real-world performance may be influenced by contextual, motivational, and behavioral factors beyond the assessed movement quality and capacity under controlled conditions [15,16]. By integrating these different domains into clinical practice, digital measures could provide information beyond observer-based ratings [17,18] and enable clinicians to identify meaningful patterns, such as a potential mismatch between increased capacity and reduced real-world arm use performance [15,19]. For instance, by capturing both *digital metrics of functioning* (ie, *kinematic* and *real-world performance metrics*) and clinical assessments, individuals requiring special attention can be identified, including those who show a mismatch between increased functional capacity and declining real-world performance. These people may benefit from personalized rehabilitation that includes monitoring arm use performance, feedback, and targeted interventions [20]. This underscores the value of multidomain assessments in detecting such discrepancies and guiding intervention planning.

Among *kinematic metrics*, those quantifying the movement quality of reach-to-grasp movements have been validated most extensively [21]. The reach-to-grasp function is critical for independence in daily life [22,23] and provides insights into recovery mechanisms [11,16]. The drinking task, a standardized and meaningful ADL, has been recommended by international expert consensus for assessing upper limb movement quality in both research and clinical contexts [11]. A set of validated *kinematic metrics* exists for individuals with stroke or SCI, as well as healthy populations [24-26].

Real-world performance metrics are particularly valuable for monitoring arm use behavior in uncontrolled, ecologically valid environments [27]. Established *real-world performance metrics* include measures of arm use duration and intensity [28,29], which are widely used in research but rarely in clinical settings [30]. Novel metrics, such as real-world reach-to-grasp quantification [31] or mapping functional reaching space [32], provide task-specific insights into the real-world performance of the upper limbs.

Both kinematic and real-world performance metrics are modulated by contextual and task-related constraints, such as health status, environmental factors, and personal factors, which influence movement execution [33,34] and everyday arm use [35] beyond underlying motor capacity. When operating in an ICF-based framework, clinicians need complementary information on health-related goals, environments, and individual behavior to integrate *digital metrics of functioning* into clinical reasoning.

Despite these advances, it remains insufficiently understood how established *digital metrics of functioning* are interpreted and integrated into everyday clinical reasoning by frontline clinicians. While international expert consensus initiatives have defined key metrics and assessment paradigms [8,9,11], less is known about how practicing rehabilitation professionals make sense of these metrics in relation to their clinical expertise and decision-making processes. To support user-centered adoption of digital assessment technologies, evidence is needed on how established *digital metrics of functioning* can be harmonized

XSL•FO
RenderX

with clinicians' empirical knowledge and contextual judgment. The objective of this study was to explore rehabilitation professionals' perspectives on the use of *digital metrics of functioning* when assessing the upper limb in neurological conditions. In this context, relevant *digital metrics of functioning* are understood as those that convey actionable meaning for clinical understanding, decision-making, and communication in practice. By examining how clinicians interpret and integrate such metrics into clinical reasoning, this perspective complements existing consensus and validation work by adding the dimension of clinician-perceived interpretability and clinical utility. Specifically, we addressed the following research questions:

1. Which measurement constructs do clinicians consider essential when interpreting *digital metrics of functioning* of the upper limb (ie, *kinematic* and *real-world performance metrics*) in neurological rehabilitation?
2. Which specific *kinematic* and *real-world performance metrics* are regarded as relevant and intuitive for daily clinical practice?
3. Which complementary information is deemed important to facilitate the integration of *kinematic* and *real-world performance metrics* into clinical reasoning?

Methods

Study Design

We conducted a focus group study using reflexive thematic analysis (realist orientation) to explore opinions on metrics for functioning in the Swiss neurorehabilitation landscape. Focus groups were held between December 2023 and July 2024. All participants received verbal and written information about the study and provided written informed consent prior to participation. Participants also consented to the audio and video recording of the focus group discussions for research purposes. Data were anonymized during transcription, and any potentially identifiable details were removed before analysis and reporting.

Ethical Considerations

According to the Swiss Federal Act on Research involving Human Beings (HRA), the study did not require formal ethical approval because it involved health care professionals only and did not collect health-related personal data. A formal determination of competence (Req-2025-009-30) was obtained after data collection, serving as administrative confirmation that the project was outside the scope of the HRA.

Participants and Recruitment

Invitations were distributed via email, along with a recruitment flyer addressed to therapy leads and medical directors at 8 neurorehabilitation clinics. Clinics were asked to nominate at least 3 clinicians who treated neurological patients with upper limb impairments and had at least 3 years of clinical experience. Prior experience with technology-based assessment was described as beneficial but not required. Nonparticipation was mainly due to organizational constraints (ie, allocating staff for a 90-minute time commitment). Out of 8 invited clinics, 3 agreed to participate in a focus group discussion (one per site).

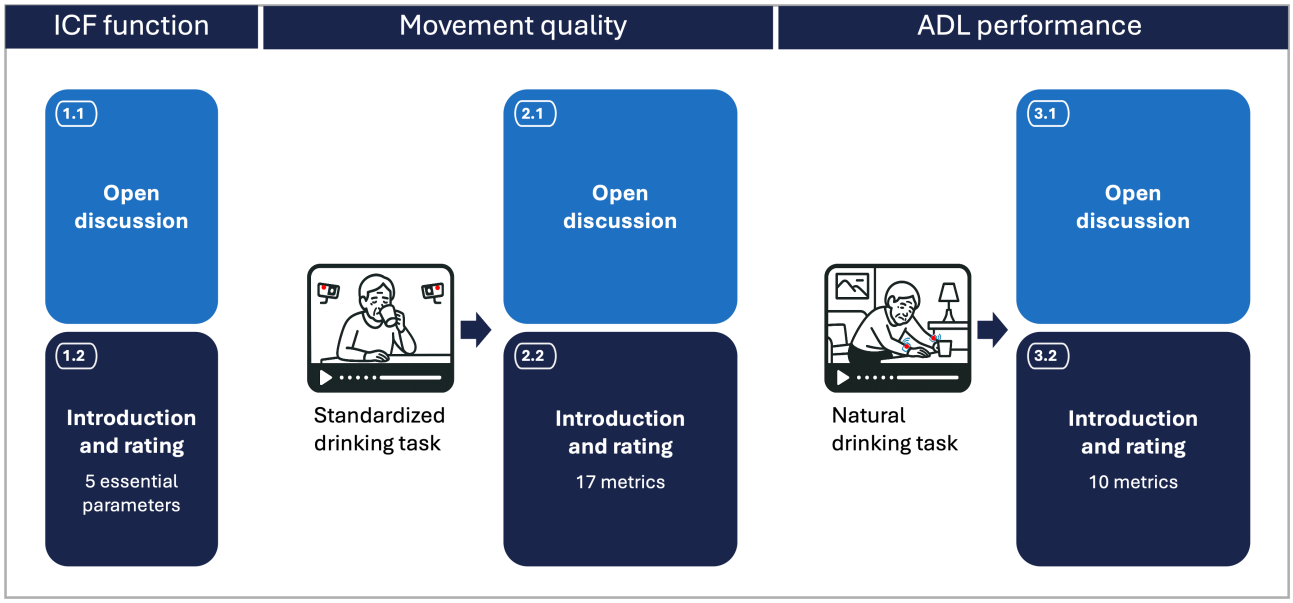
Procedures

Focus groups, each lasting 90 minutes, were held in quiet meeting rooms at the participating clinics and guided by a moderator (JP, a physiotherapist with 10 years of experience in mixed clinical and research settings), who had no prior supervisory relationship with the participants. A co-moderator (LM, a physiotherapist) managed audio recordings, field notes, and whiteboard visualizations. Focus groups consisted of 3 to 4 participants. Smaller focus groups (3 - 5 participants) are considered appropriate in health research when participants share specific expertise, as they allow more time per participant and facilitate open, more in-depth discussion [36,37].

Participants were familiarized with the principles of focus group discussions and the ICF as the conceptual framework for health assessments. The focus group discussions were conducted according to an interview guideline (Multimedia Appendix 1), which included key questions and practical patient examples, and were structured into 3 parts (Figure 1). Each part began with a discussion of parameters that are relevant and meaningful for quantifying the respective ICF domain, followed by participants indicating the importance of each parameter on individual handouts. Throughout this paper, the term "rating" denotes an indication of importance (yes/no selection) for each parameter.

XSL•FO
RenderX

Figure 1. Structure of focus group discussions, including joined components of discussions before ratings. ADL: activities of daily living; ICF: International Classification of Functioning, Disability and Health.



The first part of the session focused on identifying key metrics within the ICF body function domain (32 items, including ROM, muscle strength, and somatosensory function; Table S1 in Multimedia Appendix 2). This section served as a warm-up to familiarize participants with the domain-specific reasoning process. To stimulate discussion on prioritization and underlying reasoning at the beginning, participants were limited to selecting a maximum of 5 metrics they considered most important in their clinical routines. As this exercise was intended to stimulate reasoning rather than generate interpretable quantitative rankings, the individual votes were not analyzed as outcomes. For transparency, the full table of items is provided in Table S1 in Multimedia Appendix 2.

The second part, focusing on kinematic metrics that quantify aspects of movement quality, began with a video clip showing a patient with moderate one-sided paresis performing the standardized drinking task from both frontal and sagittal perspectives. The drinking task was designed as a representative reach-to-grasp activity to provide a shared, standardized reference for discussing movement quality constructs; however, participants were encouraged to reflect on upper limb movement more broadly, beyond this specific task. Participants then discussed important movement characteristics, clinically relevant features, and requirements for clinical reasoning. The concept of "relevance" was intentionally not predefined or operationalized using explicit criteria. Instead, participants were invited to discuss *digital metrics of functioning* within clinical scenarios, allowing perceptions of relevance to be implied through how metrics were contextualized, interpreted, and integrated into clinical reasoning. In practice, participants appeared to associate relevance with aspects such as clinical interpretability, usefulness for decision-making, familiarity, feasibility in clinical workflows, and perceived patient-centered value.

Thereafter, participants indicated the perceived importance of a list of kinematic metrics for the drinking task [24,25], which are commonly categorized into the following domains: *velocity and movement time*, *movement smoothness and coordination*, *movement strategy*, and *ROM* (Table S2 in Multimedia Appendix 2).

The third part included key discussion points on upper limb performance during ADLs, prompted by video clips of the same patient performing the drinking task in a daily-life context. Metric importance was assessed across 10 performance metrics in the domains of *arm use duration*, *arm use intensity*, and *reaching count and space*. Metrics quantifying the duration and intensity of arm use have shown good measurement properties [29] and explain real-world performance well [28] (Table S3 in Multimedia Appendix 2). Despite the limited evidence base for validity, the metrics for reaching count and movement space were included based on previous reports of clinical professionals' preferences [38].

Data Analysis

The participants' opinions were analyzed using a mixed methods analysis. In qualitative evaluations, reflexive thematic analysis was applied to transcriptions, while ratings were evaluated quantitatively by calculating, for each metric, the proportion of participants who selected it. In the ICF-function section, quantitative evaluations of ratings were excluded due to differences in rating logic.

The reflexive thematic analysis framework [39] enables an interpretative, flexible, and researcher-reflexive engagement with participants' perspectives, making it particularly well suited to capturing how clinical experts construct clinical reasoning, decision-making, and value-based evaluation of upper limb functioning. Transcripts were imported into the text annotation software Taguette (v1.4.1) [40] and coded by 2 coders (JP and LM) from a realist perspective, assuming that the participants' words represented their actual experiences and opinions. Codes were generated by adopting an inductive-deductive approach, combining data-driven coding with concepts from existing frameworks (Table 1). Both coders discussed the essential aspects, interpretations, and underlying theories of code labels

XSL•FO
RenderX

in regular meetings. Codes were then structured into clusters, forming subthemes and superordinate themes, by 2 researchers (JP and LM), who collated coherent clusters. The meaning, reflexivity, and coherence of themes were then discussed between both researchers. Differences in coding or theme boundaries were addressed through deeper analysis of the coded data. The researchers explored how different interpretations reflected different aspects of participants' opinions. Where divergent interpretations arose, both perspectives were discussed and revised as the analysis progressed.

Finally, verbatim German quotations and their English translations, together with contextual descriptions of the respective subthemes, were sent to participants for review. All participants confirmed that the translations accurately reflected the original meaning and context. This study was conducted and reported in accordance with the Standards for Reporting Qualitative Research (SRQR) [41]. The completed SRQR checklist, indicating where each item is addressed in the paper, is provided in Checklist 1.

Table . Coding process.

Step	Action
1. Inductive coding	Independent coding (JP and LM) capturing semantic meaning line-by-line. No fixed codebook or a priori agreement threshold was imposed.
	Discussion, merging, or retaining code labels whose semantics were judged "functionally identical." Differences in coding were resolved with deeper analysis of the code and discussions of divergent interpretations until consensus was reached.
2. Deductive mapping	Code mapping of agreed labels to ICF[a] and existing measurement taxonomies.
3. Theme construction	Code-clustering into preliminary themes (JP and LM) upon agreement for coherence, divergences, and reflexive commentary.
4. Integration of quantitative priorities	Alignment of priority ratings with qualitative themes in a joint display, where illustrative quotes explain *why* specific metrics were prioritized or deprioritized.

[a]ICF: International Classification of Functioning, Disability and Health.

Results

Study Sample and Focus Group Characteristics

Out of the 8 invited clinics, 3 participated in the focus group study, including a total of 11 participants across 3 focus groups (one conducted in each clinic). Participants presented a range of professional backgrounds, including physiotherapy, occupational therapy, movement science, and medical practice, and varied in clinical roles and experience levels (Table 2). Clinical experience ranged from 4 to 26 years, and research experience ranged from 0 to 20 years. Average durations of focus group discussions were 33 (SD 9.6) minutes for ICF functions, 25 (SD 1.5) minutes for kinematic metrics, and 15 (SD 3.5) minutes for real-world performance, respectively.

XSL•FO
RenderX

Table . Participant characteristics.

ID	Field	Specialization	Profession	Clinical experience in years	Research experience in years
P1-C1	Stroke, ND[a], SCI[b]	Therapy sciences	PT[c]	6	20
P2-C1	Stroke, ND, SCI	Neurology	MP[d]	15	15
P3-C1	Stroke, ND, SCI	Neurorehabilitation	PT	4	1
P4-C2	SCI	Technology-assisted therapy	MS[e]	6	4
P5-C2	SCI	Upper extremity–focused therapy	MS	8	11
P6-C2	SCI	Upper extremity–focused therapy	PT	8	1
P7-C2	SCI	Upper extremity–focused therapy	OT[f]	18	0
P8-C3	Stroke, ND, SCI	CIMT[g], PNF[h]	OT	26	3
P9-C3	Stroke, ND, SCI	Technology-assisted therapy	OT	3	0
P10-C3	Stroke, ND, SCI	Neurorehabilitation	PT	17	0
P11-C3	Stroke, ND, SCI	Neurorehabilitation	OT	4	0

[a]ND: neurodegenerative diseases.

[b]SCI: spinal cord injury.

[c]PT: physiotherapist.

[d]MP: medical practitioner.

[e]MS: movement scientist.

[f]OT: occupational therapist.

[g]CIMT: constraint-induced movement therapy.

[h]PNF: proprioceptive neuromuscular facilitation.

Five themes were identified in the analysis: (1) functional requirements to interpret movement quality and performance, (2) essential aspects of movement quality, (3) added value of real-world performance, (4) individualizing what matters, and (5) blending clinical eye and reference data. Each theme comprises between 2 and 4 subthemes, as shown in Figure 2. Illustrative quotations are referenced in the text using an identifier (eg, Q1). The corresponding quotations, including participant IDs and full wording, are provided in Tables S1-S5 in Multimedia Appendix 3. Quantitative ratings of *kinematic* and *real-world performance metrics* are presented in Figure 3 and are reported convergently within respective themes.

XSL•FO
RenderX

Figure 2. Themes and subthemes from the clinician focus group analysis illustrating the key dimensions guiding the interpretation and integration of *digital metrics of functioning*.

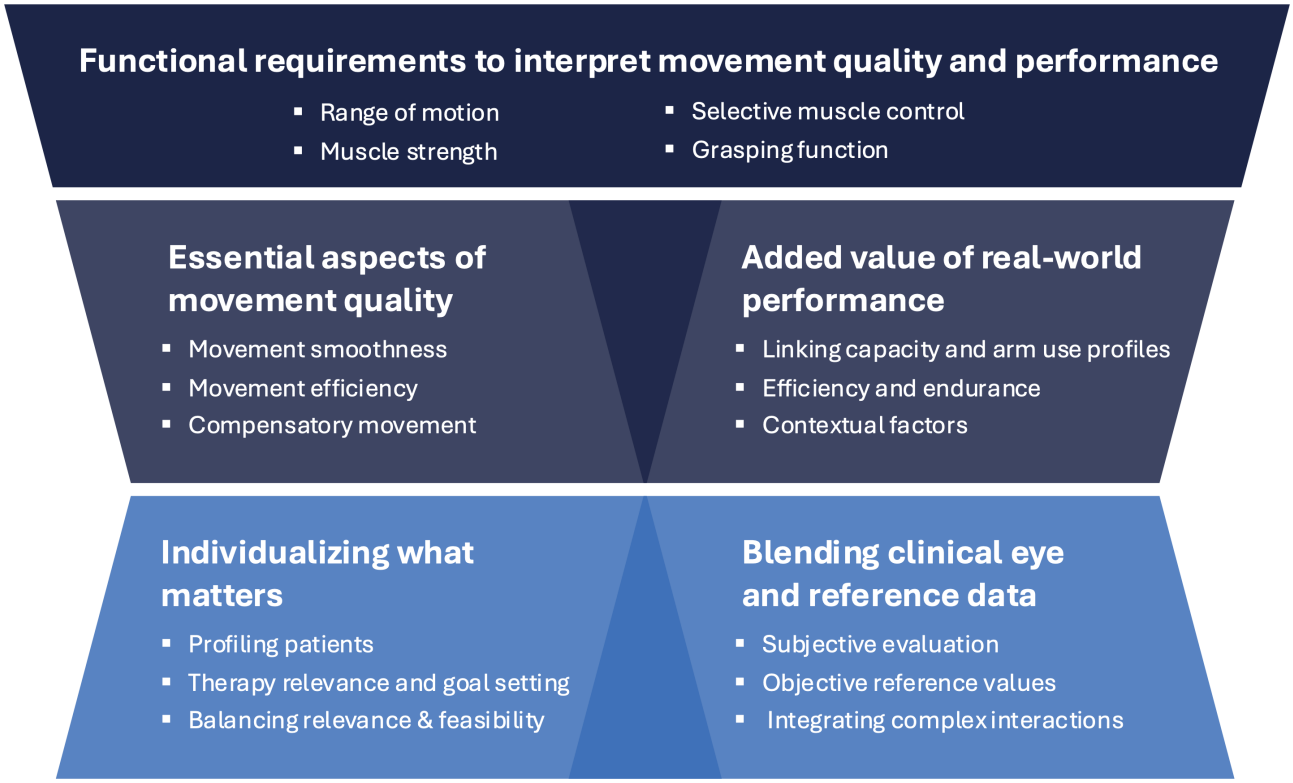
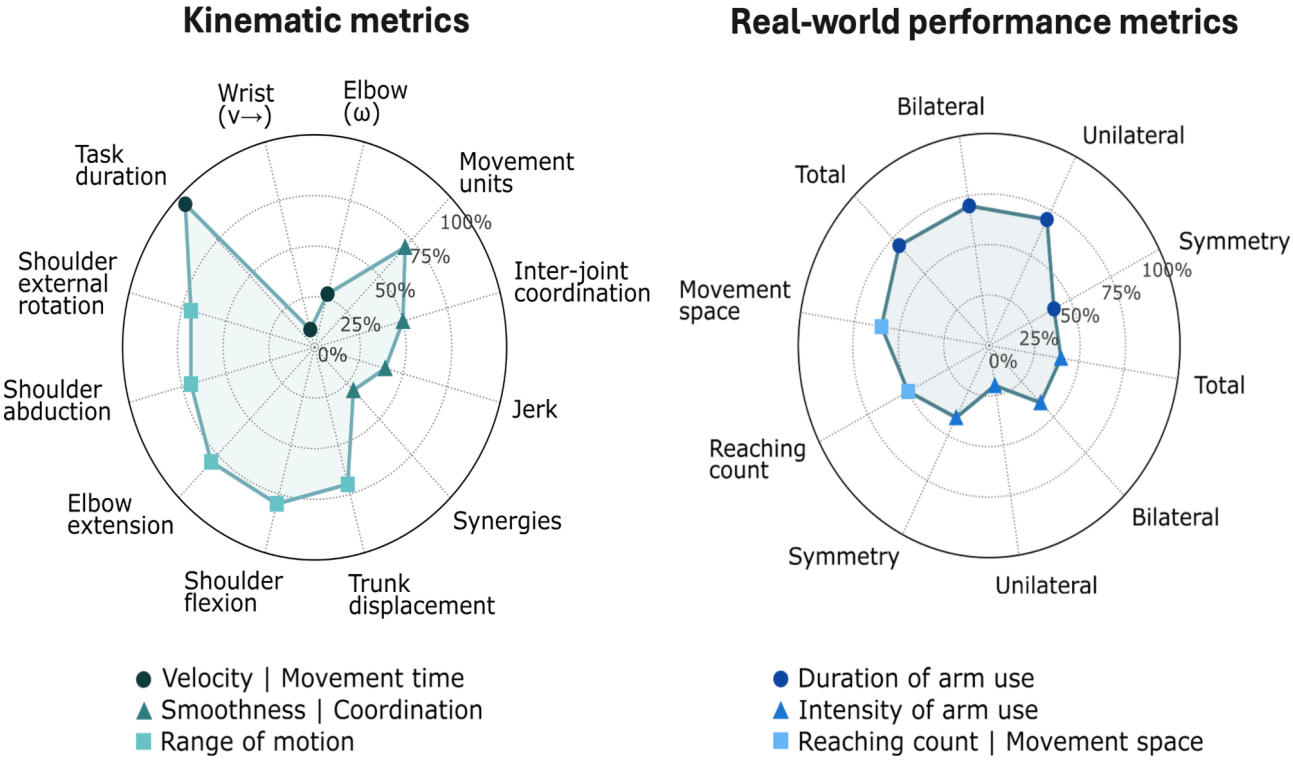


Figure 3. Radar plots representing the proportion of participants who rated each metric as meaningful and relevant for clinical practice. Kinematic metrics and real-world performance metrics, color-coded by metric domains. In the real-world performance panel, metrics with identical names (eg, symmetry) are distinguished by their domain (intensity vs duration), as indicated in the legend. Percentages represent the proportion of participants who rated each metric (0% at the center indicates no rating; 100% at the outer ring indicates all participants rated). The movement strategy domain was excluded from the *kinematic metrics* because no participant identified it as important. *v*→: velocity; ω: angular velocity. Detailed metric definitions and units are provided in Tables S2 and S3 in Multimedia Appendix 2.



XSL•FO
RenderX

Theme 1: Functional Requirements to Interpret Movement Quality and Performance

This theme outlines the functional requirements that clinicians considered essential for meaningful interpretation of movement quality, whether assessed through observational movement analysis or instrumented kinematic analysis. These prerequisites, typically assessed through standard clinical evaluations, provide the baseline context needed to interpret technology-based metrics. Participants highlighted 4 key aspects as particularly decisive for understanding upper limb performance: range of motion, muscle strength, selective muscle control, and grasping function (Table S1 in Multimedia Appendix 3).

Range of Motion

Participants consistently emphasized that both active and passive ROM must be known to interpret movement quality and determine whether upper limb ADL can be attempted. For instance, elevation above 90° was highlighted as necessary for tasks such as combing hair or reaching overhead [Q1]. Importantly, clinicians considered functional joint interaction, such as shoulder, elbow, and spinal mobility, as a whole rather than in isolation.

Muscle Strength

Assessing strength was considered an important complement to active ROM, as it provides information on stability and control during functional movements such as grasping or lifting objects. Rather than isolated maximal testing, participants preferred strength assessment in task contexts, as this better reflects real-world performance. One clinician noted that strength “is always somewhere in relation to mobility” [Q2].

Selective Muscle Control

The ability to activate muscles selectively, such as flexing the elbow while relaxing the extensors, was described as crucial for coordinated movement. Impaired selectivity, often linked with sensory-motor deficits, could lead to compensatory strategies or coexist with abnormal muscle tone [Q3, Q4]. Clinicians highlighted that such neuromuscular control issues strongly influence how movement quality should be interpreted.

Grasping Function

Finally, hand opening and grip strength were identified as the most practical tests of functional capacity. Grasping integrates ROM, strength, and selectivity into a single action and was considered indispensable for ADL tasks involving objects, such as drinking from a glass [Q5].

Taken together, these 4 requirements provide a functional foundation for interpreting movement analysis in clinical contexts. Without them, technology-driven assessments risk misrepresenting patients’ true abilities. Participants also noted that this information helps them *reason about the causes of impaired movement quality*, for example, whether limitations in ROM, strength, selectivity, or grasping contribute to observed difficulties.

Theme 2: Essential Aspects of Movement Quality

This theme highlights movement smoothness, efficiency, and compensatory movements as the 3 key attributes that clinicians use to judge movement quality in relation to the drinking task (Table S2 in Multimedia Appendix 3). These aspects capture how well movements are performed, how effortful they are, and whether patients rely on compensatory patterns.

Movement Smoothness

Participants referred to fluent versus jerky movement as meaningful indicators of movement quality [Q6.2]. Participants with research experience favored jerk as an intuitive single-value measure, whereas therapists were less familiar with such metrics. Although the concept of smoothness appeared familiar to therapists, they were not familiar with jerk or interjoint coordination, as reflected in ratings ranging from 4 (40%) to 5 (50%) across 10 participants (Figure 3). The *kinematic metrics* of the movement strategy domain for the drinking task were considered irrelevant or unintuitive and were not selected by any participant.

Movement Efficiency

Efficiency was understood as completing a task within a reasonable time and without visible strain. For instance, clinicians looked for speed and absence of effort cues, such as facial tension [Q7.1]. The metric *task duration* was indicated as important by all participants. The time required to complete a task was found to be an interesting factor in determining whether the task would be adopted into daily movement behavior. Quantitative prioritization echoed this emphasis: *task duration* received the highest mean rating among *kinematic metrics* (Figure 3).

Compensatory Movement

Compensatory movement, especially excessive trunk flexion or shoulder elevation, was considered a marker for adaptation to motor impairment [Q8.1]. Thus, clinicians were particularly interested in distinguishing between compensatory and physiologically correct movement. Participants found the analysis of compensatory movements challenging due to the dynamic interactions among multiple joints. Participants expected certain compensatory movement patterns involving the shoulder and torso [Q8.2, Q8.3], based on functional deficits highlighted in theme 1.

Theme 3: Added Value of Real-World Performance

Clinicians emphasized that *what patients do outside the therapy room* ultimately determines functional recovery. Three interrelated aspects of real-world performance framed this discussion: the amount and patterns of arm use, efficiency and endurance, and contextual factors influencing ADL behavior (Table S3 in Multimedia Appendix 3).

Linking Capacity and Arm Use Profiles

Participants valued continuous activity profiles across the day, with timing information to link activities to mealtimes or therapy sessions. Altered patterns, such as asymmetrical arm use, were flagged as essential indicators of disability and the need for intervention [Q9.1-Q9.3]. The ability to track both the amount and timing of use was seen as highly relevant for follow-up. Quantification of the functional workspace, as reflected in terms such as “volume” or “radius” of purposeful movement, was

XSL•FO
RenderX

considered highly relevant for capturing real-world performance [Q10] relative to instantaneous ROM and muscle strength [Q11].

Efficiency and Endurance

Clinicians emphasized that performance in everyday tasks depends not only on speed but also on the sustained capacity to repeat movements without excessive strain [Q10.1, Q10.2]. They noted that the effortless execution of a straightforward standardized task indicated a likely transfer into ADLs, whereas repeatedly eating or drinking at mealtimes required substantially more endurance due to the effortful nature of the movements [Q10.3]. This distinction between instantaneous efficiency and sustained endurance was considered central to judging functional outcomes.

Contextual Factors

Clinicians considered individual motivation, fatigue tolerance, and frustration levels, along with environmental influences such as assistive devices, to be crucial for real-world performance [Q11.1]. Wheelchair use was mentioned as particularly important, as it fundamentally shapes the range and types of upper limb movement available [Q11.2]. Such contextual modifiers were consistently described as necessary to interpret measures of arm-use volume and endurance in relation to each patient's capacities and circumstances.

Taken together, clinicians considered real-world performance metrics most clinically meaningful when they combined patterns of arm use with endurance and contextual information.

Theme 4: Individualizing What Matters

This theme underscores the importance of individualized selection of assessments, as standardized protocols often fail to capture patient-specific differences. Clinicians highlighted tailoring measures to priority ADLs, balancing functional relevance with feasibility, and adapting to diagnosis-specific challenges (Table S4 in Multimedia Appendix 3).

Profiling Patients

Patients with the same diagnosis often require different selective assessments depending on residual function, spasticity, or use of assistive devices. For example, transfers in paraplegic patients demand highly adaptive assessments of wheelchair transitions: One participant stated, "In our patient clientele, I would say that this is more in the direction of highly individualized functions rather than a one-size-fits-all model" [Q12.1]. As confirmed by another clinician, severely affected patients, in particular, require a problem-specific assessment in addition to standardized measurements [Q12.2]. Standardized protocols were mentioned as often not suitable, for instance, in patients with SCI, where residual function and spasms can be highly variable. Clinicians agreed that tailoring was especially important for patients with severe conditions and high support needs.

Therapy Relevance and Goal Setting

Goal setting and intervention, informed by individualized assessment, were other key components that emerged from expert discussions [Q13.1]. For example, *real-world performance metrics* were individually screened for patients prone to compensation or learned nonuse, and consequently, effective intervention programs, such as constraint-induced movement therapy or an individual tele-rehabilitation program [Q13.2], were selected. Participants would identify metrics indicating limitations in performing ADLs and evaluate potential changes in their real-world performance profile following the intervention. Therapy goals focus on enhancing specific functions that directly support meaningful activity improvements, with assessments tailored to measurable parameters and joints for active therapeutic intervention. Function alone is not considered a sufficient outcome measure if it is not translated into daily life activity [Q13.4].

Balancing Relevance and Feasibility

Active ROM and muscle strength were prioritized as the most relevant metrics, yet clinicians considered their feasibility and level of detail needed: "You also have to think about the clinic […] what can be measured in everyday clinical practice, or what is realistic here in a laboratory?" [Q14.1]. Basic information, such as whether the upper limb could be moved through full ROM against gravity, was sufficient to judge the patient's capability. Some questioned the value of standardized impairment measures, noting that "What we measure [in] muscle status does not mean that a patient can really use it in the activity" [Q14.2]. Real-world arm use was therefore considered more informative, though more challenging to capture feasibly in daily practice.

Theme 5: Blending Clinical Eye and Reference Data

This theme described how clinicians combine subjective observation, objective reference values, and reasoning across interacting constructs when interpreting movement analysis (Table S5 in Multimedia Appendix 3).

Subjective Evaluation

Clinicians agreed on the importance of visual movement analysis for primarily assessing functional impairment. Clinicians argued to "simply look" and identify how a problematic movement task begins, proceeds, and ends without formal quantification [Q15.1]. Thus, visual observation is an integral part of the therapeutic setting and a straightforward way to identify a motor problem. When exploring potential causes for abnormal movement characteristics (compensatory or jerky movement), taking a closer look at movement units and interpreting *kinematic metrics* became more interesting [Q15.4]. Despite the interest in movement quality, we identified generally cautious opinions on whether *kinematic metrics* would confirm visually flagged deficiencies or support interpretation [Q15.4].

Objective Reference Values

Reference values and normative data played important roles in interpretation and clinical implications. Participants expressed the need for reference values to compare patients with healthy individuals and to assess the meaningfulness of changes registered by *digital metrics of functioning*. Normative data from healthy individuals were consistently mentioned for interpreting movement smoothness metrics (eg, jerk or multijoint coordination). The comparison between the left and right upper limbs was considered a practical norm for patients with unilateral impairments such as stroke [Q16.1]. For bilateral

XSL•FO
RenderX

conditions like SCI, however, clinicians emphasized that reference to healthy controls was crucial [Q16.5]. Yet even when normative data were available, interpreting *kinematic metrics* remained challenging. Because daily movement behaviors are highly variable, applying and comparing such benchmarks to real-world performance was approached with caution [Q16.6].

Integrating Complex Interactions

Clinicians acknowledged the difficulty of disentangling the interrelationships among strength, ROM, coordination, and sensation, which often jointly shape real-world performance. Deviations in smoothness or interjoint coordination were seen as cues to probe underlying impairments, such as force deficits, motor control issues, or sensory loss [Q17.1-Q17.7]. They also emphasized that movement quality (*kinematic metrics*) and *real-world performance metrics* are closely interrelated and that only empirical experience with these measures in practice will build the expertise needed to interpret them reliably [Q17.7]. While movement quality and real-world performance were considered central to clinical reasoning, clinicians acknowledged limited experience applying these metrics and anticipated that their value might shift as they are used more widely in clinical contexts.

Discussion

This study examined clinicians' views on the constructs that underpin functional health, identified which *digital metrics of functioning* were considered clinically meaningful, and outlined what is needed for intuitive integration into clinical systems. Clinicians reasoned across *functional requirements*, *kinematic* and *real-world performance metrics*, while tailoring judgments to the individual and anchoring observations to reference information. Together, these themes demonstrate how *digital metrics of functioning* can become clinically interpretable when integrated into patient-specific reasoning and contextualized within conventional assessment frameworks.

Functional Requirements as the Foundation

With respect to objective 1, clinicians emphasized that ROM, strength, selective muscle control, and grasping function form the foundation for functional health and for the interpretation of movement analysis. Testing functional requirements (eg, passive ROM) is recommended as critical before conducting standardized observation-based assessments, such as the Fugl-Meyer Assessment for Upper Extremity (FMA-UE) [42] and the Action Research Arm Test (ARAT) [43]. The FMA-UE and ARAT are core assessments for stroke [8,11]. They also serve as a reference for identifying benchmark *kinematic* [21,44] and *real-world performance metrics* [29,45] in individuals with neurologic conditions. Notably, impairment measures and *digital metrics of functioning* are not only highly correlated but also exhibit similar recovery trajectories [46,47]. These relationships are robust yet not interchangeable: task instructions (eg, maximal speed vs comfortable speed) and compensatory movements can alter kinematic profiles without corresponding changes in impairment. The reliance on basic impairment and capacity assessment reflects long-standing rehabilitation research. Our findings reinforce that *digital metrics of functioning* should complement rather than replace these clinical assessments.

Digital Metrics of Functioning: Relevance and Meaning

Considering objective 2, the clinicians prioritized *digital metrics of functioning* that are simple, familiar, and intuitively interpretable. In our study, the drinking task duration, number of movement units, trunk displacement, and dynamic motion trajectories of the elbow and shoulder joints were the metrics consistently selected across professions. These metrics have previously shown good reliability [48] and responsiveness (ie, sensitivity to change over time) [47] in stroke populations, and also inform observation-based assessment of drinking tasks [49]. While these metrics were evaluated in the context of the drinking task, the underlying constructs they represent, such as movement efficiency, smoothness, and compensatory strategies, are relevant across a range of functional upper limb activities. Quantifying upper limb movement quality has been found to be critical for assessing and preventing compensatory movements [38], particularly abnormal interactions among shoulder, elbow, and trunk motion trajectories [25,49,50]. Clinicians in our sample found quantification of compensatory movements or movement intermittency/jerk particularly important, as these metrics could help explain a person's limited real-world arm use outside the therapy setting.

Performance Outside Therapy

Revealing patients' activity performance outside inpatient therapy sessions, or even in habitual environments, is considered highly relevant by clinicians in this and other cohorts [51-54]. This insight into real-world arm use behavior enables goal setting and evaluation. This performance-centered clinical decision-making may be enhanced when therapists are provided with sensor-based measures [51,53,55], whereas therapists showed a preference for activity capacity measures [56].

Advances in measurement technology and data processing enable differentiation of activity types, such as walking [57,58], distinguishing functional from nonfunctional upper limb movements [59,60], and even movement primitives, such as reaching or object transportation [61,62]. About half of the therapists in this study selected more specific metrics, such as reaching count and movement space, which were also reported in other studies on clinicians' preferences [38,63]. Therapists in our study emphasized daily arm use behavior, which is effectively captured by arm-use duration metrics, including symmetry, both of which are recognized as key indicators in stroke rehabilitation research [28,29,64]. Performance duration metrics are intuitive and straightforward because they directly measure time, making them easy to understand without additional context. Notably, therapists working in rehabilitation after SCI did not consider symmetry metrics important, which lowered the overall percentage of symmetry metrics. Consistent with theme 4 (individualizing what matters), this suggests that metric relevance is shaped by diagnosis-specific reasoning: in unilateral conditions such as stroke, between-limb symmetry is a meaningful comparison, whereas in bilateral conditions such as SCI, it loses interpretive value. Selection of real-world performance metrics should therefore be aligned with the clinical population rather than applied uniformly.

XSL•FO
RenderX

Reference and Context for Interpretation

Revealing the meaning behind naked numbers can be challenging and requires complementary information, especially for clinicians who lack experience and knowledge about metrics and their psychometric properties. With respect to objective 3, clinicians expressed a need for visual references and normative data to interpret *kinematic metrics*, as well as contextual information and normative data for *real-world performance metrics*. Daily arm use performance can be highly variable, influenced by numerous factors, including health status, activity limitations, functional impairment, personal, and environmental factors [35]. To support interpretation and informed decision-making, clinicians in this study expressed a need for population-specific reference data, alongside cutoff values (eg, percentiles, visual benchmarks, or expected recovery trajectories), to determine whether observed changes are clinically meaningful. In clinical settings, aligning data with daytime and therapy schedules is a practical way to contextualize performance metrics to activity type [51,65]. For goal setting and evaluating therapy effectiveness, the minimal detectable change (MDC) and minimal important change (MIC) are important references for within-person change. Consequently, the successful integration of *digital metrics of functioning* into clinical reasoning depends on their linkage to patient-specific contexts, diagnostic-specific expectations, and usable normative data to inform treatment planning and progression. Benchmarking schemes for upper limb *kinematic* and *real-world performance metrics* [66], together with patient-centered benchmarking approaches [67], offer concrete frameworks that translate changes in *digital metrics of functioning* into actionable targets for neurorehabilitation practice.

Future Directions

Electronic health records are the backbone of information collection and sharing among clinical professionals. An integration of continuous monitoring of *digital metrics of functioning* could enable personalized rehabilitation, provide richer representations of a patient's health status, and enable timely detection of recovery plateaus or declines in real-world performance. To ensure clinical interpretability across heterogeneous neurorehabilitation populations, robust reference datasets are needed that include healthy controls and relevant patient groups, stratified by diagnosis, age, hand dominance, and impairment level. For *kinematic metrics*, task-specific benchmarks for validated tasks, such as the drinking tasks [25,48,68], could provide z-scores or percentiles alongside MDC/MIC thresholds. For *real-world performance metrics*, normative data and MIC thresholds exist but remain limited to stroke populations [29].

Future research should explore the optimal convergent set of meaningful, intuitive information that is accessible through graphical illustration. Nearly a decade ago, Ploderer et al [69] proposed dashboard designs that bring together ROM, kinematic analyses, hourly activity patterns, and 3D functional workspace visualizations. The need for integrated, actionable displays remains, but infrastructural and implementation barriers continue to slow adoption [70,71]. Currently, measurement technologies that capture upper limb *digital metrics of functioning*, such as 3D trajectory analysis and ADL movement space and task detection, remain largely impractical for routine clinical settings. Advances in scalable sensor- or computer vision-based systems are therefore crucial for facilitating precise, clinically relevant visualizations of *kinematic metrics* and functional reaching space during ADL.

Building on the identified clinical reasoning priorities, Weikert et al [72] developed a learning health system embedded in clinical neurorehabilitation practice. This integrated learning health system enables an end-to-end evaluation of multimodal patient data capture and clinical data presentation. These systems aim to integrate kinematics [73], performance data, cognitive/mental status, patient-reported outcomes, speech, sleep, nutrition, and contextual information (including therapy schedules and temporal patterns) [72]. A key principle for clinical embedding is that clinician-facing outcomes must provide meaningful information to ensure adherence and acceptance.

Such systems directly address two priorities raised by our participants: blending objective metrics with contextual reference information (theme 5) and combining task-based capacity with real-world performance profiles (theme 3). Realizing this potential, however, requires systematic evaluation of the system components, the clinical utility [74], usability [75,76], and the real-world effectiveness of new digital outcome domains. Ultimately, a multipronged approach comprising clinical value systems [77], educational programs, and user-friendly data capture is needed to introduce precision neurorehabilitation at scale [78-80].

Training programs should align with recognized digital-health competency frameworks [81,82] to ensure clinicians acquire the knowledge, skills, and attitudes needed to interpret and act upon *digital metrics of functioning*. As demonstrated in electronic health record–based educational initiatives, targeted modules on dashboard use and metric literacy can improve clinician proficiency in real-world data-driven workflows [83]. In the present context, educational efforts should therefore focus on foundational metric interpretation and practical integration of *digital metrics of functioning* into routine decision-making. Such programs could serve as an educational scaffold for translating digital metric evidence into individualized, data-informed rehabilitation planning.

Limitations

This study was conducted in 3 Swiss rehabilitation clinics, which may limit transferability to other health care systems, settings (inpatient vs. outpatient), and languages/cultures. The low participation rate (3/8 clinics, 38%) was primarily due to organizational constraints in releasing multiple clinicians simultaneously for a 90-minute focus group discussion. Recruitment through clinic nomination may have favored clinicians who are more open to digital assessments, potentially more optimistic in their perspectives, although skeptical views also emerged during the focus groups.

The small number of focus groups and the limited number of participants per group may have influenced group interaction dynamics and the breadth of perspectives discussed. Given the

XSL•FO
RenderX

limited number and size of focus groups, thematic saturation cannot be assumed. Furthermore, the small group sizes and site-specific sampling limit the transferability of the findings to clinics with a broader mix of patient populations. Consistent with the exploratory qualitative design, we aimed to provide information-rich insights rather than comprehensive coverage. Accordingly, our findings should be interpreted as exploratory and context-dependent rather than representative or generalizable.

The concept of "relevance" was not operationalized using predefined criteria during the focus group discussions. This approach was chosen to allow clinicians to articulate their own interpretations and reasoning processes. However, participants may have interpreted the term differently, for example, in terms of clinical usefulness, familiarity, feasibility, or patient-centeredness. Such variability may introduce construct ambiguity or response-process bias. While the qualitative analysis allowed us to capture these perspectives through thematic interpretation, future studies could benefit from more explicit operationalization or structured rating criteria.

In addition, the variation in clinicians' prior experience with *digital metrics of functioning* may have influenced interpretations. Future research should examine how perspectives differ across professional roles, experience levels, and clinical contexts and assess the impact of targeted training and exposure on the interpretation of *digital metrics of functioning* in clinical routine.

Moreover, participants were shown an example of a patient performing a drinking task and were evaluated using kinematic metrics validated in this context to provide a tangible, standardized reference and ensure a common frame of reference. While this approach supported understanding of abstract kinematic concepts, presenting a specific task and task-validated metrics may have introduced contextual bias and may limit the full generalizability of the findings.

Both moderator and co-moderator were clinician-researchers, introducing potential social desirability bias despite reflexive views. To mitigate this risk, analytic decisions and theme development were documented through saved versions of code structures, cluster maps, and theme definitions, allowing us to track the evolution of codes and themes.

Finally, we focused on the clinical significance of *digital metrics of functioning* while intentionally setting aside practical constraints. The feasibility, usability, and data-quality requirements of instrumented assessments (eg, sensor setup and maintenance, nonwear detection, calibration procedures, processing resources, and data privacy demands) warrant dedicated evaluation before routine deployment.

Conclusions

This exploratory study mapped how clinicians in Swiss neurorehabilitation interpret and prioritize digital metrics of functioning of the upper limb. Participants valued *kinematic* and *real-world performance* metrics when they were embedded in patient-specific reasoning and paired with contextual information. They consistently favored intuitive metrics, such as task duration, number of movement units, and ROM, over more complex measures whose interpretation requires specialized knowledge. Across the five themes, three priorities emerged for translating digital metrics of functioning into routine practice: (1) stratified normative and reference datasets and meaningful change thresholds, integrated into co-designed dashboards that link task-based kinematic metrics to daily activity profiles; (2) targeted training that builds metric literacy and supports interpretation alongside conventional clinical reasoning; and (3) scalable measurement technologies that integrate seamlessly into clinical workflows, evaluated pragmatically for decision impact, usability, and adoption. These findings should be interpreted as context-dependent rather than generalizable, given the small number of focus groups and the specific clinical and organizational settings sampled. Perspectives are likely to vary across health care systems, professional backgrounds, and levels of experience with digital assessments. Nonetheless, the convergence across professions and clinics on the need for interpretability, contextualization, and clinical embedding offers a clear direction for future work. Centering on clinician and patient preferences will be essential for meaningful integration of *digital metrics of functioning* into clinical practice and for narrowing the evidence-practice gap.

Acknowledgments

The authors would like to thank André Boeni for his valuable support in developing the interview guidelines and conducting the pilot focus groups, and Selina Tanner for administrative support during the first focus group session. We would also like to thank all the volunteers who participated in the study, as well as Sandra Stadelmann for assistance with transcriptions and data administration.

Funding

This work was funded by the P&K Puhringer Foundation.

Data Availability

Due to the qualitative nature of the data and the inclusion of potentially identifiable video and audio material, the raw focus group data (recordings and transcripts) cannot be publicly shared. Deidentified extracts supporting the study findings are included in Multimedia Appendices 2 and 3. Additional anonymized excerpts or coding summaries may be shared by the corresponding author upon reasonable request, subject to ethical and institutional data protection requirements.

XSL•FO
RenderX

Authors' Contributions

JP conceived and led the study, including conceptualization, methodology design, and overall supervision. JP moderated the focus groups with LM as co-moderator and revised the interview transcriptions for correctness. JP and LM conducted data labeling and thematic analysis. JP created all visualizations and drafted the manuscript, with LM contributing complementary sections and revisions. CEA and OL provided critical review, interpretation of findings, and editorial feedback on the final manuscript.

Conflicts of Interest

Some authors (JP, CA) were involved in early pilot work related to the described learning health system; however, this work was conducted in a noncommercial research context and does not constitute a conflict of interest.

Multimedia Appendix 1
Interview guideline for focus groups.
[PDF File, 289 KB - rehab_v13i1e87339_app1.pdf]

Multimedia Appendix 2
Ratings of ICF parameters and digital metrics of functioning. ICF: International Classification of Functioning, Disability and Health.
[PDF File, 92 KB - rehab_v13i1e87339_app2.pdf]

Multimedia Appendix 3
Participants' quotations per theme and subtheme.
[PDF File, 133 KB - rehab_v13i1e87339_app3.pdf]

Checklist 1
SRQR checklist.
[PDF File, 63 KB - rehab_v13i1e87339_app4.pdf]

References

1. Javeed S, Greenberg JK, Zhang JK, et al. Association of upper-limb neurological recovery with functional outcomes in high cervical spinal cord injury. J Neurosurg 2023 Sep 1;39(3):355-362. [doi: 10.3171/2023.4.SPINE2382]
2. Lawrence ES, Coshall C, Dundas R, et al. Estimates of the prevalence of acute stroke impairments and disability in a multiethnic population. Stroke 2001 Jun;32(6):1279-1284. [doi: 10.1161/01.str.32.6.1279] [Medline: 11387487]
3. International Classification of Functioning, Disability and Health (ICF). World Health Organization. 2001. URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2026-06-02]
4. ten Cate O. Introduction. In: ten Cate O, Custers E, Durning S, editors. Principles and Practice of Case-Based Clinical Reasoning Education: Springer; 2018, Vol. 15:3-19. [doi: 10.1007/978-3-319-64828-6_1]
5. Brogioli M, Schneider S, Popp WL, et al. Monitoring upper limb recovery after cervical spinal cord injury: insights beyond assessment scores. Front Neurol 2016;7:142. [doi: 10.3389/fneur.2016.00142] [Medline: 27630612]
6. Kwakkel G, Kollen B, Lindeman E. Understanding the pattern of functional recovery after stroke: facts and theories. Restor Neurol Neurosci 2004;22(3-5):281-299. [doi: 10.3233/rnn-2004-00282] [Medline: 15502272]
7. Kwakkel G, Stinear C, Essers B, et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. Eur Stroke J 2023 Dec;8(4):880-894. [doi: 10.1177/23969873231191304] [Medline: 37548025]
8. Pohl J, Held JPO, Verheyden G, et al. Consensus-based core set of outcome measures for clinical motor rehabilitation after stroke—a Delphi study. Front Neurol 2020;11:875. [doi: 10.3389/fneur.2020.00875] [Medline: 33013624]
9. Kwakkel G, Lannin NA, Borschmann K, et al. Standardized measurement of sensorimotor recovery in stroke trials: consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2017 Jul;12(5):451-461. [doi: 10.1177/1747493017711813] [Medline: 28697709]
10. Demers M, Levin MF. Do activity level outcome measures commonly used in neurological practice assess upper-limb movement quality? Neurorehabil Neural Repair 2017 Jul;31(7):623-637. [doi: 10.1177/1545968317714576] [Medline: 28675943]
11. Kwakkel G, Van Wegen E, Burridge JH, et al. Standardized measurement of quality of upper limb movement after stroke: consensus-based core recommendations from the Second Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2019 Oct;14(8):783-791. [doi: 10.1177/1747493019873519] [Medline: 31510885]
12. Essers B, Meyer S, De Bruyn N, et al. Mismatch between observed and perceived upper limb function: an eye-catching phenomenon after stroke. Disabil Rehabil 2019 Jun;41(13):1545-1551. [doi: 10.1080/09638288.2018.1442504] [Medline: 29564912]

XSL•FO
RenderX

13. van Delden A, Peper CLE, Beek PJ, Kwakkel G. Match and mismatch between objective and subjective improvements in upper limb function after stroke. Disabil Rehabil 2013;35(23):1961-1967. [doi: 10.3109/09638288.2013.768303] [Medline: 23611500]
14. Doman CA, Waddell KJ, Bailey RR, Moore JL, Lang CE. Changes in upper-extremity functional capacity and daily performance during outpatient occupational therapy for people with stroke. Am J Occup Ther 2016;70(3). [doi: 10.5014/ajot.2016.020891] [Medline: 27089298]
15. Lang CE, Holleran CL, Strube MJ, et al. Improvement in the capacity for activity versus improvement in performance of activity in daily life during outpatient rehabilitation. J Neurol Phys Ther 2023 Jan 1;47(1):16-25. [doi: 10.1097/NPT.0000000000000413] [Medline: 35930404]
16. Levin MF, Kleim JA, Wolf SL. What do motor "recovery" and "compensation" mean in patients following stroke? Neurorehabil Neural Repair 2009 May;23(4):313-319. [doi: 10.1177/1545968308328727] [Medline: 19118128]
17. Kanzler CM, Lamers I, Feys P, Gassert R, Lambercy O. Personalized prediction of rehabilitation outcomes in multiple sclerosis: a proof-of-concept using clinical data, digital health metrics, and machine learning. Med Biol Eng Comput 2021 Nov 25;60(1):249-261. [doi: 10.1007/s11517-021-02467-y]
18. Schwarz A, Bhagubai MMC, Nies SHG, et al. Characterization of stroke-related upper limb motor impairments across various upper limb activities by use of kinematic core set measures. J NeuroEngineering Rehabil 2022 Dec;19(1):2. [doi: 10.1186/s12984-021-00979-0]
19. Rand D, Eng JJ. Disparity between functional recovery and daily use of the upper and lower extremities during subacute stroke rehabilitation. Neurorehabil Neural Repair 2012 Jan;26(1):76-84. [doi: 10.1177/1545968311408918]
20. Essers B, Veerbeek JM, Luft AR, Verheyden G. The feasibility of the adapted H-GRASP program for perceived and actual daily-life upper limb activity in the chronic phase post-stroke. Disabil Rehabil 2024 Feb 8;46(24):5815-5828. [doi: 10.1080/09638288.2024.2313121]
21. Schwarz A, Kanzler CM, Lambercy O, Luft AR, Veerbeek JM. Systematic review on kinematic assessments of upper limb movements after stroke. Stroke 2019 Mar;50(3):718-727. [doi: 10.1161/strokeaha.118.023531]
22. Kalsi-Ryan S, Beaton D, Curt A, Popovic MR, Verrier MC, Fehlings MG. Outcome of the upper limb in cervical spinal cord injury: profiles of recovery and insights for clinical studies. J Spinal Cord Med 2014 Sep;37(5):503-510. [doi: 10.1179/2045772314y.0000000252]
23. Rudd KD, Lawler K, Callisaya ML, Alty J. Investigating the associations between upper limb motor function and cognitive impairment: a scoping review. Geroscience 2023 Jun 20;45(6):3449-3473. [doi: 10.1007/s11357-023-00844-z]
24. Alt Murphy M, Willén C, Sunnerhagen KS. Movement kinematics during a drinking task are associated with the activity capacity level after stroke. Neurorehabil Neural Repair 2012 Nov;26(9):1106-1115. [doi: 10.1177/1545968312448234]
25. Alt Murphy M, Murphy S, Persson HC, Bergström UB, Sunnerhagen KS. Kinematic analysis using 3D motion capture of drinking task in people with and without upper-extremity impairments. J Vis Exp 2018 Mar 28(133):57228. [doi: 10.3791/57228] [Medline: 29658937]
26. Lili L, Sunnerhagen KS, Rekand T, Alt Murphy M. Quantifying an upper extremity everyday task with 3D kinematic analysis in people with spinal cord injury and non-disabled controls. Front Neurol 2021;12:755790. [doi: 10.3389/fneur.2021.755790] [Medline: 34721277]
27. Kim GJ, Parnandi A, Eva S, Schambra H. The use of wearable sensors to assess and treat the upper extremity after stroke: a scoping review. Disabil Rehabil 2021 Jul 30;44(20):6119-6138. [doi: 10.1080/09638288.2021.1957027]
28. Barth J, Lohse KR, Konrad JD, Bland MD, Lang CE. Sensor-based categorization of upper limb performance in daily life of persons with and without neurological upper limb deficits. Front Rehabil Sci 2021 Oct 20;2:741393. [doi: 10.3389/fresc.2021.741393]
29. Pohl J, Verheyden G, Held JPO, Luft AR, Easthope Awai C, Veerbeek JM. Construct validity and responsiveness of clinical upper limb measures and sensor-based arm use within the first year after stroke: a longitudinal cohort study. J Neuroeng Rehabil 2025 Jan 29;22(1):14. [doi: 10.1186/s12984-024-01512-9]
30. Braakhuis HEM, Bussmann JBJ, Ribbers GM, Berger MAM. Wearable activity monitoring in day-to-day stroke care: a promising tool but not widely used. Sensors (Basel) 2021 Jun 12;21(12):4066. [doi: 10.3390/s21124066]
31. Yang CL, Cheng YY, Lin CH, Menon C, Eng JJ. Quantifying the reach-and-grasp practice using novel wearable technology: application in a stroke rehabilitation setting. J Hand Ther 2025 Oct;38(4):841-847. [doi: 10.1016/j.jht.2025.02.017]
32. Okuyama K, Kawakami M, Tsuchimoto S, et al. Depth sensor-based assessment of reachable work space for visualizing and quantifying paretic upper extremity motor function in people with stroke. Phys Ther 2020 May 18;100(5):870-879. [doi: 10.1093/ptj/pzaa025] [Medline: 32048724]
33. Gomes CLA, Cacho RO, Nobrega VTB, et al. Effects of attentional focus on upper extremity motor performance in post stroke patients: a randomized pilot study. Medicine (Baltimore) 2021 Mar 5;100(9):e24656. [doi: 10.1097/MD.0000000000024656] [Medline: 33655929]
34. Wu CY, Trombly CA, Lin KC, Tickle-Degnen L. A kinematic study of contextual effects on reaching performance in persons with and without stroke: influences of object availability. Arch Phys Med Rehabil 2000 Jan;81(1):95-101. [doi: 10.1016/s0003-9993(00)90228-4] [Medline: 10638883]

XSL•FO
RenderX

35. Gagné-Pelletier L, Poitras I, Roig M, Mercier C. Factors associated with upper extremity use after stroke: a scoping review of accelerometry studies. J Neuroeng Rehabil 2025 Feb 24;22(1):33. [doi: 10.1186/s12984-025-01568-1] [Medline: 39994630]
36. Tausch AP, Menold N. Methodological aspects of focus groups in health research. Glob Qual Nurs Res 2016 Jan 1;3. [doi: 10.1177/2333393616630466]
37. Wilkinson S. Focus groups in health research. J Health Psychol 1998 Jul;3(3):329-348. [doi: 10.1177/135910539800300304]
38. Simpson LA, Menon C, Hodgson AJ, Ben Mortenson W, Eng JJ. Clinicians' perceptions of a potential wearable device for capturing upper limb activity post-stroke: a qualitative focus group study. J Neuroeng Rehabil 2021 Sep 8;18(1):135. [doi: 10.1186/s12984-021-00927-y] [Medline: 34496894]
39. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
40. Rampin R, Rampin V. Taguette: open-source qualitative data analysis. J Open Source Soft 2021 Dec 10;6(68):3522. [doi: 10.21105/joss.03522]
41. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388]
42. Hervé-Colas J, Newton SP, Engelter ST, et al. Standardized international manual of the Fugl-Meyer assessment of motor function after stroke. Neurorehabil Neural Repair 2026 Apr;40(4):306-319. [doi: 10.1177/15459683251412300]
43. Yozbatiran N, Der-Yeghiaian L, Cramer SC. A standardized approach to performing the Action Research Arm Test. Neurorehabil Neural Repair 2008;22(1):78-90. [doi: 10.1177/1545968307305353] [Medline: 17704352]
44. Kanzler CM, Rinderknecht MD, Schwarz A, et al. A data-driven framework for selecting and validating digital health metrics: use-case in neurological sensorimotor impairments. NPJ Digit Med 2020;3(1):80. [doi: 10.1038/s41746-020-0286-7] [Medline: 32529042]
45. Lohse KR, Miller AE, Bland MD, Lee JM, Lang CE. Validation of real-world actigraphy to capture post-stroke motor recovery. medRxiv. Preprint posted online on Nov 4, 2024. [doi: 10.1101/2024.11.03.24316674]
46. Lang CE, Waddell KJ, Barth J, Holleran CL, Strube MJ, Bland MD. Upper limb performance in daily life approaches plateau around three to six weeks post-stroke. Neurorehabil Neural Repair 2021 Oct;35(10):903-914. [doi: 10.1177/15459683211041302] [Medline: 34510934]
47. Thrane G, Sunnerhagen KS, Murphy MA. Upper limb kinematics during the first year after stroke: the stroke arm longitudinal study at the University of Gothenburg (SALGOT). J Neuroeng Rehabil 2020 Jun 15;17(1):76. [doi: 10.1186/s12984-020-00705-2] [Medline: 32539738]
48. Frykberg GE, Grip H, Alt Murphy M. How many trials are needed in kinematic analysis of reach-to-grasp? A study of the drinking task in persons with stroke and non-disabled controls. J Neuroeng Rehabil 2021 Jun 15;18(1):101. [doi: 10.1186/s12984-021-00895-3] [Medline: 34130716]
49. Jose M, Munoz-Novoa M, Alt Murphy M. A reliable and valid assessment of upper limb movement quality after stroke: the observational drinking task assessment. J Rehabil Med 2024 Oct 9;56:jrm40362. [doi: 10.2340/jrm.v56.40362] [Medline: 39382398]
50. Sauerzopf L, Panduro CGC, Luft AR, et al. Evaluating inter- and intra-rater reliability in assessing upper limb compensatory movements post-stroke: creating a ground truth through video analysis? J Neuroeng Rehabil 2024 Dec 20;21(1):217. [doi: 10.1186/s12984-024-01506-7] [Medline: 39702329]
51. Jung HT, Kim Y, Lee J, Lee SI, Choe EK. Envisioning the use of in-situ arm movement data in stroke rehabilitation: stroke survivors' and occupational therapists' perspectives. PLoS ONE 2022;17(10):e0274142. [doi: 10.1371/journal.pone.0274142] [Medline: 36264782]
52. Mayrhuber L, Roffler L, Easthope CA, Gassert R, Lambercy O. Arm activity feedback for stroke survivors – a user needs analysis for the design of a mobile application. In: 2025 International Conference on Rehabilitation Robotics (ICORR): IEEE; 2025:876-881. [doi: 10.1109/ICORR66766.2025.11062943]
53. Miller AE, Holleran CL, Bland MD, et al. Perspectives of key stakeholders on integrating wearable sensor technology into rehabilitation care: a mixed-methods analysis. medRxiv. Preprint posted online on Nov 29, 2024. [doi: 10.1101/2024.11.25.24317911]
54. Serrano LP, Maita KC, Avila FR, et al. Benefits and challenges of remote patient monitoring as perceived by health care practitioners: a systematic review. Perm J 2023 Dec 15;27(4):100-111. [doi: 10.7812/TPP/23.022] [Medline: 37735970]
55. Holleran CL, Bland MD, Lang CE. Comprehensive assessment of the activity level of the ICF using both capacity and performance measures: a case report. Arch Rehabil Res Clin Transl 2023 Sep;5(3):100277. [doi: 10.1016/j.arrct.2023.100277] [Medline: 37744190]
56. Bland MD, Holleran CL, Newman CA, et al. ICF classification of therapeutic goals for outpatient episodes of neurorehabilitation in post-stroke and Parkinson disease. Disabil Rehabil 2024 Oct;46(20):4772-4778. [doi: 10.1080/09638288.2023.2290201] [Medline: 38059563]
57. Leuenberger KD. Long-term activity and movement monitoring in neurological patients. Thesis.: ETH Zurich; 2015. [doi: 10.3929/ethz-a-010594517]

XSL•FO
RenderX

58. Pohl J, Ryser A, Veerbeek JM, et al. Accuracy of gait and posture classification using movement sensors in individuals with mobility impairment after stroke. Front Physiol 2022;13:933987. [doi: 10.3389/fphys.2022.933987] [Medline: 36225292]
59. Lum PS, Shu L, Bochniewicz EM, et al. Improving accelerometry-based measurement of functional use of the upper extremity after stroke: machine learning versus counts threshold method. Neurorehabil Neural Repair 2020 Dec;34(12):1078-1087. [doi: 10.1177/1545968320962483] [Medline: 33150830]
60. Pohl J, Ryser A, Veerbeek JM, et al. Classification of functional and non-functional arm use by inertial measurement units in individuals with upper limb impairment after stroke. Front Physiol 2022;13:952757. [doi: 10.3389/fphys.2022.952757] [Medline: 36246133]
61. Kaku A, Parnandi A, Venkatesan A, Pandit N, Schambra H, Fernandez-Granda C. Towards data-driven stroke rehabilitation via wearable sensors and deep learning. Proc Mach Learn Res 2020 Aug;126:143-171. [Medline: 34337420]
62. Parnandi A, Kaku A, Venkatesan A, et al. Data-driven quantitation of movement abnormality after stroke. Bioengineering (Basel) 2023 May 26;10(6):648. [doi: 10.3390/bioengineering10060648]
63. van Meulen FB, Reenalda J, Buurke JH, Veltink PH. Assessment of daily-life reaching performance after stroke. Ann Biomed Eng 2014 Dec 2;43(2):478-486. [doi: 10.1007/s10439-014-1198-y]
64. Macpherson CE, Bland MD, Gordon C, et al. Replication of sensor-based categorization of upper-limb performance in daily life in people post stroke and generalizability to other populations. Sensors (Basel) 2025 Jul 25;25(15):4618. [doi: 10.3390/s25154618]
65. McLaren R, Signal N, Lord S, Taylor S, Henderson J, Taylor D. The volume and timing of upper limb movement in acute stroke rehabilitation: still room for improvement. Disabil Rehabil 2019 Apr 5;42(22):3237-3242. [doi: 10.1080/09638288.2019.1590471]
66. Longatelli V, Torricelli D, Tornero J, et al. A unified scheme for the benchmarking of upper limb functions in neurological disorders. J Neuroeng Rehabil 2022 Sep 27;19(1):102. [doi: 10.1186/s12984-022-01082-8] [Medline: 36167552]
67. Manta C, Patrick-Lake B, Goldsack JC. Digital measures that matter to patients: a framework to guide the selection and development of digital measures of health. Digit Biomark 2020;4(3):69-77. [doi: 10.1159/000509725] [Medline: 33083687]
68. Unger T, de Sousa Ribeiro R, Mokni M, et al. Upper limb movement quality measures: comparing IMUs and optical motion capture in stroke patients performing a drinking task. Front Digit Health 2024;6:1359776. [doi: 10.3389/fdgth.2024.1359776] [Medline: 38606036]
69. Ploderer B, Fong J, Klaic M, et al. How therapists use visualizations of upper limb movement information from stroke patients: a qualitative study with simulated information. JMIR Rehabil Assist Technol 2016 Oct 5;3(2):e9. [doi: 10.2196/rehab.6182] [Medline: 28582257]
70. Cain A, Gunby T, Winstein C, Demers M. Advancing stroke rehabilitation: the role of wearable technology according to research experts. Disabil Rehabil Assist Technol 2025 Jul;20(5):1460-1469. [doi: 10.1080/17483107.2025.2459326] [Medline: 39874139]
71. Swain TA, McNarry MA, Mackintosh KA. Assessing and enhancing movement quality using wearables and consumer technologies: thematic analysis of expert perspectives. JMIR Form Res 2024 Sep 13;8:e56784. [doi: 10.2196/56784]
72. Weikert T, Paez-Granados D, Stoller O, Easthope CA. HealthCore: a unified digital ecosystem for collection of health data throughout the continuum of care. In: Converging Clinical and Engineering Research on Neurorehabilitation: Springer; 2024:461-465. [doi: 10.1007/978-3-031-77584-0_90]
73. Unger T, Moslehian AS, Peiffer JD, et al. Differentiable biomechanics for markerless motion capture in upper limb stroke rehabilitation: a comparison with optical motion capture 2024. arXiv. Preprint posted online on Nov 22, 2024. [doi: 10.48550/arXiv.2411.14992]
74. Weikert T, Li Y, Paez-Granados D, Easthope CA. Automated prediction of item-level ARAT scores from wearable sensors. In: 2025 International Conference on Rehabilitation Robotics (ICORR): IEEE; 2025:1239-1244. [doi: 10.1109/ICORR66766.2025.11063162]
75. Du E, D. Peiffer J, Weikert T, Awai CE, James Cotton R. Usability of iGait@Healthcore: gait analysis from a smartphone for use in clinical routine. Gait Posture 2025 Aug;121:65-66. [doi: 10.1016/j.gaitpost.2025.07.075]
76. Unger T, Yao X, Schmitt A, et al. Usability of capturing upper limb movement kinematics in clinical routine: do wearable sensors beat markerless motion capture? Gait Posture 2025 Aug;121:66-67. [doi: 10.1016/j.gaitpost.2025.07.076]
77. Domnik N, Krasnopolska K, Sylvester R, et al. User-centered development of a digital platform to provide neurorehabilitation patients with feedback on health data. JMIR Preprints. Preprint posted online on Sep 30, 2025. [doi: 10.2196/preprints.85072]
78. Azodo I, Williams R, Sheikh A, Cresswell K. Opportunities and challenges surrounding the use of data from wearable sensor devices in health care: qualitative interview study. J Med Internet Res 2020 Oct 22;22(10):e19542. [doi: 10.2196/19542] [Medline: 33090107]
79. Denny JC, Collins FS. Precision medicine in 2030—seven ways to transform healthcare. Cell 2021 Mar;184(6):1415-1419. [doi: 10.1016/j.cell.2021.01.015]
80. Helminski D, Sussman JB, Pfeiffer PN, et al. Development, implementation, and evaluation methods for dashboards in health care: scoping review. JMIR Med Inform 2024 Dec 10;12:e59828. [doi: 10.2196/59828]
81. Car J, Ong QC, Erlikh Fox T, et al. The digital health competencies in medical education framework. JAMA Netw Open 2025 Jan 2;8(1):e2453131. [doi: 10.1001/jamanetworkopen.2024.53131]

XSL•FO
RenderX

82. Lawrence K, Levine DL. The digital determinants of health: a guide for competency development in digital care delivery for health professions trainees. JMIR Med Educ 2024 Aug 29;10:e54173. [doi: 10.2196/54173]
83. Bhavaraju VL, Panchanathan S, Willis BC, Garcia-Filion P. Leveraging the electronic health record to measure resident clinical experiences and identify training gaps: development and usability study. JMIR Med Educ 2024 Nov 6;10:e53337. [doi: 10.2196/53337] [Medline: 39504418]

Abbreviations

ADL: activities of daily living
ARAT: Action Research Arm Test
FMA-UE: Fugl-Meyer Assessment for Upper Extremity
HRA: Swiss Federal Act on Research involving Human Beings
ICF: International Classification of Functioning, Disability and Health
MDC: minimal detectable change
MIC: minimal important change
ROM: range of motion
SCI: spinal cord injury
SRQR: Standards for Reporting Qualitative Research

Edited by S Munce; submitted 07.Nov.2025; peer-reviewed by E Cocco, JG Schonhammer; revised version received 12.Apr.2026; accepted 23.Apr.2026; published 24.Jun.2026.

Please cite as:
Pohl J, Mayrhuber L, Lambercy O, Easthope Awai C
From Metrics to Meaning in Neurological Rehabilitation: Clinicians' Perspectives on Digital Metrics of Upper Limb Functioning—A Focus Group Study
JMIR Rehabil Assist Technol 2026;13:e87339
URL: https://rehab.jmir.org/2026/1/e87339
doi:10.2196/87339

© Johannes Pohl, Laura Mayrhuber, Olivier Lambercy, Chris Easthope Awai. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone App–Delivered Mindfulness-Based Intervention for Concussion in Adolescents (MBI-4-mTBI): Feasibility Randomized Controlled Trial

Veronik Sicard[1], PhD; Molly Cairncross[2,3,4], PhD; Roger Zemek[1,5], MD; Noah D Silverberg[4,6,7], PhD; Gary S Goldfield[1,5,8], PhD; Nick Reed[9], PhD; Bechara J Saab[10], PhD; Andra Smith[8], PhD; Andrée-Anne Ledoux[1,11,12], PhD

[1]Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, ON, Canada
[10]Mobio Interactive, Singapore, Singapore
[11]Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada
[12]University of Ottawa Brain and Mind Research Institute, University of Ottawa, Ottawa, ON, Canada
[2]Department of Psychology, Simon Fraser University, Burnaby, BC, Canada
[3]BC Children's Hospital Research Institute, Vancouver, BC, Canada
[4]Rehabilitation Research Program, Centre for Aging SMART, Vancouver Coastal Health Research Institute, Vancouver, BC, Canada
[5]Department of Pediatrics, University of Ottawa, Ottawa, ON, Canada
[6]Department of Psychology, University of British Columbia, Vancouver, BC, Canada
[7]Djavad Mowafaghian Centre for Brain Health, Vancouver, BC, Canada
[8]School of Psychology, University of Ottawa, Ottawa, ON, Canada
[9]Department of Occupational Science & Occupational Therapy, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Andrée-Anne Ledoux, PhD
Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, ON, Canada

Related Article:

This is a corrected version. See correction statement: https://rehab.jmir.org/2026/1/e110902

Abstract

Background: Persisting symptoms affect about one-third of youth following concussion. Mental health history, distress, and coping style are key predictors of prolonged recovery. Early and scalable psychological interventions, such as mindfulness-based intervention (MBI) delivered via smartphones, may improve patients' ability to regulate their emotions and neurophysiologically recover, reducing overall symptom burden. However, no digital therapeutic (DTx) trials in adolescents experiencing concussion exist.

Objective: This study primarily aimed to assess the feasibility of conducting a larger randomized controlled trial (RCT) evaluating the effectiveness of a DTx-MBI in adolescents with a concussion compared with an attention-matched sham intervention.

Methods: This was a Health Canada-regulated, parallel-group, blinded, single-crossover feasibility RCT. Adolescents aged 12 to <18 years presenting to a Pediatric Emergency Department or interdisciplinary concussion clinic within 7 days of a physician-diagnosed concussion were approached for participation from November 2022 to June 2024. After providing consent, participants were randomized (1:1), stratified by sex, to either the experimental group (DTx-MBI) or the control group (sham, attention-matched math puzzle game). The DTx-MBI was delivered via the AmDTx platform (Mobio Interactive Pte Ltd, Singapore) as a custom-designed 4-to-8-week program of 8 standardized modules for adolescents with concussion, including audio-recorded guided mindfulness exercises, goal setting, journaling, and psychoeducation. The control intervention, delivered through the same interface, excluded mindfulness content and instead featured the open-source game "2048". Participants in both groups were encouraged to engage with the app for at least 10 minutes/day, at least 4 days/week. Feasibility criteria to support progression to a full-scale RCT included: eligibility rate >40% of those screened; recruitment rate >50% of eligible participants randomized; intervention credibility >70% scoring above the midpoint on the credibility and expectancy questionnaire (CEQ) at 1 week; retention >75% of randomized participants completing 4-week outcomes; and adherence >70% completing 10 minutes of intervention on at least 4 days/week for 4 weeks.

Results: A total of 124 out of 195 (63.6%) screened youth met eligibility criteria. Of these, 99/124 (79.8%) consented and were randomized to either the DTx-MBI group (n=49, median [IQR] age=15.28 [13.66 - 16.19] years, 30 [61.2%] female) or the Sham

XSL•FO
RenderX

group (n=50, median [IQR] age=14.92 [13.32 - 16.71] years, 30 [60.0%] female). Credibility was high, with 62/83 (74.7%) of participants scoring above the credibility midpoint (DTx-MBI: 75.0%; Sham: 74.4%). Retention was strong, with 89/99 (89.9%) of participants completing the 4-week outcomes (DTx-MBI: 89.8%; control: 90.0%). Overall adherence was moderate (54/99 [54.5%]; DTx-MBI: 59.2%; control: 50.0%), and a little higher among outcome assessment completers (53/89 [59.6%]; DTx-MBI: 63.6%; Sham: 55.6%). Feasibility indicators were similar between groups.

Conclusions: This feasibility trial supports the implementation of a larger RCT, with modifications to enhance adherence, to rigorously evaluate the clinical efficacy of the DTx-MBI. By targeting modifiable psychological risk factors through a scalable digital platform, DTx-MBI could be a low-burden, cost-effective adjunct to pediatric concussion care.

Trial Registration: ClinicalTrial.gov NCT05105802; https://clinicaltrials.gov/study/NCT05105802

International Registered Report Identifier (IRRID): RR2-10.2196/57226

(JMIR Rehabil Assist Technol 2026;13:e84623) doi:10.2196/84623

KEYWORDS

pediatric; concussion; mindfulness; clinical trial; psychological intervention; youth; digital therapeutics; eHealth; mHealth; mild traumatic brain injury

Introduction

Approximately one-in-three children sustaining a concussion develop persisting symptoms after concussion (PSAC), defined as ongoing physical, cognitive, emotional, and sleep symptoms that persist beyond 4 weeks postinjury [1], and may persist for months [2,3]. PSAC can disrupt academic performance, reduce social engagement, and increase susceptibility to substance use, depression, anxiety, and suicidal ideation [4-7]. Youth recovering from a concussion face an elevated risk of mental health challenges, including a 40% increased likelihood of developing psychiatric disorders, a 47% higher risk of hospitalization for psychiatric conditions, and a 49% increased risk of self-harm behaviors compared with age- and sex-matched youth with an orthopedic injury [4]. Both injury-related and non-injury-related factors contribute to PSAC [3,8,9]; however, the influence of injury-specific factors tends to diminish over time. In contrast, premorbid factors, such as somatic symptoms [10], migraines [3], cognitive difficulties [11], attention and mood disorders [12], anxiety [13], and maladaptive coping strategies [13], remain strong predictors of prolonged recovery [10,14]. Early interventions that bolster psychological resilience and emotional regulation may be critical in mitigating symptom burden and preventing chronicity [11,14].

Mindfulness-based interventions (MBIs) offer a promising approach by promoting present-focused practices that cultivate nonjudgmental awareness of thoughts and emotions, thereby improving affect regulation and reducing psychological distress [15]. MBIs have been shown to enhance attention, cognitive flexibility, emotional well-being, and academic performance in children, while reducing anxiety, depressive symptoms, affective reactivity, and fear [16]. In addition, MBIs have been shown to reduce default mode network (DMN) hyperactivity across clinical and pediatric populations [17-22], and since concussion similarly involves DMN disruptions [23-27], MBIs may confer neurophysiological benefits postconcussion. MBIs may target several known preinjury risk factors for PSAC, including mood disorders [28-30], somatization [28,31,32], headache disorders [33], and chronic pain [34]. MBIs may improve fatigue, depression, self-efficacy, and quality of life in individuals of different ages with persistent symptoms following mild-to-moderate traumatic brain injuries, including concussions [35-38]. However, these in-person programs are often costly, time-intensive, and not always accessible [35-38]. Digital therapeutics (DTx) have emerged as scalable tools for delivering psychological interventions. Evidence supports the effectiveness of brief web-based MBIs in reducing anxiety, perceived stress, and negative mood across different populations [39-43]. Moreover, a recent meta-analysis found that app-based MBIs show small to moderate effectiveness in improving stress, mood, and quality of life among users without a recent concussion [44]. However, the use of short, app-based DTx-MBIs in pediatric concussion remains unexplored.

This study aimed to assess the feasibility of conducting a larger randomized controlled trial (RCT) evaluating the effectiveness of DTx-MBI to reduce symptom burden and reduce risk of PSAC in youth with a concussion compared with an attention-matched sham intervention by examining recruitment efficiency, intervention credibility, intervention adherence, and participant retention rates at week 4 postenrollment. The secondary research objectives included the following. (1) Determine the participants' treatment expectations and satisfaction with both interventions. (2) Explore potential efficacy signals across clinical outcomes (ie, concussion symptoms, quality of life, fatigue, anxiety, depression, resiliency, self-efficacy, mindfulness, and cognitive function) to inform future sample size calculations. (3) Examine safety by comparing adverse events between interventions. (4) Assess the feasibility of incorporating a neuroimaging component into the larger trial by examining recruitment and retention for the neuroimaging subset, as well as intervention credibility and intervention adherence among participants who completed neuroimaging (Multimedia Appendix 1 [26,45]). (5) Explore differences in baseline characteristics, adherence, and attrition rates between the experimental group and control group. Experimental group: participants who decided to complete the study at 4 weeks and participants who continued for an additional 4 weeks of MBI for weeks 5 - 8; Control group: participants who decided to complete the study at 4 weeks and who crossed over to the experimental intervention for weeks 5 - 8 (Multimedia Appendix 1 [26,45]).

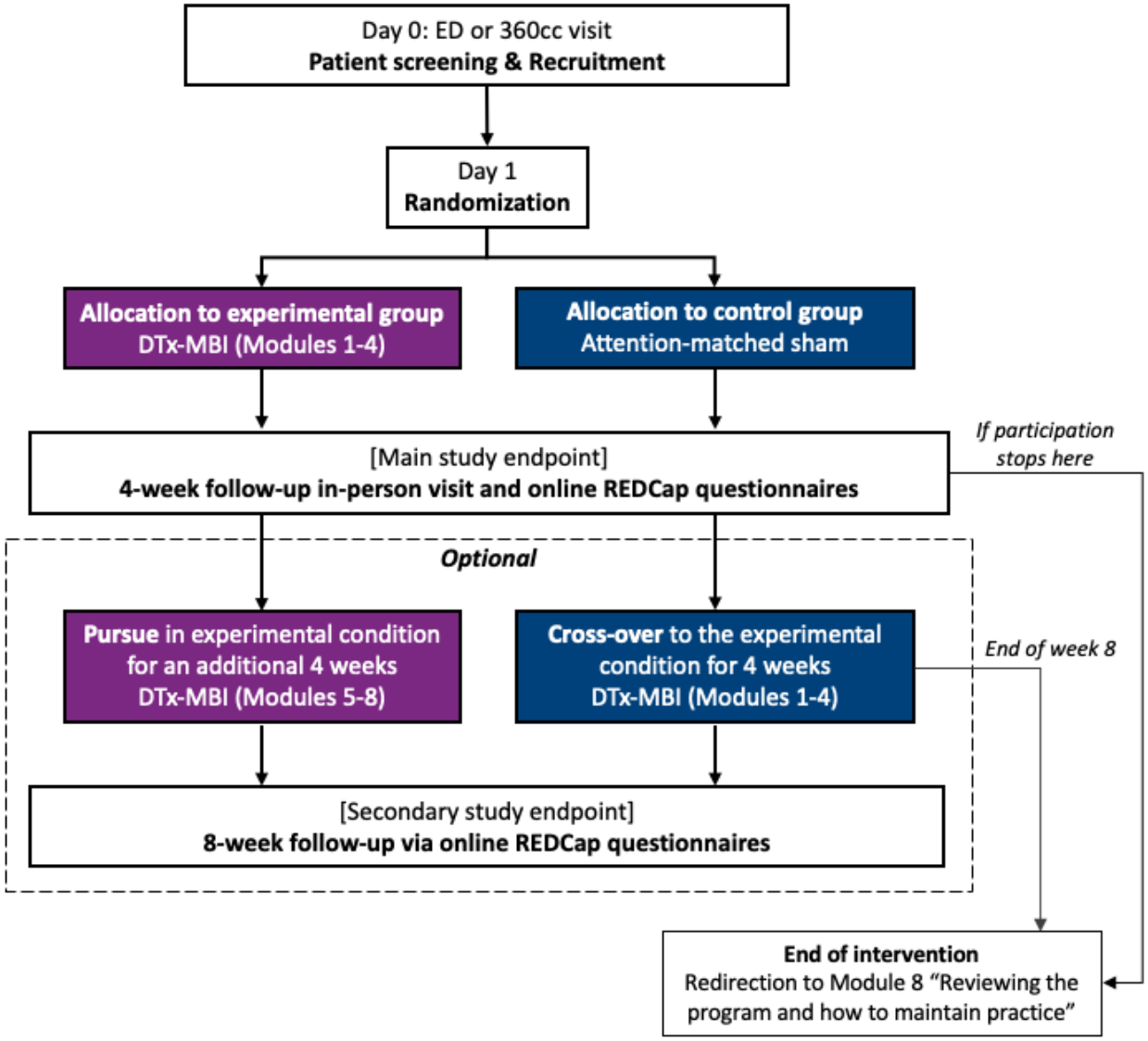
XSL•FO
RenderX

Methods

Study Design

This Health Canada-regulated feasibility RCT (ClinicalTrials.gov: NCT05105802) used a parallel-group (1:1), blinded, single-crossover design. Participants in the control group who completed the initial 4-week study and wished to pursue the experimental intervention crossed over (modules 1 - 4) during weeks 5 - 8, while those in the experimental group could extend their MBI training for an additional 4 weeks (modules 5 - 8; Figure 1). To maintain focus on the primary endpoints, only the 4-week outcomes are presented in the main text; week-8 outcomes are reported in Multimedia Appendix 1 [26,45]. Methods and results of the feasibility of the magnetic resonance imaging (MRI) component are presented in Multimedia Appendix 1 [26,45].

Figure 1. Graphical depiction of study design. 360cc: 360 Concussion Care; DTx-MBI: digital therapeutics mindfulness-based intervention; ED: emergency department; REDCap: Research Electronic Data Capture.



Ethical Considerations

This study was approved by the Research Ethics Board (protocol #20/720X) of the Children's Hospital of Eastern Ontario (CHEO) Research Institute in August 2020. Authorization from Health Canada was granted in July 2022. The study was carried out according to the principles outlined in the Declaration of Helsinki and Good Clinical Practices, and within the laws and regulations of the Tri-Council Policy Statement.

Before consent was sought, the study was introduced, and its purpose was explained verbally. The consent form also presented the purpose of the research, the procedures, potential risks and benefits, and the use and security of the data. Informed written consent was obtained from the parents, and participants were deemed capable of consenting. Informed assent was obtained from the participants deemed by the research assistant (RA) to be cognitively unable to provide informed consent.

The participants were remunerated with a CAD $20 (US $15) gift certificate (eg, Tango) for completing surveys. Participants

XSL•FO
RenderX

received an additional CAD $20 (US $15) gift certificate for completing the MRIs. The gift certificate was sent electronically to their email address. They also received a parking voucher or vouchers for in-person meetings and a letter attesting that they have completed 20 hours of volunteer work (5 hours per week completing the intervention).

Participant information was coded using study identification numbers. Research personnel took all appropriate and customary steps to ensure that data remain secure and that patient privacy and confidentiality are maintained.

Procedures

The full protocol was published previously [46]. Briefly, participants were recruited from the CHEO emergency department (ED) and the 360 Concussion Care (360cc) clinic, an interdisciplinary tertiary health clinic in Ottawa, from October 2022 to June 2024. Families of adolescents with a concussion were approached by an RA to discuss the research study. Interested families completed a screening form on a tablet with the RA, and if eligible (concussion confirmed by a physician) and enrolled, filled out electronic questionnaires via Research Electronic Data Capture (REDCap) [4748] to collect patient demographics, diagnostic history, injury characteristics, and a concussion symptoms checklist.

The day after the ED or 360cc visit, an RA called participants to randomize them into either the experimental or control group. They provided instructions for installing AmDTx on the participant's mobile device of their choice. Participants completed Day 1 surveys by phone or via REDCap and began the intervention that same day. At the 4-week mark, participants attended an in-person appointment to complete the cognitive testing at either the CHEO Research Institute or the Royal Ottawa Hospital (for those enrolled in the neuroimaging component; see below). After this visit, participants in the experimental group were offered 4 additional weeks of psychoeducation modules and mindfulness practices. Control group participants were given the option to cross over and begin the DTx-MBI. Follow-up questionnaires were completed on REDCap at 1, 2, 4, and 8 weeks. A subset from each group was prospectively approached and enrolled for a neuroimaging (Multimedia Appendix 1 [26,45]).

Randomization, allocation concealment, and blinding

Participants were randomly assigned to one of 2 groups using a 1:1 allocation ratio stratified by sex. (1) Experimental group: DTx-MBI program, or (2) control group: sham digital attention-matched puzzle game. The CHEO Clinical Research Unit provided data management services for this study and retained randomization codes. A statistician who was not involved in the study created and maintained the master list.

Only the RAs responsible for randomization and app setup were unblinded to participants' group assignments, as this was necessary for intervention setup and delivery. These RAs were blinded to the randomization sequence until the point of group assignment and were not involved in outcome assessment or data analysis.

During the consent process and follow-up, we gave controlled information to participants about both interventions, while keeping them unaware of the hypothesis, design, and treatment assignment (experimental or control). The study hypothesis and the distinction between experimental and control conditions were disclosed only at the 4-week follow-up, at which point DTx-MBI participants could choose to extend their intervention and the Sham participants could cross over to the experimental condition.

Participants

Adolescents presenting to the ED or 360cc within 7 days of sustaining a direct or indirect traumatic brain injury were eligible to participate if they: (1) were aged 12 to <18 years; (2) had a diagnosed concussion (by a physician); (3) had 1 highest level of certainty symptom (eg, loss of consciousness, posttraumatic amnesia) or 2 higher level of certainty symptoms (eg, nausea/vomiting, headache) immediately or within 1 hour of injury from the adapted version of the Centers for Disease Control and Prevention tiered framework [49]; (4) had a score of >4 on the Predicting and Preventing Postconcussive Problems in Pediatrics clinical risk score, a validated prediction score to identify those at risk of poorer outcomes at 4 weeks postinjury [3]; (5) were proficient in English.

Patients were excluded if they presented with traumatic brain injuries with any of the following: (1) Glasgow Coma Scale score of ≤13; (2) trauma-related abnormality on standard neuroimaging studies (if performed) [50]; (3) neurosurgical operative intervention, intubation, or intensive care required; (4) multisystem injuries with the treatment requiring hospital admission, operating room, or procedural sedation in ED; (5) severe neurological developmental delay resulting in communication difficulties; (6) intoxication to alcohol or drugs at the time of ED/360cc presentation as per clinician judgment; (7) no clear history of head trauma as the primary event to the concussion (eg, seizure and migraine attack); (8) prior psychiatric hospitalization; (9) inability to obtain written informed consent or assent; (10) legal guardian not present (certain forms needed to be completed by parents or legal guardians); (11) no internet and mobile or tablet access; and (12) was in psychological therapy at enrollment.

Intervention

Experimental Group: DTx-MBI

The DTx-MBI program was delivered via the AmDTx platform and consisted of a 4-to-8-week, custom-made program (containing 8 psychoeducation modules) for adolescents with concussions, including goal setting, check-ins on mood and stress, audio-recorded psychoeducation, guided meditations, and journaling. Each standardized module was unlocked as the participant progressed through the program. Participants were encouraged to participate in the DTx-MBI activities for at least 10 minutes every day, for a minimum of 4 days/week over the study period. Those who chose to end their participation at the end of week 4 were immediately directed to module 8, "Reviewing the program and how to maintain practice," while those who decided to continue moved on to modules 5 through

XSL•FO
RenderX

8. The DTx had previously undergone pilot testing demonstrating acceptability, credibility, and usability [51].

Control Group: Attention-Matched Sham Comparator

The sham digital math puzzle was delivered on the same main interface within AmDTx, including the Snapshot (see below); however, these participants did not have access to any of the MBI training or psychoeducational content. The sham DTx was the open-source math game called "2048", developed by Gabriele Cirulli [52]. The game required participants to slide numbered tiles on a grid and combine matching numbers, aiming to reach 2048 by creating larger numbers through merges. Each move adds a new tile, and the game ends when the grid is full with no possible merges. The game is programmed to only be accessible for 10 minutes per session. The selected control task has precedent as a digital comparator and has shown no effects on stress, psychological well-being, or symptoms in prior studies [42,53].

Snapshot

AmDTx includes 4 simple tools for tracking patient well-being over time, collectively called the Snapshot, in which participants in both groups were prompted to complete before and after each session:(1) 30-second "selfie" video analyzed via photoplethysmography [54] and a deep neural network to quantify psychological stress from heart rate [55] (2) a digital emotion mapping board (circumplex) to capture emotional stress levels; (3) a subjective stress slider, benchmarked against standard psychological surveys [42]; and (4) an open-text personal notes field. The snapshot feature qualified as meaningful engagement; however, completing the snapshot alone did not fulfill the requirement.

Standard of Care

All participants received standard-of-care [56] instruction pamphlets provided by RAs in the ED or at 360cc. The pamphlet includes advice to refrain from physical and cognitive exertion for the first 24 to 48 hours after injury. After the initial rest period, low-to-moderate levels of physical and mental activity can be started, including screen time, provided that the activities do not cause a recurrence or exacerbation of symptoms. It further advises that adolescents avoid activities with an increased risk of reinjury until fully asymptomatic and cleared by their primary care or other medical providers. The study intervention was therefore intended as a supplement to usual care rather than a replacement for standard concussion management.

Measures and Outcomes

Primary Outcomes

The prespecified feasibility targets for progressing to a definitive trial were informed by established benchmarks from previous feasibility trials [57,58] and our experience recruiting pediatric concussion patients from the ED [59]. We used a traffic light system to assess the feasibility of proceeding to a full-scale trial [60]. Progression is classified as feasible without modifications (green), feasible with adjustments (amber), or infeasible due to unresolvable issues (red). Advancement to a full-scale trial is determined by the criterion with the least favorable performance. The 5 prespecified targets are outlined in Table 1. Credibility was assessed using the CEQ (questions 1 - 3) at week 1 postenrollment [61], and retention was defined as completing the health behavior inventory at the 4-week follow-up [62].

Table . Feasibility benchmarks and performance.

Feasibility outcomes	Prespecified benchmark		
	Green zone (go)	Amber zone (amend)	Red zone (stop)
Eligibility	>40% of patients screened are eligible	20% - 40% of patients screened are eligible	<20% of patients screened are eligible
Recruitment	>50% of eligible patients are randomized	30% - 50% of eligible patients are randomized	<30% of eligible patients are randomized
Credibility	>70% of participants score above the scale mid-point[a]	50% - 70% of participants score above the scale mid-point[a]	<50% of participants score above the scale mid-point[a]
Retention	>75% of participants complete the follow-up assessment at 4-week[b]	50% - 75% of participants complete the follow-up assessment at 4-week[b]	<50% of participants complete the follow-up assessment at 4-week[b]
Adherence	>70% of participants complete the minimal requirement (ie, 10 minutes of activity, 4 times/week)	50% - 70% of participants complete the minimal requirement, that is, approximately (ie, 10 minutes of activity, 4 times /week)	<50% of participants who completed the outcomes complete the minimal requirement (ie, 10 minutes of activity, 4 times/week)

[a]Credibility and expectancy questionnaire questions 1 to 3.

[b]Health behavior inventory.

Secondary Outcomes

Treatment expectancy was assessed with the CEQ (questions 4 to 6) at week 1 after enrollment [61]. Expectancy was considered good if >70% participants scored above the midpoint.

Intervention satisfaction was assessed with the modified version (to align with the digital nature of the intervention) of the Client Satisfaction Questionnaire at 4 weeks [63]. The intervention was deemed satisfactory if >70% of participants scored above the midpoint on the questionnaire.

XSL•FO
RenderX

Safety of the MBI was determined by capturing and monitoring the worsening of symptoms and adverse events associated with the app or intervention. Adverse events were defined as any unscheduled visits to the ED or primary medical providers because of exacerbation of symptoms during the study participation. Events were considered intervention-related if they involved symptom provocation during or within 30 minutes of DTx use. The determination of whether an adverse event was related to the intervention was made by the Qualified Investigator, based on predefined criteria established by the Data and Safety Monitoring Board and reviewed by the institutional quality assurance and regulatory compliance personnel. Participants were asked about possible adverse events at the 1-, 2-, 4-, and 8-week follow-ups.

Other outcome measures were collected at baseline (in the ED or on day 1 for 360cc) and at 4- and 8-week follow-ups, unless otherwise specified:

- Symptom burden: Health behavior inventory, assessed retrospectively for preinjury and acute postinjury symptoms, with additional assessment at 1 and 2 weeks follow-ups [62];
- Quality of life: Pediatric Quality of Life Inventory (PedsQL, version 4.0) [64,65];
- Fatigue: PedsQL Multidimensional Fatigue Scale (PedsQL-MFS, version 4.0) [66,67];
- Depression symptoms: Center for Epidemiologic Studies Depression Scale for Children [68];
- Anxiety: Severity Measure for Generalized Anxiety Disorder–Child Age 11 - 17 (DSM-5 GAD-C) [69];
- Resilience: Connor-Davidson Resilience Scale-10 [70];
- Mindfulness: Child and Adolescent Mindfulness Measure [71,72];
- Self-efficacy: Self-Efficacy Questionnaire for Children [73];
- Cognition: National Institutes of Health Toolbox (NIH Toolbox) [74,75] assessed fluid cognition on week 4 only: Fluid Cognition composite score, and Picture Sequence Memory, List Sorting Working Memory, Flanker Inhibitory Control and Attention, Dimensional Change Card Sorting, and Pattern Comparison Processing Speed tests [76].

Adherence

App adherence was quantified based on meaningful in-app activity time, defined as active engagement with core features assigned to each group. For the DTx-MBI group, this included guided meditation exercises, psychoeducational content, and snapshots. For the Sham group, the core features were the "2048" game and snapshots, which served as the control activities for engagement tracking. Any single app session lasting <3 minutes were excluded from adherence calculations, as this duration was deemed insufficient to complete meaningful activities. Weekly adherence was calculated by summing the total seconds of the qualifying app session within every 7 days over the 4-week study period.

Statistical Analysis

Overview

Descriptive statistics were used to summarize participant characteristics, with medians reported for continuous variables and frequencies for categorical variables. Feasibility outcomes were reported as proportions. For outcome measures at 4 and 8 weeks, medians and IQRs were calculated for each group. In a post hoc exploratory analysis, a multiple regression was conducted to examine factors associated with app engagement (in seconds, square root transformed to address nonnormality and improve model assumptions) over the 4-week intervention period. The regression model included group allocation, age, sex, and preinjury ratings of anxiety, depression, mindfulness, and self-efficacy. All statistical analyses were conducted on R (version 4.4.1; R Core Team, 2024) [77].

Sample Size Calculation

Detailed sample size justifications can be found in the published protocol [46]. Briefly, sample size estimates were based on feasibility trial guidelines using the traffic light system [60,78]. To achieve >90% power for each of the 5 feasibility criteria, the highest requirement was adherence (n=55). Accounting for a 20% attrition, 70 participants were needed. Assuming 40% eligibility and 50% recruitment rates, ≈350 children needed to be screened.

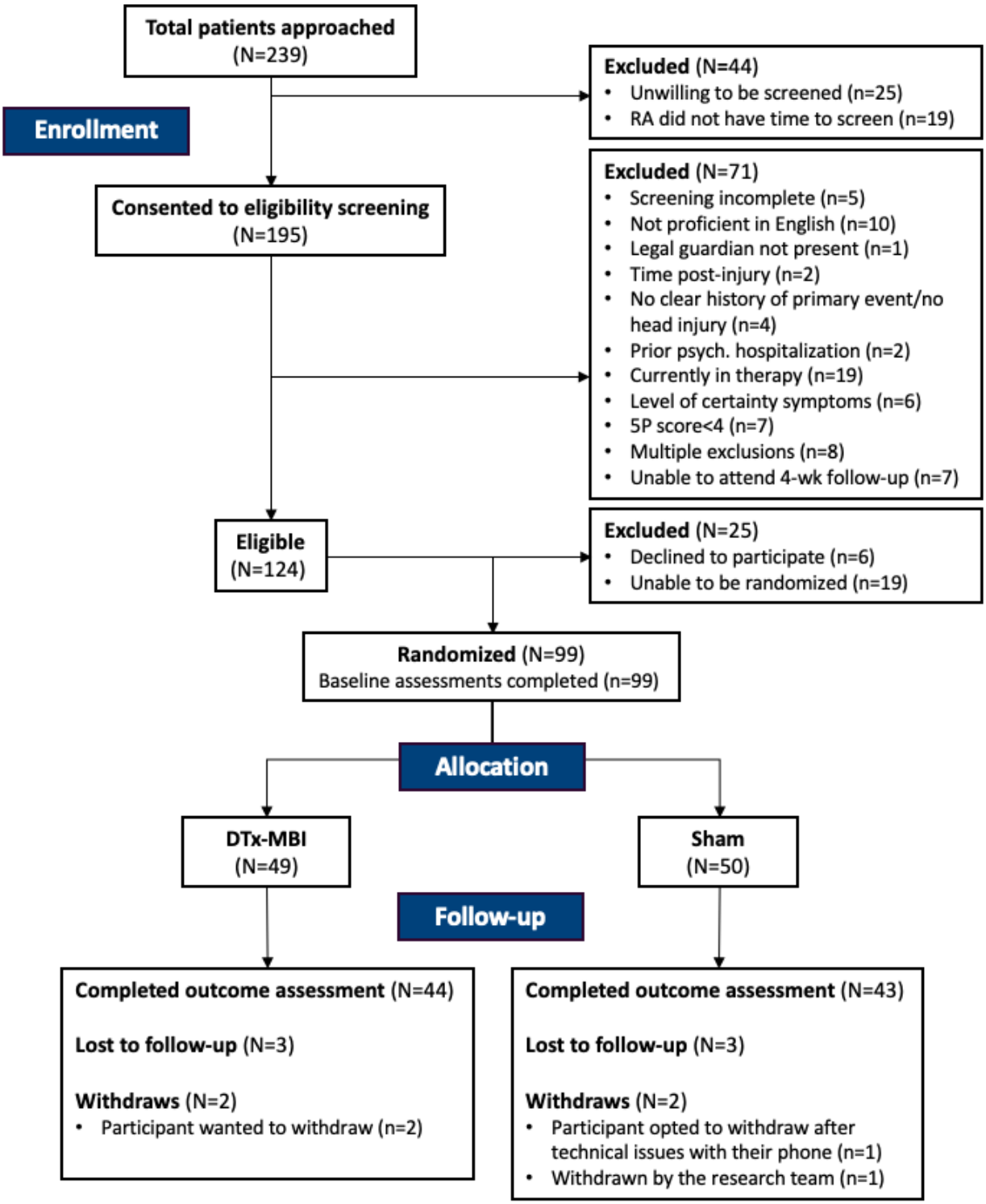
Results

Participants

A total of 99 out of 124 eligible patients were randomized to DTx-MBI (n=49; median age [IQR]=15.28 [13.66 - 16.19] years; 30 [61%] female) or Sham groups (n=50; median age [IQR]=14.92 [13.32 - 16.71] years; 30 [60%] female) (Figure 2). A total of 65 participants were recruited from the ED and 34 from 360cc. Demographics and clinical characteristics were comparable between groups among all randomized participants and those who completed the 4-week follow-up (Table 2).

XSL•FO
RenderX

Figure 2. Participant flow diagram. 5P: Predicting and Preventing Postconcussive Problems in Pediatrics study clinical risk score; DTx-MBI: digital therapeutics mindfulness-based intervention; RA: research assistant.



XSL•FO
RenderX

Table . Participants' demographic and clinical characteristics.

Variables	Randomized		Completed the 4-week intervention	
	DTx-MBI[a] (n=49)	Sham (n=50)	DTx-MBI (n=44)	Sham (n=45)
Demographics				
Age (years), median (IQR)	15.28 (13.66 - 16.19)	14.92 (13.32 - 16.71)	14.92 (13.61 - 16.17)	14.84 (13.30 - 16.63)
Female sex, n (%)	30 (61.2)	30 (60.0)	27 (61.4)	26 (57.8)
Diagnostic history				
Previous concussion, n (%)	21 (42.9)	23 (46.0)	20 (45.5)	21 (46.7)
Number of previous concussions, median (IQR)	1 (1-2)	2 (1-2)	1 (1-2.25)	2 (1-2)
Migraines, n (%)	4 (8.2)	2 (4.0)	4 (9.1)	2 (4.4)
Learning disabilities, n (%)	4 (8.3)[b]	11 (22.0)	4 (9.3)[b]	9 (20.0)
Attention-deficit/hyperactivity disorder (ADHD), n (%)	7 (14.6)[b]	6 (12.0)	7 (16.3)[b]	4 (8.9)
Other developmental disorders, n (%)	1[c] (2.1)[b]	1[d] (2.0)[b]	1[c] (2.3)[b]	0 (0)
Anxiety, n (%)	5 (10.4)[b]	9 (18.0)	5 (11.6)[b]	8 (17.8)
Depression, n (%)	5 (10.4)[b]	1 (2.0)	4 (9.3)[b]	1 (2.2)
Sleep disorders, n (%)	1 (2.1)[b]	2 (4.0)	1 (2.3)[b]	0 (0)
Other mental health diagnoses, n (%)	1[e] (2.1)[b]	0 (0)	1[e] (2.3)[b]	0 (0)
Clinical variables				
Recruitment site, n (%)				
Emergency department	30 (61.2)	35 (70.0)	27 (61.4)	30 (66.7)
360 Concussion Care	19 (38.8)	15 (30.0)	17 (38.6)	15 (33.3)
Hours post-injury at enrollment, median (IQR)	28.83 (3.23 - 99.27)	28.69 (7.82 - 96.11)	30.50 (4.39 - 96.82)	35.05 (8.63 - 108.17)
5P risk score[f], median (IQR)	7.00 (6.00 - 8.00)	7.00 (6.00 - 8.00)	7.00 (6.00 - 8.00)	7.00 (6.00 - 8.00)
Low risk, n (%)	0 (0)	1 (2.0)	0 (0)	1 (2.2)
Medium risk, n (%)	38 (77.6)	37 (74.0)	34 (77.3)	33 (73.3)
High risk, n (%)	11 (22.4)	12 (24.0)	10 (22.7)	11 (24.4)
Loss of consciousness, n (%)	9 (18.8)[b]	5 (10.2)[b]	8 (18.6)[b]	5 (11.1)
Amnesia, n (%)	16 (33.3)[b]	16 (32.0)	15 (34.1)[b]	14 (40.0)
Previous experience with mindfulness training therapy, n (%)	3 (6.3)[b]	4 (8.0)	2 (4.7)[b]	2 (4.4)
Prior experience with mindfulness, n (%)[g]	7 (14.6)[b]	7 (14.0)	5 (11.6)[b]	6 (13.3)
Mechanism of injury, n (%)				
Fall down stairs	1 (2.0)	1 (2.0)	1 (2.3)	1 (2.2)
Sports	37 (75.5)	34 (68.0)	32 (72.7)	30 (66.7)

XSL•FO
RenderX

Variables	Randomized		Completed the 4-week intervention	
	DTx-MBI[a] (n=49)	Sham (n=50)	DTx-MBI (n=44)	Sham (n=45)
Fall from standing, walking, or running	3 (6.1)	5 (10.0)	3 (6.8)	5 (11.1)
Ran into stationary object	0 (0)	2 (4.0)	0 (0)	2 (4.4)
Other mechanism	7 (14.3)	7 (14.0)	7 (15.9)	6 (13.3)
Missing	1 (2.0)	0 (0.0)	1 (2.3)	1 (2.2)
Retrospective HBI[h], median (IQR)	10.50 (4.00 - 17.00)[b]	11.00 (5.25 - 17.75)	11.00 (4.00 - 17.00)[b]	11.00 (4.00 - 16.00)

[a]DTx-MBI: digital therapeutics mindfulness-based intervention.

[b]Missing values (mv)=1.

[c]Word retrieval/language delay.

[d]Autism spectrum.

[e]Unspecified.

[f]Low risk=0 - 3, Medium risk=4 - 8, Higher risk=9 - 12.

[g]Originally coded as “None,” “A little”, “Some”, “A lot” but recoded as “Yes” or “No.”

[h]HBI: health behavior inventory.

Primary Analyses: Feasibility Outcomes at 4-Week Postenrollment

Four out of the 5 prespecified feasibility outcomes met the “green zone” benchmarks.

Eligibility

A total of 124 out of 195 (63.6%) participants were screened, exceeding the >40% eligibility criteria.

Recruitment

A total of 99 out of 124 (80%) eligible participants consented and were randomized, surpassing the 50% benchmark.

Credibility

A total of 62 out of 83 (75%) of participants scored above the scale midpoint, with similar results between the DTx-MBI (33/44, 75%) and Sham groups (29/39, 74%), exceeding the 70% target.

Retention

A total of 89 out of 99 (90%) of participants completed the 4-week follow-up assessments, with similar completion rates between the DTx-MBI (44/49, 90%) and Sham groups (45/50, 90%), meeting the 75% benchmark.

Adherence

A total of 53 out of 89 (60%) of participants completed the minimal requirement of activity, below the 70% target, falling into the amber zone. The DTx-MBI group showed slightly higher adherence (28/44, 64%) compared with the Sham group (25/45, 56%).

Secondary Analyses

Treatment Expectancy and Satisfaction

A total of 61 out of 83 (74%) participants scored above the scale mid-point on treatment expectancy (DTx-MBI: 34/44, 77%; Sham: 27/39, 69%). Moreover, intervention satisfaction was good, with 71/86 (83%) participants scoring above the scale mid-point (DTx-MBI: 37/43, 86%; Sham: 34/43, 79%).

Efficacy Signal

Given the nonnormal distribution of the outcome measures, medians and IQRs were calculated for all outcomes at baseline and week 4 (Table 3) and week 8 (Tables S1 and S2 in Multimedia Appendix 1 [26,45]). These analyses were intended to inform future sample size calculations; the study was not powered to detect statistically significant between-group differences.

XSL•FO
RenderX

Table . Efficacy signal for outcomes at week 4 postenrollment.

Outcomes	Baseline (preintervention)		4-week endpoint	
	DTx-MBI[a] (n=44)[b]	Sham (n=45)[b]	DTx-MBI (n=44)	Sham (n=45)
Retrospective HBI[c], median (IQR)	11.00 (4.00 - 17.00)[d]	11.00 (4.00 - 16.00)	—[e]	—
HBI (symptom burden), median (IQR)	26.00 (17.00 - 34.50)	25.00 (13.00 - 32.00)	16.50 (10.00 - 23.00)	13.00 (6.50 - 22.50)
PSAC,[f] n (%)	—	—	5 (11.6)[d]	5 (11.6)
PedsQL[g] (quality of life), median (IQR)	80.00 (71.72 - 90.00)[d]	80.63 (70.43 - 89.88)[d]	81.17 (67.93 - 89.92)	84.38 (70.16 - 92.11)
Physical	90.63 (76.56 - 96.88)[d]	87.50 (80.47 - 93.75)	90.63 (71.09 - 93.75)	90.63 (71.88 - 96.88)
Emotional	75.00 (65.00 - 95.00)[d]	75.00 (58.75 - 90.00)	77.50 (60.00 - 95.00)	75.00 (65.00 - 92.50)
Social	95.00 (75.00 - 100.00)[d]	90.00 (73.75 - 100.00)[d]	95.00 (73.75 - 100.00)	95.00 (85.00 - 100.00)
School	70.00 (60.00 - 85.00)[d]	70.00 (55.00 - 86.25)[d]	67.50 (53.75 - 80.00)	75.00 (52.50 - 90.00)
PedsQL-MFS[h] (fatigue), median (IQR)	70.83 (61.11 - 80.56)[d]	69.44 (62.15 - 80.90)[d]	66.67 (54.51 - 80.90)	73.61 (52.08 - 86.11)
Cognitive	70.84 (52.08 - 77.08)[d]	77.08 (58.33 - 87.50)[d]	68.75 (50.00 - 71.88)	70.83 (52.08 - 86.11)
Sleep/Rest	62.50 (52.08 - 77.08)[d]	62.50 (50.00 - 71.88)[d]	58.33 (50.00 - 71.88)	66.67 (52.08 - 75.00)
General	79.17 (66.67 - 87.50)[d]	79.17 (66.67 - 91.67)[d]	75.00 (57.29 - 91.67)	75.00 (62.50 - 91.67)
CES-DC[i] (depression), median (IQR)	4.00 (3.00 - 9.00)	5.00 (3.00 - 7.00)	6.00 (3.75 - 10.00)	6.00 (4.00 - 8.00)
DSM-5 GAD-C[j] (anxiety), median (IQR)	4.00 (0.50 - 9.00)[d]	5.00 (1.00 - 10.00)[d]	3.00 (1.00 - 10.25)	5.00 (0.00 - 9.00)
CD-RISC-10[k] (resilience), median (IQR)	28.00 (23.50 - 30.50)[d]	28.00 (24.75 - 31.00)[d]	25.50 (10.75 - 30.50)	26.00 (20.00 - 32.00)
CAMM[l] (mindfulness), median (IQR)	32.00 (26.50 - 37.00)[d]	31.00 (25.00 - 36.00)[d]	33.00 (24.00 - 38.00)	32.00 (26.50 - 36.50)
SEQ-C[m] (self-efficacy), median (IQR)	85.00 (73.00 - 94.50)[d]	82.50 (75.00 - 90.25)[d]	80.50 (67.75 - 89.25)	81.00 (70.50 - 95.00)
Emotional	28.00 (23.00 - 31.50)[d]	27.00 (24.00 - 30.00)[d]	26.50 (21.75 - 32.00)	26.00 (22.00 - 32.50)
Social	29.00 (26.00 - 33.00)[d]	29.00 (26.00 - 31.25)[d]	27.00 (24.00 - 32.00)	29.00 (25.00 - 32.00)
Academic	26.00 (22.00 - 32.00)[d]	27.00 (24.00 - 30.25)[d]	26.00 (22.00 - 33.00)	27 (21.50 - 32.00)
NIH[n] toolbox (cognition), median (IQR)				
Fluid cognition composite score (age-corrected)	—	—	103.00 (85.00 - 113.00)[o]	104.00 (91.50 - 117.00)[p]
Flanker (inhibition control, attention)	—	—	85.00 (75.50 - 97.00)[q]	89.00 (84.00 - 100.50)[p]
List sorting (working memory)	—	—	98.00 (92.50 - 103.50)[r]	100.00 (84.00 - 100.50)[p]
Dimensional change card sort (executive functions)	—	—	101.00 (85.50 - 115.00)[r]	105.00 (86.00 - 120.00)[p]
Pattern comparison (processing speed)	—	—	112.00 (94.50 - 128.00)[r]	103.00 (93.50 - 124.00)[p]
Picture sequence (memory)	—	—	105.00 (93.00 - 110.00)[r]	105 (96.00 - 115.00)[p]

XSL•FO
RenderX

Outcomes	Baseline (preintervention)		4-week endpoint	
	DTx-MBI[a] (n=44)[b]	Sham (n=45)[b]	DTx-MBI (n=44)	Sham (n=45)
Agreed to continue to 8-week, n (%)	—	—	10 (22.7)	11 (25.6)

[a]DTx-MBI: digital therapeutics mindfulness-based intervention.

[b]Include participants who completed the 4-week protocol.

[c]HBI: health behavior inventory.

[d]Missing value (mv)=1.

[e]Not applicable.

[f]Persisting Symptoms After Concussion (PSAC) was determined through reliable change z scores of≥1.65 [79]. The reliable change score compares the parent's retrospective rating of preinjury total symptoms with the child's ratings reported at the 4-week follow-up, based on a formula derived from regression analyses in children with orthopedic injuries.

[g]PedsQL: Pediatric Quality of Life Inventory, version 4.0.

[h]MFS: Multidimensional Fatigue Scale.

[i]CES-DC: Center for Epidemiologic Studies Depression Scale for Children.

[j]DSM-5 GAD-C: Severity Measure for Generalized Anxiety Disorder–Child Age 11 - 17.

[k]CD-RISC-10: Connor-Davidson Resilience Scale-10.

[l]CAMM: Child and Adolescent Mindfulness Measure.

[m]SEQ-C: Self-Efficacy Questionnaire for Children.

[n]NIH: National Institutes of Health.

[o]Missing value (mv)=11.

[p]Missing value (mv)=4.

[q]Missing value (mv)=10.

[r]Missing value (mv)=9.

Safety of the App

Two participants from the DTx-MBI group reported unscheduled medical visits during their study participation, but no visits met our predefined safety criteria for adverse events associated with the use of the application: one participant visited the ED for a second concussion in week 3, while the other visited for worsening of concussion symptoms prior to AmDTx use.

Post Hoc Exploration: Factors Associated With Activity Engagement

Given low adherence, we conducted a post hoc analysis to examine factors associated with intervention adherence. In the post hoc exploratory multiple regression, none of the demographic or preinjury factors showed statistically significant associations with activity engagement during the 4-week intervention period (Table 4).

Table . Demographic and preinjury factors associated with activity engagement (measured in seconds) over the 4-week intervention period.

Variables	Estimate	SE	Statistic	*P* value	95% CI
Intercept	59.62	121.06	0.42	.62	−181.49 to 300.73
Group	−0.45	13.84	−0.03	.97	−28.02 to 27.12
Age	1.00	4.47	0.22	.82	−7.91 to 9.91
Sex	21.08	15.90	1.33	.19	−10.60 to 52.75
HBI[a] (symptoms)	−0.81	0.69	−1.18	.24	−2.18 to −0.56
DSM-5 GAD-C[b] (anxiety)	−9.72	15.12	−0.64	.52	−39.83 to 20.39
CES-DC[c] (depression)	1.29	2.31	0.56	.58	−3.32 to 5.89
CAMM[d] (mindfulness)	−1.55	1.36	−1.14	.26	−4.26 to 1.16
SEQ-C[e] (self-efficacy)	1.02	0.61	1.68	.10	−0.19 to 2.24

[a]HBI: health behavior inventory.

[b]DSM-5 GAD-C: Severity Measure for Generalized Anxiety Disorder–Child Age 11 - 17.

[c]CES-DC: Center for Epidemiologic Studies Depression Scale for Children.

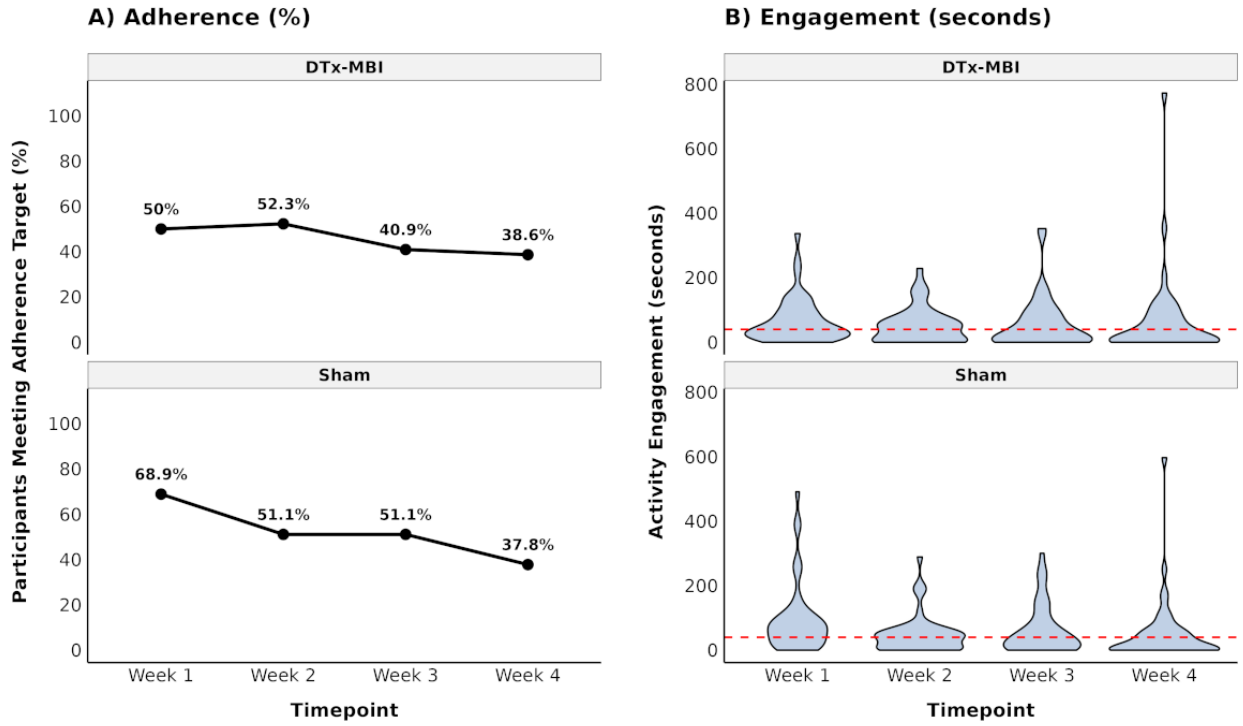
[d]CAMM: Child and Adolescent Mindfulness Measure.

[e]SEQ-C: Self-Efficacy Questionnaire for Children.

XSL•FO
RenderX

Among participants who completed the intervention, weekly adherence rates (ie, meeting the threshold of ≥10 minutes of "meaningful" app engagement for 4 days over a 7-day week) declined from week 1 (DTx-MBI: 22/44, 50%; Sham: 31/45, 69%) to week 4 (DTx-MBI: 17/44, 39%; Sham: 17/45, 38%). Similarly, median weekly engagement time decreased from 40.9 minutes (IQR 20.2 - 89.9) to 17.1 minutes (IQR 0 - 76.8) in the DTx-MBI group and from 75.2 minutes (IQR 26.2 - 108.1) to 12.6 minutes (IQR 0 - 60.0) in the Sham group over the 4 weeks (Figure 3).

Figure 3. Weekly adherence. (A) Percentage of participants meeting the weekly 40-minute app activity engagement target across the 4 weeks of protocol within the experimental (DTx-MBI) and control (attention-matched cognitive Sham comparison; Sham) groups. Data points show exact percentages for each group at each week. (B) Distribution of weekly app activity engagement (in minutes) shown as violin plots for each group. Each violin plot shows the density of engagement duration, with a larger width indicating higher data concentration. Black dots represent median values, and vertical lines span the IQR (**Q1-Q3**). The red dashed horizontal line at 40 minutes indicates the target engagement threshold. Only participants who completed the 4-week outcomes were included in the figure. DTx-MBI: digital therapeutics mindfulness-based intervention.



Discussion

Principal Findings

This feasibility RCT indicated that a larger-scale trial evaluating the effectiveness of DTx-MBI in youth with a concussion compared with an attention-matched sham intervention is viable, with 4 out of the 5 prespecified feasibility benchmarks—eligibility, recruitment, credibility, and retention—being met or exceeded. Only adherence fell slightly below the target threshold, with engagement levels comparable to other digital health interventions in youth [28], suggesting that minor adjustments could improve adherence in the large-scale efficacy RCT.

The high eligibility rate suggests that the inclusion criteria were well aligned with the characteristics of the target population, and the high recruitment rates underscore the strong relevance and suitability of this intervention for the target population. Moreover, retention at 4 weeks approached 90% in both groups, reflecting strong participant commitment and the feasibility of our follow-up strategy. The high credibility and satisfaction across both intervention and control groups support the acceptability of the app-based interventions, regardless of their content. This reinforces findings from our earlier open-label preliminary study [51,80]. In the mixed methods evaluation (n=7), credibility, expectancy, and satisfaction ratings were even higher, although this may have been influenced by the small sample size, open-label design, and participants' awareness that they would be asked for feedback during the 4-week interview [51]. This blinded RCT strengthens the evidence for the acceptability of the DTx-MBI through a more rigorous randomized design and the inclusion of an attention-matched comparator. Although treatment expectancy was above the scale midpoint for all participants, those in the MBI-DTx reported a qualitatively higher expectancy that the treatment would be helpful, suggesting greater perceived relevance or believability of the experimental intervention. The treatment expectancy ratings in the Sham group were just below the threshold considered "good" in this study, whereas the experimental group met the benchmark. In the efficacy RCT, we will address this potential imbalance by providing participants in the Sham group with additional education about how cognitive training may promote recovery through neuroplasticity during enrollment.

Notably, in this study, no adverse events related to app use were reported, supporting the intervention's safety. Furthermore, the

XSL•FO
RenderX

optional MRI component also met the feasibility benchmarks, with high enrollment and scan completion rates, suggesting that a neuroimaging component can be successfully integrated into future trials.

Adherence, while below the 70% benchmark, remained within an acceptable range (nearly 60%), with slightly higher adherence in the DTx-MBI group across the 4 weeks. Weekly adherence declined over the 4 weeks in both groups, with median weekly engagement time decreasing substantially from week 1 to week 4 in both groups, with a more pronounced decline in the Sham group. This reduction may reflect the natural waning of novelty or motivation over time. This pattern is consistent with Eysenbach's "Law of Attrition," which posits that drop-off in usage is an expected and inherent feature of eHealth interventions, as users often disengage over time regardless of intervention quality or intent [81]. The drop-off in app engagement around weeks 2 and 3 may also coincide with the typical course of concussion recovery, a point at which participants may no longer perceive the intervention as necessary [9]. While the DTx-MBI group began with lower adherence rates than the Sham group, they sustained slightly higher engagement by week 4, potentially indicating greater long-term relevance or perceived efficacy or benefit of the mindfulness-based content. Previous studies in adolescents and adults have reported wide variability in adherence to digital MBIs, ranging from 40% to 92% [28]. In this study, we did not identify any participant characteristics associated with lower app engagement. However, prior work suggests that psychological factors such as resilience and coping style may moderate response to behavioral concussion interventions, suggesting that engagement may depend on characteristics not captured here [82]. Future efficacy trials may benefit from enhanced engagement strategies, such as personalized reminders and motivational interviewing. However, such strategies may reduce scalability and accessibility, as they require additional resources and personalization. Implementing them only when early signs of disengagement emerge (as observed in this study around week 2) may help balance these trade-offs.

It is also important to note that the engagement tracking metrics may under-report actual engagement, as the developer's system was designed to capture only completed "meaningful interactions." Sessions that were started but not completed (eg, if a participant began a meditation recording but exited before finishing) were not logged. Likewise, activities and assessments (snapshots, journaling, etc) that lasted at least 3 minutes but were not explicitly submitted by the participant once they were done, were not logged. The same goes for the sham app. Based on the developer's user testing, these incomplete sessions and unsubmitted assessments may account for roughly 10% - 20% of potential meaningful interactions. As such, the engagement data presented here represent only fully completed meaningful interactions, and actual usage may have been modestly higher than reported.

Although not powered for efficacy, both groups seemed to show reductions in symptom burden and improvements in quality of life at the 4-week assessment. The Sham group appeared to report higher sleep and school quality of life at 4 weeks, while these scores declined slightly in the DTx-MBI group. At 8 weeks, among participants who continued with the DTx-MBI, sleep and school quality of life increased, whereas they showed minimal change in the Sham participants who crossed over to the MBI intervention. Moreover, performance on the NIH Toolbox Pattern Comparison task, a measure of processing speed, was modestly higher in the DTx-MBI group compared with the Sham group, although this difference may reflect pre-existing group differences rather than a potential intervention effect since no baseline measures were available for the NIH Toolbox. These descriptive comparisons were not formally tested and are presented only to provide preliminary signals to inform future fully powered trials. Accordingly, these preliminary findings should be interpreted cautiously and warrant further investigation in a fully powered trial to determine the intervention's efficacy and clarify its impact across recovery outcomes.

Limitations

Despite overall strong feasibility metrics, this study has some limitations. First, estimates of recruitment, eligibility, adherence, and retention were based on a modest sample size, which may limit precision. Second, the control group's treatment expectancy was qualitatively lower and just below the benchmark for being considered "good," which may suggest that some participants suspected they were not receiving the genuine cognitive training intervention, indicating partial unblinding or differential perceived value, despite the use of an active sham. This may have been further compounded by the nature of the control condition, which consisted of math puzzles without concussion-specific content. However, it should be noted that all participants received standard-of-care concussion guidance regardless of group allocation, which is itself expected to support natural recovery and limits the interpretation that any observed differences reflect expectancy alone. These factors will need to be carefully addressed in the design and interpretation of the efficacy trial. Last, as expected for a feasibility study, it was not powered to detect efficacy, and outcome findings should be considered exploratory and interpreted accordingly. This is particularly relevant for weeks 5 - 8, which included only a small subset of participants who elected to continue.

Conclusions

There are a few established preventive nonpharmacological interventions for PSAC in pediatrics [83]; however, psychological approaches are increasingly explored. This feasibility trial supports the implementation of a larger RCT to rigorously evaluate the clinical efficacy of the DTx-MBI in pediatric concussion when compared with an attention-matched cognitive sham delivered on the same digital platform. With refinements to improve adherence, especially in weeks 2 to 4, we will be positioned to evaluate whether this intervention reduces symptom burden and enhances multifaceted recovery outcomes, including quality of life, mood, cognition, and neurophysiology, in youth with a concussion.

XSL•FO
RenderX

Acknowledgments

We thank the Children's Hospital of Eastern Ontario emergency room research assistant volunteers and research staff at 360 Concussion Care who initially approached participants to gauge interest in the study and completed the eligibility screening. We thank Olivier Brown, Lauren Butterfield, Zoe Ivan, and Matthew Richmond for their contribution to data collection.

This study was completed in collaboration with TRANSCENDENT Concussion Integrated Discovery Research Program.

Funding

This study was funded by a CHAMO Innovation grant (PI: AAL and RZ), a University of Ottawa Brain and Mind Research Institute team grant (Principal Investigator [PI]: AAL), a Martin Osmond Research Grant Award at Children's Hospital of Eastern Ontario Research Institute (PI: AAL), and Ontario Brain Institute NERD Program (PI: AAL). VS is supported by a Canadian Institutes of Health Research postdoctoral fellowship.

Data Availability

The datasets generated and analyzed during the current feasibility RCT are available from the corresponding author on reasonable request. Requests will be considered in accordance with institutional guidelines and applicable ethical approvals.

Authors' Contributions

VS contributed to conceptualization, methodology, data curation, formal analysis, visualization, and writing of the original draft. MC contributed to the conceptualization, methodology, and review and editing of the manuscript. RZ contributed to the conceptualization, provision of resources, and review and editing of the manuscript. NDS, GSG, NR, and AS contributed to the conceptualization and review and editing of the manuscript. BJS contributed to software development and review and editing of the manuscript. AAL contributed to the conceptualization, methodology, project administration, funding acquisition, supervision, visualization, and writing of the original draft.

Conflicts of Interest

VS is supported by a Canadian Institutes of Health Research postdoctoral fellowship and receives royalties for the publication of a book on concussion by Flammarion. MC is one of the authors of the Mindfulness-Based Intervention for Concussion protocol used in this study, but she did not receive any financial benefit. She has received competitively funded research grants from the Canadian Institutes of Health Research (CIHR) and Michael Smith Health Research BC. RZ received financial support through competitively funded research grants from the Canadian Institutes of Health Research, Ontario Neurotrauma Foundation, Physician Services Incorporated Foundation, Children's Hospital of Eastern Ontario Foundation, University of Ottawa Brain and Mind Research Institute, Ontario Brain Institute, National Football League, Ontario Ministry of Health, Public Health Agency of Canada, Health Canada, Parachute Canada, and Ontario SPOR Support Unit. RZ was supported by a Tier 1 Clinical Research Chair in Pediatric Concussion from the University of Ottawa. All grant funding goes directly to the institution. RZ sits on the board of directors for the North American Brain Injury Society, which is a volunteer (unpaid) role. RZ is a founding partner and a minority shareholder of 360 Concussion Care (a learning health system and network of interdisciplinary concussion clinics in Ontario); no proceeds have been transferred to RZ. NR holds a Canada Research Chair (Tier 2) in Pediatric Concussion and has received research funding from the Canadian Institutes of Health Research (CIHR), Social Sciences and Humanities Research Council (SSHRC), Ontario Neurotrauma Foundation (ONF), Public Health Agency of Canada (PHAC), Parachute Canada, Special Olympics Canada, Greater Toronto Hockey League, Dr. Tom Pashby Sport Safety Fund, Holland Bloorview Kids Rehabilitation Hospital and Scotiabank. NR is an Executive Board Member for the International Pediatric Brain Injury Society (IPIBS), which is an unpaid and volunteer role. All other authors declare no other conflicts of interest. BJS is the Chief Scientist and CEO of Mobio Interactive PTE LTD (Singapore), the parent company of Mobio Interactive Inc. (Toronto, Canada), and he owned approximately 22% of the company at the time of manuscript submission. BJS exclusively served as a technical liaison for the study and did not contribute to the specific study design, select the outcome measures, nor did he directly contribute to, or have influence over, data collection or analysis. AAL is one of the authors of the Mindfulness-Based Intervention for Concussion protocol used in this study, but she did not receive any financial benefit. Andrée-Anne Ledoux has received competitively funded research grants from the Canadian Institutes of Health Research (CIHR), Ontario Neurotrauma Foundation (ONF), Physician Services Incorporated (PSI) Foundation, CHEO Foundation, Ontario Brain Institute (OBI), and University of Ottawa Brain and Mind Research. AAL received funding from the Ontario Brain Institute Neurotech Early Research & Development Program for a separate component of the study. AAL has no other conflict of interest to disclose.

Multimedia Appendix 1
Additional results.
[DOCX File, 2126 KB - rehab_v13i1e84623_app1.docx]

XSL•FO
RenderX

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 328 KB - rehab_v13i1e84623_app2.pdf]

References

1. Patricios JS, Schneider KJ, Dvorak J, et al. Consensus statement on concussion in sport: the 6th international conference on concussion in sport-Amsterdam, October 2022. Br J Sports Med 2023 Jun;57(11):695-711. [doi: 10.1136/bjsports-2023-106898] [Medline: 37316210]
2. Chadwick L, Sharma MJ, Madigan S, Callahan BL, Owen Yeates K. Classification criteria and rates of persistent postconcussive symptoms in children: a systematic review and meta-analysis. J Pediatr 2022 Jul;246:131-137. [doi: 10.1016/j.jpeds.2022.03.039] [Medline: 35358589]
3. Zemek R, Barrowman N, Freedman SB, et al. Clinical risk score for persistent postconcussion symptoms among children with acute concussion in the ED. In: DiGirolamo V, J M, Lagace M, R C, E F, MacIntyre J, et al, editors. JAMA 2016 Mar 8;315(10):1014-1025. [doi: 10.1001/jama.2016.1203] [Medline: 26954410]
4. Ledoux AA, Webster RJ, Clarke AE, et al. Risk of mental health problems in children and youths following concussion. JAMA Netw Open 2022 Mar 1;5(3):e221235. [doi: 10.1001/jamanetworkopen.2022.1235] [Medline: 35254429]
5. Gornall A, Takagi M, Morawakage T, Liu X, Anderson V. Mental health after paediatric concussion: a systematic review and meta-analysis. Br J Sports Med 2021 Sep;55(18):1048-1058. [doi: 10.1136/bjsports-2020-103548] [Medline: 33926965]
6. Gornall A, Takagi M, Clarke C, et al. Behavioral and emotional difficulties after pediatric concussion. J Neurotrauma 2020 Jan 1;37(1):163-169. [doi: 10.1089/neu.2018.6235] [Medline: 31072265]
7. Mantey DS, Omega-Njemnobi O, Barroso CS, Kelder SH. Self-reported history of concussions is associated with risk factors for suicide completion among high school students. J Affect Disord 2020 Feb;263:684-691. [doi: 10.1016/j.jad.2019.11.047]
8. Iverson GL, Gardner AJ, Terry DP, et al. Predictors of clinical recovery from concussion: a systematic review. Br J Sports Med 2017 Jun;51(12):941-948. [doi: 10.1136/bjsports-2017-097729]
9. Ledoux AA, Tang K, Yeates KO, et al. Natural progression of symptom change and recovery from concussion in a pediatric population. JAMA Pediatr 2019 Jan 1;173(1):e183820. [doi: 10.1001/jamapediatrics.2018.3820] [Medline: 31577037]
10. Nelson LD, Tarima S, LaRoche AA, et al. Preinjury somatization symptoms contribute to clinical recovery after sport-related concussion. Neurology (ECronicon) 2016 May 17;86(20):1856-1863. [doi: 10.1212/WNL.0000000000002679]
11. McNally KA, Bangert B, Dietrich A, et al. Injury versus noninjury factors as predictors of postconcussive symptoms following mild traumatic brain injury in children. Neuropsychology 2013 Jan;27(1):1-12. [doi: 10.1037/a0031370] [Medline: 23356592]
12. Plourde V, Yeates KO, Brooks BL. Predictors of long-term psychosocial functioning and health-related quality of life in children and adolescents with prior concussions. J Int Neuropsychol Soc 2018 Jul;24(6):540-548. [doi: 10.1017/S1355617718000061] [Medline: 29560837]
13. Brooks BL, Plourde V, Beauchamp MH, et al. Predicting psychological distress after pediatric concussion. J Neurotrauma 2019 Mar;36(5):679-685. [doi: 10.1089/neu.2018.5792]
14. Barlow KM. Postconcussion syndrome: a review. J Child Neurol 2016 Jan;31(1):57-67. [doi: 10.1177/0883073814543305] [Medline: 25330797]
15. Kabat-Zinn J. Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life: Hyperion; 1994.
16. Dunning D, Tudor K, Radley L, et al. Do mindfulness-based programmes improve the cognitive skills, behaviour and mental health of children and adolescents? an updated meta-analysis of randomised controlled trials. Evid Based Ment Health 2022 Jul 12;25(3):135-142. [doi: 10.1136/ebmental-2022-300464] [Medline: 35820989]
17. Bauer CCC, Rozenkrantz L, Caballero C, et al. Mindfulness training preserves sustained attention and resting state anticorrelation between default-mode network and dorsolateral prefrontal cortex: a randomized controlled trial. Hum Brain Mapp 2020 Dec 15;41(18):5356-5369. [doi: 10.1002/hbm.25197] [Medline: 32969562]
18. Smith AM, Leeming A, Fang Z, et al. Mindfulness-based stress reduction alters brain activity for breast cancer survivors with chronic neuropathic pain: preliminary evidence from resting-state fMRI. J Cancer Surviv 2021 Aug;15(4):518-525. [doi: 10.1007/s11764-020-00945-0]
19. Zhao XR, Chen ZF, Kang CY, et al. Mindfulness‐based cognitive therapy is associated with distinct resting‐state neural patterns in patients with generalized anxiety disorder. Asia Pac Psychiatry 2019 Dec;11(4). [doi: 10.1111/appy.12368]
20. Fahmy R, Wasfi M, Mamdouh R, et al. Mindfulness-based therapy modulates default-mode network connectivity in patients with opioid dependence. Eur Neuropsychopharmacol 2019 May;29(5):662-671. [doi: 10.1016/j.euroneuro.2019.03.002]
21. Carmona I Farrés C, Elices M, Soler J, et al. Effects of mindfulness training on the default mode network in borderline personality disorder. Clin Psychol Psychother 2019 Sep;26(5):562-571. [doi: 10.1002/cpp.2382] [Medline: 31132302]
22. Ives-Deliperi VL, Howells F, Stein DJ, Meintjes EM, Horn N. The effects of mindfulness-based cognitive therapy in patients with bipolar disorder: a controlled functional MRI investigation. J Affect Disord 2013 Sep 25;150(3):1152-1157. [doi: 10.1016/j.jad.2013.05.074] [Medline: 23790741]

XSL•FO
RenderX

23. Puig J, Ellis MJ, Kornelsen J, et al. Magnetic resonance imaging biomarkers of brain connectivity in predicting outcome after mild traumatic brain injury: a systematic review. J Neurotrauma 2020 Aug 15;37(16):1761-1776. [doi: 10.1089/neu.2019.6623] [Medline: 32228145]
24. Sours C, Zhuo J, Roys S, Shanmuganathan K, Gullapalli RP. Disruptions in resting state functional connectivity and cerebral blood flow in mild traumatic brain injury patients. PLoS One 2015;10(8):e0134019. [doi: 10.1371/journal.pone.0134019] [Medline: 26241476]
25. Stephens JA, Salorio CF, Barber AD, Risen SR, Mostofsky SH, Suskauer SJ. Preliminary findings of altered functional connectivity of the default mode network linked to functional outcomes one year after pediatric traumatic brain injury. Dev Neurorehabil 2018 Oct 3;21(7):423-430. [doi: 10.1080/17518423.2017.1338777]
26. Healey K, Fang Z, Smith A, Zemek R, Ledoux AA. Adolescents with a concussion have altered brain network functional connectivity one month following injury when compared to adolescents with orthopedic injuries. Neuroimage Clin 2022;36:103211. [doi: 10.1016/j.nicl.2022.103211] [Medline: 36182818]
27. Brown O, Fang Z, Smith A, Healey K, Zemek R, Ledoux AA. Associations between changes in psychological resilience and resting-state functional connectivity throughout pediatric concussion recovery. Brain Connect 2024 Sep;14(7):357-368. [doi: 10.1089/brain.2023.0096] [Medline: 38874977]
28. Spijkerman MPJ, Pots WTM, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114. [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
29. Boettcher J, Aström V, Påhlsson D, Schenström O, Andersson G, Carlbring P. Internet-based mindfulness treatment for anxiety disorders: a randomized controlled trial. Behav Ther 2014 Mar;45(2):241-253. [doi: 10.1016/j.beth.2013.11.003] [Medline: 24491199]
30. Ly KH, Trüschel A, Jarl L, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014 Jan 9;4(1):e003440. [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
31. Fjorback LO, Arendt M, Ørnbøl E, et al. Mindfulness therapy for somatization disorder and functional somatic syndromes—randomized trial with one-year follow-up. J Psychosom Res 2013 Jan;74(1):31-40. [doi: 10.1016/j.jpsychores.2012.09.006]
32. Lakhan SE, Schofield KL. Mindfulness-based therapies in the treatment of somatization disorders: a systematic review and meta-analysis. PLoS One 2013;8(8):e71834. [doi: 10.1371/journal.pone.0071834] [Medline: 23990997]
33. Gu Q, Hou JC, Fang XM. Mindfulness meditation for primary headache pain. Chin Med J 2018;131(7):829-838. [doi: 10.4103/0366-6999.228242]
34. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. J Consult Clin Psychol 2010 Apr;78(2):169-183. [doi: 10.1037/a0018555]
35. Acabchuk RL, Brisson JM, Park CL, Babbott‐Bryan N, Parmelee OA, Johnson BT. Therapeutic effects of meditation, yoga, and mindfulness‐based interventions for chronic symptoms of mild traumatic brain injury: a systematic review and meta‐analysis. Appl Psych Health Well 2021 Feb;13(1):34-62 [FREE Full text] [doi: 10.1111/aphw.12244]
36. Azulay J, Smart CM, Mott T, Cicerone KD. A pilot study examining the effect of mindfulness-based stress reduction on symptoms of chronic mild traumatic brain injury/postconcussive syndrome. J Head Trauma Rehabil 2013;28(4):323-331. [doi: 10.1097/HTR.0b013e318250ebda] [Medline: 22688212]
37. Devore E, Morgenlander A, Krucoff C, Morgenlander J. Yoga and mindfulness intervention for concussion recovery. Int J Yoga Therap 2023 Jan 1;33(2023):Article 1. [doi: 10.17761/2023-D-21-00074] [Medline: 37011045]
38. Paniccia M, Knafo R, Thomas S, et al. Mindfulness-based yoga for youth with persistent concussion: a pilot study. Am J Occup Ther 2019;73(1):7301205040p1-7301205040p11. [doi: 10.5014/ajot.2019.027672] [Medline: 30839259]
39. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: a pilot randomised controlled trial. PLoS One 2018;13(12):e0209482. [doi: 10.1371/journal.pone.0209482] [Medline: 30596696]
40. Zeidan F, Johnson SK, Gordon NS, Goolkasian P. Effects of brief and sham mindfulness meditation on mood and cardiovascular variables. J Altern Complement Med 2010 Aug;16(8):867-873. [doi: 10.1089/acm.2009.0321] [Medline: 20666590]
41. Zeidan F, Johnson SK, Diamond BJ, David Z, Goolkasian P. Mindfulness meditation improves cognition: evidence of brief mental training. Conscious Cogn 2010 Jun;19(2):597-605. [doi: 10.1016/j.concog.2010.03.014] [Medline: 20363650]
42. Walsh KM, Saab BJ, Farb NA. Effects of a mindfulness meditation app on subjective well-being: active randomized controlled trial and experience sampling study. JMIR Ment Health 2019 Jan 8;6(1):e10844. [doi: 10.2196/10844] [Medline: 30622094]
43. Manotas M, Segura C, Eraso M, Oggins J, McGovern K. Association of brief mindfulness training with reductions in perceived stress and distress in Colombian health care professionals. Int J Stress Manag 2014;21(2):207-225. [doi: 10.1037/a0035150]
44. Gál É, tefan S, Cristea IA. The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. J Affect Disord 2021 Jan 15;279:131-142. [doi: 10.1016/j.jad.2020.09.134] [Medline: 33049431]

XSL•FO
RenderX

45. Desmond JE, Glover GH. Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. J Neurosci Methods 2002 Aug 30;118(2):115-128. [doi: 10.1016/s0165-0270(02)00121-8] [Medline: 12204303]
46. Ledoux AA, Zemek R, Cairncross M, et al. Smartphone app-delivered mindfulness-based intervention for mild traumatic brain injury in adolescents: protocol for a feasibility randomized controlled trial. JMIR Res Protoc 2024 Apr 11;13:e57226. [doi: 10.2196/57226] [Medline: 38602770]
47. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208]
48. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010]
49. Peterson A, Gabella BA, Johnson J, et al. Multisite medical record review of emergency department visits for unspecified injury of head following the ICD-10-CM coding transition. Inj Prev 2021 Mar;27(Suppl 1):i13-i18. [doi: 10.1136/injuryprev-2019-043517]
50. McCrory P, Meeuwisse W, Aubry M, et al. Consensus statement on concussion in sport—the 4th international conference on concussion in sport held in Zurich, November 2012. Br J Sports Med 2013;23(2):89-117. [doi: 10.1097/JSM.0b013e31828b67cf]
51. Sicard V, O'Kane K, Brown O, et al. Acceptability, usability, and credibility of a mindfulness-based digital therapeutic for pediatric concussion: a mixed-method study (preprint). JMIR Mental Health. Preprint posted online on 2022. [doi: 10.2196/preprints.44757]
52. 2048. Play 2048. URL: https://play2048.co [accessed 2024-11-18]
53. Teel E, Brossard-Racine M, Corbin-Berrigan LA, Gagnon I. Perceptual cognitive training does not improve clinical outcomes at 4 and 12 weeks following concussion in children and adolescents: a randomized controlled trial. J Head Trauma Rehabil 2021;36(2):E97-E107. [doi: 10.1097/HTR.0000000000000633]
54. Chwyl B, Chung AG, Amelard R, Deglint J, Clausi DA, Wong A. SAPPHIRE: stochastically acquired photoplethysmogram for heart rate inference in realistic environments. 2016 Presented at: 2016 IEEE International Conference on Image Processing (ICIP); Sep 25-28, 2016; Phoenix, AZ, USA p. 1234. [doi: 10.1109/ICIP.2016.7532554]
55. Al-Jebrni AH, Chwyl B, Wang XY, Wong A, Saab BJ. AI-enabled remote and objective quantification of stress at scale. Biomed Signal Process Control 2020 May;59:101929. [doi: 10.1016/j.bspc.2020.101929]
56. Reed N, Zemek R, Dawson J, et al. Living guideline for pediatric concussion care. OSF. 2021. URL: https://doi.org/10.17605/OSF.IO/3VWN9 [accessed 2026-04-22]
57. Silverberg ND, Cairncross M, Brasher PMA, et al. Feasibility of concussion rehabilitation approaches tailored to psychological coping styles: a randomized controlled trial. Arch Phys Med Rehabil 2022 Aug;103(8):1565-1573. [doi: 10.1016/j.apmr.2021.12.005]
58. Mace RA, Greenberg J, Lemaster N, et al. Live video mind-body program for patients with knee osteoarthritis, comorbid depression, and obesity: development and feasibility pilot study. JMIR Form Res 2022 Apr 27;6(4):e34654. [doi: 10.2196/34654] [Medline: 35475787]
59. Ledoux AA, Barrowman N, Bijelić V, et al. Is early activity resumption after paediatric concussion safe and does it reduce symptom burden at 2 weeks post injury? the pediatric concussion assessment of rest and exertion (PedCARE) multicentre randomised clinical trial. Br J Sports Med 2022 Mar;56(5):271-278. [doi: 10.1136/bjsports-2021-105030] [Medline: 34836880]
60. Lewis M, Bromley K, Sutton CJ, McCray G, Myers HL, Lancaster GA. Determining sample size for progression criteria for pragmatic pilot RCTs: the hypothesis test strikes back!. Pilot Feasibility Stud 2021 Dec;7(1). [doi: 10.1186/s40814-021-00770-x]
61. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000 Jun;31(2):73-86. [doi: 10.1016/s0005-7916(00)00012-4] [Medline: 11132119]
62. Ayr LK, Yeates KO, Taylor HG, Browne M. Dimensions of postconcussive symptoms in children with mild traumatic brain injuries. J Int Neuropsychol Soc 2009 Jan;15(1):19-30. [doi: 10.1017/S1355617708090188] [Medline: 19128525]
63. Boß L, Lehr D, Reis D, et al. Reliability and validity of assessing user satisfaction with web-based health interventions. J Med Internet Res 2016 Aug 31;18(8):e234. [doi: 10.2196/jmir.5952]
64. Varni JW, Burwinkle TM, Seid M, Skarr D. * 4.0 as a pediatric population health measure: feasibility, reliability, and validity. Ambul Pediatr 2003 Nov;3(6):329-341. [doi: 10.1367/1539-4409(2003)003<0329:TPAAPP>2.0.CO;2]
65. Desai AD, Zhou C, Stanford S, Haaland W, Varni JW, Mangione-Smith RM. Validity and responsiveness of the pediatric quality of life inventory (PedsQL) 4.0 generic core scales in the pediatric inpatient setting. JAMA Pediatr 2014 Dec;168(12):1114-1121. [doi: 10.1001/jamapediatrics.2014.1600] [Medline: 25347549]
66. Varni JW, Burwinkle TM, Szer IS. The PedsQL multidimensional fatigue scale in pediatric rheumatology: reliability and validity. J Rheumatol 2004 Dec;31(12):2494-2500. [Medline: 15570657]
67. Varni JW, Beaujean AA, Limbers CA. Factorial invariance of pediatric patient self-reported fatigue across age and gender: a multigroup confirmatory factor analysis approach utilizing the PedsQLTM Multidimensional Fatigue Scale. Qual Life Res 2013 Nov;22(9):2581-2594. [doi: 10.1007/s11136-013-0370-4]

XSL•FO
RenderX

68. Radloff LS. The CES-D Scale. Appl Psychol Meas 1977 Jun;1(3):385-401. [doi: 10.1177/014662167700100306]
69. Craske M, Wittchen U, Bogels S, Stein M, Andrews G, Lebeu R. Severity measure for generalized anxiety disorder—child age 11–17. American Psychiatric Association. 2013. URL: https://www.psychiatry.org/getmedia/73e293d0-6355-4258-9cd3-2fc8a5c5b830/APA-DSM5TR-SeverityMeasureForGeneralizedAnxietyDisorderChildAge11To17.pdf [accessed 2026-08-31]
70. Connor KM, Davidson JRT. Development of a new resilience scale: the connor-davidson resilience scale (CD-RISC). Depress Anxiety 2003;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
71. Greco LA, Baer RA, Smith GT. Assessing mindfulness in children and adolescents: development and validation of the child and adolescent mindfulness measure (CAMM). Psychol Assess 2011 Sep;23(3):606-614. [doi: 10.1037/a0022819] [Medline: 21480722]
72. Kuby AK, McLean N, Allen K. Validation of the child and adolescent mindfulness measure (CAMM) with non-clinical adolescents. Mindfulness (N Y) 2015 Dec;6(6):1448-1455. [doi: 10.1007/s12671-015-0418-3]
73. Muris P. A brief questionnaire for measuring self-efficacy in youths. J Psychopathol Behav Assess 2001 Sep;23(3):145-149. [doi: 10.1023/A:1010961119608]
74. Denboer JW, Nicholls C, Corte C, Chestnut K. National institutes of health toolbox cognition battery. Arch Clin Neuropsychol 2014 Nov 1;29(7):692-694. [doi: 10.1093/arclin/acu033]
75. Heaton RK, Akshoomoff N, Tulsky D, et al. Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. J Int Neuropsychol Soc 2014 Jul;20(6):588-598. [doi: 10.1017/S1355617714000241]
76. Akshoomoff N, Beaumont JL, Bauer PJ, et al. VIII. NIH toolbox cognition battery (CB): composite scores of crystallized, fluid, and overall cognition. Monogr Soc Res Child Dev 2013 Aug 16;78(4):119-132. [doi: 10.1111/mono.12038]
77. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2022. URL: http://www.R-project.org/ [accessed 2026-04-21]
78. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol 2012 Mar;65(3):301-308. [doi: 10.1016/j.jclinepi.2011.07.011] [Medline: 22169081]
79. O'Brien H, Minich NM, Langevin LM, et al. Normative and psychometric characteristics of the health and behavior inventory among children with mild orthopedic injury presenting to the emergency department: implications for assessing postconcussive symptoms using the child sport concussion assessment tool 5th edition (Child SCAT5). Clin J Sport Med 2021;31(5):e221-e228. [doi: 10.1097/JSM.0000000000000943]
80. O'Kane KMK, Otamendi T, Silverberg ND, et al. Development of therapeutic alliance and social presence in a digital intervention for pediatric concussion: qualitative exploratory study. JMIR Form Res 2024 Mar 22;8:e49133. [doi: 10.2196/49133] [Medline: 38517472]
81. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11. [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
82. Wade SL, Sidol C, Babcock L, et al. Findings from a randomized controlled trial of SMART: an eHealth intervention for mild traumatic brain injury. J Pediatr Psychol 2023 Mar 20;48(3):241-253. [doi: 10.1093/jpepsy/jsac086]
83. Furlan JC, Radan MM, Tator CH. A scoping review of registered clinical studies on mild traumatic brain injury and concussion (2000 to 2019). Neurosurg 2020;87(5):891-899. [doi: 10.1093/neuros/nyaa151]

Abbreviations

360cc: 360 Concussion Care
CEQ: credibility and expectancy questionnaire
CHEO: Children's Hospital of Eastern Ontario
DTx: digital therapeutics
ED: emergency department
MBI: mindfulness-based intervention
MRI: magnetic resonance imaging
NIH: National Institutes of Health
PedsQL: Pediatric Quality of Life Inventory
PSAC: persisting symptoms after concussion
RA: research assistant
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by S Munce; submitted 22.Sep.2025; peer-reviewed by J Greenberg, L Babcock; revised version received 12.Mar.2026; accepted 23.Mar.2026; published 11.May.2026.

Please cite as:
Sicard V, Cairncross M, Zemek R, Silverberg ND, Goldfield GS, Reed N, Saab BJ, Smith A, Ledoux AA
Smartphone App–Delivered Mindfulness-Based Intervention for Concussion in Adolescents (MBI-4-mTBI): Feasibility Randomized Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e84623
URL: https://rehab.jmir.org/2026/1/e84623
doi:10.2196/84623

© Veronik Sicard, Molly Cairncross, Roger Zemek, Noah D Silverberg, Gary S Goldfield, Nick Reed, Bechara J Saab, Andra Smith, Andrée-Anne Ledoux. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 11.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Experience of Extended Reality Treatment for Visuospatial Neglect Among Patients and Informal Caregivers: Qualitative Interview Study

Eileen Bousché[1], MSc; Hanne Huygelier[1,2], PhD; Tanja CW Nijboer[1], PhD; Knowledge Broker Neglect Study Group[3,4]

[1]Experimental Psychology, Helmholtz Institute, Utrecht University, Heidelberglaan 1, Utrecht, The Netherlands

[2]Department of Brain and Cognition, KU Leuven, Leuven, Belgium

[3]See Acknowledgments

[4]

Corresponding Author:
Eileen Bousché, MSc
Experimental Psychology, Helmholtz Institute, Utrecht University, Heidelberglaan 1, Utrecht, The Netherlands

Abstract

Background: Visuospatial neglect (VSN) is a cognitive disorder following a stroke, where individuals fail to perceive or respond to stimuli on the contralesional side of space. Visual scanning training (VST) is the recommended treatment in clinical guidelines.

Objective: This qualitative study explored how patients (N=10) and informal caregivers (N=8) perceived the usability and potential implementation of 3 extended reality (XR)–based serious games to enhance VST—1 in virtual reality and 2 in augmented reality—as tools for VST. These technologies not only enhance patient engagement but also enable detailed data collection to monitor therapeutic progress.

Methods: The themes and feedback were compared with themes and feedback from VSN therapists from a previous study: (1) suitability for VSN rehabilitation, (2) applicability, (3) motivation, (4) guidance, (5) versatility, and (6) detailed insights in game performance.

Results: Highlights were that patients reported high engagement and enjoyment, with many expressing willingness to use the games in both clinical and home settings. Informal caregivers supported these findings and emphasized the motivational value of the games. Additionally, both groups noted the importance of accessible instructions and technical support.

Conclusions: Although XR technology offers potential in neurorehabilitation, a uniform solution may not suit all users. This study showed the importance of including diverse end-user groups in development for usability, acceptance, and implementation. Successful integration of XR in rehabilitation requires customizable features, structured support, and attention to the differing roles of caregivers. Further research is needed to evaluate the clinical effectiveness and optimize patient-tailored applications of XR in VSN treatment.

(JMIR Rehabil Assist Technol 2026;13:e80136) doi:10.2196/80136

KEYWORDS

visuospatial neglect; visual scanning training; extended reality; virtual reality; augmented reality; qualitative data

Introduction

One of the fundamental cognitive skills required for daily activities, such as locating personal belongings, navigating traffic, and engaging in social interactions, is the ability to scan the environment. Unfortunately, for individuals who have experienced a stroke, this ability is often impaired due to visuospatial neglect (VSN). VSN is a common poststroke cognitive disorder characterized by a lateralized attentional deficit [1-4]. This deficit manifests in behaviors such as bumping into objects, overlooking food on one side of the plate, or grooming only one side of the body. The consequences of VSN negatively affect functional independence, activities of daily living, and societal participation [5-8]. Informal caregivers often bear a significant burden, compensating for the functional deficits caused by VSN [9]. Clinical guidelines recommend visual scanning training (VST) as a primary intervention [10]. VST is a compensatory therapy designed to train systematic eye and head movements toward the contralesional side [11-13]. The therapy dose is quite intensive, as it requires time commitment up to 40 hours over 5 weeks [11]. Moreover, patients' lack of awareness of their VSN, a condition known as anosognosia [14-16], may diminish their motivation and engagement in therapy. This is particularly problematic when the effectiveness or quality of the learned compensatory strategy is difficult to assess visually using current forms of VST. If a device could be used that not only provides individualized,

XSL•FO
RenderX

interactive training but also objectively indicates how effectively the strategy is being applied, this could significantly enhance both monitoring and therapeutic outcomes.

Advancements in extended reality (XR) technology, especially virtual reality (VR) and augmented reality (AR), offer promising new approaches to VSN rehabilitation [17-19]. VR immerses individuals in fully simulated environments via specialized headsets, whereas AR integrates digital elements into real-world settings, allowing for interaction within familiar surroundings. When combined with serious games [20], a gamified form of therapy, XR not only enhances patient engagement, enjoyment, and usability [21-23] but also enables detailed data collection to monitor therapeutic progress [24,25]. Studies further support the potential of XR for both assessing and improving visuospatial attention after stroke. Immersive and semi-immersive VR environments have been shown to elicit and measure neglect-like behavior [26-28]. Moreover, a VR-based exploration task has demonstrated positive effects on spatial attention and engagement [29], and AR tasks have been used to strengthen compensatory scanning behavior in functional contexts [30,31]. In parallel, growing evidence shows the importance of adaptability and personalization in neglect rehabilitation. For instance, an adaptive cueing protocol, in which visual cues were gradually reduced as patients improved, resulted in meaningful gains in VSN-related functional outcomes [32]. Although this paradigm was developed outside XR, it provides a strong conceptual foundation for implementing performance-dependent, adaptive task adjustments within XR-based VST.

Despite these encouraging developments, current XR interventions for VSN often do not proceed to the implementation process. A crucial step required to eventually move from the development lab toward clinical practice is to involve end users. Thorough evaluation of end-user perspectives, especially those of patients and caregivers, is required to secure the implementation of VR and AR therapy [33,34].

Building on previous therapist evaluations [35], this study incorporated patient and informal caregiver feedback into the iterative development of 3 serious games for VSN rehabilitation (1 VR-based and 2 AR-based). In this study, we collected perspectives of patients and informal caregivers and assessed to what extent they were distinct from the perspectives of therapists on the same 3 games reported by Bousché et al [35]. We aimed to assess whether these 2 end-user groups also mentioned the same 6 key findings from therapist evaluations: (1) suitability for VSN rehabilitation, (2) applicability, (3) motivation, (4) guidance, (5) versatility (ie, patient's needs), and (6) detailed insights in game performance. Variance in responses among groups will emphasize the need for inclusion of different end-user groups in the development process.

Methods

Participants

The first group of end users consisted of patients with VSN admitted to De Hoogstraat Rehabilitation Center for inpatient rehabilitation between April and October 2021. Inclusion criteria were as follows: (1) admitted to the neurology ward, (2) presence of VSN diagnosed by a neuropsychologist or screened by a therapist using the Balloons Test [36] or Catherine Bergego Scale (CBS) [37], and (3) age 18 years or older. The second group of end users consisted of informal caregivers of participating patients. Inclusion criteria were as follows: (1) age 18 years or older and (2) a familial or social relationship with the patient. Epilepsy was the sole exclusion criterion for all participants because of potential triggers for seizures in the serious games.

XR Serious Games

Three XR serious games were administered. First, the AR Virtual Museum used a Microsoft HoloLens 1 (1268×720 pixel per eye; 60 Hz; 35° FOV). This game promotes visual search by projecting virtual paintings onto real-world walls. Participants moved around the room to locate paintings, which transformed into short video clips upon discovery [30]. Second, VR HEMIRehApp employed an Oculus Rift CV1 (1080×1200 pixel per eye; 90 Hz; 110° FOV) with handheld controllers. Designed to retrain spatial attention toward the contralesional side, the game places participants in a virtual farm setting with tasks such as harvesting crops and feeding animals. It was initially developed for right-hemispheric stroke patients with intact language abilities [38, 39], and its validity and usability were confirmed in a pilot study [24]. Third, the Balloon Popping game, developed by Holomoves, used a HoloLens 2 (2048×1080 pixel per eye; 60 Hz; 52° FOV). This AR-based intervention integrates visual search training with physical movement. Virtual balloons appear around the player, who must locate and "pop" them using finger gestures. Auditory (verbal instructions) and visual (green directional arrows) cues aid the task. Participants played 3 mini games: simple balloon popping, image pairing, and word formation.

Outcome Measures: Interviews

A semistructured interview (Textbox 1) was designed by the authors, which consisted of 15 qualitative and 2 quantitative questions where the participant could rate an aspect of the XR game. The design was informed by the System Usability Scale [40], a widely used instrument for rapid quantitative assessment of usability. To obtain a deeper understanding of participants' experiences, we supplemented the scale with additional targeted questions addressing VST, user experience (eg, own performance, equipment use, cybersickness), usability (eg, playing the game autonomously, during therapy), and preference (eg, liking, improvement) of the games. The interview was recorded for analysis. Informal caregivers received a similar interview, without item 16 (practicing visual search outside therapy). Additionally, the items on practicing autonomously were rephrased for helping the patient at home or at the rehabilitation center.

XSL•FO
RenderX

Textbox 1. Interview on user experience, usability, and preference questionnaire. Participants rated difficulty and enjoyment on a scale from 0 to 10, where higher scores indicate greater difficulty or enjoyment.

User experience

1. What did you think of the game?

2a. What went well?

2b. What went less well (game-related or personal performance)?

3. How did the headset feel on your head?

4. Was there anything about the headset or controllers that you found uncomfortable?

5. How did you experience moving around in the virtual environment?

6. Some people feel warm or need time to adjust. How was this for you? How did it feel to play the game?

Usability

7a. Did you find the game difficult or easy (rate on a scale from 0=very easy to 10=very difficult)?

7b. What influenced this?

8. What could make playing the game easier or more difficult for you?

9. What did you think about the duration of the game?

10a. Can you imagine using this game on your own for extra practice?

10b. Why or why not?

11. If we were to offer a 2-week training program to practice daily with this game, would you be willing to participate?

12a. Would you recommend this game to fellow rehabilitation patients?

12b. Why or why not?

Preference

13. How enjoyable did you find the game (rate on a scale from 0=not enjoyable at all to 10=very enjoyable)?

14. What aspects made the game particularly enjoyable to play?

15. What do you think could make the game even more enjoyable?

16. Do you currently practice searching and scanning for objects outside of therapy?

17. Is there anything that would prevent you from wanting to play this game?

Procedure

Patients were first informed about the study by researchers and invited to participate. Interested patients were scheduled to play their first game within 2 weeks. Before gameplay, the researcher reviewed the information letter and study procedure with the patient. For patients with aphasia, a speech therapist facilitated communication using simplified phrasing, repetition techniques, questioning, and pictorial techniques. The AR games were played either while walking, sitting, or rolling in a wheelchair or office chair. The VR game was played while participants were seated. After the patient completed the games, their informal caregiver was invited to participate and play the games as well. Each participant played all games once during one-on-one sessions with the researcher within a period of 3 weeks in a quiet room. Researchers assisted with headset placement, provided instructions, and answered questions. Throughout gameplay, the researcher closely monitored participants for signs of discomfort, fatigue, or cybersickness, such as dizziness, nausea, or disorientation. Participants were explicitly instructed to report any discomfort at any time, and gameplay was paused or stopped if adverse effects were observed or reported. Game order was counterbalanced, with each session lasting 10 to 15 minutes, including instructions. Each game was played once. After each game, a 15-minute semistructured interview was conducted with each participant.

Patient Data

The following patient data were collected from electronic patient files: age, sex, time poststroke onset (in days), stroke type (ischemic, hemorrhagic), and lesion side (left, right, bilateral). Functional scores were also collected for strength in the upper and lower extremities using the Motricity Index (MI) [41], global cognitive function using the Montreal Cognitive Assessment [42], level of communication using the Dutch Aphasia Foundation (Stichting Afasie Nederland) [43], and independence in activities of daily living using the Barthel Index [44]. Two neuropsychological tests were used to assess VSN severity: the Balloons Test [36,45] and the CBS [37].

Analyses

To explore the user experience, applicability, and suggestions of the patients and informal caregivers, the ratings were described and the qualitative data were analyzed. Audio files of the interviews were transcribed using Microsoft Word and later adjusted by the researchers to make a coherent story. Then, a deductive coding scheme was formed based on 6 key findings from therapists' evaluations: (1) suitability for VSN

XSL•FO
RenderX

rehabilitation, (2) applicability, (3) motivation, (4) guidance, (5) versatility (ie, patient's needs), and (6) detailed insights in game performance [35], following the approach described by Boeije et al [46]. These themes served as a preliminary analytic framework, based on the assumption that variation in perspectives across stakeholder groups would show the importance of incorporating multiple end-user voices in XR development. For the present study, each of the 6 themes was further differentiated into "positive," "negative," and "neutral" subcategories to capture the valence of participants' remarks and "suggestions" to facilitate a more nuanced comparison across user groups (Multimedia Appendix 1). All interviews were coded by 1 researcher, with 2 other researchers taking samples from both groups (ie, patients and informal caregivers). The samples were then further discussed in the research team to reach consensus on the coding process. After reaching consensus, all interviews were coded again based on the agreed-upon criteria. Qualitative findings were descriptively triangulated with test observations and available clinical data (eg, neglect severity, lesion side) to support interpretation and identify converging or diverging patterns, without formal quantitative hypothesis testing.

Ethical Considerations

This study was approved by the Medical Ethics Committee of the University Medical Center Utrecht and the Ethical Committee of De Hoogstraat Rehabilitation Center (reference number: 21/706). All procedures were conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent prior to participation. The data consisted of anonymized patient and interview data, which were collected and processed in a manner that ensured participant privacy and confidentiality. Participants did not receive financial or other compensation.

Results

Patient Data

A total of 12 patients with VSN were recruited and provided informed consent to participate in the study. However, 2 patients withdrew before participation due to fatigue, reduced cognitive capacity, lack of motivation, and difficulty recalling study details or scheduled sessions. Of the 10 remaining patients (age 49 - 77 y; 70% male), 3 played all games, 6 played 2 games, and 1 played 1 game. Additionally, 10 informal caregivers were recruited (age 45 - 75 y; 25% male), of whom 8 participated in this study (n=5 played all games; n=2 played 2 games; n=1 played 1 game). Interview excerpts were translated from Dutch to English to serve as examples in this publication.

Table 1 presents individual patient demographic and clinical characteristics. The majority (n=7) had experienced ischemic stroke, 2 had experienced a hemorrhagic stroke, and 1 patient had a cerebral abscess. Lesion lateralization was right hemispheric in 6 patients, left hemispheric in 3, and bilateral in 1 patient. Most (n=8) were in the subacute phase of recovery; the rest were in the acute phase (n=2). The patient group presented with VSN that was mild or not detectable using conventional paper-and-pencil tests. However, clinical observations as scored with the CBS revealed the full range of VSN severity. CBS data for 7 patients were available, indicating mild VSN for 10% (CBS<10), moderate VSN for 30% (CBS<20), and severe VSN for 30% (CBS>20). For those patients without CBS scores, 1 patient exhibited moderate neglect symptoms on the Balloons Test (Subtest B score<17; lateralization balance score<40%), and for 2 other patients, clinical observations indicated the presence of VSN. Although VSN was underrepresented on standardized assessments, the group nonetheless included patients across the entire spectrum of VSN. Cognitive screening indicated that all patients had cognitive impairment, ranging from impaired to severely impaired levels, with all scores falling below the clinical cutoff (Montreal Cognitive Assessment score<26). Motor function, using the MI, showed that patients on average had moderate motor impairment (0 - 33 severe; 34 - 66 moderate; and 67 - 99 mild/normal). In terms of functional independence, patients were generally highly dependent on others for daily care. Based on their scores on the Barthel Index, the group as a whole fell within the range corresponding to "severely dependent" (5-9), although individual scores ranged from "completely dependent" (0 - 4) to "moderately dependent" (10-14). This reflects a patient population with substantial functional needs and limited autonomy in daily living.

XSL•FO
RenderX

Table . Patient demographic and clinical characteristics at admission to the rehabilitation center.[a]

ID	Age (y)	Sex	Time post onset (d)	Hemisphere	Etiology	MI[b] Arm (0 - 99)	MI Leg (0 - 99)	MoCA[c] (0 - 30)	SAN[d] (0 - 7)	BI[e] (0 - 20)	CBS[f] (0 - 30)	Balloons (A and B=L0-10/R0-10; LB[g] 0% - 50%)
1	53	Male	25	Left	Ischemic	18	59	17/30	5	8/20	17.5	—[h]
2	60	Male	22	Left	Hemorrhagic	0	0	Partial (aphasia) visuospatial 2/5; orientation 3/6	2	2/20	—	A=L10R10; B=L10R10; LB=50%
3	49	Male	13	Right	Hemorrhagic	56	75	18/30	7	4/20	—	A=L10R10; B=L5/R8; LB=38.5%
4	64	Male	72	Right	Ischemic	29	75	23/30	7	8/20	22.5	A=L4/R10; B=L0/R3; LB=0%
5	56	Male	9	Right	Ischemic	0	83	25/30	7	12/20	13.3	A=L10R10; B=L7/R10; LB=41.2%
6	68	Male	4	Right	Ischemic	78	78	21/30	5	13/20	27	A=L9/R10; B=L4/R8; LB=34%
7	77	Female	7	Right	Recurrent ischemic	76	64	23/30	5 - 6	18/30	16.25	A=L10R10; B=L8/R7; LB=46.7%
8	52	Female	9	Left	Ischemic	0	—	Not assessed (aphasia)	2	7/20	9	—
9	61	Male	14	Right	Cerebral abscess	83	91	24/30	7	13/20	—	A=L10R10; B=L10R10; LB=50%
10	74	Female	14	Bilateral	Ischemic	76	83	4/30	3 - 4	12/20	24.4	—

[a]Values are presented as raw scores unless otherwise specified. Aphasia was assessed using the SAN test. Complete SAN data were available for all participants.

[b]MI: Motricity Index.

[c]MoCA: Montreal Cognitive Assessment.

[d]SAN: Stichting Afasie Nederland (aphasia).

[e]BI: Barthel Index.

[f]CBS: Catherine Bergego Scale.

[g]LB: lateralization balance.

[h]Missing data.

Suitability for VSN Rehabilitation

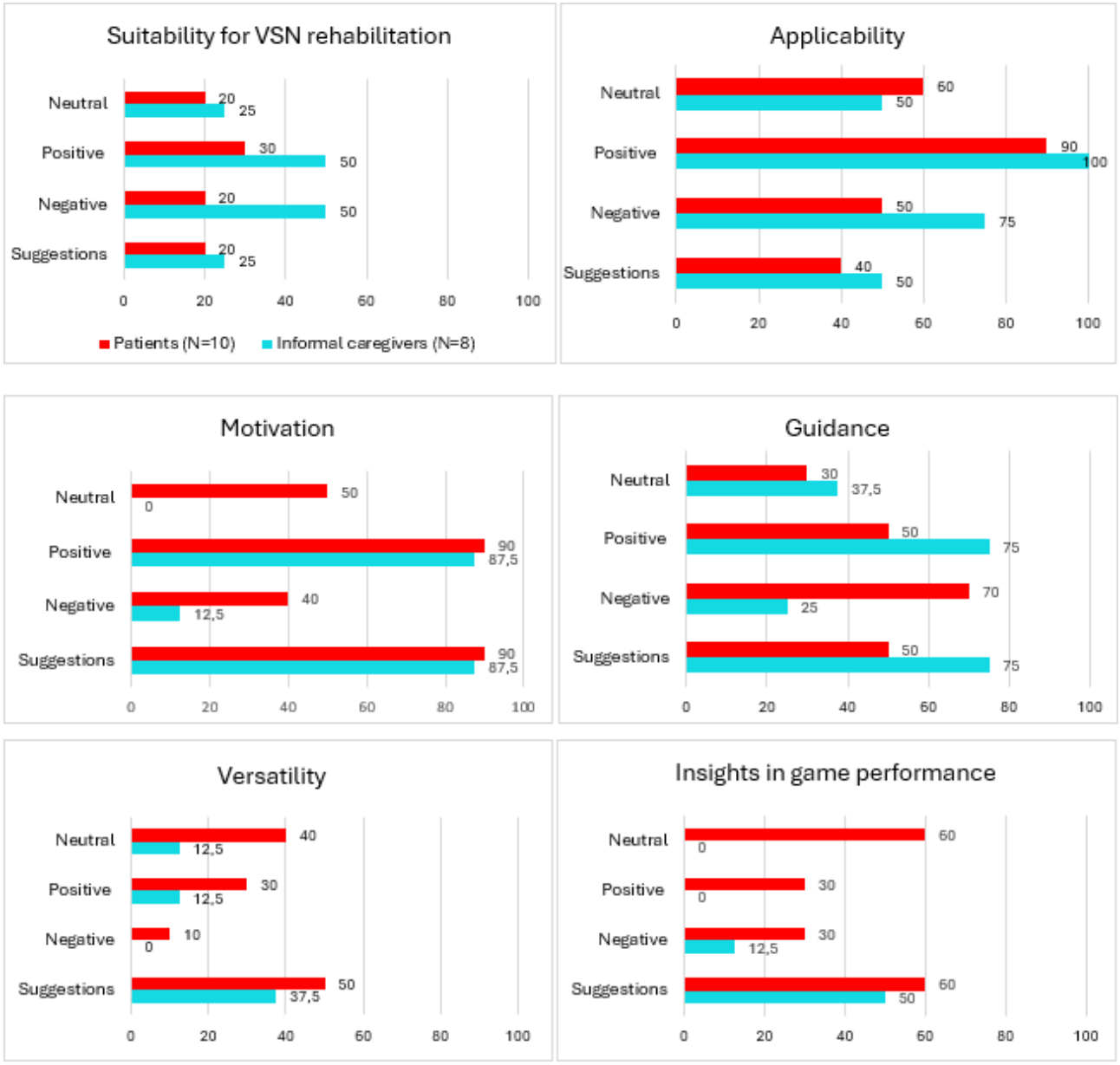
The participants' perceptions of the games' suitability for VSN rehabilitation were varied (Figure 1). Patients provided fewer comments on this aspect compared with other categories and even fewer than their informal caregivers. Among informal caregivers, the distribution of positive and negative responses was equal. An informal caregiver mentioned that for their loved one, training would be appropriate, while it was too easy for a healthy person such as themselves. It was also suggested to improve clarity regarding the game's main objective and ensure that patients understand the added value of playing. There were also suggestions for improving the suitability of the games. For example, one informal caregiver mentioned that adjustments need to be made for suitability for neglect, because currently, the targets were only right in front of the player.

> *…I thought: now something is going to happen, but most of the targets were actually right in front of me. Those are things to which more attention is paid to: who will be playing next? Okay, this patient has left-sided neglect, then we need to place more targets*

XSL•FO
RenderX

on the left. Sometimes also behind the patient, so they really have to turn their head and search for where it is. [Informal caregiver03]

Figure 1. The coding scheme reflected in percentages of patients (N=10) and their informal caregivers (N=8) who gave neutral, positive, and negative responses and suggestions. VSN: visuospatial neglect.



Applicability

The applicability of VR and AR games to the target patient population was an important topic for both patients and their informal caregivers. Ninety percent (9/10) of patients and 100% (8/8) of informal caregivers made positive remarks about the feasibility and applicability of these games (Figure 1). Patients reported that the games were easy to play and that the duration of gameplay was sufficient. Patients also indicated feeling comfortable while playing and were able to move freely while seated in an office chair or wheelchair.

Similarly, most informal caregivers found the games intuitive and of appropriate duration when played themselves; however, some expressed concerns regarding the potential difficulty of gameplay duration for patients with VSN, thus making a distinction between their own experiences and the anticipated challenges faced by patients.

Despite overall positive responses, 50% (5/10) of patients also expressed concerns regarding the applicability of the games. One patient wondered whether older individuals would be able to engage with the games effectively.

> *I really like the idea. But I can imagine that for older people it might be quite difficult to use.* [Patient05]

Even more informal caregivers were concerned (6/8, 75%), with issues related to game content (eg, difficulty with multitasking), safety (eg, potential distraction leading to reduced awareness of surroundings), and patient-specific factors (eg, individuals with memory impairments might struggle with certain games).

XSL•FO
RenderX

> *…You get lots of information to read, so you have a lot to process. This could be hard for people who have had a stroke.…*. [Informal caregiver08]

> *No I wasn't distracted by the paintings, but the patient might bump into something.* [Informal caregiver07]

> *…But with the patient it will not be that easy. I think when making a turn, the patient might fall.* [Informal caregiver03]

> *In my case, my memory is good. But for people with a stroke, memory is often affected, right? If you make the game more difficult, I think it might become too demanding for them.* [Informal caregiver05]

Furthermore, it was observed that 1 patient encountered difficulties while engaging with AR Balloons and VR HEMIRehApp, due to the specific motor demands of these games (eg, pointing and using a controller). Conversely, AR Museum proved to be more suitable for this patient, as it only required eye fixations. Although the patient was not able to finish the games, the patient still had positive views on this type of technology.

Additionally, 40% (4/10) of the patients provided suggestions for improving applicability. Patients questioned whether it was feasible to integrate the games within the busy schedule of the rehabilitation center, proposing that a single session might be appropriate, without being specific about an exact therapy dose.

> *…But it has to be feasible alongside other therapy sessions. Well…it doesn't need to be a problem, but when you have many [therapy sessions] it could become difficult. But when you would receive it [the equipment] for a single day, it would be okay.* [Patient08]

Among informal caregivers, 50% (4/8) suggested modifications to enhance applicability, primarily focusing on safety (eg, using a large, empty space to minimize the risk of collisions) and interface adjustments (eg, a more salient cursor to keep track of one's actions in the game).

> *…I noticed quickly that you really have to poke through the balloon, but I wonder how patients will handle this, that you really need to poke through the balloon.…Maybe the game could make the cursor or pointer more salient so it's clearer what they are actually interacting with. For me it wasn't a problem, but it might be for others.* [Informal caregiver03]

Motivation

Both groups were asked about the perceived likability of the games and their motivation to continue playing or reengage with the games. Although participants expressed a range of motivation-related sentiments (neutral, positive, and negative), the content of these responses showed little variation. Overall, patients (9/10, 90%) and informal caregivers (7/8, 87.5%) reported similarly positive perceptions of the games' motivational aspects. The most frequently mentioned motivational factor was positive affect, as participants indicated that they enjoyed playing the games. For instance, 1 patient noted that, beyond the therapeutic necessity of the games, engaging with them provided a welcome distraction from routine therapy.

> *…as for the game element, I think many people would enjoy it simply because it's something different and engaging.* [Patient05]

Patients appreciated the physical engagement involved, such as making head movements, searching, and scanning, which the games facilitated.

> *I really appreciated the movements involved: searching, moving the head, and scanning. I saw many positive aspects in it.* [Patient 09]

> *Moving around was easy. You knew when the music stopped the game was finished, so you had to keep going until you found another target.* [Patient01]

Informal caregivers mentioned the appeal of the games' colors and dynamic elements, as well as their informative nature.

> *In the last game, I especially liked that the deer was moving that made it more engaging. Overall, I thought the game was well designed.* [Informal caregiver10]

> *I really enjoyed it, you receive information and things are explained, for example about a painting. I didn't necessarily recognize anything personally, but I still liked that something was being told and that you could learn from it.* [Informal caregiver08]

Negative feedback was more prevalent among patients (4/10, 40%) than among informal caregivers (1/8, 12.5%). Informal caregivers primarily mentioned the difficulty level of the games. Patients, however, reported concerns such as inadequate gamification of therapeutic elements within the game design and difficulties in locating objects.

> *At this moment I would not yet recommend it, because it is still in the research phase. First, people should practice with everyday situations, things they normally encounter and tend to neglect in daily life. That should be mastered first.* [Patient03]

> *…Yes, because I couldn't find them easily. If the targets were easier to locate, your attention would be drawn to them more, and you would probably be able to continue for longer.* [Patient04]

> *…I think it has to do with the fact that you just have to click away two things each time. Not much else happens, you know?* [Informal caregiver07]

Suggestions for enhancing motivation were largely consistent across groups, with 90% (9/10) of patients and 87% (7/8) of informal caregivers proposing modifications such as incorporating diverse themes (eg, zoo, art), increasing difficulty levels (eg, more targets, time constraints), and introducing additional objectives (eg, puzzles, object-chasing tasks).

> *Maybe there could be multiple themes, or a slightly more difficult level. For example, combining two themes in one task. That would add more challenge, because you would need to look more actively for different types of objects.* [Patient09]

XSL•FO
RenderX

> *Maybe there could be a time limit, so you know in advance that you have two or three minutes. But it has to be realistic, it should still be achievable for someone with a condition like mine.* [Patient05]

Guidance

The theme regarding guidance encompassed tutorials or support from another person, from the game itself, or from documents. Positive remarks from both patients and informal caregivers were that the instructions within the games were clear. The graphics and targets were visible, and the instructional voice was audible.

> *The instructions were explained clearly while I was seated, and everything was said out loud and shown on screen, so I didn't even have to read them. I could just start playing right away.* [Patient01]

According to the patients (7/10, 70%), negative aspects of guidance were the instructions for gameplay, controllers, or the act of selecting targets, as well as the lack of feedback on their gaming performance.

> *Maybe the next time, once you know what the intention is, it will be easier.* [Patient10]

> *What you have to remember is that with the green leaf you need to use the lower button instead of the upper one and with the orange leaf it's the other way around.* [Patient05]

Negative aspects according to informal caregivers (2/8, 25%) were, for example, that the in-game instructions were not sufficiently noticeable or understandable for patients.

> *I think that for people who have difficulties understanding due to a brain injury, a bit more guidance at the beginning would be helpful, clearly explaining what to do. After that, they would probably be able to manage.* [Informal caregiver8]

Suggestions from patients (5/10, 50%) included allowing participants to practice more to accurately remember the instructions. Another interesting suggestion was to enhance cues and feedback related to game performance.

> *I don't know if it's technically possible, but if you could see what I'm seeing, it might be helpful if you could say something like "your gaze is in the right place" or "look up" or "look down," especially since our eyes and pupils are apparently being measured.* [Patient06]

Additionally, 75% (6/8) of the informal caregivers had suggestions for improving the guidance. A recommendation was to give the instructions a more therapeutic focus, for example, by suggesting the use of a different hand to select targets.

> *…You really need to indicate "use your other hand." It's not that you will automatically use your left. If you have left-sided neglect, use your left hand. Maybe there could be text in between, indicating which finger or hand to use, to stimulate that.* [Informal caregiver09]

Versatility (ie, Patient's Needs)

Next to suitability, versatility or tailoring the games to individual patients' needs was less often mentioned, compared with other themes from the coding scheme. Versatility of the serious games was perceived positively by 30% (3/10) of patients and 12% (1/8) of informal caregivers. One negative aspect was mentioned by a single patient.

> *For me, the music could be different. I can imagine that younger people might really like this, but we are from a different generation, so I would prefer another type of music.* [Patient06]

Suggestions regarding versatility were made by half of the patients (5/10, 50%) and just over a third of the informal caregivers (3/8, 38%). Patients and informal caregivers primarily suggested adapting game themes to personal preferences, such as allowing users to choose their own virtual environment.

> *Maybe there could be multiple themes, or a slightly higher level of difficulty, for example by combining two themes in one task. That would add more challenge, because you would have to search for different types of objects.* [Patient09]

> *There should be more variation so people can choose what they want to do, for example, different environments or more animals appearing in the game.* [Informal caregiver05]

> *For example, my brother loves farm life. If the game included a video about a farm and the animals, he would probably find it much more interesting.* [Informal caregiver01]

Detailed Insights in Game Performance

A difference was found between the 2 groups in mentioning the insights derived from gameplay. Whereas patients expressed a clear desire to know how they performed in the games, informal caregivers showed less interest, with only 13% (1/8) mentioning it as a negative aspect. Suggestions for game performance insights primarily involved displaying accuracy or total score. A notable suggestion was to include an option that allows users to compare their current performance with previous results or the performance of other players.

> *Just showing the number of points or the percentage of correct answers is enough for me. Something like a graph would also work.* [Patient 03]

> *…What matters is whether you did it correctly or not.* [Patient 02]

> ***Researcher:*** *Would you like to know how well you did in the game?*

> ***Patient:*** *Yes, of course.*

> ***Researcher:*** *Would you also like to see how others performed, or just compare your own results next time?*

> ***Patient:*** *Yes actually, I'd like to see everyone's scores. That would be fun.*

> ***Researcher:*** *Then you could gradually try to catch up with the top score, like "Jan," who did better.*

XSL•FO
RenderX

Patient: *Yes, I like that.* [Patient10]

Other Findings

Other findings revealed that informal caregivers evaluated the games from both their own perspective and that of the patients. Informal caregivers often elaborated on imagining how playing the game would be for the patient.

> *Yes, I think it went well. I had no problems, but I can imagine that a patient, like the patient with left-sided neglect, would find it harder to distinguish, for instance, where a balloon is located…I can imagine that a patient might have more difficulty and would need to think constantly.* [Informal caregiver03]

> *I can imagine that if someone isn't really into games, like my brother, who doesn't like them at all, they might quickly lose interest.* [Informal caregiver01]

This dual perspective is comparable to the approach taken by therapists who primarily considered the patient's viewpoint. Interestingly, therapists leaned more toward patient-centered observations, while informal caregivers balanced their own experience with their perception of the patient's experience.

To increase the total time for dedicated training, it was explored whether patients would use the serious games at home or at the rehabilitation center, preferably with the help of an informal caregiver. Patients expressed willingness to practice at home if provided with the necessary equipment. Similarly, they were open to using the headsets at rehabilitation centers if readily available. However, reasons for not doing so included having other matters to attend to at home or preferring to spend time with friends and family at the rehabilitation center. Informal caregivers (6/8, 75%) expressed a willingness to support home practice, provided it fit within their daily schedules. However, their willingness to assist in rehabilitation centers was more divided (4/8, 50%), often dependent on practical considerations such as work commitments or the belief that therapy should primarily be managed by health care professionals.

> *It depends on the location. Not here at the Hoogstraat Rehabilitation, that would take too much time. But if he had one at home, I would be willing to do it. If I had to come here every day with him, then I think the care staff would need to do this more often.* [Informal caregiver01]

Lastly, no adverse effects were reported. The hardware and software were generally well received (7/10, 70% patients; 6/8, 75% informal caregivers). However, a few participants (2/10, 20% patients; 2/8, 25% informal caregivers) reported initial difficulty adjusting to the VR headset.

> *It took a moment to get used to it, a bit of dizziness, and then I returned and,"phew."* [Patient02]

Realistic VR scenarios, such as fish catching over a virtual lake, induced vertigo for one informal caregiver who reported fear of heights after the session.

> *I personally have a bit of a fear of heights, which may sound strange, but in that water scene it really felt as if I was above the water. When I looked down, I thought, "wow."* [Informal caregiver07]

Discussion

This study assessed the user experience of patients with VSN and their informal caregivers and compared the findings to the themes frequently mentioned by VSN therapists in a previous study by Bousché et al [35]. The participants engaged in 3 serious games, 1 VR and 2 AR games, based on VST for VSN.

Principal Findings

Participants' perceptions regarding the suitability of the games for VSN rehabilitation were varied. Patients generally made fewer comments on this aspect compared with informal caregivers, who showed an equal distribution of positive and negative feedback. Some informal caregivers noted that the training would be suitable for their loved ones but too easy for healthy individuals, and others suggested improving the clarity of game objectives and tailoring the games specifically for neglect (eg, adjusting target placement). In terms of applicability, both groups found the games feasible and user-friendly, although some expressed concerns regarding the duration and potential difficulties for patients with VSN. Suggestions for improvement included better customization for older individuals and ensuring safe use in clinical settings. It seemed to be easier for both groups to talk about applicability and feasibility than the more therapeutic properties of the games, something that is obviously not a problem for VSN therapists.

Next to the therapeutic aspects, motivational aspects were also evaluated. Incorporating gamification elements into XR rehabilitation can increase patient motivation and engagement [31,47-49]. Gamification VST was positively perceived by both patients and informal caregivers. The most frequently mentioned motivation indicator was positive affect, as participants enjoyed the gameplay and appreciated the physical engagement. However, some patients expressed concerns about insufficient gamification of therapeutic elements and difficulty locating objects. Informal caregivers suggested increasing the difficulty level to sustain interest. This is consistent with therapists' views, which were positive regarding motivational aspects but questioned whether players would pick up the game again. Suggestions for enhancing motivation were similar across groups (patients and informal caregivers), including incorporating diverse themes (eg, zoo, art), increasing difficulty levels, and introducing additional objectives (eg, puzzles).

Both groups appreciated the clear in-game instructions and the visibility of graphics that accompanied the instructions. However, more patients than informal caregivers reported challenges related to guidance, such as difficulty with controller usage and the lack of performance feedback. Morse et al [50] pointed out the importance of integrating clear guidance within XR interventions, as well as including feedback mechanisms to support user engagement. Therapists took it a step further; almost half of the therapists suggested improving the in-game tutorials, instructions for therapists, and protocols for handling the equipment.

Additionally, tailoring the games to the patient's individual needs was mentioned less frequently compared with other themes, suggesting that further adaptation may enhance the

XSL•FO
RenderX

therapeutic experience. VR interventions can be effective when the level of immersion is adapted to patient needs [28]. Tailoring by both patients and informal caregivers was targeted towards suggestions on adjusting themes to the liking of the patient, in that they could choose their own virtual environment, while therapists leaned more toward adjusting the games to different patient types and clinical requirements (eg, patient's mobility or sensitivity to sensory input).

Interestingly, we found differences between patients and informal caregivers regarding game performance insights. Patients expressed a desire to receive feedback on their gameplay, while informal caregivers showed limited interest, with one mentioning not being interested. Suggestions to improve game performance insights included displaying accuracy or total scores and incorporating a feature that allows users to compare their current performance with previous results or with others. The interest and suggestions were limited compared with the evaluations of therapists. They mentioned accuracy, temporal measures, strategy, motor measures, or combinations thereof [35].

This pattern is noteworthy in light of the fact that patients with VSN frequently present with anosognosia or reduced awareness of their deficits [14-16]. One might therefore expect limited interest in performance feedback if patients believed they were performing adequately. The opposite pattern observed in this study suggests that interactive XR-based tasks may facilitate a degree of online awareness of performance during task execution, even when more generalized or offline insight remains impaired [14]. The immediate and goal-directed nature of XR gameplay may make errors or omissions more salient to patients than conventional assessment tasks. Although this interpretation remains speculative, it aligns with theoretical accounts distinguishing between online performance monitoring and global awareness of deficits, and it points to the potential of XR environments to support self-monitoring processes without necessarily increasing cognitive burden.

Our study revealed that informal caregivers evaluated the games from both their own perspective and that of the patients. This dual perspective is different from the approach taken by therapists [35], where therapists primarily considered the patient's viewpoint. Interestingly, therapists leaned more toward patient-centered observations, while informal caregivers balanced their own experience with their perception of the patient's experience. This difference emphasizes the distinct roles and perspectives of informal caregivers compared with health care professionals. The importance of involving informal caregivers in rehabilitation is further supported by a randomized controlled trial, which found that caregiver participation in VR training improved patient outcomes in unilateral neglect [51]. They therefore recommended trained informal caregivers to provide VR guidance and supervision to patients who have unilateral neglect.

We explored whether patients could independently use the serious games at home, aiming to intensify VST by increasing the therapy dose as described by Pizzamiglio et al [11]. According to Morse et al [50], end users were positive and interested in this type of telerehabilitation for VSN. In the current study, patients expressed willingness to practice at home if provided with the necessary equipment. Similarly, they were open to using the headsets at rehabilitation centers if readily available. Since informal caregiver participation can improve therapeutic outcomes [51,52], they were asked whether they could aid the patients to play at home. Informal caregivers generally expressed a willingness to support home practice, provided it fit within their daily schedules. However, their willingness to assist in rehabilitation centers was more divided, often dependent on practical considerations such as work commitments or the belief that therapy should primarily be managed by health care professionals.

In a systematic review by Lundin et al [53], the importance of monitoring and reporting adverse effects in VR interventions was advocated. In our study, no adverse effects were reported by both patients and informal caregivers. The headsets and VR controllers were generally well received, although some participants initially struggled with headset adaptation. One informal caregiver with a fear of heights experienced mild vertigo during a realistic scenario, calling attention to the need for monitoring cybersickness in XR interventions, specifically in VR.

To conclude, consistent themes across user groups were identified. These findings demonstrate the value of involving diverse end-user groups when designing XR serious games for health care, particularly for VSN therapy. Including patients, their informal caregivers, and therapists throughout the development process [54] can enhance usability and therapeutic relevance, as supported by studies on VR-based rehabilitation [51,52,55].

Strengths and Limitations

A strength of this study is that it involved patients and their informal caregivers in the development process of different types of XR technology for VSN therapy. The aim was not only to eventually optimize the technology and tailor the games to the patient's needs but also to investigate the potential of including informal caregivers in the therapy and letting the patient work autonomously, as well as make the feedback comparable to the previous study with therapists. The diversity of end users ensures that all relevant groups were involved in the development and evaluation process. Another strength of this study is the inclusion of a representative patient sample, capturing the full range of relevant clinical impairments. The sample included patients with cognitive impairment ranging from mild to severe, VSN severity spanning from very mild to severe as measured by the CBS, and moderate motor impairments as assessed by the MI. This broad clinical spectrum enhances the ecological validity and generalizability of the findings to broader rehabilitation populations. Additionally, the sample size was somewhat larger and with a broader age range than comparable studies [30,31,50].

A potential limitation of this study concerns the assessment of VSN for participant inclusion. Conventional standardized measures were not systematically incorporated, which makes it possible that some participants were already relatively well trained or that neglect symptoms were no longer apparent during static, paper-and-pencil tests typically used in clinical practice.

XSL•FO
RenderX

Consequently, variability in neglect severity across participants may have influenced their interaction with the XR games. While no clear patterns emerged, there were indications that participants with more pronounced neglect or right-hemispheric lesions required more guidance during gameplay, which may have shaped their user experience. In addition, the presence of aphasia in some participants may have affected the depth and clarity of their verbal feedback, despite facilitation by a speech therapist. Although care was taken to support communication, responses from participants with aphasia may have differed in richness or nuance compared to those without language impairments, potentially influencing the qualitative findings. Another potential limitation is the monocenter design of the study. The rehabilitation center involved is progressive in its use of science and technology, which may have positively influenced participants' attitudes toward XR-based interventions. Furthermore, no questionnaire was administered to assess prior experience with digital technology, which limits insight into possible familiarity effects and introduces the risk of selection or contextual bias related to the study setting. Additionally, gameplay sessions were relatively short, which likely limited the occurrence of attentional fluctuations, fatigue, or more pronounced manifestations of VSN during use. As a result, the findings should be interpreted as a snapshot of user experience at a single point in time. Future research should therefore investigate patient and caregiver experiences across longer or repeated gameplay sessions to better assess sustained engagement, usability, and the potential therapeutic value of XR-based interventions over time.

Clinical Implications

This study contributed to the first steps toward implementing XR technology in rehabilitation centers. Our findings suggest that integrating XR-based therapeutic interventions into rehabilitation programs for VSN can significantly improve patient engagement and therapy adherence. Given the positive user experience reported by both patients and informal caregivers, the use of VR and AR games appears promising as a supplementary tool for VST. However, to enhance clinical applicability, it is crucial to address challenges related to guidance, customization, and motivation. Specifically, rehabilitation centers should consider integrating these technologies in a way that accommodates both patients and their informal caregivers, providing clear instructions and adaptable game settings. Furthermore, recognizing the varying levels of willingness among informal caregivers to support therapy sessions in different settings is essential for successful implementation.

Conclusion

This study reported the need for a nuanced understanding of how various end users perceive and interact with XR-based therapeutic interventions. It showed the importance of involving different end users for the development process to become more inclusive and tailored, ultimately enhancing the usability and therapeutic potential of XR applications for VSN. Future research should continue to explore the efficacy and patient-tailored approaches to optimize XR interventions in neurorehabilitation.

Acknowledgments

We thank the Knowledge Broker Neglect Study Group from De Hoogstraat Rehabilitation Center and Judy Bakker from Stichting Omring for their valuable input and assistance in recruiting participants. The members of the Knowledge Broker Neglect Study Group are as follows: Eugenie Brinkhof (Physical therapist), De Hoogstraat Rehabilitation, Utrecht, The Netherlands and Dutch Institute of Allied Health Care/Nederlands Paramedisch instituut, Amersfoort, The Netherlands; Oscar Haver (Rehabilitation physician, MD), De Hoogstraat Rehabilitation, Utrecht, The Netherlands; Marike Jansen (Occupational therapist), De Hoogstraat Rehabilitation, Utrecht, The Netherlands; Cindy van de Moosdijk (Psychologist), De Hoogstraat Rehabilitation, Utrecht, The Netherlands; Anja Timmerarends-Van Eck (Nurse specialist), De Hoogstraat Rehabilitation, Utrecht, The Netherlands.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: TCWN (lead), EB (supporting)
Data curation: EB
Formal analysis: EB (lead), HH (supporting), TCWN (supporting)
Funding acquisition: TCWN
Investigation: EB (lead), HH (supporting)
Methodology: EB (equal), HH (equal), TCWN (equal)
Project administration: TCWN
Resources: TCWN (lead), KBNSG (supporting)
Supervision: TCWN (lead), HH (supporting)
Validation: EB, HH, TCWN

XSL•FO
RenderX

Visualization: EB
Writing – original draft: EB (lead), HH (supporting), TCWN (supporting)
Writing – review & editing: EB (lead), HH (supporting), TCWN (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Coding scheme.
[DOCX File, 17 KB - rehab_v13i1e80136_app1.docx]

References

1. Chen P, Chen CC, Hreha K, Goedert KM, Barrett AM. Kessler Foundation Neglect Assessment Process uniquely measures spatial neglect during activities of daily living. Arch Phys Med Rehabil 2015 May;96(5):869-876.e1. [doi: 10.1016/j.apmr.2014.10.023] [Medline: 25461827]
2. Esposito E, Shekhtman G, Chen P. Prevalence of spatial neglect post-stroke: a systematic review. Ann Phys Rehabil Med 2021 Sep;64(5):101459. [doi: 10.1016/j.rehab.2020.10.010] [Medline: 33246185]
3. Nijboer TCW, Kollen BJ, Kwakkel G. The impact of recovery of visuo-spatial neglect on motor recovery of the upper paretic limb after stroke. PLoS One 2014;9(6):e100584. [doi: 10.1371/journal.pone.0100584] [Medline: 24950224]
4. Ten Brink AF, Biesbroek JM, Kuijf HJ, et al. The right hemisphere is dominant in organization of visual search-a study in stroke patients. Behav Brain Res 2016 May 1;304:71-79. [doi: 10.1016/j.bbr.2016.02.004] [Medline: 26876010]
5. Embrechts E, Loureiro-Chaves R, Nijboer TCW, Lafosse C, Truijen S, Saeys W. The association of personal neglect with motor, activities of daily living, and participation outcomes after stroke: a systematic review. Arch Clin Neuropsychol 2024 Feb 19;39(2):249-264. [doi: 10.1093/arclin/acad063] [Medline: 37591497]
6. Nijboer T, van de Port I, Schepers V, Post M, Visser-Meily A. Predicting functional outcome after stroke: the influence of neglect on basic activities in daily living. Front Hum Neurosci 2013;7:182. [doi: 10.3389/fnhum.2013.00182] [Medline: 23675336]
7. Hammerbeck U, Gittins M, Vail A, Paley L, Tyson SF, Bowen A. Spatial neglect in stroke: identification, disease process and association with outcome during inpatient rehabilitation. Brain Sci 2019 Dec 13;9(12):374. [doi: 10.3390/brainsci9120374] [Medline: 31847166]
8. Di Monaco M, Schintu S, Dotta M, Barba S, Tappero R, Gindri P. Severity of unilateral spatial neglect is an independent predictor of functional outcome after acute inpatient rehabilitation in individuals with right hemispheric stroke. Arch Phys Med Rehabil 2011 Aug;92(8):1250-1256. [doi: 10.1016/j.apmr.2011.03.018] [Medline: 21807144]
9. Bosma MS, Nijboer TCW, Caljouw MAA, Achterberg WP. Impact of visuospatial neglect post-stroke on daily activities, participation and informal caregiver burden: a systematic review. Ann Phys Rehabil Med 2020 Jul;63(4):344-358. [doi: 10.1016/j.rehab.2019.05.006] [Medline: 31200080]
10. Cappa SF, Benke T, Clarke S, et al. EFNS guidelines on cognitive rehabilitation: report of an EFNS task force. Eur J Neurol 2005 Sep;12(9):665-680. [doi: 10.1111/j.1468-1331.2005.01330.x] [Medline: 16128867]
11. Pizzamiglio L, Antonucci G, Judica A, Montenero P, Razzano C, Zoccolotti P. Cognitive rehabilitation of the hemineglect disorder in chronic patients with unilateral right brain damage. J Clin Exp Neuropsychol 1992 Nov;14(6):901-923. [doi: 10.1080/01688639208402543] [Medline: 1452637]
12. van Kessel ME, Geurts ACH, Brouwer WH, Fasotti L. Visual scanning training for neglect after stroke with and without a computerized lane tracking dual task. Front Hum Neurosci 2013;7:358. [doi: 10.3389/fnhum.2013.00358] [Medline: 23847519]
13. Priftis K, Passarini L, Pilosio C, Meneghello F, Pitteri M. Visual scanning training, limb activation treatment, and prism adaptation for rehabilitating left neglect: who is the winner? Front Hum Neurosci 2013;7:360. [doi: 10.3389/fnhum.2013.00360] [Medline: 23847520]
14. Chen P, Toglia J. Online and offline awareness deficits: anosognosia for spatial neglect. Rehabil Psychol 2019 Feb;64(1):50-64. [doi: 10.1037/rep0000207] [Medline: 29648845]
15. Tham K, Borell L, Gustavsson A. The discovery of disability: a phenomenological study of unilateral neglect. Am J Occup Ther 2000;54(4):398-406. [doi: 10.5014/ajot.54.4.398] [Medline: 10932310]
16. Tham K, Ginsburg E, Fisher AG, Tegnér R. Training to improve awareness of disabilities in clients with unilateral neglect. Am J Occup Ther 2001;55(1):46-54. [doi: 10.5014/ajot.55.1.46] [Medline: 11216366]
17. Maggio MG, Maresca G, De Luca R, et al. The growing use of virtual reality in cognitive rehabilitation: fact, fake or vision? A scoping review. J Natl Med Assoc 2019 Aug;111(4):457-463. [doi: 10.1016/j.jnma.2019.01.003] [Medline: 30739728]
18. Gamito P, Oliveira J, Coelho C, et al. Cognitive training on stroke patients via virtual reality-based serious games. Disabil Rehabil 2017 Feb;39(4):385-388. [doi: 10.3109/09638288.2014.934925] [Medline: 25739412]

XSL•FO
RenderX

19. Bohil CJ, Alicea B, Biocca FA. Virtual reality in neuroscience research and therapy. Nat Rev Neurosci 2011 Nov 3;12(12):752-762. [doi: 10.1038/nrn3122] [Medline: 22048061]
20. Alvarez J, Djaouti D. An introduction to serious game definitions and concepts. Presented at: Proceedings of the Serious Games & Simulation for Risks Management Workshop; Dec 5, 2011 URL: https://hal.science/hal-04675725 [accessed 2025-07-04]
21. Ding K, Ma Y, Zhang L, et al. Patient-centered insights into virtual reality rehabilitation for stroke: a systematic review and qualitative meta-synthesis. J Neuroeng Rehabil 2025 Jun 2;22(1):124. [doi: 10.1186/s12984-025-01641-9] [Medline: 40452007]
22. Chen J, Or CK, Li Z, Yeung EHK, Chen T. Perceptions of patients with stroke regarding an immersive virtual reality-based exercise system for upper limb rehabilitation: questionnaire and interview study. JMIR Serious Games 2025 Jan 1;13:e49847. [doi: 10.2196/49847] [Medline: 39742513]
23. Tobler-Ammann BC, Surer E, Knols RH, Borghese NA, de Bruin ED. User perspectives on exergames designed to explore the hemineglected space for stroke patients with visuospatial neglect: usability study. JMIR Serious Games 2017 Aug 25;5(3):e18. [doi: 10.2196/games.8013] [Medline: 28842390]
24. Howard MC, Davis MM. A meta-analysis and systematic literature review of mixed reality rehabilitation programs: Investigating design characteristics of augmented reality and augmented virtuality. Comput Human Behav 2022 May;130:107197. [doi: 10.1016/j.chb.2022.107197]
25. Pridmore T, Cobb S, Hilton D, Green J, Eastgate R. Mixed reality environments in stroke rehabilitation: interfaces across the real/virtual divide. Int J Disabil Hum Dev 2007;6(1). [doi: 10.1515/IJDHD.2007.6.1.87]
26. Huygelier H, Tuts N, Michiels K, et al. The efficacy and feasibility of an immersive virtual reality game to train spatial attention orientation after stroke: a stage 2 report. J Neuropsychol 2025 Jun;19(2):338-389. [doi: 10.1111/jnp.12403] [Medline: 39668467]
27. Palmans E, Tuts N, Michiels K, et al. Does immersive VR improve diagnostic accuracy of post-stroke spatial neglect relative to conventional and digital tests? medRxiv. Preprint posted online on Feb 25, 2025. [doi: 10.1101/2025.02.24.25322775]
28. Salatino A, Zavattaro C, Gammeri R, et al. Virtual reality rehabilitation for unilateral spatial neglect: a systematic review of immersive, semi-immersive and non-immersive techniques. Neurosci Biobehav Rev 2023 Sep;152:105248. [doi: 10.1016/j.neubiorev.2023.105248] [Medline: 37247829]
29. Motomura K, Amimoto K, Numao T, Kaneko F. Effects of a stimulus response task using virtual reality on unilateral spatial neglect: a randomized controlled trial. Arch Phys Med Rehabil 2024 Aug;105(8):1449-1457. [doi: 10.1016/j.apmr.2024.05.009] [Medline: 38750715]
30. Bakker MDJ, Boonstra N, Nijboer TCW, Holstege MS, Achterberg WP, Chavannes NH. The design choices for the development of an augmented reality game for people with visuospatial neglect. Clin eHealth 2020;3:82-88. [doi: 10.1016/j.ceh.2020.11.003]
31. Stammler B, Flammer K, Schuster T, Lambert M, Karnath HO. Negami: an augmented reality app for the treatment of spatial neglect after stroke. JMIR Serious Games 2023 Feb 27;11:e40651. [doi: 10.2196/40651] [Medline: 36848215]
32. Turgut N, Möller L, Dengler K, et al. Adaptive cueing treatment of neglect in stroke patients leads to improvements in activities of daily living: a randomized controlled, crossover trial. Neurorehabil Neural Repair 2018 Nov;32(11):988-998. [doi: 10.1177/1545968318807054] [Medline: 30328767]
33. Donaldson A, Finch CF. Planning for implementation and translation: seek first to understand the end-users' perspectives. Br J Sports Med 2012 Apr;46(5):306-307. [doi: 10.1136/bjsports-2011-090461] [Medline: 22006931]
34. Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696. [doi: 10.1136/bmj.321.7262.694] [Medline: 10987780]
35. Bousché E, Bakker MDJ, Holstege MS, et al. Virtual and augmented reality gamification of visuospatial neglect treatment: therapists' user experience. BMC Digit Health 2024;2(1). [doi: 10.1186/s44247-024-00061-4]
36. Edgeworth JA, Robertson IH, McMillan TM. The Balloons Test Manual: Thames Valley Test Company Limited; 1998.
37. Ten Brink A, Nijboer T, Van Beekum L, et al. De Nederlandse Catherine Bergego schaal: een bruikbaar en valide instrument in de CVA zorg [Article in Dutch]. Wetenschappelijk Tijdschrift Voor Ergotherapie 2013;6(3):27-35 [FREE Full text]
38. Huygelier H, Gillebert CR, van Ee R, Vanden Abeele V. The design of a virtual reality game for stroke-induced attention deficits. Presented at: CHI PLAY '17 Extended Abstracts: Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in PLAY; Oct 15-18, 2017. [doi: 10.1145/3130859.3131308]
39. Huygelier H, Schraepen B, Lafosse C, et al. An immersive virtual reality game to train spatial attention orientation after stroke: a feasibility study. Appl Neuropsychol Adult 2022;29(5):915-935. [doi: 10.1080/23279095.2020.1821030] [Medline: 32945702]
40. Brooke J. SUS: a "quick and dirty" usability scale. In: Usability Evaluation in Industry: CRC Press; 1996:189-194.
41. Demeurisse G, Demol O, Robaye E. Motor evaluation in vascular hemiplegia. Eur Neurol 1980;19(6):382-389. [doi: 10.1159/000115178] [Medline: 7439211]
42. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MOCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]

XSL•FO
RenderX

43. SAN Test, Een Afasietest Voor Auditief Taalbegrip En Mondeling Taalgebruik [Book in Dutch], 2nd edition: Swets & Zeitlinger; 2000. URL: https://research.rug.nl/en/publications/san-test-een-afasietest-voor-auditief-taalbegrip-en-mondeling-taa [accessed 2025-07-04]
44. Della Pietra GL, Savio K, Oddone E, Reggiani M, Monaco F, Leone MA. Validity and reliability of the Barthel index administered by telephone. Stroke 2011 Jul;42(7):2077-2079. [doi: 10.1161/STROKEAHA.111.613521] [Medline: 21527755]
45. Diesfeldt HFA. Aandacht voor visuele inattentie. De Balloons Test in de psychogeriatrie [Article in Dutch]. Tijdschrift voor Gerontologie en Geriatrie 2012 Feb;43(1):7-17. [doi: 10.1007/s12439-012-0002-4]
46. Boeije H, Bleijenbergh IL. Analyseren in Kwalitatief Onderzoek: Denken En Doen [Book in Dutch], 4th edition: Boom Uitgevers; 2019. URL: https://www.boom.nl/media/21/9789024425945_inkijkexemplaar_analyseren_in_kwalitatief_onderzoek.pdf [accessed 2026-03-23]
47. Winter C, Kern F, Gall D, Latoschik ME, Pauli P, Käthner I. Immersive virtual reality during gait rehabilitation increases walking speed and motivation: a usability evaluation with healthy participants and patients with multiple sclerosis and stroke. J Neuroeng Rehabil 2021 Apr 22;18(1):68. [doi: 10.1186/s12984-021-00848-w] [Medline: 33888148]
48. Apăvăloaiei RF, Achirei . Gamification of post-stroke neuromotor rehabilitation exercises using hand tracking in VR. Buletinul Inst Politeh Ia i – Electrotehnică, Energetică, Electronică 2023 Mar 1;69(1):23-45. [doi: 10.2478/bipie-2023-0002]
49. Sánchez-Gil JJ, Sáez-Manzano A, López-Luque R, Ochoa-Sepúlveda JJ, Cañete-Carmona E. Gamified devices for stroke rehabilitation: a systematic review. Comput Methods Programs Biomed 2025 Jan;258:108476. [doi: 10.1016/j.cmpb.2024.108476] [Medline: 39520875]
50. Morse H, Biggart L, Pomeroy V, Rossit S. Exploring perspectives from stroke survivors, carers and clinicians on virtual reality as a precursor to using telerehabilitation for spatial neglect post-stroke. Neuropsychol Rehabil 2022 Jun;32(5):707-731. [doi: 10.1080/09602011.2020.1819827] [Medline: 32942950]
51. Dai CY, Huang YH, Chou LW, Wu SC, Wang RY, Lin LC. Effects of primary caregiver participation in vestibular rehabilitation for unilateral neglect patients with right hemispheric stroke: a randomized controlled trial. Neuropsychiatr Dis Treat 2013;9:477-484. [doi: 10.2147/NDT.S42426] [Medline: 23630423]
52. Osawa A, Maeshima S. Family participation can improve unilateral spatial neglect in patients with acute right hemispheric stroke. Eur Neurol 2010;63(3):170-175. [doi: 10.1159/000286517] [Medline: 20185919]
53. Lundin RM, Yeap Y, Menkes DB. Adverse effects of virtual and augmented reality interventions in psychiatry: systematic review. JMIR Ment Health 2023 May 5;10:e43240. [doi: 10.2196/43240] [Medline: 37145841]
54. Smits DW, van Meeteren K, Klem M, Alsem M, Ketelaar M. Designing a tool to support patient and public involvement in research projects: the Involvement Matrix. Res Involv Engagem 2020;6:30. [doi: 10.1186/s40900-020-00188-4] [Medline: 32550002]
55. Danso A, Nijhuis P, Ansani A, et al. Development and user experiences of a novel virtual reality task for poststroke visuospatial neglect: an exploratory pilot study. arXiv. Preprint posted online on Dec 19, 2023. [doi: 10.48550/arxiv.2312.12399]

Abbreviations

AR: augmented reality
CBS: Catherine Bergego Scale
MI: Motricity Index
VR : virtual reality
VSN: visuospatial neglect
VST: visual scanning training
XR: extended reality

Edited by S Munce; submitted 04.Jul.2025; peer-reviewed by J Thrift, M Mondellini, N Turgut; revised version received 14.Jan.2026; accepted 03.Feb.2026; published 20.Apr.2026.

Please cite as:
Bousché E, Huygelier H, Nijboer TCW, Knowledge Broker Neglect Study Group
User Experience of Extended Reality Treatment for Visuospatial Neglect Among Patients and Informal Caregivers: Qualitative Interview Study
JMIR Rehabil Assist Technol 2026;13:e80136
URL: https://rehab.jmir.org/2026/1/e80136
doi:10.2196/80136

XSL•FO
RenderX

© Eileen Bousché, Hanne Huygelier, Tanja C W Nijboer, Knowledge Broker Neglect Study Group. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Platform to Provide Health Data Feedback for Neurorehabilitation Patients: User-Centered Development and Proof-of-Concept Usability Study

Nadine Domnik[1], MSc; Katarzyna Krasnopolska[1], MSc; Ramona Sylvester[2], MSc; Jens Bansi[2,3], PhD; Jaeyong Song[1], MSc; Roman Gonzenbach[2], MD, PhD; Olivier Lambercy[1,4], PhD

[1]Department of Health Science and Technology, Rehabilitation Engineering Laboratory, ETH Zurich, Gloriastrasse 37/39, Zurich, Switzerland

[2]Neurology Department, Kliniken Valens, Valens, Switzerland

[3]Department of Health, OST - Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland

[4]Future Health Technologies Programme, Singapore-ETH Centre, Singapore, Singapore

Corresponding Author:
Nadine Domnik, MSc
Department of Health Science and Technology, Rehabilitation Engineering Laboratory, ETH Zurich, Gloriastrasse 37/39, Zurich, Switzerland

Abstract

Background: An increasing amount of digital health data are being collected across rehabilitation settings, but their integration into routine clinical practice remains limited, despite its potential to motivate patients or inform clinical decision-making. Specifically, effective visualization and communication of assessment outcomes to both patients and health care practitioners (HCPs) represent a key gap in the neurorehabilitation practice.

Objective: This study describes the development and evaluation of RehaLink (author ND, ETH Zürich), a proof-of-concept mobile app that delivers structured, interpretable feedback from conventional and technology-based assessments to neurorehabilitation patients and HCPs.

Methods: The app was developed through a 3-step iterative co-design process involving 17 inpatients with multiple sclerosis and 15 HCPs from a single rehabilitation center. The app integrates a full battery of conventional assessments routinely conducted at the clinic, as well as digital health metrics from the Virtual Peg Insertion Test, a validated technology-based assessment of upper limb function, as a proof of concept for integrating technology-based assessment data into clinical workflows. Three structured feedback sessions were conducted, in which participants evaluated feedback types, visualization formats, and app usability using Likert-scale ratings, preference rankings, open-ended questions, and the System Usability Scale. Data were analyzed using descriptive statistics and directed content analysis.

Results: Across all 3 sessions, progress bars and color-coded indicators were consistently preferred over text-heavy or abstract formats by both patients and HCPs. A persistent set of competing demands was observed, with participants requesting both visual simplicity and access to absolute values and normative comparisons. HCPs tended to underestimate patients' preference for informative visualizations. The perceived value of structured feedback increased over the course of the study; patients' median ratings rose from 4.0 to 5.0 and HCPs' from 4.0 to 4.5 on a 5-point Likert scale. The resulting mobile app prototype demonstrated high usability, with patients achieving a mean System Usability Scale score of 93.6 (mean 6.4; best imaginable) and HCPs 80.9 (SD 8.1; good), according to established benchmarks.

Conclusions: These findings demonstrate the feasibility and value of a co-designed digital feedback tool for neurorehabilitation. By combining conventional and technology-based assessment outcomes in an accessible, user-centered format, the app has the potential to enhance patient engagement, support clinical decision-making, and advance the implementation of value-based, personalized care.

(JMIR Rehabil Assist Technol 2026;13:e85072) doi:10.2196/85072

KEYWORDS

neurorehabilitation; digital health; user-centered design; clinical decision-making; health data visualization; mobile health (mHealth); value based health care (VBHC); digital feedback application

XSL•FO
RenderX

Introduction

During neurorehabilitation, both conventional and technology-based assessments are essential for diagnosing impairments, tracking progress, and tailoring treatments [1-5]. Standardized tools such as the Nine Hole Peg Test [6], the Timed Up and Go [7], and the Functional Independence Measure [8] are widely used [9]. Although they capture different domains of the International Classification of Functioning, Disability, and Health framework [10], they are subjective, time-consuming, and lack sensitivity [11-13].

Technology-based assessments address these limitations through automated data collection, improved objectivity, and greater sensitivity. Examples include (markerless) motion capture [14-16], wearable sensors [17], and robotic tools such as the Kinesiological Instrument for Normal and Altered Reaching Movement (KINARM) [18,19] or the Virtual Peg Insertion Test (VPIT) [20,21]. However, their outputs can be abstract, poorly validated, and difficult to integrate into clinical workflows [22].

Despite their potential to support the implementation of precision medicine, integrating both conventional and digital assessments systematically into clinical routine decision-making remains challenging. In many cases, assessment outcomes are difficult to synthesize and communicate effectively during patient consultations, which can limit their ability to enhance patient motivation, satisfaction, and adherence to treatment. This limited adoption is largely driven by the high effort required to collect, summarize, and interpret the data, which can outweigh the perceived benefit in many clinical contexts. Moreover, while the level of detail offered by technology-based assessments is valuable in research [23-25], its clinical relevance remains uncertain, and there is a lack of clear guidelines for selecting or applying metrics in practice [26].

Regular and well-communicated feedback on assessment outcomes has been shown to enhance patient motivation [27,28], support engagement in shared decision-making [29], and improve rehabilitation outcomes [30,31]. Likewise, transparent and interpretable data can support clinical decisions that are often otherwise driven by prior experience and/or subjective judgment [32]. To leverage the full potential of both conventional and technology-based assessments, their results must be accessible, interpretable, and actionable for both patients and health care practitioners (HCPs). Achieving this requires a systematic and collaborative approach to identify clinically meaningful outcome measures and establish health communication strategies that enhance their practical use in clinical routine [5,33,34].

Digital apps have the potential to bridge the gap between researchers, HCPs, and patients by enabling access to individualized health data and translating complex information into intuitive, user-friendly formats. By streamlining data collection and interpretation, these tools have the potential to support more frequent assessments and promote continuous communication of outcomes throughout the rehabilitation process. In recent years, the number of released health apps in areas such as mental health [35,36], consumer health [37], and fitness [38,39] increased rapidly. While many of these apps lack sufficient validation or fail to demonstrate clear benefits, others show strong potential to increase patient motivation and adherence [35,36,38,40]. A key factor in making digital health tools effective and engaging for both patients and HCPs is the active involvement of end users in the development process [41-46]. Applying user-centered design principles helps ensure that these tools are better aligned with the real needs of all stakeholders [3-5,40,41,46-50]. However, a key limitation is that many existing digital health apps are developed primarily for the consumer market rather than for integration into clinical workflows. As a result, their design often overlooks the specific requirements of clinical environments, limiting their adoption and utility in routine care. This gap is particularly evident in neurorehabilitation, where existing digital health record systems still lack effective features for summarizing and visualizing health-related data. Consequently, both HCPs and patients may lack access to user-friendly tools to monitor and understand rehabilitation progress.

This study describes a 3-step user-centered design process for developing a digital feedback app that displays digital health metrics from conventional and technology-based assessments to patients and HCPs along the neurorehabilitation journey. As the first step, we defined the need and requirements and evaluated expectations for such feedback together with HCPs. Through iterative rounds of discussions and structured interviews with both neurological patients and HCPs, we identified an optimal type and amount of information, along with the appropriate visualizations to display health data in an intuitive and engaging way. Our evaluation focused on 3 key aspects, including informativeness, understandability, and visual appeal. Building on this, we developed RehaLink (author ND, ETH Zürich), a proof-of-concept app that visualizes typical rehabilitation data collected during clinical routine assessments, including digital health metrics from the VPIT, and evaluated its usability. This work provides insights from end users—patients and HCPs—on how to meaningfully present assessment outcomes and include digital health metrics in neurorehabilitation settings that will open new avenues for empowering patients with their health data and support HCPs in decision-making towards precision medicine.

Methods

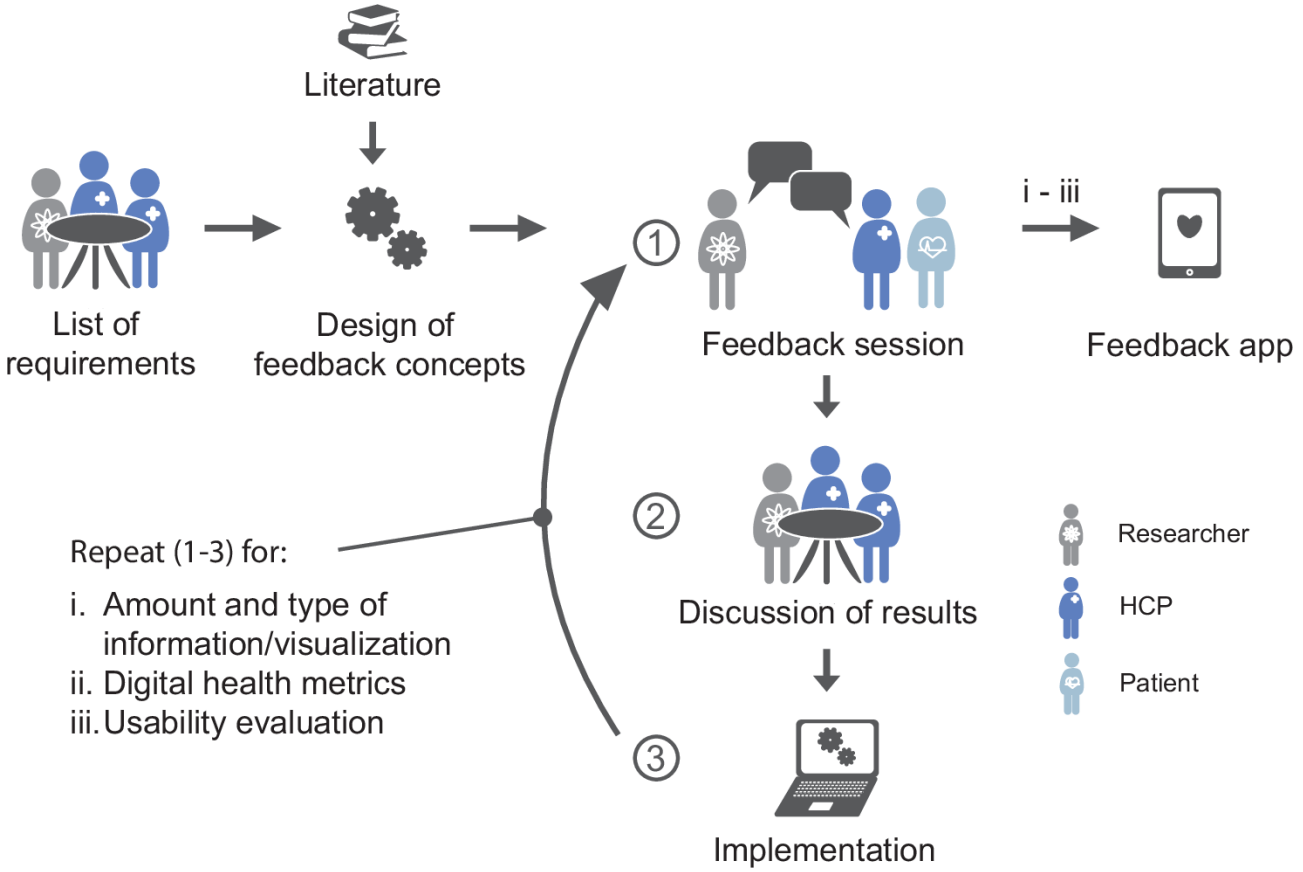
Conceptual Approach

We used a structured co-design process to develop a digital app aimed at enhancing patient motivation and supporting clinical decision-making in neurorehabilitation. The initial requirements were identified through a multidisciplinary roundtable involving researchers from ETH Zurich and clinical staff from Kliniken Valens, including the heads of neurology, research, and therapy, as well as therapists with different backgrounds. The discussions focused on defining clinically relevant assessments, selecting the appropriate platform, and ensuring integration into routine care. The development process included three iterative feedback sessions with HCPs, designed to evaluate (1) the content and visualization of patient information, (2) the integration of digital health metrics, and (3) overall usability. Between sessions, targeted follow-up meetings with researchers and HCPs enabled

XSL•FO
RenderX

systematic incorporation of feedback into subsequent development stages (Figure 1).

Figure 1. Overview of the co-design process for a digital feedback app for patients and health care practitioners (HCPs). The illustration shows a 3-step iterative approach guided by co-design principles. HCP: health care practitioner.



To ensure clinical relevance and align with clinical practice, the app was designed to include the full battery of conventional assessments routinely conducted at admission and discharge at the Kliniken Valens. To explore integration of digital health data, we included the VPIT, a well-validated tool for assessing upper limb function. The metrics derived from the VPIT are representative of the digital health data typically obtained from wearable sensors or robotic devices used in neurorehabilitation assessments. We implemented the app as a mobile solution for iOS, with an optional future compatibility across additional operating systems. At the time of this writing, the app is intended for use on tablets, as this offers greater practicality during patient consultations compared to stationary terminals.

Participant Recruitment

Patients were recruited through purposive sampling from the inpatient neurology unit at Kliniken Valens. Although the app is intended for use across neurorehabilitation populations, patients living with multiple sclerosis (MS) were selected as the initial target group for this proof-of-concept study given the clinic's large inpatient living with MS rehabilitation program that provides a readily accessible and well-characterized patient pool for this co-design process. Recruitment targeted eligible participants advised to inpatient rehabilitation with a diagnosis of MS and sufficient cognitive and communicative capacity to participate in structured feedback sessions. Eligible patients were identified by the treating clinical team through review of the inpatient list, selecting those with an MS diagnosis and sufficient cognitive and communicative capacity as judged by the attending clinicians. Participants provided informed consent under the institutional general consent framework. No formal inclusion or exclusion criteria were predefined, consistent with the exploratory, co-design nature of this work. Interested patients were approached by clinical staff and invited to participate voluntarily. HCPs were recruited through announcements during regular team meetings within the therapy and neurology departments of the Kliniken Valens.

The number of participants per session was determined pragmatically, as the primary aim of the feedback sessions was an iterative design refinement within a co-design process, rather than achieving qualitative data saturation in the traditional sense. Session recruitment concluded when the research team judged that sufficient stakeholder input had been obtained to inform the next design iteration, consistent with descriptions of co-design as an iterative, resource-bounded process [51-53]. Session sizes were consistent with recommendations for iterative, design-oriented evaluation, in which small samples have been shown to yield sufficient feedback for informing successive design iterations [54,55]. Nonetheless, qualitative responses in later sessions showed increasing thematic overlap with earlier feedback, suggesting that the key design-relevant perspectives had been captured within the sample.

XSL•FO
RenderX

Conventional Assessments

The conventional assessments at Kliniken Valens were grouped into 6 categories, including endurance, mobility, strength, coordination, speed, and global health/function to support clearer interpretation in the app (Table S1 in Multimedia Appendix 1). These categories were based on the International Classification of Functioning, Disability, and Health framework and refined with HCP input.

VPIT

The VPIT is a technology-aided assessment of upper limb function that integrates a haptic device with a virtual environment [20]. The task involves goal-directed reaching, in which users are instructed to insert nine virtual pegs into 9 holes displayed on a board as quickly and precisely as possible. During task execution, kinetic and kinematic data are continuously recorded. From this data, 10 digital health metrics are computed to quantify various aspects of movement quality, including smoothness, efficiency, accuracy, speed, and grip force control [21]. These metrics enable sensitive and objective characterization of upper limb impairments and have been validated in diverse patient populations, including individuals with MS [56,57], stroke [58], and ataxia [59,60]. Importantly, the VPIT served as a model for integrating digital health metrics into clinical feedback systems, offering a template for incorporating validated, objective assessments into user-facing tools. As such, it offers a clinically relevant and generalizable use case for exploring how complex sensor-based outputs can be effectively visualized and communicated to both patients and HCPs. Its integration into the co-designed digital feedback app supports our broader goal of making digital health data more accessible and actionable, fostering patient engagement and shared decision-making in neurorehabilitation. In collaboration with HCPs from Kliniken Valens, the VPIT metrics were grouped into the same 5 functional categories as the conventional assessments. An overview of the 10 core VPIT metrics, including their descriptions and associated categories, is provided in Table S2 in Multimedia Appendix 1.

Structured Feedback Sessions

Overview

Three structured feedback sessions were conducted to iteratively evaluate key aspects of the app design, including information presentation, metric representation, and overall usability. Each session followed a standardized procedure lasting approximately 30-45 minutes, with 2 participants—either HCPs or patients—participating simultaneously. To minimize bias, all participants first received a neutral written introduction outlining the session's purpose. In sessions 2 and 3, this was supplemented by a short video and an explanation of the VPIT.

Participants were then shown a series of feedback concepts and completed the structured questionnaires in Multimedia Appendices 2-7. These included Likert-scale items (1=strongly disagree-5=strongly agree) [61], open-ended questions, and preference rankings of the concepts. Concepts were rated on motivational value, visual appeal, informativeness, and understandability. Additional items assessed the perceived usefulness of performance feedback—either for enhancing therapy motivation (patients) or supporting clinical decision-making (HCPs)—as well as attitudes toward technology-based metrics, using the VPIT as an example. Participants were encouraged to ask both technical and content-related questions throughout, ensuring a full understanding of the material.

After completing the questionnaire, participants could share additional thoughts in a brief open discussion. All reflections were manually transcribed, and quotes included in this manuscript were translated from German into English.

The feedback concepts were developed according to established user-centered design principles [41,45,48], balancing clarity and representativeness while minimizing bias during evaluation.

Feedback Session 1: Type and Amount of Information

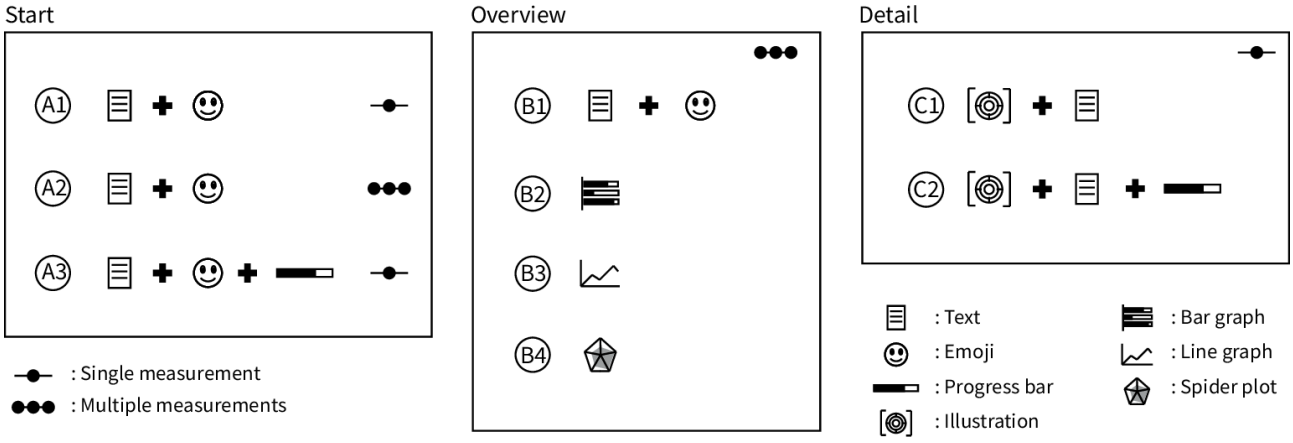
The first structured feedback session focused on evaluating how different types and levels of information, as well as their visual presentation, influenced user perceptions of usefulness, clarity, and engagement. To reflect a realistic interaction flow, we structured the session around three conceptual screens: (1) a start screen summarizing overall therapy progress, (2) an assessment overview screen displaying all assessment outcomes, and (3) an assessment detail screen providing more in-depth information on a selected assessment outcome. The 3-level screen structure was designed to balance conciseness and depth, allowing users to quickly access key insights or explore detailed results as needed. Limiting the structure to 3 screens aimed to maintain a shallow navigation hierarchy, thereby enhancing usability in rehabilitation contexts, where cognitive and physical demands on users may be heightened [62].

For each screen, participants were shown a set of alternative feedback concepts that varied in content complexity and visual format. These included combinations of textual summaries, color-coded timelines, bar and line charts, spider plots, emojis, and progress bars. A text- and emoji-only version was also included in the first set to serve as a low-complexity baseline. These types of visualizations were selected for their common use in digital health or consumer apps, with bar and line graphs being the most prevalent in digital health contexts [63,64].

An overview of the 3 screens and their respective feedback concepts is presented in Figure 2. This layered approach was designed to assess preferences across varying levels of information granularity and visual representation.

XSL•FO
RenderX

Figure 2. Illustration of feedback concepts shown during the first structured feedback session. The figure presents the 3 conceptual screens used to evaluate feedback design: the start screen, which provides overall performance summaries; the metric overview screen, which compares visualization types across multiple assessments; and the assessment detail screen, which presents in-depth explanations of single assessment outcomes. Each screen included multiple concepts (labeled 1-4) that varied in complexity, visual format, and content presentation.



In addition to the concept-specific ratings, patients were asked whether feedback on assessments could enhance their motivation for therapy and which types of feedback they would prefer to receive. HCPs were asked whether feedback could support clinical decision-making and help motivate patients. Further, HCPs were asked if they think that the VPIT could be helpful for them to evaluate the upper limb function of the patients. The questions of whether patients think that structured feedback could enhance their motivation and whether HCPs think that it could support decision-making were asked before providing feedback concepts and again afterward, so that we get an unbiased opinion first. Both groups were also invited to indicate the kinds of technology-based assessment feedback they considered most meaningful.

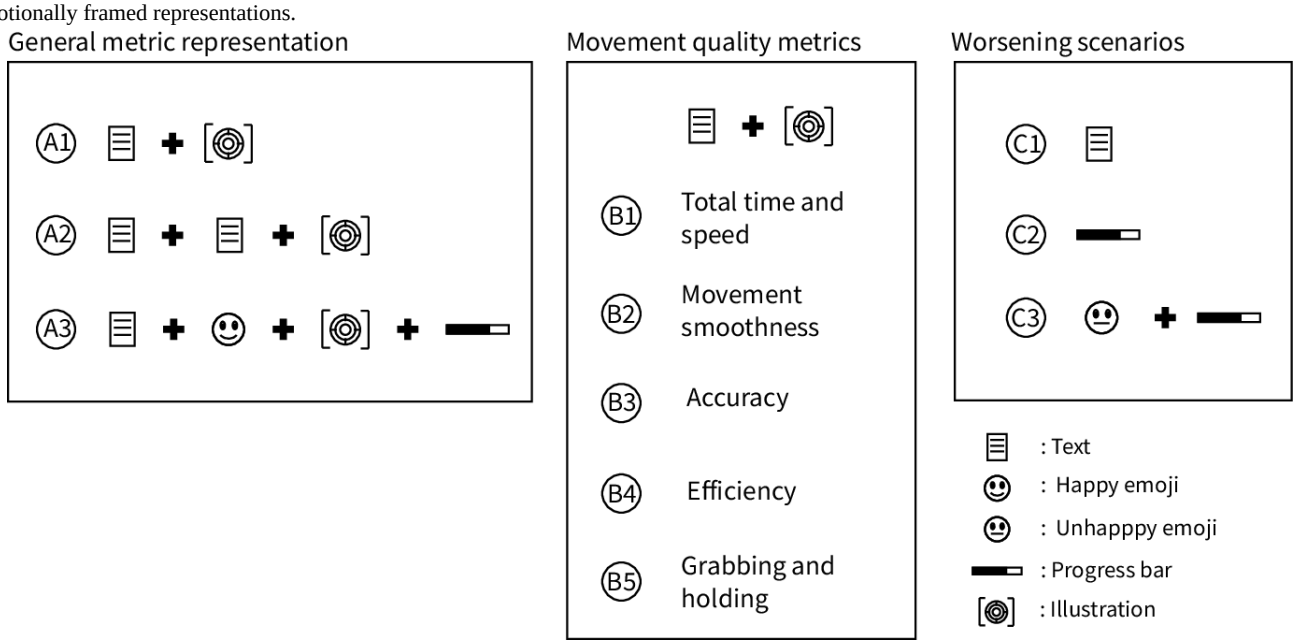
Feedback Session 2: Presentation of Digital Health Metrics

The second structured feedback session focused on how digital health metrics could be effectively communicated to patients and HCPs through intuitive, engaging visual designs. To enhance understandability, the 10 VPIT metrics (Table S2 in Multimedia Appendix 1) along with overall task completion time were grouped into 5 lay-language metrics, including total time and speed, movement smoothness, accuracy, efficiency, and grabbing and holding. These categories were derived from the original research-based classifications and were adapted in collaboration with HCPs to ensure that terminology was intuitive and meaningful for both patients and HCPs. To support aggregation of metrics with different units and easier interpretation by nonexpert users, each VPIT metric was first normalized to a 0 - 100 scale, where 0 represents the most impaired neurological subject in our database (excluding outliers), and 100 corresponds to the median performance of age-matched healthy controls. Without normalization, the raw values are too abstract and/or technical to convey meaningful feedback to patients or HCPs. The normalized metrics within each category were then averaged, without weighting, to create composite scores representing broader functional domains (eg, dimensionless log jerk and spectral arc length were combined under movement smoothness). This approach preserved metric validity while enhancing interpretability for end users. Details of the metric groupings are provided in the description column of Table S2 in Multimedia Appendix 1.

Participants were presented with three sets of feedback concepts (Figure 3), each designed to explore a different aspect of digital health metric feedback delivery.

XSL•FO
RenderX

Figure 3. Overview of feedback concepts presented in the second structured feedback session. The figure summarizes 3 sets of visual feedback concepts shown to participants: Set 1 presents different designs for a single digital health metric (movement smoothness); Set 2 shows all 5 aggregated metrics with metric names, icons, visual illustrations, and explanations; and Set 3 explores how to communicate worsening outcomes, ranging from neutral to emotionally framed representations.



Set 1 explored the general presentation of a single metric, using movement smoothness as an example. The 3 concepts varied in their level of explanation and visual complexity. Each included the metric name, an icon, and different combinations of visual elements such as color-coded bars (green for healthy and red for impaired), progress bars, emojis, and explanatory text.

Set 2 provided an overview of all 5 aggregated VPIT metrics. Each metric was shown with its name, icon, short descriptive text, and a corresponding illustration. This set aimed to assess whether the metrics were perceived as clear, useful, and engaging.

Set 3 addressed how to communicate worsening outcomes and included 3 visualizations with varying emotional tone and visual emphasis. The first concept displayed a neutral message ("No improvement") without emojis or colored bars. The second used a green progress bar to indicate decline but excluded emojis and text. The third showed worsening results with an orange progress bar and an unhappy orange emoji. Each concept included an information button linking to an HCP-validated disclaimer explaining that performance changes may reflect normal day-to-day variation and should not be interpreted as clinically significant.

Participants rated each concept based on visual appeal, informativeness, and understandability. Additional open-ended questions captured qualitative impressions. Patients and HCPs were also asked whether they found the disclaimer clear and reassuring.

Feedback Session 3: Usability Evaluation of the App Prototype

In the third feedback session, participants evaluated the usability of a high-fidelity app prototype created using Figma (Figma, Inc) [65]. The prototype featured randomly generated data and allowed participants to freely explore its interactive components, including screen navigation and button functionalities. The use of randomly generated data for prototype usability testing is consistent with standard iterative user-centered design practice [41,42], in which interface usability is established prior to clinical validation. Further, instead of just illustrations for the VPIT metrics, we included animations.

Participants were instructed to interact with the app independently. HCPs were asked to imagine they were preparing for a consultation and tasked with reviewing a patient's most recent assessment results, identifying any worsening trends, and checking the summary screen. Patients were instructed to imagine they had just completed a new assessment and were asked to explore their feedback, compare it with previous results, and interpret the meaning of their scores. Participants were encouraged to ask questions during the interaction if clarification was needed.

Following the exploration phase, all participants completed the System Usability Scale (SUS) [66] to evaluate overall usability. In addition, they responded to several targeted questions using 5-point Likert scales and open-ended questions. These assessed whether participants perceived structured feedback as valuable and if there is something to improve in the app design and implementation. Items also addressed the desire for comparative feedback (eg, performance relative to healthy individuals or other patients) and whether participants would like to celebrate personal milestones. Participants were also asked to suggest changes or propose features to add or remove.

Data Analysis

Quantitative Analysis

Likert-scale responses were collected on paper and subsequently transferred to Python (version 3.13.2; Python Software Foundation) for analysis. No data exclusion or cleaning criteria

XSL•FO
RenderX

were applied. For each item, the median and IQR were calculated separately for patients and HCPs. Median and IQR were chosen due to the ordinal nature of the data and their robustness to outliers, particularly in the context of small sample sizes. Accordingly, no inferential statistical testing was performed. Participants also ranked the feedback concepts within each concept set by assigning a unique rank to each concept, with no ties permitted. Median ranks were computed to identify overall user preferences for different content types and visual formats. Visualizations such as heatmaps and summary tables were created in Python to support interpretation.

Usability Assessment

SUS questionnaires were collected on paper and subsequently transferred to Python (version 3.13.2) for analysis. Raw SUS responses were converted into standardized scores (0 - 100) following the standard scoring procedure [66]. Group-level means and standard deviations were calculated separately for patients and HCPs. Scores were interpreted using the adjective rating scale proposed by Bangor et al [67], ranging from "worst imaginable" (<50) to "best imaginable" (>90), to contextualize the usability findings.

Qualitative Analysis

Open-ended responses from all 3 feedback sessions were analyzed using direct content analysis. All responses were collected and analyzed in German. A single coder conducted multiple passes through the data, an initial pass to assign open codes to individual responses, followed by subsequent passes to group related codes into higher-level categories and identify recurring themes. An initial codebook was developed deductively based on the predefined questionnaire dimensions (eg, visual appeal, informativeness, understandability, and motivational value). Open-ended responses were then coded iteratively, with additional codes emerging inductively from the data and grouped into higher-level categories as recurring patterns were identified across participants. Coded themes were subsequently reviewed and discussed with all co-authors to ensure their plausibility and consistency with the broader dataset. Representative quotes were selected to illustrate key findings and translated from German into English by the first author (ND) for inclusion in the manuscript.

Ethical Considerations

All participants were informed about the use of their data for scientific purposes and provided consent for further use of their data at entry to the Valens clinic. The responsibilities for this analysis were obtained by the Regional Ethics Committee of Eastern Switzerland (EKOS; EKOS 25/209). All procedures were conducted in accordance with the Declaration of Helsinki and the Swiss Human Research Act. All data collected were fully anonymized prior to analysis, and no personally identifiable information is reported in this manuscript. Participants received no financial compensation for their involvement.

Results

Participant Overview

A total of 17 neurological inpatients diagnosed with MS and 15 distinct HCPs from the Kliniken Valens participated in one of 3 structured feedback sessions. Several HCPs participated in more than one session, resulting in a total of 26 HCP participations. An overview of participant demographics and roles is provided in Table 1.

XSL•FO
RenderX

Table . Participant demographics across the 3 structured feedback sessions. The table summarizes the number of participants, their roles, and basic demographic characteristics for both patients and health care practitioners (HCPs) involved in each of the 3 structured feedback sessions.

Session and group	Participants, n	Age (years), mean (SD)	Gender (n)'	Role
Session 1				
Patients	5	40.8 (15.6)	• Female: 2 • Male: 3	—[a]
HCPs[b]	8	35.8 (10.4)	• Female: 5 • Male: 3	PT[c] (n=3), ST[d] (n=2), OT[e] (n=2), HST[f] (n=1)
Session 2				
Patients	6	56.5 (7.8)	• Female: 1 • Male: 5	—
HCPs	9	35.2 (9.3)	• Female: 3 • Male: 6	PT (n=5), ST (n=2), OT (n=2)
Session 3				
Patients	6	46.1 (14.8)	• Female: 5 • Male: 1	—
HCPs	8	33.1 (10.3)	• Female: 1 • Male: 7	PT (n=3), ST (n=2), OT (n=2), HST (n=1)

[a]Not available.

[b]HCP: health care practitioner.

[c]PT: physiotherapist.

[d]ST: sports therapist/scientist.

[e]OT: occupational therapist.

[f]HST: Health Science and Technology degree.

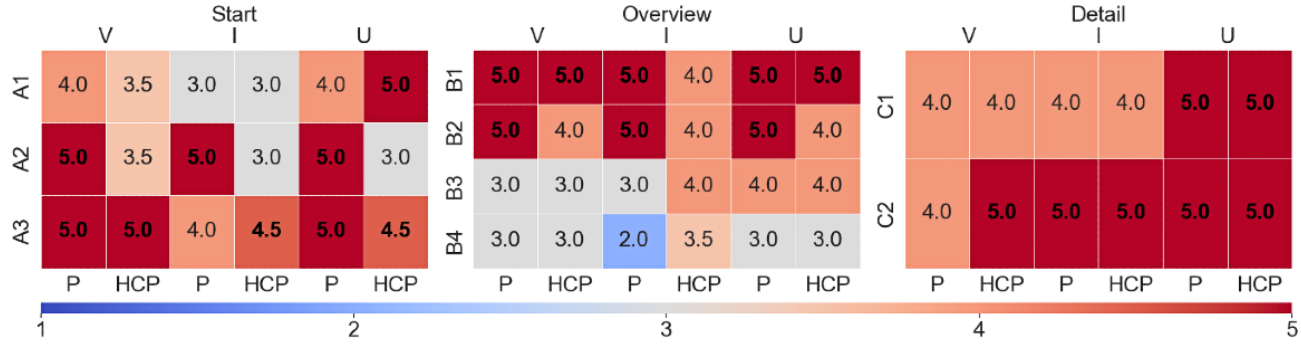
Feedback Session 1: Type and Amount of Information

Overview

The primary aim of this feedback session was to assess the preferred type and quantity of information to be presented to patients and HCPs within a digital feedback app. To reduce potential bias, participants were asked general questions about structured feedback both prior (pre) and then following their exposure to the different feedback concepts (post). The feedback concepts included various mockup screens designed to give participants an idea of how feedback might be presented; for example, through different data visualization formats such as bar charts, line graphs, spider plots, text, or even emojis. These concepts were created for a range of topics. In addition to presenting and evaluating feedback concepts, HCPs were asked whether feedback on completed assessments could aid clinical decision-making (pre: median 4.0, IQR 3.5 - 4.0; post: median 4.0, IQR 4.0 - 5.0), while patients were asked whether such feedback might increase their motivation to engage more consistently in therapy (pre: median 4.0, IQR 4.0 - 5.0; post: median 4.0, IQR 4.0 - 5.0). Median responses therefore remained consistent before and after viewing the feedback concepts. Additionally, HCPs were asked whether providing feedback to patients could enhance their motivation (median 4.5, IQR 4.0 - 5.0) and whether digital health metrics such as VPIT-derived feedback would be useful in evaluating upper limb function (median 4.0, IQR 4.0 - 5.0). Participants evaluated 9 feedback concepts across 3 screen types: the start screen, which summarized overall performance; the overview screen, which presented multiple assessment outcomes; and the detail screen, which displayed results for a single metric. All concepts were rated on visual appeal, informativeness, and understandability using Likert scales (Figure 4; Table S3 in Multimedia Appendix 1) and were ranked by preference.

XSL•FO
RenderX

Figure 4. Median Likert-scale ratings for feedback concepts in feedback session 1. Participants rated 9 feedback concepts on 3 dimensions, visual appeal (V), informativeness (I), and understandability (U), using a 5-point Likert scale. Concepts are grouped by screen: start screen, overview screen, and metric detail screen. For the start screen and the detail screen, concepts are ordered by increasing informational complexity (A1-A3; C1-C2); for the overview screen, by increasing visual complexity (B1-B4). Ratings are shown separately for patients (P) and health care practitioners (HCP). Color intensity reflects the median score (red = favorable and blue = less favorable). Numbers in bold mark ratings above 4.5. I: informativeness; HCP: health care practitioner; P: patient; U: understandability; V: visual appeal.



Start Screen

The concepts evolved from a basic emoji-based design (concept A1: text+emoji) to a timeline-enhanced version (concept A2) and finally to a version featuring a progress bar (concept A3). Both Likert scale ratings and rankings showed that concept A3 (progress bar) received the highest overall scores. Patients ranked concepts A3 and A2 (timeline) the highest (median 2.0, IQR 1.0 - 2.5 and median 2.0, IQR 2.0 - 3.0, respectively), while HCPs rated concept A3 as their top choice (median 1.0, IQR 1.0 - 1.5). Concept A3 also scored highest among both groups in terms of visual appeal, informativeness, and understandability. One patient remarked, “The progress bar makes it clear at a glance if I got better or worse,” and an HCP commented, “It is easily understandable for patients.”

Overview Screen

Concepts in this screen increased in visual complexity from text+emoji (B1) to bar plots (B2), line plots (B3), and spider plots (B4). HCPs ranked concept B3 slightly higher (median 2.0, IQR 1.0 - 2.5), while patients gave similar rankings to concepts B2 (median 2.0, IQR 2.0 - 2.0) and B3 (median 2.0, IQR 2.0 - 4.0). Both groups ranked concept B4 the lowest. Several participants commented on the need for absolute values, stating, “We want to see how much improvement or worsening there is in absolute values or percentages.”

> *If patients see norm values, they can compare what is good and where they need to improve. The same for us, we can then better compare where the patient is standing.* [A health care practitioner (HCP)]

Detail Screen

Concept C2 (illustration+ text+progress bar) was ranked highest by both patients and HCPs (median 1.0, IQR 1.0 - 1.0 and median 1.0, IQR 1.0 - 2.0, respectively). Likert scores similarly favored this concept across all 3 dimensions. Participants made several comments on workflow, such as, “During clinical routine we don’t have time to study results, we want to directly be able to extract relevant information” and “Too many screens make the system not usable, it is too much effort.”

Open Feedback Themes

Participants provided additional comments regarding feedback preferences. Several HCPs mentioned the importance of wording; for example, “The language you use is important, it can be motivating for patients or not.” Concerns were raised about potential negative feedback effects; for example, “What about if their results get worse? Then it is not motivating anymore.” Patients commented on the importance of defining goals; for example, “What is my goal? I want it defined.” HCPs also suggested linking feedback to therapeutic recommendations, such as, “Add which therapy we can do to improve the specific results.” When asked what type of feedback they would like to receive about digital health metrics (eg, VPIT), HCPs emphasized graphical visualizations and clinically relevant parameters such as endurance, time, movement precision (eg, smoothness), and daily-life applicability. Several HCPs requested simple graphical trends, such as line charts, for tracking changes. Patients primarily focused on viewing their own progress over time, especially regarding speed and differences between admission and posttherapy sessions.

Feedback Session 2: Presentation of Digital Health Metrics

Overview

The primary goal of this feedback session was to assess which digital health metrics patients and HCPs consider valuable and how these metrics could be visualized in an intuitive and engaging manner. In the second feedback session, participants evaluated 3 sets of feedback concepts, namely Set 1 focused on how to generally present digital health metrics using the example of movement smoothness; Set 2 focused on which metrics to present based on the condensed VPIT metrics; and Set 3 addressed how to communicate worsening of assessment outcomes. Each concept in Sets 1 and 2 was rated on visual appeal, informativeness, and understandability; each concept in Set 3 was rated on visual appeal and motivational aspects (Figure 5; Table S4 in Multimedia Appendix 1), while concepts in Sets 1 and 3 were additionally ranked by preference (Multimedia Appendix 1).

XSL•FO
RenderX

Figure 5. Median Likert-scale ratings for feedback on all concepts in feedback session 2. Participants rated 11 feedback concepts (A1-A3; B1-B5; C1-C3) from 3 sets: Set 1 (presentation of a single digital health metric), Set 2 (overview of all 5 grouped metrics), and Set 3 (how to present worsening results). Each concept was evaluated on visual appeal (V), each concept for Set 1 and 2 on informativeness (I), understandability (U), and on motivational (M) for Set 3 using a 5-point Likert scale. Ratings are shown separately for patients (P) and health care practitioners (HCP). Color intensity reflects the median score (red = favorable and blue = less favorable). Numbers in bold mark ratings above 4.5. I: informativeness; HCP: health care practitioner; M: motivational; P: patient; U: understandability; V: visual appeal.



Set 1: General Metric Representation

Patients assigned the highest median rank to concept A2 (extended text+illustration; median 2.0, IQR 1.5 - 2.0) and A3 (progress bar; median 2.0, IQR 1.0 - 3.0), while HCPs most frequently ranked concept A3 (progress bar) as their top choice (median 1.0, IQR 1.0 - 2.0). Among HCPs, concept A3 also received the highest Likert-scale ratings across visual appeal, informativeness, and understandability. Patients rated concepts A2 and A3 similarly on the understandability and informativeness dimensions. Qualitative feedback indicated varied preferences. One HCP described concept A2 as containing "too much information," whereas concept A3 was described as "short, concise, and to the point." Four additional HCPs commented that concept A3 was "simple, but understandable and very clear" and noted that it was "very well suited for patients." Two HCPs suggested a combination of concepts A2 and A3. One patient commented that concept B2 included "too much text," while concept A3 was "very well understandable." Two other patients similarly stated that concept A3 was "very clear."

Set 2: Digital Health Metrics

Participants evaluated 5 grouped digital health metrics, including time and speed (B1), smoothness (B2), accuracy (B3), efficiency (B4), and grabbing and holding (B5). Across all evaluated dimensions, patients generally provided higher Likert-scale ratings compared to HCPs. For visual appeal, accuracy (B3) and efficiency (B4) received the highest median ratings from both groups, with scores of 4.5 among patients and 4.0 among HCPs. In contrast, time and speed (B1) was rated lower, particularly by HCPs, with median scores of 4.0 for patients and 3.0 for HCPs. Regarding informativeness, efficiency (B4) and grabbing and holding (B5) were rated most favorably by patients, each achieving median scores of 5.0. HCPs rated all metrics slightly lower, with median scores around 4.0 and a lower rating for grabbing and holding (approximately 3.5). In terms of understandability, time and speed (B1), efficiency (B4), and grabbing and holding (B5) received the highest median scores among patients (5.0), while these metrics showed the lowest median scores among HCPs (3.0). Participant comments were specific to individual visual elements, with several suggestions for refinement. For example, one HCP proposed adding a tortoise and rabbit to the speed visualization to increase clarity.

Set 3: Worsening Scenarios

Participants evaluated 3 concepts for communicating worsening outcomes: C1 (text only), C2 (progress bar), and C3 (emoji+progress bar). Among patients, C2 received the highest median ranking (median 1.5, IQR 1.0 - 2.0), while HCPs preferred C3 (median 2.0, IQR 1.0 - 2.0). C1 was ranked lowest by both groups (median 3.0, IQR 2.0 - 3.0). Likert-scale ratings showed that concepts with visual elements (C2 and C3) were generally rated higher in visual appeal, informativeness, and understandability compared to the text-only C1. Across most dimensions, patients gave higher ratings than HCPs. Participant comments added context; HCPs described C3 as a "good visualization," appreciating the emphasis on positive aspects, though some felt the worsening was shown too strongly. Several suggested revising the emoji design, eg, using "neutral-colored, smiling emojis." For C1, HCPs stressed the need for quantitative data, "always show values," and "we always want to see exact values." Patients also preferred visuals. For C3, they liked the color-coded improvements but recommended neutral smiles. One noted, "too much indicated the negative parts." Regarding C1, a patient said, "I don't want text, it is too negative. I prefer emojis and colors." Participants also rated a text explanation stating that worsening results may reflect normal daily fluctuations. Understandability was rated at a median of 4.0 (IQR 3.0 - 4.5) by HCPs and a median of 4.5 (IQR 3.5 - 5.0) by patients; the median motivational value was 4.0 (IQR 3.0 - 4.0) and 4.5 (IQR 4.0 - 5.0), respectively. HCPs suggested making the message more encouraging and replacing "measurement variability" with relatable terms like environmental or daily performance factors. Patients asked for simpler language and generally found the text helpful. General feedback about the app development was also positive. One HCP noted, "The interface is a very nice opportunity for everyone; I am glad to be part of the optimizing process," while

XSL•FO
RenderX

a patient commented, “exciting project with a lot of informative insights into the therapy progress.”

Feedback Session 3: Usability Evaluation of the App Prototype

The main goal of this feedback session was to evaluate the usability of the first functional app prototype. During this final structured feedback session, participants interacted with the fully functional app prototype, which featured short animations in place of static illustrations to visualize digital health metrics. Usability was evaluated using the SUS. Patients reported a mean SUS score of 93.6 (SD 6.4), corresponding to “best imaginable” usability, while HCPs reported a mean score of 80.9 (SD 8.1), corresponding to “good” usability, according to established benchmarks [67] (Figure 6). Participants were also invited to suggest improvements to the prototype. One HCP requested the inclusion of absolute values alongside percentages in performance bars: “I would like to see an absolute number with units in every bar, not only the percentage indicating improvement or decline.” Another HCP highlighted the clarity of the interface, stating:

> *The implementation of the user interface is really nice. It shows just the right amount of key information, it’s super easy to get a good overview.* [An HCP]

Further suggestions included:

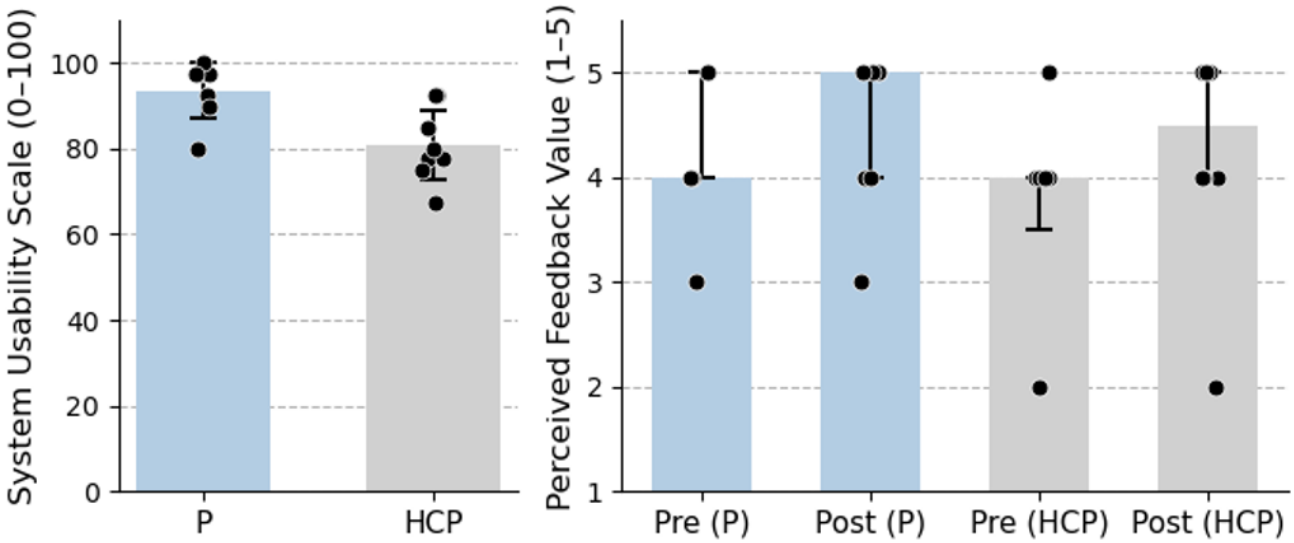
> *Consider adding brief explanations when milestones are not reached, or positive reinforcement when progress is made. I’d also like to see training recommendations as videos, such as: ‘Watch this video again to improve parameter A/B/C.’* [An HCP]

Another comment was:

> *Could the system automatically assign milestones based on initial assessments? And for chronic patients, could past data from earlier years be integrated?* [An HCP]

Patients also contributed feedback. One commented, “Compliment! There’s nothing I would want to change.” Another noted, “Can you include both the absolute values and percentage improvements in the bars?” Participants were asked to rate their perceived belief in the potential value of structured feedback in a clinical setting, reflecting attitudinal measures of perceived utility rather than demonstrated clinical benefit. Patients were specifically questioned on whether they believed such feedback could enhance their motivation during therapy, while HCPs were asked if it could aid clinical decision-making. This question was previously posed to participants of feedback session 1, before any feedback concepts were introduced, enabling a pre-post comparison. Median ratings from patients increased from 4.0 (IQR 4.0 - 5.0) to 5.0 (IQR 4.0 - 5.0), while HCPs’ median ratings rose from 4.0 (IQR 3.5 - 4.0) to 4.5 (IQR 4.0 - 5.0). Additionally, HCPs rated the potential of structured feedback to enhance patient motivation at a median of 4.5 (IQR 4.0 - 5.0). Participants were also asked whether they would find value in performance comparisons to (1) age-matched healthy individuals or (2) other patients with similar conditions. Patients rated comparisons to healthy individuals at a median of 5.0 (IQR 5.0 - 5.0) and to other patients at a median of 4.0 (IQR 3.0 - 5.0). HCPs assigned corresponding median ratings of 4.0 (IQR 4.0 - 4.5) and 3.0 (IQR 2.0 - 4.0). Finally, the perceived benefit of celebrating therapeutic milestones within clinical settings was evaluated. Patients rated this concept at a median of 4.0 (IQR 3.0 - 5.0), while HCPs provided a slightly lower median rating of 3.5 (IQR 3.0 - 4.5).

Figure 6. Evaluation of app usability and perceived value of structured feedback. The left panel shows System Usability Scale (SUS) scores for the app prototype as rated by patients (P) and health care practitioners (HCPs) in feedback session 3. The right panel presents pre–post ratings on the perceived value of structured assessment feedback in clinical routine, separately for patients and HCPs. Specifically, patients were asked if structured feedback about performed assessments could enhance their motivation to perform therapy on a more regular basis, and HCPs were asked if such feedback could aid them in decision-making processes. Patients and HCPs rated feedback value on a 5-point Likert scale before exposure to the feedback concepts (feedback session 1) and again after app testing (feedback session 3). Bars for the SUS score represent group means; error bars indicate SDs. Bars for the perceived value of structured feedback represent medians, error bars represent IQRs. Individual responses are shown as jittered black dots. HCP: health care practitioner; P: patient.



XSL•FO
RenderX

Cross-Session Themes

Across all 3 feedback sessions, several cross-cutting themes emerged. First, participants consistently preferred visually guided formats, particularly progress bars and color-coded indicators, over text-heavy or abstract representations such as spider plots or emoji-only designs, a pattern that was evident across all sessions and both stakeholder groups. Second, a set of competing demands was observed throughout: both patients and HCPs repeatedly requested access to absolute values and normative comparisons alongside simplified visuals. Third, the appropriate communication of worsening outcomes was a persistent concern across all sessions, raised by both groups from the earliest conceptual discussions through to the final prototype evaluation, ultimately informing the adoption of a neutral, color-coded feedback strategy without emotionally loaded symbols. Fourth, the perceived value of structured feedback grew across the course of the study, with both patients and HCPs rating its potential benefit more highly after exposure to the app prototype than before. Fifth, a consistent divergence in assumptions between patients and HCPs was observed. HCPs tended to recommend simpler formats out of concern that more detailed visualizations might overwhelm patients, while patients themselves consistently expressed a preference for richer, more informative representations.

Final Implementation of the Digital Feedback App

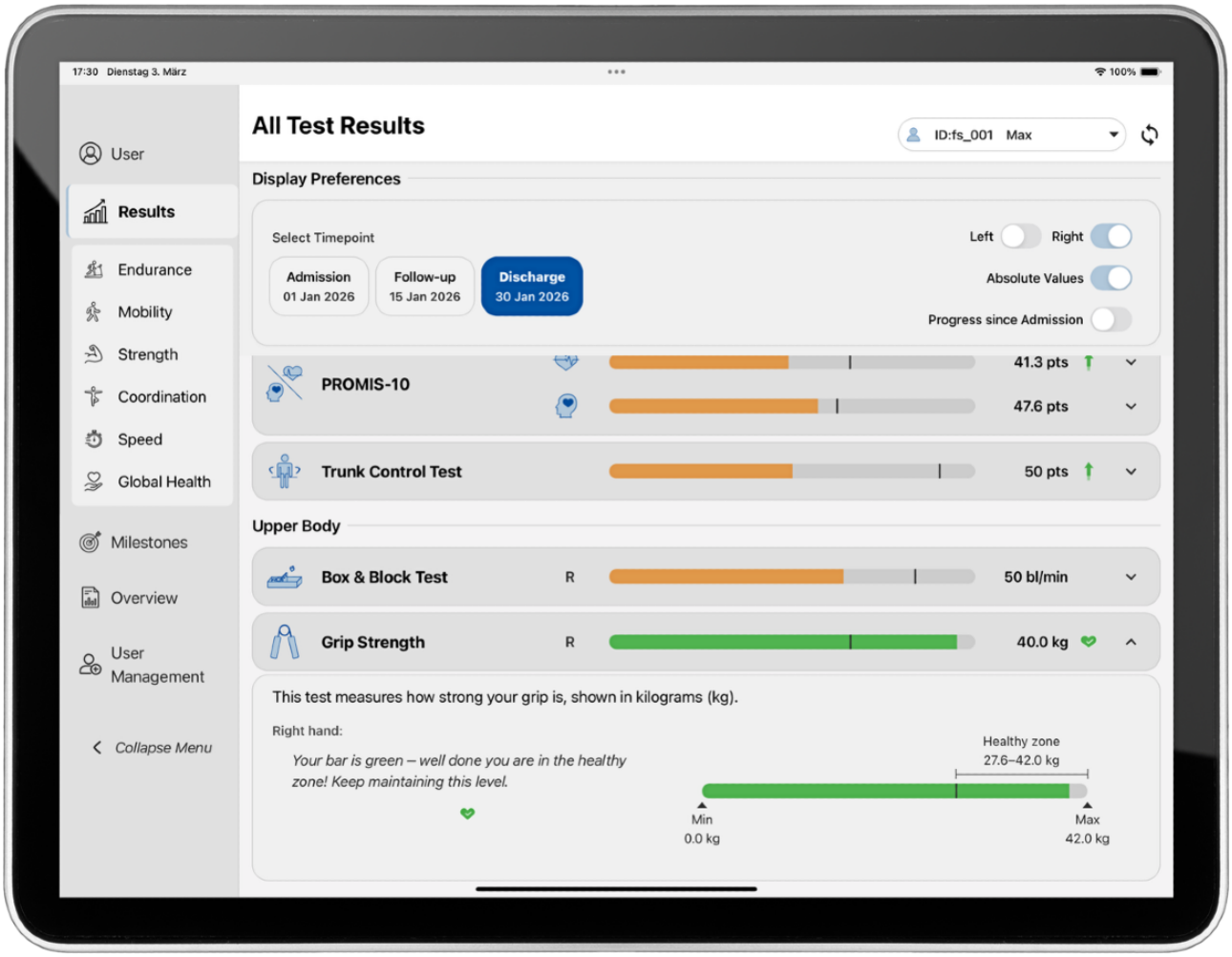
The final RehaLink app incorporated input from both patients and HCPs across all 3 structured feedback sessions. It was implemented as a mobile iOS app with 2 distinct user interfaces, one for patients and one for HCPs. The primary difference lies in data management functionality, which is available only to HCPs. This includes the ability to create new patient accounts and manage user credentials, such as resetting passwords.

The app was developed using React Native (version 15.1.3; Meta Platforms) [68], leveraging the Expo framework (version 52.0.23; Expo) [69] to facilitate efficient cross-platform development, deployment, and maintenance. It was optimized for tablet use to ensure seamless integration into clinical workflows, including bedside consultations. The backend infrastructure was implemented with Django (version 5.1.5; Django Software Foundation) [70] and a structured SQLite3 (version 3.45.3; Hwaci) database [71], providing robust mechanisms for secure data storage, user authentication, and real-time data synchronization.

Data from technology-based assessments (eg, the VPIT) is automatically analyzed and then visualized in the app interface. By contrast, data from conventional assessments is currently extracted manually from the hospital information system and exported to the app. The final digital feedback app integrates all iterative refinements from the 3 performed structured feedback sessions (Figure 7).

XSL•FO
RenderX

Figure 7. Final feedback app. Screenshot of the final RehaLink interface. The app is in normative mode; that is, shown progress bars indicate the users' functionality compared to an assessment-specific normative population, including age, gender, and handedness factors. Drop-down menus provide assessment and visualization explanations. Assessments are grouped by body region and can be filtered by functional domain (left panel). The left panel can be collapsed to show icons only.



Following session 1, findings were reviewed in a joint discussion with HCPs, resulting in a shared visual interface for both patients and HCPs to support use during consultations, with the HCP interface additionally including a patient selector and a user management tab. Navigation depth was reduced from 3 to 2 levels, with a login screen leading directly to an assessment overview. Assessments were grouped into upper body, lower body, and general categories, with a sidebar panel enabling filtering by functional domain (eg, strength and coordination) and access to personal health data (eg, BMI and weight) and the 3 primary therapy goals defined at admission. Color-coded bar charts were selected as the primary visualization format, supplemented by emoji-based indicators where appropriate.

Following session 2, a standardized format was introduced for digital health metrics, combining a recognizable icon, concise descriptive text, and an on-demand animated visualization illustrating key performance features (eg, jerky vs smooth movements). The VPIT metrics were visually aligned with conventional assessments, and all entries were supplemented with icons to support faster navigation. A left-right toggle was added for bimanual assessments to reduce visual clutter. Performance declines are indicated solely through color changes, without emotive symbols, accompanied by an explanatory note clarifying that fluctuations may reflect normal day-to-day variability. An overview panel was added displaying all assessment results in tabular format, including the patient's absolute values, the corresponding normative reference values, and the minimal detectable change (MDC) or minimal clinically important difference (MCID) thresholds as explicit columns. Values falling outside the normative range are highlighted in bold red; directional arrows indicate whether changes exceed the MDC or MCID threshold, flagging clinically significant improvements or deteriorations. This panel supports rapid clinical review and is designed to serve as a printable summary.

Following session 3, information buttons were replaced with expandable drop-down menus providing on-demand contextual explanations. VPIT metrics were grouped into a single assessment block with distinct subcategories to prevent visual overemphasis. Value labels were added next to performance bars, available in both absolute and percentage formats with a user toggle between the two. A tab-based navigation system was introduced to enable longitudinal browsing across assessment time points, with the most recent results displayed by default.

Two display modes were implemented. The progress mode, shown by default, displays changes in performance relative to

XSL•FO
RenderX

the patient's baseline at admission, using a color-coding scheme based on the MDC or, where unavailable, the MCID. Improvements beyond the threshold appear as green bars extending right, changes within the threshold range in orange, and deteriorations beyond the threshold in red extending left. The normative mode maps each outcome to a 0 - 100 scale referenced against an age-, gender-, condition-, and handedness-matched healthy population, where 100 represents the median healthy performance and the lower bound of the healthy range is defined as the median minus 2 SDs. Scores within the healthy range appear in green, values below the healthy threshold but above 50% in orange, and scores below 50% in red. Normative values for the VPIT were derived from a healthy reference database (n=120); norm values for conventional assessments were taken from previously published reference data [6,9,72-79]

To promote accessibility and personalization, the app includes user-adjustable settings such as dark mode, scalable font sizes, and high-contrast color schemes. Users can additionally hide assessment entries for tests that have not yet been performed, reducing visual clutter and cognitive load. Further, the default displayed unit (absolute vs percentage), hand side, and mode (normative vs progress) can be set for each user account. The interface is available in 4 languages, German, French, Italian, and English, to reflect the multilingual reality of Swiss rehabilitation settings.

Discussion

This study presents a 3-step, iterative, user-centered approach to designing a digital app that delivers structured feedback on assessment results to patients and HCPs in neurorehabilitation. After defining clinical and user requirements, we conducted three structured feedback sessions to evaluate: (1) the type, amount, and visual format of information presented; (2) the integration and representation of digital health metrics from technology-based assessments; and (3) the usability of a digital app prototype delivering this feedback. The final application presents assessment outcomes in a co-designed format tailored to both stakeholder groups. Likert-scale ratings indicate a growing perceived value across all sessions.

The motivational and clinical benefits of personalized feedback are well supported in rehabilitation literature. Moreover, feedback that is timed, personalized, and performance-related improves engagement and motor outcomes and reduces patient anxiety [27,28,30,31]. While prior work mainly focused on real-time feedback during therapy, our app delivers post-assessment feedback. Nevertheless, similar benefits are expected, including increased motivation, better understanding of progress, and enhanced shared decision-making. Structured feedback has also been recognized as a tool for more evidence-based, value-based health care (VBHC) [32], which our co-designed digital app supports.

To ensure these expected benefits translate effectively into practice, our main focus was to evaluate how feedback is presented to users. Users preferred clear visuals such as bar charts with optional, layered explanations over more complex formats like line or spider graphs, which can be harder to interpret [63,64]. Beyond confirming general preferences for visual simplicity, our findings provide context-specific insights into why particular visualization formats failed in this neurorehabilitation setting. Spider plots were consistently ranked lowest by both groups, likely because radial layouts mask the directionality and magnitude of change, information that is particularly relevant for patients tracking rehabilitation progress over time. Text-only formats were rejected not only because they lacked visual appeal, but also because participants associated plain text with negativity, particularly when communicating worsening outcomes, consistent with prior findings on the psychological impact of health data framing [80,81]. Emoji-only formats failed because both groups considered quantitative information essential for meaningful interpretation, a finding that contrasts with their common use in consumer health apps. Extended text formats were specifically rejected by HCPs due to information overload in time-constrained clinical workflows, while patients were more tolerant of detailed explanations, reinforcing the divergence in informational needs between the 2 groups. Importantly, this divergence itself represents a finding that extends beyond established heuristics; HCPs consistently underestimated patients' wishes for detailed, quantitative visualizations, suggesting that clinician judgments about patient interface preferences may be systematically biased toward oversimplification. More broadly, while existing literature addresses specific aspects of health data visualization in isolation, such as general patient preferences for chart types or usability of consumer health apps, our study takes a more comprehensive approach. Specifically, it is distinctive in evaluating the full pipeline from information selection and visualization design through to prototype usability, simultaneously engaging both patients and HCPs in a rehabilitation-specific context centered on the presentation of assessment outcomes. To our knowledge, no prior work has examined this complete process in this clinical setting, particularly with the integration of technology-based assessment metrics alongside conventional clinical assessments.

In contrast to many digital health tools that offer static or device-specific reports, our digital feedback app applies strategies like expandable content design and visual simplification [64] to preserve clinical utility without sacrificing usability. Compared to typical device-generated reports, our solution emphasizes personalization, clarity, and emotional neutrality. These preferences align with usability literature, especially for patients and older adults, key populations given the increasing prevalence of age-related neurological disorders. They tend to favor icon-driven navigation, high-contrast layouts, and intuitive visuals [46,49,50,82]. In line with these recommendations, RehaLink incorporates adjustable font sizes, high-contrast color options, and dark mode support, features shown to improve readability and reduce visual fatigue in older and neurologically impaired users [49,50]. The observed divergence between HCP assumptions and patient preferences underscores the importance of iterative, inclusive stakeholder engagement throughout the design process. Features like toggleable views—between absolute and normalized scores or between progress and normative benchmarks—allowed personalization without cognitive overload. These solutions

XSL•FO
RenderX

addressed users' varied informational needs while preserving interpretability and clinical relevance. Ramesh et al [83] used a symbolic, metaphor-driven approach that omitted numerical data and involved patients only at the final evaluation stage. In contrast, our participants were engaged throughout the full design process and explicitly requested access to raw scores. By combining quantitative metrics with intuitive visuals, our tool meets diverse informational needs and promotes transparency. Moreover, our larger, more diverse sample supports a broader generalizability. A key design insight was the need to balance simplicity with sufficient depth.

These design preferences not only influenced visual choices but also highlighted the importance of emotional tone and clinical context in delivering feedback. Participants were not just passive recipients of information but active stakeholders who were concerned with how their performance data were presented, especially when it involved sensitive or potentially discouraging results. As such, our co-design process extended beyond aesthetics to encompass the psychological and clinical dimensions of feedback delivery. While emojis were appreciated for reinforcing positive progress, participants favored more neutral, color-coded indicators and supportive language when feedback indicated worsening. This reflects prior findings that uncontextualized or overly negative feedback can increase patient anxiety [80,81] and should be carefully framed. Co-design allowed visual elements to be emotionally contextualized—emojis for improvement, neutral tones for decline—supporting engagement and comprehension. To ensure appropriate support, this digital feedback app is designed for use during patient consultations rather than in isolation.

In addition to emotional sensitivity, participants expressed a desire for actionable guidance and goal-setting tools. Patients valued reassurance during setbacks and the ability to track their own progress, while HCPs stressed the importance of clinically meaningful metrics and normative comparisons. This aligns with evidence supporting goal setting as a key motivator [27,84] and broader recommendations for embedding digital tools within VBHC frameworks [26]. Despite VBHC's potential, its implementation remains challenging due to fragmented infrastructure and limited outcome tracking [3,33].

By combining conventional and digital health metrics in a co-designed tool, our approach addresses these gaps and has the potential to support data-driven, patient-centered care. A central aim of this work was to demonstrate how digital health metrics can be meaningfully integrated into clinical routines. Digital health metrics, such as those exemplified by the VPIT, contribute objective, fine-grained, and repeatable measurements that can enhance the sensitivity and clinical utility of outcome tracking [11,21,56,59]. Their standardized format allows for automated longitudinal monitoring, eliminating the need for manual digitization or additional data processing, and supports meaningful comparisons across patients and time points. While the VPIT specifically targets upper limb sensorimotor function, its metric design is conceptually aligned with other technology-based assessments, such as those using wearable sensors for upper limb monitoring [85,86]. These approaches share the goal of capturing movement quality and functional performance through quantifiable, reproducible measures. Moreover, the principles guiding the presentation and interpretation of VPIT metrics are broadly applicable to other rehabilitation domains, for example, including lower limb function and cognitive assessments, where a wealth of digital health metrics has been proposed [87-90]. As such, the use of digital health metrics in feedback tools offers significant potential to advance VBHC by improving outcome transparency, personalization, and clinical efficiency.

The final digital app, refined through iterative feedback, achieved high usability scores from both stakeholder groups (SUS: patients 93.6, SD 6.4; HCPs 80.9, SD 8.1). Users highlighted ease of navigation, clear interface design, and appropriate feedback granularity. Importantly, even participants with low digital literacy reported no major interaction issues, supporting the tool's accessibility. Implementation feedback prioritized interoperability with electronic health records, inclusion of therapy recommendations, and milestone tracking. These additions, including automatic goal setting and benchmarking, enhance clinical integration and self-management.

While promising, this study is subject to several limitations. The patient cohort consisted only of persons with MS, recruited from a single neurorehabilitation center, which represents a limitation with respect to the direct generalizability of findings to other neurorehabilitation populations such as stroke or traumatic brain injury. However, several aspects of the study support broader applicability. First, no design decisions were specific to MS or restricted by condition-specific cognitive or motor profiles; the only practical prerequisite for app use is sufficient cognitive capacity to interact with a tablet-based interface, which applies across neurorehabilitation populations. Second, the HCPs involved were diverse neurorehabilitation therapists, and the conventional assessments integrated into the app are routinely used across neurorehabilitation conditions. Third, the VPIT, used here as a proof of concept for technology-based assessment integration, has been validated in individuals with different neurological disorders, such as MS, stroke, and ataxia. Moreover, while this evaluation was conducted using mock data, future studies should assess the app's real-world impact on clinical decision-making, patient adherence, and therapeutic outcomes. The clinical utility of the selected conventional and technology-based assessments also remains to be established through routine use and empirical testing. Additionally, interpreting the digital health metrics, particularly their prognostic relevance and implications for therapy adaptation, requires larger datasets and further investigation. Understanding how these metrics translate into actionable clinical decisions is crucial for the tool's long-term adoption and effectiveness. A limitation of the implementation is that data from conventional assessments requires manual extraction from the hospital information system to be exported to the app. While this introduces an additional workflow step, direct integration with clinical management systems was out of the scope of this work and not feasible at this stage of development due to data protection and information technology security constraints, which commonly apply in clinical settings. Should the ongoing evaluation demonstrate sufficient clinical benefit, integration of the app into the clinic's existing

XSL•FO
RenderX

information technology infrastructure is planned, which would enable automated data transfer and eliminate this manual step.

In summary, we demonstrate the feasibility and value of a co-designed tool for rehabilitation. Through structured feedback and iterative design, we developed a digital app that aligns with clinical workflows, enhances patient engagement, and has the potential to advance digital health integration. While these findings are promising, especially in terms of usability and stakeholder acceptance, further research is needed to evaluate the tool's impact in real-world settings, across broader patient populations, and in fully integrated clinical infrastructures. Our results offer a foundational step toward embedding patient-centered feedback systems into clinical routine, with potential to support data-driven decision-making and VBHC.

Acknowledgments

We thank all patients and health care practitioners who participated in this study for their valuable contributions. We also thank Franziska Riegel and Daniel Riese for their input during discussions on the study design. ChatGPT (OpenAI) was used during the preparation of this manuscript for rephrasing and language editing of selected passages. The authors take full responsibility for the integrity and accuracy of all content, and all AI-assisted text was reviewed and verified by the authors

Funding

This study was funded by: the Swiss Multiple Sclerosis (MS) Society, Grant Number 2021-06, the Swiss National Science Foundation as partner of the European Joint Programme on Rare Diseases (EJP RD), under grant agreement No02424 – 000, and the National Research Foundation Singapore (NRF) under its Campus for Research Excellence and Technological Enterprise (CREATE) program.The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available due to data protection reasons but are available from the corresponding author on reasonable request. The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Authors' Contributions

Study design: ND, OL, RG, JB, RS
Data collection: ND, KK
Participant Recruitment: RS
Data analysis: ND
Data interpretation: ND, OL, RG, JB
App Development ND, JS
Manuscript writing: ND
Manuscript review: OL, RG, JB, JS
All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of all performed clinical assessments, VPIT metrics and median and IQR for participant ratings.
[PDF File, 98 KB - rehab_v13i1e85072_app1.pdf]

Multimedia Appendix 2
Patient feedback questionnaire 1.
[PDF File, 197 KB - rehab_v13i1e85072_app2.pdf]

Multimedia Appendix 3
Therapist feedback questionnaire 1.
[PDF File, 208 KB - rehab_v13i1e85072_app3.pdf]

Multimedia Appendix 4
Patient feedback questionnaire 2.
[PDF File, 197 KB - rehab_v13i1e85072_app4.pdf]

XSL•FO
RenderX

Multimedia Appendix 5
Therapist feedback questionnaire 2.
[PDF File, 198 KB - rehab_v13i1e85072_app5.pdf]

Multimedia Appendix 6
Patient feedback questionnaire 3.
[PDF File, 143 KB - rehab_v13i1e85072_app6.pdf]

Multimedia Appendix 7
Therapist feedback questionnaire 3.
[PDF File, 145 KB - rehab_v13i1e85072_app7.pdf]

References

1. Khan F, Turner-Stokes L, Ng L, Kilpatrick T, Amatya B. Multidisciplinary rehabilitation for adults with multiple sclerosis. Cochrane Database Syst Rev 2007 Apr 18;2007(2):CD006036. [doi: 10.1002/14651858.CD006036.pub2] [Medline: 17443610]
2. Westerink HJ, Bresser CC, Garvelink MM, et al. The use of outcome data in patient consultations from the healthcare professionals' and patients' perspectives: a mixed methods study. Patient Educ Couns 2024 Jan;118:108043. [doi: 10.1016/j.pec.2023.108043] [Medline: 37925975]
3. van der Nat PB. The new strategic agenda for value transformation. Health Serv Manage Res 2022 Aug;35(3):189-193. [doi: 10.1177/09514848211011739] [Medline: 33900128]
4. Porter ME. Defining and introducing value in health care. In: Evidence-Based Medicine and the Changing Nature of Health Care: 2007 IOM Meeting Summary: National Academies Press; 2008:161-172.
5. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff (Millwood) 2008;27(3):759-769. [doi: 10.1377/hlthaff.27.3.759] [Medline: 18474969]
6. Mathiowetz V, Weber K, Kashman N, Volland G. Adult norms for the nine hole peg test of finger dexterity. Occup Ther J Res 1985 Jan;5(1):24-38. [doi: 10.1177/153944928500500102]
7. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
8. Keith RA, Granger CV, Hamilton BB, Sherwin FS. The functional independence measure: a new tool for rehabilitation. Adv Clin Rehabil 1987;1:6-18. [Medline: 3503663]
9. Schädler S, Kool J, Lüthi H, Marks D, Oesch P. Assessments in Der Rehabilitation Band 1: Neurologie, 4th edition: Hogrefe AG; 2019:133-239. [doi: 10.1024/85889-000]
10. World health organization international classification of functioning, disability and health (ICF). World Health Organization. 2001. URL: https://www.who.int/classifications/icf/en/ [accessed 2026-05-26]
11. Kanzler CM, Schwarz A, Held JPO, Luft AR, Gassert R, Lambercy O. Technology-aided assessment of functionally relevant sensorimotor impairments in arm and hand of post-stroke individuals. J Neuroeng Rehabil 2020 Sep 25;17(1):128. [doi: 10.1186/s12984-020-00748-5] [Medline: 32977810]
12. Gladstone DJ, Danells CJ, Black SE. The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair 2002 Sep;16(3):232-240. [doi: 10.1177/154596802401105171] [Medline: 12234086]
13. Hawe RL, Scott SH, Dukelow SP. Taking proportional out of stroke recovery. Stroke 2019 Jan;50(1):204-211. [doi: 10.1161/STROKEAHA.118.023006] [Medline: 30580742]
14. Murphy MA, Willén C, Sunnerhagen KS. Kinematic variables quantifying upper-extremity performance after stroke during reaching and drinking from a glass. Neurorehabil Neural Repair 2011 Jan;25(1):71-80. [doi: 10.1177/1545968310370748]
15. Alt Murphy M, Murphy S, Persson HC, Bergström UB, Sunnerhagen KS. Kinematic analysis using 3D motion capture of drinking task in people with and without upper-extremity impairments. J Vis Exp 2018 Mar 28(133):57228. [doi: 10.3791/57228] [Medline: 29658937]
16. Unger T, Sal Moslehian A, Peiffer JD, et al. Differentiable biomechanics for markerless motion capture in upper limb stroke rehabilitation: a comparison with optical motion capture. IEEE Trans Med Robot Bionics 2026;8(1):90-97. [doi: 10.1109/TMRB.2025.3605962]
17. Unger T, de Sousa Ribeiro R, Mokni M, et al. Upper limb movement quality measures: comparing IMUs and optical motion capture in stroke patients performing a drinking task. Front Digit Health 2024;6:1359776. [doi: 10.3389/fdgth.2024.1359776] [Medline: 38606036]
18. Dukelow SP, Herter TM, Moore KD, et al. Quantitative assessment of limb position sense following stroke. Neurorehabil Neural Repair 2010 Feb;24(2):178-187. [doi: 10.1177/1545968309345267] [Medline: 19794134]

XSL•FO
RenderX

19. Graser JV, Prospero L, Liesch M, Keller U, van Hedel HJA. Test-retest reliability of upper limb robotic exoskeleton assessments in children and youths with brain lesions. Sci Rep 2022 Oct 6;12(1):16685. [doi: 10.1038/s41598-022-20588-8] [Medline: 36202825]
20. Fluet MC, Lambercy O, Gassert R. Upper limb assessment using a virtual peg insertion test. 2011 Presented at: 2011 IEEE 12th International Conference on Rehabilitation Robotics (ICORR); Jun 29 to Jul 1, 2011; Zurich p. 192-197. [doi: 10.1109/ICORR.2011.5975348]
21. Kanzler CM, Rinderknecht MD, Schwarz A, et al. A data-driven framework for selecting and validating digital health metrics: use-case in neurological sensorimotor impairments. NPJ Digit Med 2020;3:80. [doi: 10.1038/s41746-020-0286-7] [Medline: 32529042]
22. Schwarz A, Kanzler CM, Lambercy O, Luft AR, Veerbeek JM. Systematic review on kinematic assessments of upper limb movements after stroke. Stroke 2019 Mar;50(3):718-727. [doi: 10.1161/STROKEAHA.118.023531] [Medline: 30776997]
23. Balasubramanian S, Melendez-Calderon A, Burdet E. A robust and sensitive metric for quantifying movement smoothness. IEEE Trans Biomed Eng 2012 Aug;59(8):2126-2136. [doi: 10.1109/TBME.2011.2179545] [Medline: 22180502]
24. Rohrer B, Fasoli S, Krebs HI, et al. Movement smoothness changes during stroke recovery. J Neurosci 2002 Sep 15;22(18):8297-8304. [doi: 10.1523/JNEUROSCI.22-18-08297.2002] [Medline: 12223584]
25. Hope TMH, Friston K, Price CJ, Leff AP, Rotshtein P, Bowman H. Recovery after stroke: not so proportional after all? Brain (Bacau) 2019 Jan 1;142(1):15-22. [doi: 10.1093/brain/awy302] [Medline: 30535098]
26. van Staalduinen DJ, van den Bekerom P, Groeneveld S, Kidanemariam M, Stiggelbout AM, van den Akker-van Marle ME. The implementation of value-based healthcare: a scoping review. BMC Health Serv Res 2022 Mar 1;22(1):270. [doi: 10.1186/s12913-022-07489-2] [Medline: 35227279]
27. Oyake K, Suzuki M, Otaka Y, Momose K, Tanaka S. Motivational strategies for stroke rehabilitation: a Delphi study. Arch Phys Med Rehabil 2020 Nov;101(11):1929-1936. [doi: 10.1016/j.apmr.2020.06.007] [Medline: 32753111]
28. Muraoka K, Takimoto Y, Nakazawa E, Tsuji T, Liu M. Stroke survivors' experiences and needs during the decision-making process considering rehabilitation options: a pilot descriptive study in Japan. Prog Rehabil Med 2022;7:20220024. [doi: 10.2490/prm.20220024] [Medline: 35633760]
29. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. Cochrane Database Syst Rev 2017 Nov 20;11(11):CD008349. [doi: 10.1002/14651858.CD008349.pub4] [Medline: 29156493]
30. Dobkin BH, Plummer-D'Amato P, Elashoff R, Lee J, SIRROWS Group. International randomized clinical trial, stroke inpatient rehabilitation with reinforcement of walking speed (SIRROWS), improves outcomes. Neurorehabil Neural Repair 2010;24(3):235-242. [doi: 10.1177/1545968309357558] [Medline: 20164411]
31. Popović MD, Kostić MD, Rodić SZ, Konstantinović LM. Feedback-mediated upper extremities exercise: increasing patient motivation in poststroke rehabilitation. Biomed Res Int 2014;2014:520374. [doi: 10.1155/2014/520374] [Medline: 24991557]
32. McGlinchey MP, Faulkner-Gurstein R, Sackley CM, McKevitt C. Factors guiding therapist decision making in the rehabilitation of physical function after severely disabling stroke - an ethnographic study. Disabil Rehabil 2024 Feb;46(4):672-684. [doi: 10.1080/09638288.2023.2172463] [Medline: 36734838]
33. van Veghel D, Daeter EJ, Bax M, et al. Organization of outcome-based quality improvement in Dutch heart centres. Eur Heart J Qual Care Clin Outcomes 2019 Apr 26. [doi: 10.1093/ehjqcco/qcz021]
34. Vijverberg JRG, Daniels K, Steinmann G, et al. Mapping the extent, range and nature of research activity on value-based healthcare in the 15 years following its introduction (2006-2021): a scoping review. BMJ Open 2022 Aug 16;12(8):e064983. [doi: 10.1136/bmjopen-2022-064983] [Medline: 35973705]
35. Zielasek J, Reinhardt I, Schmidt L, Gouzoulis-Mayfrank E. Adapting and implementing apps for mental healthcare. Curr Psychiatry Rep 2022 Sep;24(9):407-417. [doi: 10.1007/s11920-022-01350-3] [Medline: 35835898]
36. King DR, Emerson MR, Tartaglia J, Nanda G, Tatro NA. Methods for navigating the mobile mental health app landscape for clinical use. Curr Treat Options Psychiatry 2023 May 24:1-15. [doi: 10.1007/s40501-023-00288-4] [Medline: 37360961]
37. Phillips JC, Alfano AR, Barfield LC, et al. Exploring maternal and infant health app development and effectiveness research: scoping review. JMIR Pediatr Parent 2024 Jan 26;7:e46973. [doi: 10.2196/46973] [Medline: 38055330]
38. Schoeppe S, Alley S, Van Lippevelde W, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 7;13(1):127. [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
39. Higgins JP. Smartphone applications for patients' health and fitness. Am J Med 2016 Jan;129(1):11-19. [doi: 10.1016/j.amjmed.2015.05.038] [Medline: 26091764]
40. Torous J, Andersson G, Bertagnoli A, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98. [doi: 10.1002/wps.20592] [Medline: 30600619]
41. International Organization for Standardization. ISO 9241-210:2019 ergonomics of human-system interaction - part 210: human-centred design for interactive systems. : ISO; 2019 URL: https://www.iso.org/standard/77520.html [accessed 2026-05-26]
42. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]

XSL•FO
RenderX

43. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20. [doi: 10.2196/mhealth.3335] [Medline: 25803705]
44. Matthew-Maich N, Harris L, Ploeg J, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 9;4(2):e29. [doi: 10.2196/mhealth.5127] [Medline: 27282195]
45. IDEO.org. The Field Guide to Human-Centered Design: Design Kit; 2015.
46. Stonbraker S, Porras T, Schnall R. Patient preferences for visualization of longitudinal patient-reported outcomes data. J Am Med Inform Assoc 2020 Feb 1;27(2):212-224. [doi: 10.1093/jamia/ocz189] [Medline: 31670816]
47. Porter ME. What is value in health care? N Engl J Med 2010 Dec 23;363(26):2477-2481. [doi: 10.1056/NEJMp1011024] [Medline: 21142528]
48. Draper SW, Norman DA. User Centered System Design: New Perspectives on Human-Computer Interaction: CRC Press; 1986:526.
49. Zhang S, Song J. An empirical investigation into the preferences of the elderly for user interface design in personal electronic health record systems. Front Digit Health 2023;5:1289904. [doi: 10.3389/fdgth.2023.1289904] [Medline: 38348367]
50. Asfandiyar S, Hough S, Agugo V, Brixey JJ, Gong Y. Intuitive data visualization for enhanced patient comprehension. Stud Health Technol Inform 2024 Jul 24;315:37-42. [doi: 10.3233/SHTI240102] [Medline: 39049222]
51. Vargas C, Zorbas C, Longworth GR, et al. Exploring co-design: a systematic review of concepts, processes, models, and frameworks used in public health research. J Public Health (Bangkok) 2025 Dec 1;47(4):e616-e639. [doi: 10.1093/pubmed/fdaf084]
52. Wang Z, Jiang T, Huang J, Tai Y, Trapani P. How might we evaluate co-design? A literature review on existing practices. In: Lockton D, Lenzi S, Hekkert P, Oak A, Sádaba J, Lloyd P, editors. : Design Research Society; 2022 Presented at: DRS2022; Jun 25 to Jul 3, 2022; Bilbao, Spain. [doi: 10.21606/drs.2022.774]
53. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
54. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. : ACM; 1993 Presented at: The SIGCHI conference; Apr 24-29, 1993; Amsterdam, The Netherlands p. 206-213 URL: http://portal.acm.org/citation.cfm?doid=169059 [accessed 2026-05-27] [doi: 10.1145/169059.169166]
55. Jakob N. Usability Engineering: Morgan Kaufmann Publishers; 1994.
56. Kanzler CM, Lamers I, Feys P, Gassert R, Lambercy O. Personalized prediction of rehabilitation outcomes in multiple sclerosis: a proof-of-concept using clinical data, digital health metrics, and machine learning. Med Biol Eng Comput 2022 Jan;60(1):249-261. [doi: 10.1007/s11517-021-02467-y] [Medline: 34822120]
57. Kanzler CM, Sylvester R, Gassert R, Kool J, Lambercy O, Gonzenbach R. Goal-directed upper limb movement patterns and hand grip forces in multiple sclerosis. Mult Scler J Exp Transl Clin 2022;8(3):20552173221116272. [doi: 10.1177/20552173221116272] [Medline: 35982915]
58. Kanzler CM, Averta G, Schwarz A, et al. A low-dimensional representation of arm movements and hand grip forces in post-stroke individuals. Sci Rep 2022 May 9;12(1):7601. [doi: 10.1038/s41598-022-11806-4] [Medline: 35534629]
59. Kanzler CM, Lessard I, Gassert R, Brais B, Gagnon C, Lambercy O. Reliability and validity of digital health metrics for assessing arm and hand impairments in an ataxic disorder. Ann Clin Transl Neurol 2022 Apr;9(4):432-443. [doi: 10.1002/acn3.51493] [Medline: 35224896]
60. Kanzler CM, Lessard I, Gassert R, Brais B, Gagnon C, Lambercy O. Digital health metrics reveal upper limb impairment profiles in ARSACS. J Neurol Sci 2023 May 15;448:120621. [doi: 10.1016/j.jns.2023.120621] [Medline: 37004405]
61. Likert R. A technique for the measurement of attitudes. Arch Psychol 1932;22(140):1-55 [FREE Full text]
62. Roman LC, Ancker JS, Johnson SB, Senathirajah Y. Navigation in the electronic health record: a review of the safety and usability literature. J Biomed Inform 2017 Mar;67:69-79. [doi: 10.1016/j.jbi.2017.01.005] [Medline: 28088527]
63. Turchioe MR, Myers A, Isaac S, et al. A systematic review of patient-facing visualizations of personal health data. Appl Clin Inform 2019 Aug;10(4):751-770. [doi: 10.1055/s-0039-1697592] [Medline: 31597182]
64. Snyder C, Smith K, Holzner B, et al. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. Qual Life Res 2019 Feb;28(2):345-356. [doi: 10.1007/s11136-018-2020-3] [Medline: 30306533]
65. Figma: collaborative interface design tool. Figma. 2025. URL: https://www.figma.com [accessed 2026-05-27]
66. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194.
67. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
68. React native. Facebook. 2025. URL: https://reactnative.dev [accessed 2026-05-27]
69. Expo documentation. Expo. 2025. URL: https://docs.expo.dev/ [accessed 2026-05-27]
70. Holovaty A. The Django web framework. 2005 Presented at: Proceedings of the 2005 Python Conference; Mar 23-25, 2005 URL: https://www.djangoproject.com [accessed 2026-05-27]

XSL•FO
RenderX

71. ISO/IEC 9075-1:2023 – information technology — database languages sqlpart 1: framework (SQL/framework). ISO/IEC. 2023. URL: https://www.iso.org/standard/76583.html [accessed 2026-06-08]
72. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. Phys Ther 2002 Feb;82(2):128-137. [doi: 10.1093/ptj/82.2.128] [Medline: 11856064]
73. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. Am J Respir Crit Care Med 1998 Nov;158(5 Pt 1):1384-1387. [doi: 10.1164/ajrccm.158.5.9710086] [Medline: 9817683]
74. Perneger TV, Combescure C, Courvoisier DS. General population reference values for the French version of the EuroQol EQ-5D health utility instrument. Value Health 2010 Aug;13(5):631-635. [doi: 10.1111/j.1524-4733.2010.00727.x] [Medline: 20412541]
75. Bohannon RW, Williams Andrews A. Normal walking speed: a descriptive meta-analysis. Physiotherapy 2011 Sep;97(3):182-189. [doi: 10.1016/j.physio.2010.12.004] [Medline: 21820535]
76. Bohannon RW, Wang YC, Gershon RC. Two-minute walk test performance by adults 18 to 85 years: normative values, reliability, and responsiveness. Arch Phys Med Rehabil 2015 Mar;96(3):472-477. [doi: 10.1016/j.apmr.2014.10.006] [Medline: 25450135]
77. Bohannon RW. Reference values for the Timed Up and Go test. J Geriatr Phys Ther 2006 Aug;29(2):64-68. [doi: 10.1519/00139143-200608000-00004]
78. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block test of manual dexterity. Am J Occup Ther 1985 Jun;39(6):386-391. [doi: 10.5014/ajot.39.6.386] [Medline: 3160243]
79. Mathiowetz V, Kashman N, Volland G, Weber K, Dowe M, Rogers S. Grip and pinch strength: normative data for adults. Arch Phys Med Rehabil 1985 Feb;66(2):69-74. [Medline: 3970660]
80. Fallowfield L, Jenkins V. Communicating sad, bad, and difficult news in medicine. The Lancet 2004 Jan;363(9405):312-319. [doi: 10.1016/S0140-6736(03)15392-5]
81. Choe EK, Duarte ME, Suh H, Pratt W, Kientz JA. Communicating bad news: insights for the design of consumer health technologies. JMIR Hum Factors 2019 May 17;6(2):e8885. [doi: 10.2196/humanfactors.8885] [Medline: 31102374]
82. Dumurgier J, Tzourio C. Epidemiology of neurological diseases in older adults. Rev Neurol (Paris) 2020 Nov;176(9):642-648. [doi: 10.1016/j.neurol.2020.01.356] [Medline: 32145981]
83. Ramesh SH, Ouskine A, Khorasani E, et al. A data visualization tool for patients and healthcare providers to communicate during inpatient stroke rehabilitation. 2024 Jun 3 Presented at: Proceedings of Graphics Interface 2024; Jun 3-6, 2024; Halifax, NS, Canada p. 1-14 URL: https://dl.acm.org/doi/proceedings/10.1145/3670947 [accessed 2026-05-27] [doi: 10.1145/3670947.3670978]
84. Oyake K, Yamauchi K, Inoue S, et al. A multicenter explanatory survey of patients' and clinicians' perceptions of motivational factors in rehabilitation. Commun Med (Lond) 2023 Jun 6;3(1):78. [doi: 10.1038/s43856-023-00308-7] [Medline: 37280319]
85. Zadeh SM, MacDermid J, Johnson J, Birmingham TB, Shafiee E. Applications of wearable sensors in upper extremity MSK conditions: a scoping review. J Neuroeng Rehabil 2023 Nov 18;20(1):158. [doi: 10.1186/s12984-023-01274-w] [Medline: 37980497]
86. Kim GJ, Parnandi A, Eva S, Schambra H. The use of wearable sensors to assess and treat the upper extremity after stroke: a scoping review. Disabil Rehabil 2022 Oct;44(20):6119-6138. [doi: 10.1080/09638288.2021.1957027] [Medline: 34328803]
87. Peters DM, O'Brien ES, Kamrud KE, et al. Utilization of wearable technology to assess gait and mobility post-stroke: a systematic review. J Neuroeng Rehabil 2021 Apr 21;18(1):67. [doi: 10.1186/s12984-021-00863-x] [Medline: 33882948]
88. Hulleck AA, Menoth Mohan D, Abdallah N, El Rich M, Khalaf K. Present and future of gait assessment in clinical practice: towards the application of novel trends and technologies. Front Med Technol 2022;4:901331. [doi: 10.3389/fmedt.2022.901331] [Medline: 36590154]
89. Salchow-Hömmen C, Skrobot M, Jochner MCE, Schauer T, Kühn AA, Wenger N. Review-emerging portable technologies for gait analysis in neurological disorders. Front Hum Neurosci 2022;16:768575. [doi: 10.3389/fnhum.2022.768575] [Medline: 35185496]
90. Polk SE, Öhman F, Hassenstab J, et al. A scoping review of remote and unsupervised digital cognitive assessments in preclinical Alzheimer's disease. NPJ Digit Med 2025 May 10;8(1):266. [doi: 10.1038/s41746-025-01583-5] [Medline: 40348870]

Abbreviations

EKOS: Ethics Committee of Eastern Switzerland
HCP: health care practitioner
KINARM: Kinesiological Instrument for Normal and Altered Reaching Movement
MCID: minimal clinically important difference
MDC: minimal detectable change
MS: multiple sclerosis
SUS: System Usability Scale

XSL•FO
RenderX

VBHC: value-based health care
VPIT: Virtual Peg Insertion Test

Edited by S Munce; submitted 30.Sep.2025; peer-reviewed by E Lendaro, F Elkourdi; revised version received 09.Apr.2026; accepted 05.May.2026; published 17.Jun.2026.

Please cite as:
Domnik N, Krasnopolska K, Sylvester R, Bansi J, Song J, Gonzenbach R, Lambercy O
Digital Platform to Provide Health Data Feedback for Neurorehabilitation Patients: User-Centered Development and Proof-of-Concept Usability Study
JMIR Rehabil Assist Technol 2026;13:e85072
URL: https://rehab.jmir.org/2026/1/e85072
doi:10.2196/85072

© Nadine Domnik, Katarzyna Krasnopolska, Ramona Sylvester, Jens Bansi, Jaeyong Song, Roman Gonzenbach, Olivier Lambercy. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 17.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Therapeutic Use of Virtual Reality for Patients With Fibromyalgia and Chronic Neck Pain: Randomized Controlled Trial

Edurne Úbeda-D'Ocasar[1,2*], PhD, Prof Dr; Yaiza Moreno-Crespo[1,2*], PT; Eduardo Cimadevilla-Fernández-Pola[1,2*], PhD, Prof Dr; Juan Hernández-Lougedo[1,2*], PhD, Prof Dr; Álvaro Navas-Mosqueda[1,2*], PT; Mario Caballero-Corella[1,2*], PT; Noemí Mayoral-Gonzalo[1,2*], PT; Blanca Pedauyé-Rueda[1,2*], RDN; María Jesús Fernández-Aceñero[3*], Dr med, PhD; Juan Pablo Hervás-Pérez[4*], PhD, Prof Dr; Cristina Ojedo-Martín[1,2*], PT

[1]Physiotherapy and Health Research Group (FYSA), Faculty of Health Sciences – HM Hospitals, Camilo José Cela University, Villanueva de la Cañada, Madrid, Spain

[2]Instituto de Investigación Sanitaria HM Hospitales, Madrid, Spain

[3]Department of Legal Medicine, Psychiatry and Pathology, Facultad de Medicina, Universidad Complutense de Madrid, Fundación de Investigación Clínico San Carlos (IDiSCC), Madrid, Spain

[4]Department of Chemistry in Pharmaceutical Sciences, Analytical Chemistry Unit, Faculty of Pharmacy, Complutense University of Madrid, Plaza Ramón y Cajal s/n, Ciudad Universitaria, Madrid, Spain

[*]all authors contributed equally

Corresponding Author:
Juan Pablo Hervás-Pérez, PhD, Prof Dr
Department of Chemistry in Pharmaceutical Sciences, Analytical Chemistry Unit, Faculty of Pharmacy, Complutense University of Madrid, Plaza Ramón y Cajal s/n, Ciudad Universitaria, Madrid, Spain

Abstract

Background: Fibromyalgia (FM) causes widespread pain, fatigue, and cognitive abnormalities. Cervical pain is a common and debilitating symptom.

Objective: This study aims to evaluate the effectiveness of virtual reality (VR) as a treatment for chronic cervical pain experienced by patients with FM.

Methods: A single-blind randomized clinical trial was conducted. A total of 56 women were randomly assigned to 3 groups: G1 (VR+cervical mobility exercises), G2 (cervical mobility exercises), and the control group. Therapy was administered twice a week for 4 weeks. Variables such as disease impact, quality of life, kinesiophobia, pain, range of motion, fatigue, and treatment adherence were measured.

Results: The mean age of the participants was 54.26 (SD 7.7) years. Participants were overweight, with a mean BMI of 28.7 (SD 7.8). The mean visual analog scale value was 6.72 (SD 1.8). The baseline values for age, BMI, visual analog scale, algometric measures, and functional capacity (measured using the Timed Up and Go test, cervical rotation, and lateral displacement) were similar across the 3 groups. Following the intervention therapy, the control group did not exhibit notable improvement (mean 3.5, SD 1.4; differences of mean values –0.46, 95% CI –1.1 to 0.2; *P*=.15), particularly in pain perception, while both therapy groups did show improvements (G1: mean 3.8, SD 1.1; differences of mean values 1.2, 95% CI 0.78-1.54; *P*<.001; G2: mean 2.8, SD 1.8; differences of mean values 1.2, 95% CI 0.66-1.7; *P*<.001). Both intervention groups improved significantly compared to control postintervention in FM impact (CG vs G1: differences of mean values 9.31, 95% CI 14.7-3.8; *P*<.001; CG vs G2: differences of mean values 8.4, 95% CI 13.84-3.06; *P*<.001), central sensitization (CG vs G2: differences of mean values 7.53, 95% CI 12.12-2.95; *P*<.001), and cervical disability (CG vs G2: differences of mean values 6.44, 95% CI 9.93-2.94; *P*<.001). However, at 1 month, only G1 maintained superior improvements across all measures, including a reduction in kinesiophobia (G2: differences of mean values 6.2, 95% CI 4.7-9.8; *P*<.001), indicating a more sustained effect of the combined approach.

Conclusions: The combination of VR with cervical mobility and strengthening exercises produced superior and sustained improvements in women with FM compared to exercise alone or control. Significant benefits were observed in disease impact, central sensitization, cervical disability, and kinesiophobia, with effects maintained at 1 month only in the VR group. These findings support VR as an effective adjunct to enhance symptom management and treatment adherence in FM.

Trial Registration: ClinicalTrials.gov NCT05933941; https://clinicaltrials.gov/study/NCT05933941

(JMIR Rehabil Assist Technol 2026;13:e81158) doi:10.2196/81158

KEYWORDS

chronic neck pain; exercise; fibromyalgia; virtual reality; randomized controlled trial

XSL•FO
RenderX

Introduction

Fibromyalgia (FM) is a disease that causes widespread chronic pain and intense fatigue along with many other symptoms [1-3]. Historically, there was some skepticism among health care professionals regarding the existence of this disease and its classification as a mental disorder. Such skepticism may have arisen from concerns about the optimal use of resources in providing care for individuals with FM [3,4]. It is estimated that this disease affects between 2% and 4% of the Spanish population, making it the most common cause of diffuse chronic musculoskeletal pain. Notably, the prevalence is much higher in women, with a ratio of 8:1 [5-7]. Its incidence is increasing because of the improvements in diagnostic criteria and advances in research that are bringing this pathology to light [7,8]. Given its high prevalence, this disease represents a significant public health concern [9,10].

Currently, there are no specific diagnosis tests available for confirming an FM diagnosis. Historically, an FM diagnosis was based on the presence of widespread pain in at least 11 of 18 designated tender points for a duration of 3 months. However, this criterion proved insufficient to capture the complexity of the disorder. Consequently, more comprehensive diagnostic frameworks have been developed, including a range of associated symptoms such as fatigue, joint stiffness, cervical and lumbar pain, muscle weakness, depressive symptoms, gastrointestinal issues, cognitive impairments, and balance alterations. Collectively, these manifestations significantly compromise the quality of life of individuals affected by FM. This pain can become disabling during flare-ups and even become chronic, often being a cause of recurrent absences from work [11].

Unlike pain associated with a sedentary lifestyle or poor posture, neck pain in FM results from a complex interplay between physiological, neurological, and psychological factors. Chronic pain often involves the cervical spine, leading to acute and prolonged pain episodes, along with muscle tension in the suboccipital, trapezius, and elevator scapulae muscles. It is thought that several factors may contribute to this chronic pain, including central pain sensitization, nervous system dysregulation, sleep disturbances, musculoskeletal changes, and psychological and emotional factors [11-13].

In recent years, new therapeutic approaches have been explored to improve symptomatology and thus patients' quality of life. One of the main challenges in treating patients with chronic pain, particularly those with FM, is their low adherence to physical therapy exercises. This may be due to factors such as severe pain and fatigue, a fear of worsening symptoms, a lack of understanding about exercise benefits, inadequate support and supervision, and frustration with slow improvement. To address these issues, it would be beneficial to adopt a comprehensive approach that involves educating patients regarding the benefits of exercise, customizing programs according to individual needs, providing ongoing support via health care professionals, managing pain and fatigue during exercise, setting realistic expectations, and recognizing achievements [14-16]. It may be beneficial to consider a treatment plan that incorporates aerobic activities, strength training, and stretching exercises [17-19]. However, this type of therapy may be limited by kinesiophobia. One possible solution to this issue could be the use of virtual reality (VR) technology. Exercise interventions delivered through VR have shown superior efficacy in alleviating the core symptoms of FM, including pain, fatigue, and stiffness, while also promoting improvements in balance and postural control [20]. The affordability of portable VR devices, combined with their sustained effectiveness as a nonpharmacological intervention for chronic pain management, positions VR as a promising tool with potential future applications as an analgesic modality [21]. Numerous randomized controlled trials have demonstrated that VR constitutes an alternative and accessible therapeutic approach for pain management [22].

Immersive VR is an innovative and interactive technology that generates 3D scenarios, enabling patients to experience highly realistic simulations while interacting with the virtual environment using their own hands. VR has been researched and applied in various contexts for pain management, demonstrating considerable potential as a therapeutic tool. The effectiveness of this therapeutic approach has been demonstrated for patients with chronic pain, as it addresses several factors, such as fatigue, kinesiophobia, and range of motion (ROM) [23]. This therapy can be adapted to suit the specific requirements of each patient, providing the option to adjust the level of difficulty from low to intermediate or high. Additionally, it offers a variety of game modes that could be beneficial in targeting different clinical conditions, such as low back pain, neck pain, and balance disorders [24,25].

The therapeutic mechanisms and components of VR are distraction, activity management and behavioral activation, skills-based cognitive behavioral therapy, relaxation training and biofeedback, positive emotion induction, neuromodulation, physical rehabilitation, and reduction of kinesiophobia [8,23,26].

VR can be categorized into 2 main types: immersive and nonimmersive. Immersive VR involves the use of head-mounted displays or VR goggles that fully cover the visual field and may include motion sensors for the hands or feet. In contrast, nonimmersive VR is accessed through conventional screens or computers, without any device that isolates the user from the external environment [27] .

Chronic pain has been extensively investigated over the past decades, and research efforts persist in identifying novel strategies to mitigate its impact and enhance patients' quality of life. VR in both modalities has demonstrated efficacy in reducing chronic musculoskeletal pain. These findings support the integration of VR as a therapeutic tool in clinical populations affected by conditions associated with this type of pain [28].

Patients may perceive amplified motion in the virtual world, depending on the configuration of the headset used. This could result in small neck flexion in real life being rewarded with greater apparent motion in VR, which might increase kinesiophobia [26]. Studies have indicated that for patients with chronic low-back pain, kinesiophobia may be addressed through VR interventions [20,29,30]. VR helps to distract an individual from painful stimuli while exercising in a way that improves

XSL•FO
RenderX

their perception of pain [31]. In addition, there have been indications that improvements in fatigue, sleep quality, and ROM may also be achieved [9].

In this study, we aimed to evaluate the effectiveness of VR as a therapeutic option for neck pain in patients with FM. We hypothesized that VR could decrease cervical pain and kinesiophobia and increase ROM, facilitating improvements in quality of life and adherence to treatment.

Methods

Trial Design

A single-blind, randomized experimental clinical trial was conducted among women diagnosed with FM recruited through the Association of Fibromyalgia, "AFINSFACRO," in Móstoles, Spain. Following the selection of patients who met the inclusion criteria and the collection of informed consent forms and information sheets, according to ethical standards, the patients were randomly assigned to 1 of the 3 groups. The first group performed 20 minutes of exercise plus 10 minutes of VR (G1), the second group performed 30 minutes of exercise (G2), and the control group (CG) did not perform any exercise or VR treatment.

This study was registered at ClinicalTrials.gov (NCT05933941) and approved by the Clinical Research Committee of the Hospital Clínico San Carlos (Madrid, Spain; ID: VRTCNPPFM-13/07/2023). The research was conducted from April 1, 2024, to January 30, 2025. All participants gave informed consent, and the general scope of this study was explained to them via a participant information sheet. This clinical trial was designed and reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) guidelines, with specific adherence to the CONSORT-Harms (Consolidated Standards of Reporting Trials Harms Extension) 2022 statement (Checklist 1) to ensure transparent and comprehensive reporting of adverse events and safety outcomes.

Participants

The QuestionPro survey software (QuestionPro, Inc) was used to calculate the required sample size, considering a 95% CI and a 5% α error. The sample size calculation was based on the study by Gulsen et al [8]. A total of 56 women were enrolled, of whom 2 withdrew for personal reasons unrelated to the study. The inclusion criteria comprised female patients aged 20 to 65 years with a confirmed diagnosis of FM and chronic cervical pain lasting more than 3 months. The exclusion criteria included vertigo, claustrophobia, epilepsy, pregnancy, or refusal to provide informed consent.

Type of Sampling

A consecutive, nonprobabilistic convenience sampling strategy was applied to participants who met the inclusion criteria.

Variables and Outcomes

Sociodemographic variables included age, height, weight, BMI, marital status, pain medication use, employment status, smoking history, and comorbid conditions such as restless legs syndrome, chronic fatigue syndrome, temporomandibular joint dysfunction, migraines, irritable bowel syndrome, multiple chemical sensitivity, anxiety, and depression. All these sociodemographic variables were collected using a self-administered survey.

The following tools were used in this study:

1. Pain: we used a visual analog scale (VAS) to assess self-reported pain intensity, ensuring maximum reproducibility among different observers [32]. Furthermore, we used an analog pressure FPK 20 algometer (Wagner algometer, Force Dial FDK 20; Wagner Instruments) to evaluate pain among individuals diagnosed with FM. This instrument allows the assessment of a patient's central sensitivity to pain. Measurements focused on tender points in the occipital and upper trapezius regions bilaterally [33-35].
2. Subjective intensity of effort was assessed using the Borg Category-Ratio Scale (CR-10), which quantifies the subjective intensity of effort experienced during physical exercise or functional testing. Participants were asked to rate their exertion immediately after each test or exercise session on a 0 to 10 scale, where 0 indicates "no exertion at all" and 10 represents "maximal exertion" [36].
3. Neck Disability Index (NDI): the validated Spanish version of this questionnaire was used to assess pain and neck-related disability [37].
4. Fear of movement: the Tampa Scale for Kinesiophobia (TSK) was used to assess the fear of movement [38].
5. Exercise adherence: we used the Exercise Adherence Rating Scale, which is a validated questionnaire with 2 sections, one assessing exercise performance and another evaluating frequency, motivation, and consistency [39].
6. Impact of FM: the Fibromyalgia Impact Questionnaire (FIQ) was used to assess the impact of FM on health-related quality of life [40-42].
7. Quality of life: we used EQ-5D (EuroQol 5-Dimensions) questionnaire, which is a tool that allows the evaluation of a patient's overall quality of life in primary care settings [43-45].
8. Symptoms of central sensitization: we used the validated Spanish version of the Central Sensitization Inventory (CSI). This questionnaire assesses symptoms using a scale of 0 ("never") to 4 ("always"). The total score ranges from 0 to 100. A score above 40 indicates the presence of central sensitization [46].
9. ROM: cervical flexion, extension, lateral flexion, and rotation were assessed using a goniometer [47]. This instrument is used to evaluate one's degree of joint mobility, thereby facilitating the determination of an individual's restrictions.
10. Time Up Go test: the test consisted of measuring the time it took participants to get up from a chair with a height of 46 cm, walking 3 meters, turning around a cone, and sitting down again. This test was performed to assess physical performance, gait, and dynamic balance [48,49].

Procedure

The intervention was conducted over a period of 1 month. G1 and G2 participated in 2 sessions per week. G1 performed 20 minutes of cervical mobility and strengthening exercises

XSL•FO
RenderX

(Multimedia Appendix 1), followed by 10 minutes of immersive VR therapy. G2 completed 30 minutes of cervical mobility and strengthening exercises. The CG did not perform any cervical mobility or strengthening exercises nor participate in immersive VR therapy. The immersive VR therapy was delivered using Meta Quest 2000 headsets, using the game "Interkosmos 2000." In the game, participants assumed the role of a spacecraft pilot navigating through a series of rings while avoiding meteoroids by performing neck movements according to the game's instructions. The difficulty level was adjusted as follows: the first 3 sessions were carried out in "easy" mode, allowing a maximum cervical mobility of 30°. The subsequent 3 sessions were conducted in "medium" mode, allowing up to 60° of cervical mobility. The final 2 sessions were performed in "hard" mode, which not only increased the spacecraft's speed but also enabled a full cervical range of motion. All groups underwent assessments of cervical ROM including flexion, extension, right and left lateral inclination, and right and left rotation. In addition, pain intensity in the upper trapezius (bilaterally) and in the right and left occipital regions was quantified using an algometer. Evaluations were conducted at 4 time points: baseline (preintervention), immediately after the intervention, 15 days postintervention, and 1 month postintervention. All follow-up assessments were performed at the association's premises. After each session, participants from G1 and G2 reported their perceived levels of cervical pain and fatigue. All outcome measures were collected by a research assistant blinded to the study objectives and group allocations.

Evaluations were conducted at 4 time points: baseline (preintervention), immediately postintervention, 15 days postintervention, and 1-month postintervention. All follow-up assessments were conducted at the association's headquarters. After each session, participants in G1 and G2 reported perceived levels of cervical pain and fatigue.

All outcome measures were collected by a research assistant who was blinded to both the study's objectives and the participants' group assignments.

Statistical Analysis

Statistical analyses were performed using SPSS 29.0 (IBM Corp). The normality of quantitative variables was assessed with the Kolmogorov-Smirnov test. Depending on distribution, data were described using means and SDs or medians, IQRs, and ranges. Qualitative variables were expressed as percentages and absolute values. Baseline comparisons between the groups were conducted using chi-squared tests or ANOVA with Bonferroni post hoc analysis. Temporal changes in outcomes were analyzed with paired tests, and intergroup differences were analyzed using ANOVA, Bonferroni post hoc, or Kruskal-Wallis tests, as appropriate. Correlations between quantitative variables were assessed using the Pearson or Spearman tests, and associations between qualitative variables were assessed using chi-squared tests. Statistical significance was set at $P<.05$.

Ethical Considerations

This study received approval from the institutional review board of Hospital Clínico San Carlos (23 - 458 EC) and was conducted in accordance with the Spanish legislation, including Law 41/2002 on patient autonomy and Organic Law 3/2018 on data protection and digital rights. These laws prohibit the processing of sensitive personal data such as racial or ethnic origin, political or religious beliefs, biometric identifiers, health information, and sexual orientation. The study also adhered to the ethical principles outlined in the World Medical Association's Declaration of Helsinki 2014 [50]. All procedures conducted in this study complied with applicable ethical standards. Informed consent was obtained from all participants prior to their inclusion in the study. Participant privacy and confidentiality were strictly protected, and all data were handled and stored in accordance with relevant data protection regulations. No personally identifiable information was collected or disclosed. Participants did not receive any financial or material compensation for their participation.

Results

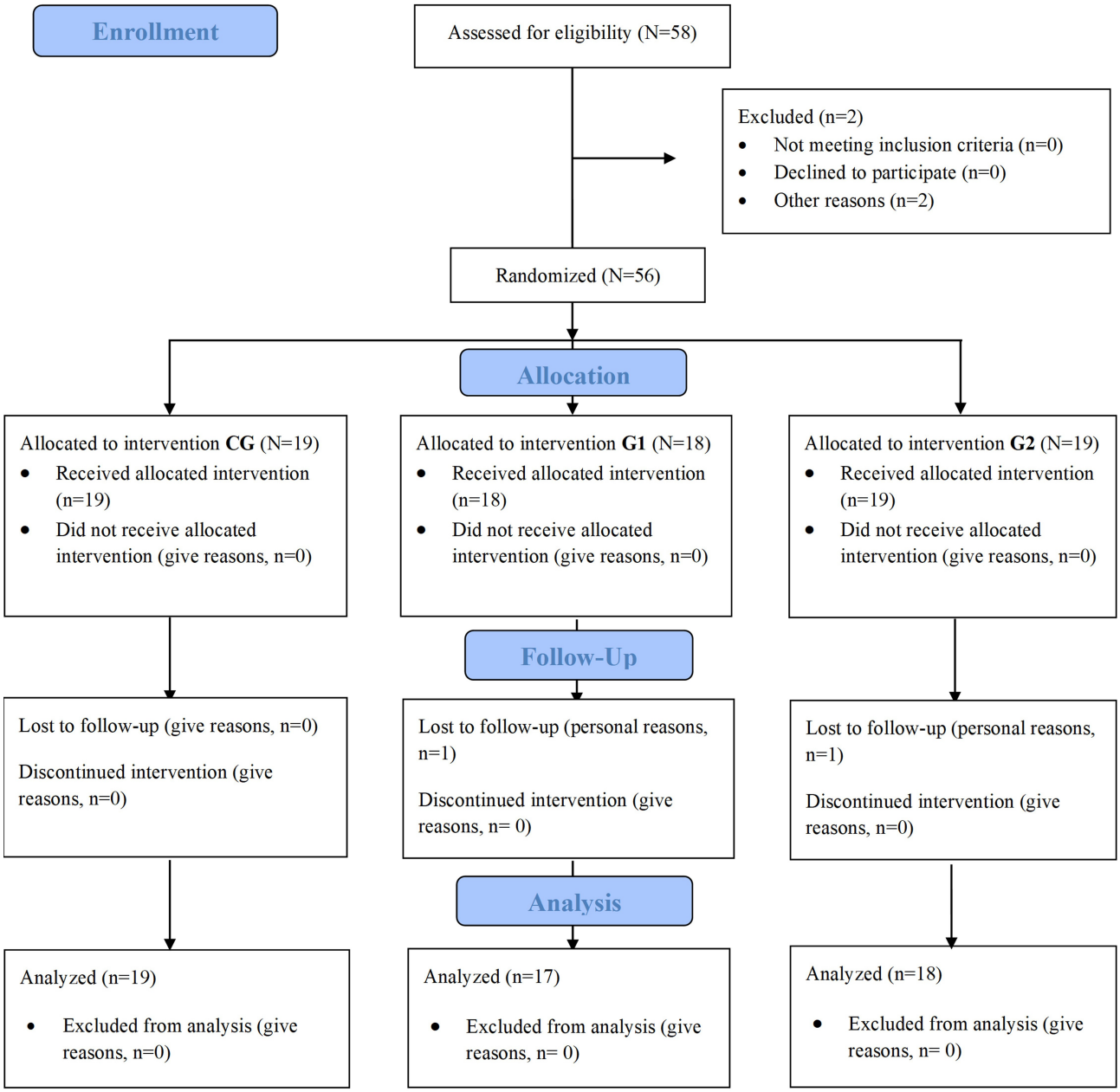
Participant and Baseline Characteristics

The final sample included 54 women with a mean age of 54.26 (SD 7.7) years. Participants were distributed among the groups by computer-generated random number sequence (Figure 1).

Among the participants, 33 (61%) presented chronic fatigue syndrome, 37 (68%) had temporomandibular joint dysfunction, and 16 (33%) reported restless legs syndrome. Regarding the psychological comorbidities, 33 (61%) participants presented anxiety and 13 (24%) depression. Baseline comparisons revealed no statistically significant differences among groups in age, BMI, or other demographic variables. However, subjective measures showed group-level differences in several baseline scales, including FIQ, EQ-5D, CSI, NDI, and Borg scale. Almost all variables at baseline start from similar values, although G1 has a value of 1 point lower on the VAS scale, indicating that they have less pain at the outset. However, in variables such as algometry, the averages reflect similar values compared to the other 2 groups (Multimedia Appendix 2).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. CG: control group.



Postintervention Outcomes

Table 1 shows the differences in the mean values measured before therapy and immediately after its end. We found significant differences in all groups that received intervention in FIQ, CSI, NDI, right and left trapezius, occipital algometer, and in all movements in which ROM was evaluated. For the right and left trapezius and occipital muscles, algometric measurements were obtained, and ROM was assessed across all evaluated movements. About the treatment adherence variable, G1 obtained an average of 94 points out of 100, compared to G2, whose post-treatment average was 74.72. The summary of the comparison between baseline measures and measures immediately after intervention (intragroup comparisons). Shown are the differences in the mean values between both time points (Multimedia Appendix 3).

XSL•FO
RenderX

Table . Intragroup (G1, G2, and CG[a]) analysis of variables measured using questionnaires between baseline and immediately after intervention.

Value and intervention group	Differences of mean values (95% CI)	*P* value
FIQ[b] (1-100)		
Whole series	5.4 (3.2 to 7.5)	<.001
G1	9 (4.6 to 13.3)	<.001
G2	8.1 (5.5 to 10.6)	<.001
CG	0.36 (−3.2 to 2.51)	.80
EQ-5D[c] (–0.59 to 1)		
Whole series	4.1 (2.6 to 5.6)	<.001
G1	3.1 (0.6 to 5.6)	.01
G2	7.7 (5.6 to 9.8)	<.001
CG	−1.7 (−4.4 to 1)	.20
TSK[d] (17-37)		
Whole series	3.2 (0.8 to 5.6)	.01
G1	1.2 (−5.2 to 7.6)	.70
G2	6.2 (4.7 to 9.8)	<.001
CG	−2.04 (−5.9 to 1.86)	.30
CSI[e] (0-100)		
Whole series	4.2 (2.4 to 6)	<.001
G1	6 (3.5 to 8.5)	<.001
G2	7.2 (4.7 to 9.8)	<.001
CG	0.32 (−3.5 to 2.9)	.80
NDI[f] (0-50)		
Whole series	4.5 (3.1 to 5.9)	<.001
G1	5.35 (3.2 to 7.5)	<.001
G2	7.4 (5.1 to 9.7)	<.001
CG	−0.95 (−2.9 to 1.05)	.33

[a]CG: control group.

[b]FIQ: Fibromyalgia Impact Questionnaire.

[c]EQ-5D: EuroQol 5-Dimensions.

[d]TSK: Tampa Scale for Kinesiophobia.

[e]CSI: Central Sensitization Inventory.

[f]NDI: Neck Disability Index Questionnaire.

Follow-Up at 1 Month

Summary of the comparison between baseline and 1-month post-intervention scores for FIQ, EQ-5D, TSK, CSI, and NDI (intragroup comparisons). Shown are the differences in the mean values between both time points (Table 2). The differences in the measurement times of the other variables can be observed in Multimedia Appendix 4.

XSL•FO
RenderX

Table . Intragroup (G1, G2, and CG[a]) analysis of variables measured using questionnaires between baseline and 1 month after intervention.

Value and intervention group	Differences of mean values (95% CI)	*P* value
FIQ[b]		
Whole series	4.4 (2.5 to 76.3)	<.001
G1	10.71 (6.6 to 14.8)	<.001
G2	2.71 (0.55 to 4.87)	.02
CG	0.33 (–2.08 to 1.42)	.70
EQ-5D[c]		
Whole series	4.7 (2.6 to 5.7)	<.001
G1	7.7 (5.46 to 9.97)	<.001
G2	4.17 (1.7 to 6.63)	.002
CG	1.07 (–3.87 to 1.73)	.43
TSK[d]		
Whole series	3.3 (1.4 to 5.1)	.001
G1	8.41 (5.25 to 11.58)	<.001
G2	2.1 (0.06 to 4.13)	.04
CG	0.2 (–2.87 to 3.28)	.89
CSI[e]		
Whole series	3.9 (2.18 to 5.8)	<.001
G1	9.78 (7.35 to 12.22)	<.001
G2	2.94 (1.25 to 4.64)	.02
CG	0.26 (–2.76 to 3.3)	.85
NDI[f]		
Whole series	4.03 (2.5 to 6.31)	<.001
G1	10.65 (8.35 to 12.94)	<.001
G2	4.22 (2.39 to 6.06)	<.001
CG	2.05 (–0.33 to 4.43)	.08

[a]CG: control group.

[b]FIQ: Fibromyalgia Impact Questionnaire.

[c]EQ-5D: EuroQol 5-Dimensions.

[d]TSK: Tampa Scale for Kinesiophobia.

[e]CSI: Central Sensitization Inventory.

[f]NDI: Neck Disability Index Questionnaire.

After confirming the effect of both interventions, we compared the intervention groups to see whether VR had a significant influence on the effect of the therapy. Table 3 shows that the ANOVA between groups was significant for FIQ, CSI, and NDI at both postintervention and 1-month follow-up (*P*<.001). Immediately after treatment, both intervention groups (G1 and G2) improved significantly compared with the CG, with no differences between G1 and G2. At 1 month, G1 maintained superior outcomes across all measures, showing significantly better scores than both GC and G2 (*P*<.001), while G2 no longer differed from GC for FIQ and CSI. For NDI, all pairwise comparisons were significant (*P*<.001), confirming that both interventions reduce neck disability, with the VR-enhanced program providing the greatest and most sustained benefit.

XSL•FO
RenderX

Table . Summary of the comparison between the groups for the change in FIQ[a], CSI[b], and NDI[c], both immediately after intervention and 1 month after.

Variable and intervention group	*P* value for ANOVA	Mean difference between time points (95% CI)	*P* value
FIQ			
Immediate	<.001		
CG[d] vs G1		9.31 (14.7 to 3.8)	<.001
CG vs G2		8.4 (13.84 to 3.06)	.001
G2 vs G1		0.85 (6.39 to −4.68)	>.99
1 month	<.001		
CG vs G1		10.38 (5.8 to 14.9)	<.001
CG vs G2		2.38 (–2.12 to 6.9)	.59
G2 vs G1		7.99 (3.35 to 12.63)	<.001
CSI			
Immediate	<.001		
CG vs G1		6.31 (10.96 to 1.66)	.04
CG vs G2		7.53 (12.12 to 2.95)	<.001
G2 vs G1		–1.22 (3.4 to −5.93)	.51
1 month	<.001		
CG vs G1		10.04 (14.16 to 5.93)	<.001
CG vs G2		3.20 (7.26 to −0.84)	.16
G2 vs G1		6.84 (11.01 to 2.67)	.001
NDI			
Immediate	<.001		
CG vs G1		4.40 (7.95 to 0.85)	.01
CG vs G2		6.44 (9.93 to 2.94)	<.001
G2 vs G1		–2.03 (1.55 to −5.62)	.50
1 month	<.001		
CG vs G1		12.69 (16.33 to 9.06)	<.001
CG vs G2		6.27 (9.85 to 2.69)	<.001
G2 vs G1		6.42 (10.10 to 2.74)	<.001

[a]FIQ: Fibromyalgia Impact Questionnaire.

[b]CSI: Central Sensitization Inventory.

[c]NDI: Neck Disability Index Questionnaire.

[d]CG: control group.

Discussion

Principal Findings

This study evaluated the effectiveness of VR as an adjunctive therapy for cervical pain in women with FM. The main findings revealed that combining VR with mobility and strengthening exercises produced significant improvements in pain perception, ROM, and functional performance compared with physical therapy alone. These benefits were also sustained at 1-month follow-up, emphasizing the potential of immersive VR interventions to enhance therapeutic outcomes in individuals with FM.

In the intergroup analysis, the G1 consistently outperformed both G2 and CG in those outcomes most closely related to global disease impact and NDI. Specifically, G1 showed significantly greater and more sustained improvements in FIQ and CSI scores than both CG and G2, whereas improvements in G2 tended to converge toward control values at 1-month follow-up. Similarly, NDI scores decreased in both intervention groups immediately after treatment, but at follow-up, a clear gradient GC>G2>G1 was observed, indicating the largest and most persistent reduction in neck-related disability in the combined intervention. These effects were paralleled by a recurrent pattern of superior gains in cervical ROM and higher-pressure pain thresholds in the trapezius and occipital in G1, while the Time Up Go test

XSL•FO
RenderX

improved in both intervention groups compared with CG and TSK decreased more markedly in G1 at 1 month. By contrast, the VAS and Borg scale showed only limited discrimination between the groups, suggesting that the added value of VR was more evident in multidimensional impact and sensitization-related and functional domains than in isolated pain ratings.

This intergroup pattern is consistent with previous evidence in FM and chronic neck pain, indicating that VR, particularly when used as an adjunct to active exercise, preferentially enhances global impact, disability, and pain-modulation outcomes. In patients with FM, a study reported that fully immersive VR combined with aerobic and pilates training produced greater improvements in pain, kinesiophobia, fatigue, physical activity levels, and mental quality of life than exercise alone, while both groups improved in FM impact, supporting an adjunctive role of immersive VR in comprehensive rehabilitation programs [8]. Other studies showed that, in individuals with chronic neck pain, VR-based cervical training led to larger gains in pressure pain thresholds at the upper cervical levels and greater reductions in functional limitation compared with motor control exercises, despite the absence of between-group differences in pain intensity and quality of life, which parallels our finding of more robust intergroup differences in pressure pain thresholds, ROM, and disability than in VAS score [51]. Furthermore, a randomized crossover trial in women with FM found that VR increased cold pain thresholds and tolerance in both FM patients and pain-free controls, whereas effects on pain intensity were limited, reinforcing the notion that VR may primarily modulate pain processing rather than consistently decreasing reported pain intensity [52]. Taken together, these converging results suggest that combining immersive VR with targeted cervical exercise may primarily potentiate mechanisms related to central sensitization, motor control, and the global impact of FM rather than merely amplifying short-term analgesic effects.

VR has emerged as a promising nonpharmacological intervention for the symptom management of FM. Its immersive and interactive features provide cognitive distraction, emotional engagement, and motor stimulation, which together enhance adherence and reduce pain intensity. Previous studies have demonstrated that VR can improve mood, motivation, and functional capacity when applied alongside traditional rehabilitation or psychological therapies. This approach not only benefits patients with FM but may also be applicable to other chronic pain conditions [49,50,53].

The use of VR has been demonstrated to significantly influence pain relief, motor function, and joint mobility among patients with a range of chronic pathologies [53]. The primary applications of VR in health care include the management of pain and anxiety as well as the enhancement of patient motivation [47,54,55].

While it is not possible to make direct comparisons between this study and research conducted by other authors, the described implementation of an 8-session intervention resulted in statistically significant outcomes for the variables under investigation. This suggests VR is effective when combined with active exercise therapy.

The comparative analysis concerning VR and physical exercise indicated that there was minimal divergence in the measured variables. In a separate study [56], 44 patients underwent an intervention for 4 weeks, with 2 sessions per week. The intervention was divided into 2 groups: G1 (whose members performed only cervical mobility exercises) and G2 (whose members conducted a treatment based solely on VR exercises). The findings indicated that there were no statistically significant differences between the 2 groups across the different variables, including pain, ROM, neck disability, incapacitating pain, and anxiety. However, in this study, the combination of VR with cervical mobility exercises yielded a significantly better outcome. Furthermore, these improvements were maintained over time (1 month) to a greater degree in the group subjected to both therapies combined when compared to the group that performed only mobility exercises. It is important to highlight that the study with which this investigation was compared implemented its intervention in a population without FM. Consequently, the observed discrepancy in outcomes could be due to this factor. A similar study was conducted [51], with an equivalent number of sessions implemented. The outcomes observed in this research were analogous to those obtained in our research, as the investigators combined active therapy and VR exercises for patients with FM [55]. However, the FM cohort in that report comprised only 20 patients. In addition, the active component focused on aerobic training and Pilates instead of the cervical exercises used in our research.

Another study assessed the efficacy of VR for patients with chronic neck pain versus a CG that performed motor control exercises. In this case, the study spanned a period of 6 weeks, comprising 3 weekly sessions. Their findings revealed that the utilization of VR was advantageous when measuring pain due to pressure with an algometer at C1-C2 and C5-C6, and it was also beneficial when considering functional limitations. However, no significant differences were observed between the groups with respect to pain intensity, muscle performance, or quality of life (36-Item Short Form Health Survey). It is evident that the results of this study deviate from those previously observed. However, it should be noted that the mentioned study did not include the population with FM, and no group underwent combined VR and active therapy [51].

Following a thorough review of the scientific literature, we found no studies addressing the use of VR in the cervical region using this protocol. Other studies have applied VR in FM use, games, and global exercises, focusing on observing aspects such as anaerobic capacity, balance, or fatigue [7,54,55]. In this study, the focus was pain in the cervical region and other more global variables such as fatigue, VAS, and FIQ. Therefore, our study offers a novel contribution to this field of research by addressing this gap in the existing literature. Given that patients with FM experience cervical pain, headaches, and tender points in the occipital regions, it is imperative to reduce the associated symptoms.

The limitations of this study lie in the fact that the results are not representative of a heterogeneous population with FM, since the cohort of participants in this research consisted solely of women. Due to the challenging nature of the disease, it was not possible to consider medication used by participants, which

XSL•FO
RenderX

may have influenced the results. Finally, the study did not include a group that used only VR, so we cannot conclude that the results are attributable to the VR component or its combination with exercise.

This approach not only benefits patients with FM but may also be applicable to other chronic pain conditions [49,50,53].

Conclusions

Virtual reality, when added to cervical mobility and strengthening exercises, produced greater and more sustained improvements in disease impact, neck disability, central sensitization, kinesiophobia, and functional outcomes in women with FM than exercise alone or no intervention. The improvements were maintained for 1 month in the G1 group in the variables evaluated. These findings support VR as an effective adjunct to active therapy, with potential benefits for symptom reduction and warrant further research on long-term effects, cost-effectiveness, and applicability to broader chronic pain populations.

Funding

This research received external funding for the FIBYSAR project of Camilo Jose Cela University.

Data Availability

The data presented in this study are available upon reasonable request from the corresponding author.

Authors' Contributions

Conceptualization: AN-M, CO-M, EÚ-D, YM-C
Data curation: BP-R, EC-F-P, EÚ-D, JPH-P
Formal analysis: JH-L, JPH-P, MJF-A, NM-G
Investigation: BP-R, EC-F-P, EÚ-D, JH-L, JPH-P, NM-G, YM-C
Methodology: EC-F-P, JH-L, JPH-P, MC-C
Resources: CO-M, EC-F-P, EÚ-D, JPH-P
Software: AN-M, JH-L, YM-C
Supervision: CO-M, EÚ-D, JPH-P
Validation: EÚ-D, EC-F-P, JPH-P, MC-C, NM-G
Visualization: EÚ-D, JPH-P, MC-C, YM-C
Writing – original draft: AN-M, CO-M, EÚ-D, NM-G, YM-C
Writing – review & editing: BP-R, EC-F-P, EÚ-D, JH-L, JPH-P, MJF-A

Conflicts of Interest

None declared.

Multimedia Appendix 1
Recommended exercises.
[DOCX File, 15 KB - rehab_v13i1e81158_app1.docx]

Multimedia Appendix 2
Baseline measurements of pain and functional capacity (mean and SD).
[DOCX File, 23 KB - rehab_v13i1e81158_app2.docx]

Multimedia Appendix 3
Intragroup analysis of baseline and immediately after intervention.
[DOCX File, 20 KB - rehab_v13i1e81158_app3.docx]

Multimedia Appendix 4
Intragroup analysis of baseline and 1 month after intervention.
[DOCX File, 26 KB - rehab_v13i1e81158_app4.docx]

Checklist 1
CONSORT-Harms-2022 checklist.
[PDF File, 208 KB - rehab_v13i1e81158_app5.pdf]

References

XSL•FO
RenderX

1. Richard JY, Hurley RA, Taber KH. Fibromyalgia: Centralized pain processing and neuroimaging. J Neuropsychiatry Clin Neurosci 2019;31(3):A6-187. [doi: 10.1176/appi.neuropsych.19050107] [Medline: 31322995]
2. Häuser W, Fitzcharles MA. Facts and myths pertaining to fibromyalgia. Dialogues Clin Neurosci 2018 Mar;20(1):53-62. [doi: 10.31887/DCNS.2018.20.1/whauser] [Medline: 29946212]
3. Solà JF. El síndrome de fatiga crónica [Article in Spanish]. Medicina Integral 2002;40(2):56-63 [FREE Full text]
4. Sumpton JE, Moulin DE. Fibromyalgia. Handb Clin Neurol 2014;119:513-527. [doi: 10.1016/B978-0-7020-4086-3.00033-3] [Medline: 24365316]
5. Ruschak I, Montesó-Curto P, Rosselló L, Aguilar Martín C, Sánchez-Montesó L, Toussaint L. Fibromyalgia syndrome pain in men and women: a scoping review. Healthcare (Basel) 2023 Jan 11;11(2):223. [doi: 10.3390/healthcare11020223] [Medline: 36673591]
6. Liang D, Guo X, Zhang J, Jiang R. Pain characteristics of patients with fibromyalgia: a comparison between gender and different emotional states. Pain Physician 2024 Jan 20;27(1):E109-E118. [doi: 10.36076/ppj.2024.27.E109] [Medline: 38285043]
7. Cabo-Meseguer A, Cerdá-Olmedo G, Trillo-Mata JL. Fibromyalgia: prevalence, epidemiologic profiles and economic costs. Med Clin 2017 Nov;149(10):441-448. [doi: 10.1016/j.medcli.2017.06.008]
8. Gulsen C PT, MSc, Soke F PT, PhD, Eldemir K PT, MSc, et al. Effect of fully immersive virtual reality treatment combined with exercise in fibromyalgia patients: a randomized controlled trial. Assist Technol 2022 May 4;34(3):256-263. [doi: 10.1080/10400435.2020.1772900] [Medline: 32543290]
9. Villanueva VL, Valía JC, Cerdá G, et al. Fibromialgia: diagnóstico y tratamiento. El estado de la cuestión [Article in Spanish]. Rev Soc Esp Dolor 2004;11(7):430-443 [FREE Full text]
10. Borchers AT, Gershwin ME. Fibromyalgia: A critical and comprehensive review. Clinic Rev Allerg Immunol 2015 Oct;49(2):100-151. [doi: 10.1007/s12016-015-8509-4]
11. Wang JJ, Tam KW, Hsiao HY, Liou TH, Rau CL, Hsu TH. Effect of resistance exercises on function and pain in fibromyalgia. Am J Phys Med Rehabil 2024;103(4):275-283. [doi: 10.1097/PHM.0000000000002318]
12. Giesecke T, Gracely RH, Grant MAB, et al. Evidence of augmented central pain processing in idiopathic chronic low back pain. Arthritis Rheum 2004 Feb;50(2):613-623. [doi: 10.1002/art.20063]
13. Yunus MB. The prevalence of fibromyalgia in other chronic pain conditions. Pain Res Treat 2012;2012:584573. [doi: 10.1155/2012/584573] [Medline: 22191024]
14. Prikhodkina M, Melnikov S. Factors that influence medication adherence in women with fibromyalgia: a path analysis. J Clin Nurs 2024 Oct;33(10):3943-3953. [doi: 10.1111/jocn.17044] [Medline: 38284436]
15. Desai R, Marlow N. Effect of non-adherence on healthcare resource utilization in fibromyalgia patients: a retrospective cohort study. Value Health 2018 May;21:S255. [doi: 10.1016/j.jval.2018.04.1776]
16. Dobkin PL, Sita A, Sewitch MJ. Predictors of adherence to treatment in women with fibromyalgia. Clin J Pain 2006;22(3):286-294. [doi: 10.1097/01.ajp.0000173016.87612.4b] [Medline: 16514330]
17. Iglesias AR, Soria Ayuda RE, Martinez AB, Sánchez AJ, Villaroya Bielsa E, Callejero Guillen AJ. La efectividad de un programa de ejercicio terapéutico en pacientes con fibromialgia [Web page in Spanish]. Revista Sanitaria de Investigacion. URL: https://revistasanitariadeinvestigacion.com/la-efectividad-de-un-programa-de-ejercicio-terapeutico-en-pacientes-con-fibromialgia/ [accessed 2025-12-30]
18. Winkelmann A, Häuser W, Friedel E, et al. Physiotherapy and physical therapies for fibromyalgia syndrome. Systematic review, meta-analysis and guideline [Title in German]. Schmerz 2012 Jun;26(3):276-286. [doi: 10.1007/s00482-012-1171-3] [Medline: 22760460]
19. Sarac AJ, Gur A. Complementary and alternative medical therapies in fibromyalgia. Curr Pharm Des 2006;12(1):47-57. [doi: 10.2174/138161206775193262] [Medline: 16454724]
20. Collado-Mateo D, Dominguez-Muñoz FJ, Adsuar JC, Merellano-Navarro E, Gusi N. Exergames for women with fibromyalgia: a randomised controlled trial to evaluate the effects on mobility skills, balance and fear of falling. PeerJ 2017;5:e3211. [doi: 10.7717/peerj.3211] [Medline: 28439471]
21. Ahmadpour N, Randall H, Choksi H, Gao A, Vaughan C, Poronnik P. Virtual reality interventions for acute and chronic pain management. Int J Biochem Cell Biol 2019 Sep;114(September):105568. [doi: 10.1016/j.biocel.2019.105568] [Medline: 31306747]
22. Huang Q, Lin J, Han R, Peng C, Huang A. Using virtual reality exposure therapy in pain management: a systematic review and meta-analysis of randomized controlled trials. Value Health 2022 Feb;25(2):288-301. [doi: 10.1016/j.jval.2021.04.1285] [Medline: 35094802]
23. Cortés-Pérez I, Zagalaz-Anula N, Ibancos-Losada MDR, Nieto-Escámez FA, Obrero-Gaitán E, Osuna-Pérez MC. Virtual reality-based therapy reduces the disabling impact of fibromyalgia syndrome in women: systematic review with meta-analysis of randomized controlled trials. J Pers Med 2021 Nov 9;11(11):1167. [doi: 10.3390/jpm11111167] [Medline: 34834518]
24. Carvalho MSD, Carvalho LC, Menezes FDS, Frazin A, Gomes EDC, Iunes DH. Effects of exergames in women with fibromyalgia: a randomized controlled study. Games Health J 2020 Oct 1;9(5):358-367. [doi: 10.1089/g4h.2019.0108]
25. Gómez-Pérez L, López-Martínez AE, Ruiz-Párraga GT. Psychometric properties of the Spanish version of the Tampa Scale for Kinesiophobia (TSK). J Pain 2011 Apr;12(4):425-435. [doi: 10.1016/j.jpain.2010.08.004] [Medline: 20926355]

XSL•FO
RenderX

26. Sato K, Fukumori S, Matsusaki T, et al. Nonimmersive virtual reality mirror visual feedback therapy and its application for the treatment of complex regional pain syndrome: an open-label pilot study. Pain Med 2010 Apr;11(4):622-629. [doi: 10.1111/j.1526-4637.2010.00819.x] [Medline: 20202141]
27. Freeman D, Reeve S, Robinson A, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017 Oct;47(14):2393-2400. [doi: 10.1017/S003329171700040X] [Medline: 28325167]
28. Lo HHM, Zhu M, Zou Z, et al. Immersive and nonimmersive virtual reality-assisted active training in chronic musculoskeletal pain: systematic review and meta-analysis. J Med Internet Res 2024 Aug 19;26:e48787. [doi: 10.2196/48787] [Medline: 39159449]
29. Harvie DS, Broecker M, Smith RT, Meulders A, Madden VJ, Moseley GL. Bogus visual feedback alters onset of movement-evoked pain in people with neck pain. Psychol Sci 2015 Apr;26(4):385-392. [doi: 10.1177/0956797614563339] [Medline: 25691362]
30. Trost Z, Zielke M, Guck A, et al. The promise and challenge of virtual gaming technologies for chronic pain: the case of graded exposure for low back pain. Pain Manag 2015;5(3):197-206. [doi: 10.2217/pmt.15.6] [Medline: 25971643]
31. Sarig Bahat H, Takasaki H, Chen X, Bet-Or Y, Treleaven J. Cervical kinematic training with and without interactive VR training for chronic neck pain—a randomized clinical trial. Man Ther 2015 Feb;20(1):68-78. [doi: 10.1016/j.math.2014.06.008] [Medline: 25066503]
32. Ubillos-Landa S, García-Otero R, Puente-Martínez A. Validation of an instrument for measuring chronic pain in nursing homes. An Sist Sanit Navar 2019 Apr 25;42(1):19-30. [doi: 10.23938/ASSN.0390] [Medline: 30706908]
33. Úbeda-D'Ocasar E, Valera-Calero JA, Hervás-Pérez JP, Caballero-Corella M, Ojedo-Martín C, Gallego-Sendarrubias GM. Pain intensity and sensory perception of tender points in female patients with fibromyalgia: a pilot study. Int J Environ Res Public Health 2021 Feb 4;18(4):1461. [doi: 10.3390/ijerph18041461] [Medline: 33557288]
34. Ayala Pastorino S, Varaldi Azcoytia G, Illescas Calegari L, Castroman Espasandín Espasandin PJ. Test cuantitativo sensorial con algometría de presión: experiencia preliminar en un servicio de anestesiología [Article in Spanish]. Rev Soc Esp Dolor 2022;28(6):311-318. [doi: 10.20986/resed.2022.3947/2021]
35. Ge HY, Nie H, Madeleine P, Danneskiold-Samsøe B, Graven-Nielsen T, Arendt-Nielsen L. Contribution of the local and referred pain from active myofascial trigger points in fibromyalgia syndrome. Pain 2009 Dec 15;147(1-3):233-240. [doi: 10.1016/j.pain.2009.09.019] [Medline: 19819074]
36. Walthall H, Jenkinson C, Boulton M. Living with breathlessness in chronic heart failure: a qualitative study. J Clin Nurs 2017 Jul;26(13-14):2036-2044. [doi: 10.1111/jocn.13615] [Medline: 27731919]
37. Jovicic MD, Konstantinovic LM, Grgurevic AD, et al. Validation of the Neck Disability Index in Serbian patients with cervical radiculopathy. J Manipulative Physiol Ther 2018;41(6):496-502. [doi: 10.1016/j.jmpt.2017.10.018] [Medline: 30107938]
38. Roelofs J, Sluiter JK, Frings-Dresen MHW, et al. Fear of movement and (re)injury in chronic musculoskeletal pain: evidence for an invariant two-factor model of the Tampa Scale for Kinesiophobia across pain diagnoses and Dutch, Swedish, and Canadian samples. Pain 2007 Sep;131(1-2):181-190. [doi: 10.1016/j.pain.2007.01.008] [Medline: 17317011]
39. Newman-Beinart NA, Norton S, Dowling D, et al. The development and initial psychometric evaluation of a measure assessing adherence to prescribed exercise: the Exercise Adherence Rating Scale (EARS). Physiotherapy 2017 Jun;103(2):180-185. [doi: 10.1016/j.physio.2016.11.001] [Medline: 27913064]
40. Gauffin J, Hankama T, Kautiainen H, Arkela-Kautiainen M, Hannonen P, Haanpää M. Validation of a Finnish version of the Fibromyalgia Impact Questionnaire (Finn-FIQ). Scand J Pain 2012 Jan 1;3(1):15-20. [doi: 10.1016/j.sjpain.2011.10.004] [Medline: 29913759]
41. Monterdea S, Salvata I, Montulla S, Fernández-Ballart J. Validación de la versión española del Fibromyalgia Impact Questionnaire [Article in Spanish]. Rev Esp Reumatol 2004;31:507-513 [FREE Full text]
42. Esteve-Vives J, Rivera Redondo J, Isabel Salvat Salvat M, de Gracia Blanco M, de Miquel CA. Propuesta de una versión de consenso del Fibromyalgia Impact Questionnaire (FIQ) para la población española. Reumatol Clín 2007 Jan;3(1):21-24. [doi: 10.1016/S1699-258X(07)73594-5] [Medline: 21794391]
43. Herdman M, Badia X, Berra S. El EuroQol-5D: una alternativa sencilla para la medición de la calidad de vida relacionada con la salud en atención primaria. Aten Primaria 2001;28(6):425-429. [doi: 10.1016/S0212-6567(01)70406-4]
44. Wolfe F, Michaud K, Li T, Katz RS. EQ-5D and SF-36 quality of life measures in systemic lupus erythematosus: comparisons with rheumatoid arthritis, noninflammatory rheumatic disorders, and fibromyalgia. J Rheumatol 2010 Feb;37(2):296-304. [doi: 10.3899/jrheum.090778] [Medline: 20032098]
45. Carralero García P, Rosa Hoyos Miranda F, Deblas Sandoval Á, López García M. Calidad del sueño según el Pittsburgh Sleep Quality Index en una muestra de pacientes recibiendo cuidados paliativos. Med Paliat 2013;20(2):44-48. [doi: 10.1016/j.medipa.2012.05.005]
46. Salaffi F, Farah S, Mariani C, Sarzi-Puttini P, Di Carlo M. Validity of the Central Sensitization Inventory compared with traditional measures of disease severity in fibromyalgia. Pain Pract 2022 Nov;22(8):702-710. [doi: 10.1111/papr.13162] [Medline: 36097821]
47. Prushansky T, Dvir Z. La prueba de la movilidad cervical: metodología e implicaciones clínicas [Article in Spanish]. Osteopat Cient 2008;3:108-114. [doi: 10.1016/S1886-9297(08)75759-X]

XSL•FO
RenderX

48. Collado-Mateo D, Domínguez-Muñoz FJ, Adsuar JC, Merellano-Navarro E, Olivares PR, Gusi N. Reliability of the timed up and go test in fibromyalgia. Rehabil Nurs 2018;43(1):35-39. [doi: 10.1002/rnj.307] [Medline: 27781288]
49. Beaudart C, Rolland Y, Cruz-Jentoft AJ, et al. Assessment of muscle function and physical performance in daily clinical practice: a position paper endorsed by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO). Calcif Tissue Int 2019;105(1):1-14. [doi: 10.1007/s00223-019-00545-w]
50. Declaración de Helsinki de la AMM [Web page in Spanish]. Asociación Médica Mundial. 2024. URL: https://www.wma.net/es/policies-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/ [accessed 2026-01-05]
51. Cetin H, Kose N, Oge HK. Virtual reality and motor control exercises to treat chronic neck pain: a randomized controlled trial. Musculoskelet Sci Pract 2022 Dec;62(August):102636. [doi: 10.1016/j.msksp.2022.102636] [Medline: 35952621]
52. Christensen SWM, Almsborg H, Vain TS, Vaegter HB. The effect of virtual reality on cold pain sensitivity in patients with fibromyalgia and pain-free individuals: a randomized crossover study. Games Health J 2023 Aug;12(4):295-301. [doi: 10.1089/g4h.2022.0138] [Medline: 36454199]
53. Garcia-Palacios A, Herrero R, Vizcaíno Y, et al. Integrating virtual reality with activity management for the treatment of fibromyalgia: acceptability and preliminary efficacy. Clin J Pain 2015 Jun;31(6):564-572. [doi: 10.1097/AJP.0000000000000196] [Medline: 25551475]
54. Darnall BD, Krishnamurthy P, Tsuei J, Minor JD. Self-administered skills-based virtual reality intervention for chronic pain: randomized controlled pilot study. JMIR Form Res 2020 Jul 7;4(7):e17293. [doi: 10.2196/17293] [Medline: 32374272]
55. Herrero R, García-Palacios A, Castilla D, Molinari G, Botella C. Virtual reality for the induction of positive emotions in the treatment of fibromyalgia: a pilot study over acceptability, satisfaction, and the effect of virtual reality on mood. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):379-384. [doi: 10.1089/cyber.2014.0052] [Medline: 24892201]
56. Tejera DM, Beltran-Alacreu H, Cano-de-la-Cuerda R, et al. Effects of virtual reality versus exercise on pain, functional, somatosensory and psychosocial outcomes in patients with non-specific chronic neck pain: a randomized clinical trial. Int J Environ Res Public Health 2020 Aug 16;17(16):5950. [doi: 10.3390/ijerph17165950] [Medline: 32824394]

Abbreviations

CG: control group
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-Harms: Consolidated Standards of Reporting Trials—Harms Extension
CSI: Central Sensitization Inventory
EQ-5D: EuroQol Dimensions
FIQ: Fibromyalgia Impact Questionnaire
FM: fibromyalgia
NDI: Neck Disability Index Questionnaire
ROM: range of motion
TSK: Tampa Scale for Kinesiophobia
VAS: visual analog scale
VR: virtual reality

Edited by S Munce; submitted 23.Jul.2025; peer-reviewed by D Parmak, GGD Alencar; revised version received 10.Dec.2025; accepted 19.Dec.2025; published 23.Jan.2026.

Please cite as:
Úbeda-D'Ocasar E, Moreno-Crespo Y, Cimadevilla-Fernández-Pola E, Hernández-Lougedo J, Navas-Mosqueda Á, Caballero-Corella M, Mayoral-Gonzalo N, Pedauyé-Rueda B, Fernández-Aceñero MJ, Hervás-Pérez JP, Ojedo-Martín C
Therapeutic Use of Virtual Reality for Patients With Fibromyalgia and Chronic Neck Pain: Randomized Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e81158
URL: https://rehab.jmir.org/2026/1/e81158
doi:10.2196/81158

© Edurne Úbeda-D'Ocasar, Yaiza Moreno-Crespo, Eduardo Cimadevilla-Fernández-Pola, Juan Hernández-Lougedo, Álvaro Navas-Mosqueda, Mario Caballero-Corella, Noemí Mayoral-Gonzalo, Blanca Pedauyé-Rueda, María Jesús Fernández-Aceñero, Juan Pablo Hervás-Pérez, Cristina Ojedo-Martín. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Gaze Allocation and Performance Across Task-Demand Conditions During Squat-Based Exergaming: Pilot Study Using Eye Tracking

Keiichi Takei[1], PhD

Department of Physical Therapy, Faculty of Rehabilitation, Tokyo Professional University of Health Sciences, Koto-ku, Tokyo, Japan

Corresponding Author:
Keiichi Takei, PhD
Department of Physical Therapy
Faculty of Rehabilitation
Tokyo Professional University of Health Sciences
2-22-10 Shiohama
Koto-ku, Tokyo, 135-0043
Japan
Phone: 81 3 6272 5671
Email: takei.k@tpu.ac.jp

Abstract

Background: Exergames integrate motor, cognitive, and postural-control demands; however, how specific task-demand manipulations influence gaze allocation during exergame performance remains insufficiently characterized.

Objective: This exploratory pilot study used a within-participant design to examine whether gaze allocation and task performance differed across baseline, concurrent cognitive-task, and unilateral-squat conditions during squat-based exergaming in healthy young adults.

Methods: Eight healthy adults (mean age 20, SD 1 years; 7 men and 1 woman) used a squat-based exergame (*Ring Fit Adventure*; Nintendo) under 3 randomized conditions: baseline, concurrent cognitive task, and unilateral squats. In the concurrent cognitive task condition, participants performed serial subtraction during squatting. In the unilateral-squat condition, participants performed single-leg squats, which were intended to increase support-leg muscular demand as well as postural and motor-control requirements. Execution time, squat score, arithmetic performance, and eye-tracking metrics were recorded. Primary gaze outcomes were the proportion of fixation time and fixation counts allocated to predefined areas of interest (AOIs; command, avatar, and score) and to regions outside these areas across the entire trial. Differences from the baseline condition were examined using Dunnett tests or paired Wilcoxon signed-rank tests with Holm adjustment, and effect sizes were reported.

Results: Exergame execution time increased from 31.0 (SD 3.3) seconds in the baseline condition to 38.2 (SD 6.1) seconds in the concurrent cognitive task condition and to 34.3 (SD 6.0) seconds in the unilateral-squat condition. In contrast, the squat score decreased from 98.8 (SD 2.8) in the baseline condition to 69.6 (SD 18.3) in the unilateral-squat condition. For fixation counts, allocation to the command AOI decreased from 44.8% (SD 11.4%) in the baseline condition to 32.1% (SD 8.9%) in the concurrent cognitive task condition and 36.2% (SD 8.5%) in the unilateral-squat condition. Outside AOI fixation counts increased in the concurrent cognitive task condition (55.7%, SD 8.6%) relative to the baseline condition (36%, SD 16.9%).

Conclusions: In this exploratory within-participant pilot study, adding a concurrent cognitive task and performing unilateral squats resulted in different patterns of performance change and fixation count redistribution during squat-based exergaming. These preliminary findings suggest that AOI-based gaze allocation metrics may help characterize task demand–related attentional shifts in this setting. Larger confirmatory studies with more diverse samples and individual-level validation are needed before these metrics can be considered for adaptive or clinical applications.

(JMIR Rehabil Assist Technol 2026;13:e81667) doi:10.2196/81667

KEYWORDS

exergame; eye tracking; gaze allocation; dual task; cognitive load

XSL•FO
RenderX

Introduction

Background

Some studies suggest that certain video game experiences may enhance attention, memory, and executive function [1,2]. Exergames inherently involve dual-task demands because they require simultaneous motor execution and visually guided cognitive processing [3-6]. Evidence from systematic reviews, meta-analyses, and controlled studies suggests that exergame-based dual-task training can improve balance, executive function, and dual-task performance in older adults [7-10]. However, the cognitive demands imposed by exergames can vary substantially depending on specific game features, such as feedback density, pacing, and concurrent cognitive tasks, as well as individual user characteristics. Without adequately characterizing these demands, it is difficult to interpret performance changes or to adjust task difficulty rationally. Quantifying in-game cognitive load may therefore help clarify how specific task manipulations influence attentional allocation during gameplay.

Despite its importance, the internal cognitive load imposed by specific exergame task configurations remains poorly characterized [11]. Most previous studies have relied on performance metrics or subjective workload measures [12,13], which do not directly capture moment-to-moment attentional allocation during gameplay [14]. In serious game and exergaming research, cognitive load has been assessed using complementary approaches, including subjective workload ratings, performance-based indices (eg, dual-task costs and in-game accuracy or speed), and physiological measures (eg, heart rate, pupillometry, and electroencephalogram or functional near-infrared spectroscopy). However, each approach has practical limitations in dynamic whole-body tasks, including limited temporal resolution, reliance on outcome-level inference, and susceptibility to motion artifacts. Eye tracking offers a behavioral marker of moment-to-moment overt visual attentional allocation and may help characterize how specific game manipulations redistribute visual attention across task-relevant display elements.

In dual-task paradigms [15], cognitive load is often inferred from performance decrements such as slower execution time or reduced motor accuracy [16]. Although informative, these measures reflect outcome-level interference rather than the underlying dynamics of attention [17]. Objective real-time indicators of cognitive load during complex motor tasks remain an active area of investigation [18-21].

Eye tracking has increasingly been used to assess attentional allocation under cognitive load in dynamic contexts such as driving [22-24], walking [19,21], and real-world task performance [20,25]. Previous studies have reported load-dependent changes in fixation duration [20], gaze dispersion [26], and the allocation of gaze to task-relevant versus irrelevant stimuli [25]. However, findings have been inconsistent: some studies report longer fixation durations under higher cognitive load, whereas others report reduced fixation time and greater gaze variability [20,27]. In addition, relatively few studies have examined gaze allocation during whole-body exercise tasks or in exergaming contexts [21,28]. Thus, it remains unclear whether different task demand manipulations within an exergame, such as adding a concurrent cognitive task or increasing unilateral motor and postural-control requirements, are associated with distinct and measurable patterns of gaze redistribution across task-relevant display elements.

Objectives

This pilot study aimed to examine whether area of interest (AOI)–specific gaze allocation differs across baseline, concurrent cognitive-task, and unilateral-squat conditions within a squat-based exergame. The study was not designed to isolate purely cognitive versus purely physical load. Rather, it was intended to explore whether AOI-based gaze allocation is sensitive to different task-demand manipulations within a single exergame context. By characterizing condition-dependent patterns of gaze redistribution in healthy young adults, this study sought to provide preliminary effect estimates and hypothesis-generating evidence to inform future confirmatory studies.

Hypotheses

It was expected that the concurrent cognitive-task condition would be associated with redistribution of gaze away from task-relevant AOIs and/or increased viewing outside predefined AOIs. In contrast, the unilateral-squat condition was expected to be associated with altered sampling of task-relevant feedback, particularly the score AOI, because single-leg squatting may increase support-leg muscular demand as well as postural- and motor-control requirements. Given the exploratory pilot design, these hypotheses were directional rather than confirmatory.

Methods

Participants

This study was designed as an exploratory pilot study to obtain preliminary effect size estimates rather than to test hypotheses definitively. Eight young adults (mean age 20, SD 1 years; 7 men and 1 woman) were recruited from Tokyo Professional University of Health Sciences. Participants were healthy university students with no self-reported history of neurological, vestibular, or musculoskeletal disorders that could affect balance or squat performance and no known cognitive impairment. On the testing day, participants confirmed that they were free from acute illness, pain, dizziness, or excessive fatigue that could substantially affect task performance. To ensure eye-tracking data quality, contact lens users were excluded. Participants who usually wore eyeglasses performed the task without glasses because of eye-tracking requirements; those unable to clearly view the display under this condition were excluded.

Ethical Considerations

This study was approved by the Ethics Committee of Tokyo Professional University of Health Sciences (TPU-25-002). All participants provided written informed consent before participation. No direct personal identifiers, such as names or contact information, were included in the study dataset. Participants were assigned study identification numbers, and all study data were managed and analyzed using these numbers.

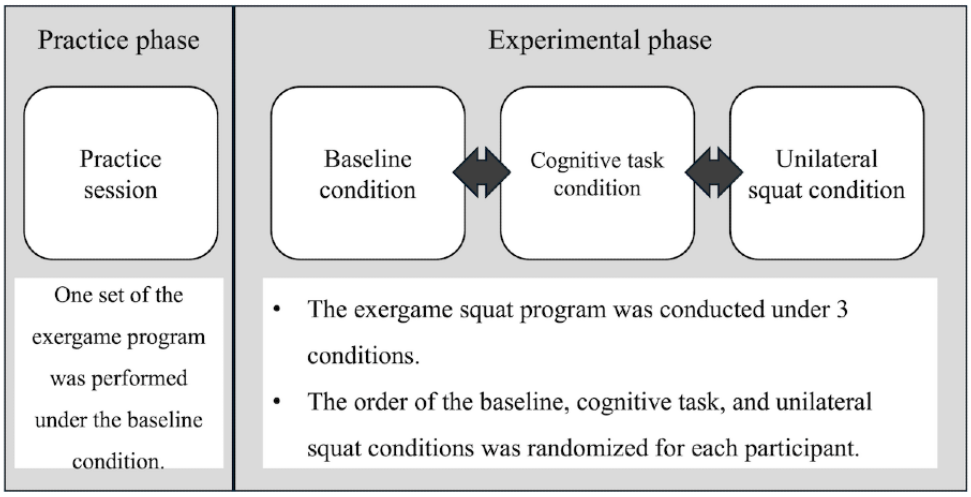
XSL•FO
RenderX

Only aggregate data are reported in this manuscript. Each participant received a gift card valued at ¥1000 (US $6.24 as of June 15, 2026) as compensation for participation. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Experimental Protocol

The experimental procedure is shown in Figure 1. The study used a repeated measures design with condition as the within-participant factor. Participants first completed 1 familiarization set under the baseline condition to become accustomed to the exergame task and interface. After familiarization, they performed 1 experimental set under each of 3 conditions: baseline, concurrent cognitive-task, and unilateral-squat conditions. The order of the 3 experimental conditions was randomized using a counterbalanced schedule, rather than administering the baseline condition first, to minimize potential order effects, including practice or familiarization effects, fatigue, and expectancy, that could systematically bias condition comparisons. A seated rest period of 3 minutes was provided between conditions to reduce immediate fatigue and potential carryover effects from the preceding condition. This interval was selected because each exergame set was brief, consisting of 10 squat repetitions, and was consistent with the use of short rest intervals in previous exergame-based rehabilitation protocols involving squat and single-leg balance tasks [29].

Figure 1. Study protocol. Baseline condition: exergame performed in a bilateral stance without an additional task; cognitive task condition: exergame performed with serial subtraction during squatting; and unilateral squat condition: exergame performed in a single-leg stance.



Across all conditions, squat execution was standardized. Participants held the Ring-Con (Nintendo) with both hands in front of the chest throughout the task. Stance width was set at approximately shoulder width. Squat tempo was guided by in-game visual cues (“lower,” “keep,” and “return”), resulting in a constant externally paced movement rhythm. Participants were instructed to aim for a squat depth corresponding to approximately 90° of knee flexion.

In the concurrent cognitive-task condition, participants performed a mental arithmetic task requiring serial subtraction of 7 from a 2-digit starting number [30,31]. For each set, a different starting number (range 70-99) was selected from a prepared list (eg, 83 and 76). Participants responded aloud in a self-paced manner as quickly and continuously as possible during squatting. Task prioritization followed a motor-first approach: participants were instructed to prioritize squat execution while continuing the arithmetic task as much as possible without compromising squat performance. No feedback on response accuracy was provided. If a participant was unable to provide a response within 5 seconds, they were instructed to skip the current subtraction and continue with the next calculation.

In the unilateral-squat condition, participants performed the same exergame squat set (10 repetitions) in a single-leg stance. This condition was intended to increase support-leg muscular demand as well as the postural stability and motor-control requirements inherent to unilateral squatting [32,33]. The supporting leg was fixed on the left side because the leg sensor was attached to the left thigh, and this standardization ensured consistency in motion capture and scoring. The nonsupporting leg was maintained in a flexed position behind the body, and no external balance support was allowed.

Apparatus and Exergame Program

The experimental setup is shown in Figure S1 in Multimedia Appendix 1. The exergame was conducted using *Ring Fit Adventure* for the Nintendo Switch (Nintendo) [34-37], with the squat task selected from the custom mode. The Ring-Con, a ring-shaped controller held with both hands, and a leg sensor attached to the left thigh enabled participants’ movements to be reflected by an on-screen avatar. The squat program required repeated squats in response to visual cues and provided a squat score (0-100 points) that primarily reflected squat depth and timing relative to the cues.

The squat program in the custom mode of *Ring Fit Adventure* was selected for 3 reasons. First, the user interface layout remains stable in this mode, facilitating consistent definition of AOIs for eye-tracking analysis. Second, the task structure is simple and repetitive, allowing the serial-subtraction and unilateral-squat manipulations to be implemented within the same exergame task while minimizing changes in task rules and visual content. Third, the program was set to 10 repetitions

XSL•FO
RenderX

to keep each trial brief, thereby reducing eye-tracking data loss and fatigue, which was appropriate for this pilot study. With respect to difficulty settings, the custom mode allows selection of the number of squat repetitions but does not provide additional difficulty parameters. Therefore, task demands were standardized across participants by using 10 repetitions in all conditions.

The game was projected onto a 130 × 230 cm screen, with participants positioned 2.5 m away. The projection was adjusted so that the center of the display was approximately at eye level during standing. Ambient lighting conditions were standardized using indoor lighting, and window light was minimized by drawing curtains to reduce glare and fluctuations in screen luminance. To reduce extraneous visual stimuli that could influence gaze dispersion, posters and other nonessential visual materials around the screen were removed, and the area surrounding the display was kept as visually uncluttered as possible.

Eye movements were recorded using a wearable eye tracker (Tobii Pro Glasses 3, Tobii Technology; sampling rate 100 Hz), and the device was recalibrated before each test condition. Gaze data were processed offline using Tobii Pro Lab (Tobii Technology) to extract fixation metrics.

Outcome Measures

Feasibility outcomes were recorded to characterize the practical implementation of mobile eye tracking during squat-based exergaming. These outcomes included (1) completion rate, defined as the proportion of participants who completed all 3 conditions; (2) calibration success, defined as successful eye tracker calibration before each condition; and (3) technical issues, defined as any device or software malfunction or recording failure requiring restart, repeat calibration, or trial exclusion.

The primary gaze outcomes were the proportion of fixation time and the proportion of fixation counts allocated to each predefined AOI (command, avatar, and score) and to regions outside these AOIs across the entire exergame trial. These proportions were calculated by dividing fixation time or fixation count within each AOI by the total fixation time or count recorded across the whole screen during the trial. This normalization was used to characterize relative gaze allocation while reducing potential bias from condition-dependent differences in trial duration and total fixation behavior. Accordingly, proportional AOI metrics should be interpreted as indexes of relative allocation rather than absolute gaze exposure. Fixations were defined using the velocity-threshold identification (Attention) filter based on eye angular velocity (threshold 100°/s) [38].

AOI metrics (fixation duration and fixation count) were exported for each AOI. Invalid samples due to blinks or temporary tracking loss were treated as missing by the software and were not manually interpolated. No additional data cleaning beyond the software's standard validity handling was applied. No outlier removal (eg, winsorization or participant-level exclusion based on gaze metrics) was performed.

Three AOIs were defined in Tobii Pro Lab: the command area displaying the next movement cue, the avatar area mirroring the participant's movement, and the score area presenting squat performance points (Figure S2 in Multimedia Appendix 1). AOI dimensions were estimated from photographed screen measurements and converted to degrees of visual angle using the viewing distance of 2.5 m. The resulting AOI sizes were approximately 9.0°×21.6° for the command AOI, 15.5°×10.7° for the avatar AOI, and 11.2°×12.2° for the score AOI.

To provide an intuitive visualization of gaze distribution, scanpaths and heat maps were generated from aggregated gaze data across participants. However, these served only as qualitative supplementary materials and were not subjected to statistical testing.

Secondary outcomes were exergame execution time (seconds) and squat score (0-100 points), recorded at the end of the squat program.

Manipulation Check

To examine whether the concurrent cognitive task and unilateral-squat manipulations altered task performance, additional performance indexes were assessed.

For the concurrent cognitive task condition, arithmetic performance was used as an objective manipulation check of the added cognitive task. Accuracy was calculated as percentage correct, defined as the number of correct responses divided by the total number of spoken responses during the set. Scoring was performed sequentially: a response was counted as correct if it was exactly 7 less than the immediately preceding response, regardless of whether the participant had made an earlier error. As individual response time stamps were not recorded, response latencies could not be calculated. Instead, the total number of responses during the set was reported together with execution time as indirect performance-based indicators related to response timing. Squat score and execution time were also analyzed as task performance indicators during the concurrent cognitive task.

For the unilateral-squat condition, squat score and execution time were analyzed as indicators of movement performance under increased unilateral motor- and postural-control requirements. In this pilot study, direct physiological indexes of exertion (eg, heart rate, ratings of perceived exertion, and electromyography) were not collected. Accordingly, the unilateral-squat condition should not be interpreted as a directly quantified increase in physiological load or as a purely physical load manipulation. Instead, changes in execution time and squat score were used as objective performance-based markers indicating that the unilateral-squat manipulation altered task difficulty relative to baseline.

Statistical Analysis

Normality for each outcome was assessed using the Shapiro-Wilk test. For outcomes meeting the normality assumption, a 1-factor repeated measures analysis of variance with condition as the within-participant factor was performed, followed by Dunnett multiple comparison tests using the baseline condition as the reference (concurrent cognitive task

XSL•FO
RenderX

condition vs baseline condition and unilateral-squat condition vs baseline condition). For outcomes violating the normality assumption, paired Wilcoxon signed-rank tests were used to compare each condition with baseline condition, and *P* values were adjusted using the Holm procedure for the 2 planned comparisons. Descriptive statistics are presented as mean (SD). Effect sizes for comparisons with the baseline condition are reported as Hedges *g* for paired samples. For each planned within-participant comparison, Hedges *g* was calculated using complete paired observations as the standardized mean change, defined as the mean within-participant difference score divided by the SD of the corresponding difference scores. Difference scores were calculated as the task demand condition minus the baseline condition, and Hedges small-sample correction was applied. Thus, positive values indicate higher values in the task demand condition than in the baseline condition, whereas negative values indicate lower values. Individual participant data were visualized alongside summary statistics (mean and SD) to enhance transparency.

Exploratory correlation analyses were conducted to examine associations between gaze allocation metrics and performance outcomes. Given the small sample size, these analyses were considered hypothesis generating; Spearman ρ and corresponding *P* values are reported, and findings were interpreted cautiously.

All analyses were performed in R (version 4.4.2; R Foundation for Statistical Computing) [39] using the effectsize package and related statistical functions [40-42].

Results

Feasibility Outcomes

All participants (n=8) completed all 3 conditions. Eye tracker calibration was successful in 100% (24/24) of attempts, with calibration performed before each condition for each participant. No technical issues (eg, device or software malfunction, or recording failure) occurred, and no trials were excluded for technical reasons.

Manipulation Check

Manipulation check outcomes are summarized in Table 1. As shown in Figure 2, participants performed the arithmetic task with high accuracy (mean 87.9, SD 16.1) and a consistent number of responses (mean 7.3, SD 1.8). Execution time was longer in the concurrent cognitive task condition than in the baseline condition (mean difference 7.1 seconds, 95% CI 3.7-10.5; *P*=.002), indicating measurable dual-task interference.

Table 1. Manipulation check outcomes for performance measures across conditions (n=8)[a].

Outcome	Baseline, mean (SD)	Cognitive task, mean (SD)	Unilateral squat, mean (SD)	Cognitive task–baseline, mean difference (95% CI)	*P* value	Hedges *g*	Unilateral squat–baseline, mean difference (95% CI)	*P* value	Hedges *g*
Exergame execution time (s)	31.0 (3.3)	38.2 (6.1)	34.3 (6.0)	7.1 (3.7 to 10.5)	.002	1.55	3.2 (0.5 to 5.9)	.02	0.88
Squat score (0 to 100)	98.8 (2.8)	97.2 (4.0)	69.6 (18.3)	−1.5 (−4.3 to 1.3)	.23	−0.4	−29.1 (−43.9 to −14.3)	.002	−1.46
Arithmetic accuracy (%)	N/A[b]	87.9 (16.1)	N/A	—[c]	—	—	—	—	—
Arithmetic responses (n)	N/A	7.3 (1.8)	N/A	—	—	—	—	—	—

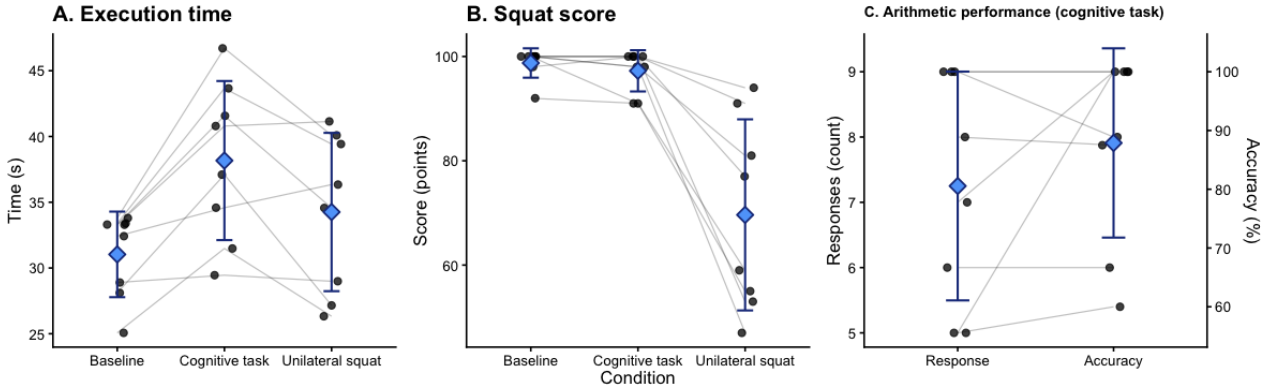
[a]Baseline: exergame performed in a bilateral stance without an additional task. Cognitive task: exergame performed with serial subtraction during squatting. Unilateral squat: exergame performed in a single-leg stance. Comparisons with baseline were performed using Dunnett test when parametric assumptions were met; otherwise, paired Wilcoxon signed-rank tests with Holm adjustment were used. Effect sizes represent Hedges *g* for paired comparisons. Arithmetic measures were collected only during cognitive task.

[b]N/A: not applicable (arithmetic outcome was not applicable because it was measured only in the cognitive task condition).

[c]No comparison with baseline was performed.

XSL•FO
RenderX

Figure 2. Manipulation check outcomes across conditions: (A) exergame execution time (seconds), (B) squat score (0-100 points), and (C) arithmetic task performance under the concurrent cognitive task condition. Individual participant data are shown as jittered points connected by thin lines to illustrate within-participant changes across conditions (A and B) and between arithmetic metrics (C). Diamonds indicate mean values, and error bars represent the mean+SD and mean−SD. Baseline: exergame performed in a bilateral stance without an additional task; cognitive task: exergame performed with serial subtraction during squatting; and unilateral squat: exergame performed in a single-leg stance.



In the unilateral-squat condition, squat score was markedly lower than in the baseline condition (mean difference −29.1 points, 95% CI −43.9 to −14.3; *P*=.002), and execution time was also longer (mean difference 3.2 seconds, 95% CI 0.5-5.9; *P*=.02), although the magnitude of this increase was smaller than that observed in the concurrent cognitive task condition. Together, the substantial reduction in squat score and the smaller but significant increase in execution time suggest that the unilateral-squat condition increased movement difficulty relative to baseline.

Gaze Allocation Across AOIs

Proportion of Fixation Time per AOI

To further characterize the redistribution of visual attention under different task demand conditions, the proportion of fixation time allocated to each predefined AOI (command, avatar, and score) and to regions outside these AOIs was analyzed (Table 2; Figure 3). In the concurrent cognitive task condition, the proportion of fixation time directed outside the predefined AOIs showed a descriptive increase relative to the baseline condition, whereas the proportions allocated to the command and score AOIs showed descriptive decreases. However, these changes were modest and did not reach statistical significance. In contrast, in the unilateral-squat condition, the proportion of fixation time allocated to the score AOI showed a descriptive increase relative to the baseline condition, accompanied by a slight descriptive decrease in fixation time directed outside the predefined AOIs. The plotted data suggested greater interindividual variability in fixation time allocation in both task demand conditions than in the baseline condition.

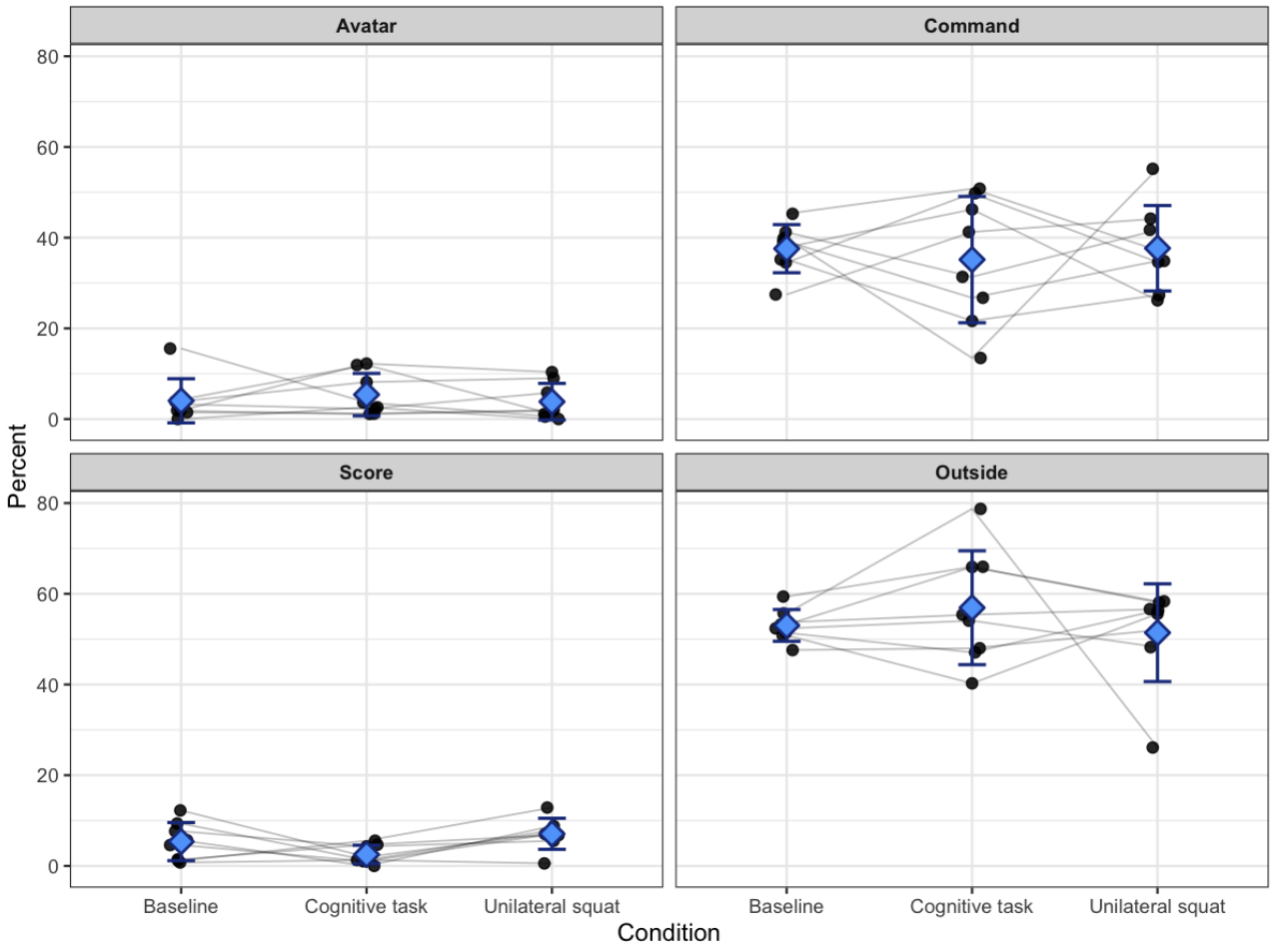
Table 2. Proportion of fixation time allocated to each area of interest (AOI) across task demand conditions (n=8)[a].

AOI	Baseline, mean (SD)	Cognitive task, mean (SD)	Unilateral squat, mean (SD)	Cognitive task minus baseline, mean difference (95% CI)	*P* value	Hedges *g*	Unilateral squat minus baseline, mean difference (95% CI)	*P* value	Hedges *g*
Command	37.6 (5.3)	35.2 (13.9)	37.7 (9.4)	−2.4 (−15.1 to 10.3)	.99	−0.14	0.1 (−8.8 to 9.0)	.99	0.01
Avatar	4.0 (4.9)	5.4 (4.7)	3.8 (4.0)	1.4 (−4.3 to 7.0)	.99	0.18	−0.2 (−6.0 to 5.6)	.99	−0.03
Score	5.4 (4.2)	2.5 (2.0)	7.1 (3.4)	−2.8 (−7.2 to 1.5)	.34	−0.48	1.7 (−2.6 to 6.1)	.38	0.29
Outside	53.0 (3.5)	56.9 (12.6)	51.4 (10.8)	3.9 (−4.8 to 12.6)	.64	0.33	−1.6 (−11.5 to 8.2)	.71	−0.12

[a]Baseline: exergame performed in a bilateral stance without an additional task. Cognitive task: exergame performed with serial subtraction during squatting. Unilateral squat: exergame performed in a single-leg stance. Values are presented as mean (SD) of the proportion (%) of fixation time within each AOI across the entire exergame trial. Proportions were calculated by dividing whole fixation duration within each AOI by the total fixation duration recorded across the entire exergame trial. Mean differences were calculated as each task demand condition minus the baseline condition. The 95% CIs were calculated from within-participant difference scores. Comparisons with baseline were performed using Dunnett test when parametric assumptions were met; otherwise, paired Wilcoxon signed-rank tests with Holm adjustment were used. Effect sizes are reported as Hedges *g* for within-participant comparisons. The “outside” category represents fixation time allocated outside the 3 predefined AOIs (command, avatar, and score).

XSL•FO
RenderX

Figure 3. Proportion of fixation time allocated to each area of interest under baseline, concurrent cognitive task, and unilateral squat conditions. Baseline: exergame performed in a bilateral stance without an additional task; cognitive task: exergame performed with serial subtraction during squatting; and unilateral squat: exergame performed in a single-leg stance. Individual participant data are shown with mean (SD). Proportions were calculated relative to total fixation time across the entire exergame trial.



Proportion of Fixation Counts per AOI

The proportion of fixation counts allocated to each AOI varied across conditions (Table 3; Figure 4). Compared with the baseline condition, the concurrent cognitive task condition showed a lower proportion of fixation counts directed to the command AOI (mean 44.8%, SD 11.4% vs mean 32.1%, SD 8.9%; *P*=.04) and score AOI (mean 14.4%, SD 9.8% vs mean 4.9%, SD 4.3%; *P*=.04), and a higher proportion directed outside the predefined AOIs (mean 36%, SD 16.9% vs mean 55.7%, SD 8.6%; *P*=.02). The unilateral-squat condition also showed a lower proportion of fixation counts directed to the command AOI than the baseline condition (mean 36.2%, SD 8.5% vs mean 44.8%, SD 11.4%; *P*=.04), but it did not show the same increase in outside viewing observed under the concurrent cognitive task condition. Instead, fixation count allocation in the unilateral-squat condition showed small descriptive increases in the score and avatar AOIs.

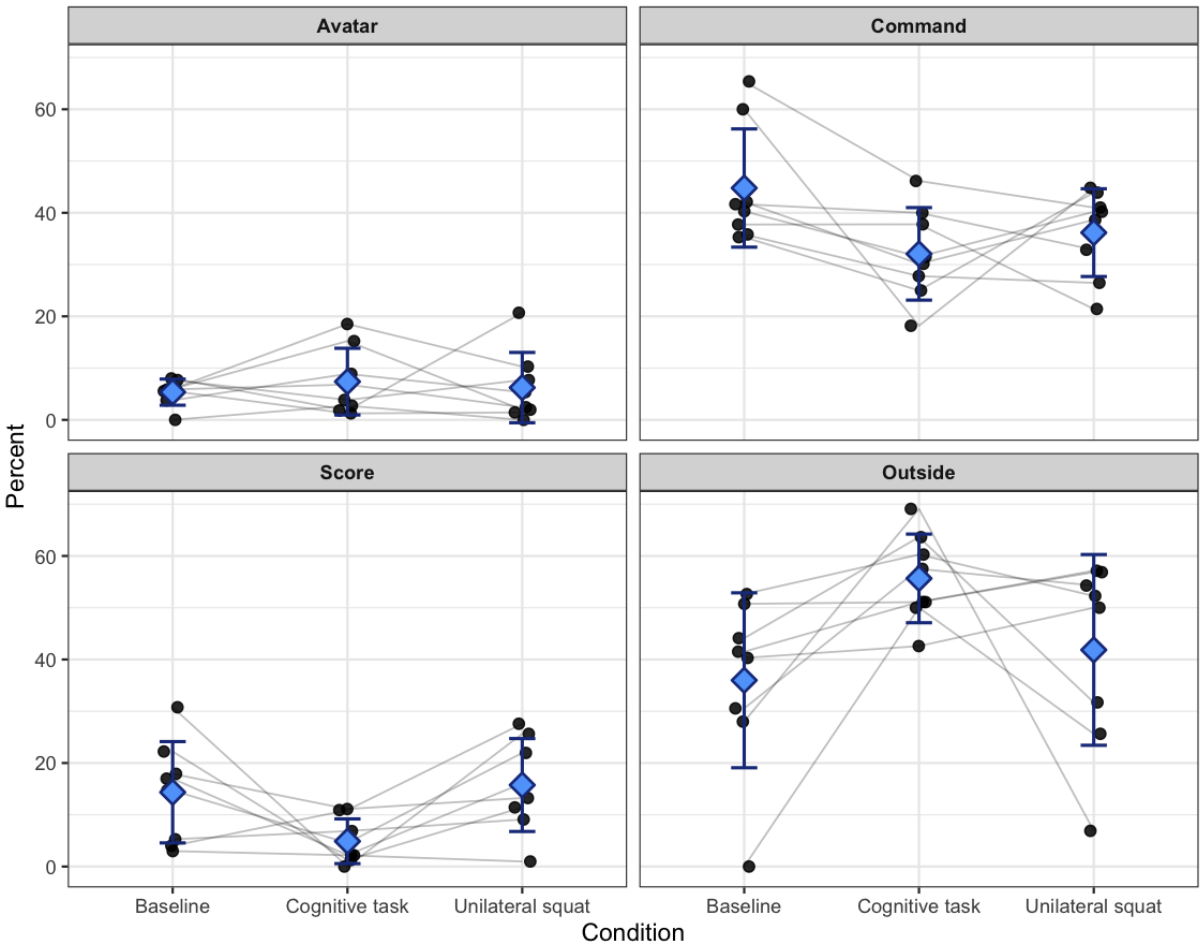
XSL•FO
RenderX

Table 3. Proportion of fixation counts allocated to each area of interest (AOI) across task demand conditions (n=8)[a].

AOI	Baseline, mean (SD)	Cognitive task, mean (SD)	Unilateral squat, mean (SD)	Cognitive task minus baseline, mean difference (95% CI)	*P* value	Hedges *g*	Unilateral squat minus baseline, mean difference (95% CI)	*P* value	Hedges *g*
Command	44.8 (11.4)	32.1 (8.9)	36.2 (8.5)	−12.7 (−23.7 to −1.7)	.04	−0.86	−8.6 (−17.3 to 0.0)	.04	−0.74
Avatar	5.4 (2.5)	7.4 (6.4)	6.2 (6.8)	2.0 (−3.6 to 7.6)	.66	0.27	0.9 (−3.8 to 5.6)	.92	0.14
Score	14.4 (9.8)	4.9 (4.3)	15.7 (9.0)	−9.5 (−19.9 to 0.9)	.04	−0.68	1.4 (−7.4 to 10.2)	.90	0.12
Outside	36.0 (16.9)	55.7 (8.6)	41.9 (18.4)	19.7 (4.4 to 35.0)	.02	0.95	5.9 (−8.0 to 19.7)	.62	0.32

[a]Baseline: exergame performed in a bilateral stance without an additional task. Cognitive task: exergame performed with serial subtraction during squatting. Unilateral squat: exergame performed in a single-leg stance. Values are presented as mean (SD) of the proportion (%) of fixation counts within each AOI across the entire exergame trial. Proportions were calculated as the number of fixations within each AOI divided by the total number of fixations recorded across the entire exergame trial. Mean differences were calculated as each task demand condition minus the baseline condition. The 95% CIs were calculated from within-participant difference scores. Comparisons with baseline were performed using Dunnett test when parametric assumptions were met; otherwise, paired Wilcoxon signed-rank tests with Holm adjustment were used. For outcomes analyzed with paired Wilcoxon signed-rank tests, *P* values are based on ranked paired differences and may not correspond directly to the mean difference 95% CIs, which are presented descriptively. Effect sizes are reported as Hedges *g* for within-participant comparisons. The "outside" category represents fixation counts allocated outside the 3 predefined AOIs (command, avatar, and score).

Figure 4. Proportion of fixation counts allocated to each area of interest (AOI) across task demand conditions. The proportion of fixation counts directed to the command, avatar, score, and outside AOIs across the entire exergame trial is shown for the baseline, concurrent cognitive task, and unilateral squat conditions. Baseline: exergame performed in a bilateral stance without an additional task; cognitive task: exergame performed with serial subtraction during squatting; and unilateral squat: exergame performed in a single-leg stance. Each proportion was calculated as the number of fixations within an AOI divided by the total number of fixations recorded during the trial.



XSL•FO
RenderX

Supportive Analyses

Exploratory Correlation Analyses

Correlation analyses between gaze allocation metrics and performance outcomes were conducted on an exploratory basis to generate hypotheses for future studies (Figure S3 in Multimedia Appendix 1). Given the small sample size, these findings should not be interpreted as confirmatory evidence.

Descriptive Gaze Distribution

Heat maps and scanpaths are presented as descriptive visualizations to contextualize the quantitative AOI allocation results (Figures S4 and S5 in Multimedia Appendix 1). As these qualitative displays are sensitive to visualization parameters and do not directly quantify gaze allocation, they were not used as stand-alone evidence. Qualitative patterns were interpreted only when consistent with the AOI-based metrics reported in Tables 2 and 3.

Discussion

Principal Results

This pilot study suggests that different task demand manipulations within a squat-based exergame may be associated with distinct patterns of gaze redistribution across task-relevant display elements, rather than a uniform reduction in visual attention. Specifically, adding a concurrent cognitive task and performing the task in a unilateral-squat condition produced different descriptive patterns of performance and gaze allocation. By quantifying AOI-specific proportions of fixation time and fixation counts standardized to the entire trial, this approach was intended to reduce potential bias arising from condition-dependent differences in execution time and to generate preliminary, hypothesis-generating estimates for future confirmatory exergame studies.

The manipulation check suggested that the concurrent cognitive task and unilateral-squat conditions affected performance in partially different ways. Specifically, the concurrent cognitive task condition was associated primarily with prolonged execution time while squat performance (score) was largely preserved, whereas the unilateral-squat condition was associated with a more pronounced reduction in squat score together with a modest but significant increase in execution time relative to baseline. Taken together, these patterns suggest that the 2 manipulations may have altered task demands through different mechanisms rather than simply producing a uniform increase in overall difficulty. However, the unilateral-squat condition should not be interpreted as a pure physical load condition because single-leg squatting inherently involves balance control, postural stability, and motor-control processes that may also require cognitive engagement.

The pattern observed in the concurrent cognitive task condition—significantly prolonged execution time, greater fixation allocation outside predefined AOIs, and reduced sampling of the command and score AOIs—is broadly consistent with real-world dual-task evidence showing increased off-task or task-irrelevant fixations when mental arithmetic is superimposed on everyday actions [21,43]. In the present exergame context, this pattern may indicate that the added cognitive task reduced visual sampling of goal-relevant instructional and feedback information while shifting gaze toward non-AOI regions. However, because the Outside category aggregated all fixations outside the 3 predefined AOIs, this interpretation remains tentative. This analysis cannot determine whether these fixations were directed toward peripheral screen areas, visually neutral regions, or other nondefined targets. Therefore, the increase in outside viewing warrants nuanced interpretation. On the one hand, it may reflect weakened top-down attentional selection, whereby limited processing resources reduce the ability to maintain gaze on goal-relevant AOIs. On the other hand, it may reflect strategic gaze aversion. As suggested by Müller et al [21] and Glenberg et al [44], individuals may avert their gaze toward visually neutral or noninformative regions to deliberately reduce external visual input, thereby facilitating demanding internal cognitive processes such as mental arithmetic. In contrast, some laboratory studies using static or tightly controlled stimuli have reported longer fixation durations and fewer fixations under higher cognitive load [45,46], suggesting that gaze responses to cognitive task demands may depend on task dynamics and constraints. One plausible explanation is that squat-based exergaming is both visually dynamic and motorically constrained, requiring intermittent postural-control and feedback sampling, which may promote context-dependent gaze redistribution rather than a single uniform gaze response [47]. Accordingly, these findings are best interpreted as preliminary evidence of context-dependent gaze reallocation during a cognitive dual-task exergame condition rather than as a universal gaze marker of cognitive load [11,20].

By contrast, the unilateral-squat condition showed a different pattern: squat performance decreased significantly, whereas gaze allocation did not show the same increase in outside viewing observed in the concurrent cognitive task condition. Instead, fixation counts directed to the command AOI were reduced, with modest descriptive increases in allocation to other task-relevant AOIs, such as the score and avatar regions. Although these shifts were small and should be interpreted cautiously, they may suggest that unilateral squatting prompted redistribution of visual attention within task-relevant display elements rather than away from them [18,47]. This interpretation is compatible with a tentative self-monitoring explanation, whereby participants may have sampled performance-relevant feedback more frequently as movement difficulty increased. However, this pattern should not be attributed solely to increased physical load because single-leg squatting involves a composite set of demands, including support leg muscular demand, balance control, postural stability, and motor planning. Therefore, the qualitative contrast between the concurrent cognitive task and unilateral-squat conditions should be interpreted as evidence of condition-dependent gaze reallocation under different task demands, not as definitive evidence distinguishing pure cognitive from pure physical load.

These pilot data suggest that gaze metrics may help characterize how users reallocate attention under different task demand conditions during exergaming, which may be useful for hypothesis generation and for identifying candidate features for

XSL•FO
RenderX

adaptive systems. However, this study did not evaluate real-time gaze-based adaptation metrics, such as latency, reliability, or within-trial detection, nor did it examine decision thresholds or downstream benefits of adaptation. Therefore, any implications for personalized difficulty adjustment remain speculative and should be reserved for future work.

Limitations

This study has several limitations. First, the sample was very small and homogeneous, with a marked gender imbalance, which limits generalizability and may influence estimates of multitasking performance and squat execution in some contexts. Replication in a larger, more gender-balanced sample is therefore warranted. Second, statistical power was limited and CIs were often wide. Accordingly, nonsignificant findings should be interpreted cautiously and should not be considered definitive evidence of no effect. The use of proportional AOI metrics is another important limitation. Although normalization by total fixation time or count helped reduce the influence of unequal trial durations across conditions, proportional metrics change the interpretation of the outcome from absolute gaze exposure to relative gaze allocation. Consequently, a decrease in the proportion of fixations directed to a given AOI does not necessarily indicate a decrease in the absolute amount of gaze directed to that AOI. In addition, the AOI-specific findings should be interpreted as exploratory evidence of relative gaze redistribution rather than as independent AOI-specific effects.

Third, the unilateral-squat condition represents a central interpretive constraint of this study. Although single-leg squatting likely increased support leg muscular demand, it cannot be interpreted as a pure physical load manipulation. Single-leg squatting inherently requires balance control, postural stability, motor planning, and continuous movement regulation, all of which may involve cognitive engagement. Therefore, the study design cannot disentangle muscular or physiological load from postural and motor-control demands. In addition, physiological indexes such as heart rate, ratings of perceived exertion, and electromyography were not collected; thus, physiological load differences between conditions could not be directly quantified. Future studies should incorporate objective physiological markers and use designs that can better distinguish muscular effort, postural-control demands, and concurrent cognitive load.

Fourth, the supporting leg in the unilateral-squat condition was fixed to the left side to maintain consistency with the leg sensor and game scoring. This may have introduced unilateral effects because single-leg squat performance can differ between dominant and nondominant limbs. If the left leg was nondominant for some participants, the unilateral-squat condition may have disproportionately increased movement difficulty, potentially contributing to larger decrements in squat score and altered execution time. Such unilateral performance differences could also have secondarily influenced gaze behavior, for example, by increasing monitoring of feedback AOIs such as the score region or by increasing off-AOI viewing, thereby limiting generalizability. Future studies should assess leg dominance and unilateral asymmetries and either counterbalance the supporting leg or include dominance as a covariate.

Finally, standard feasibility indexes such as eye-tracking data loss (eg, fixation coverage) were not systematically quantified in this pilot study. Future studies should prospectively record and report these metrics to better characterize measurement robustness, recruit larger and more diverse cohorts, and examine whether gaze-derived features reliably predict performance or learning at the individual level.

Conclusions

These pilot data provide preliminary evidence that different task-demand manipulations may be associated with distinct patterns of gaze redistribution during squat-based exergaming in healthy young adults. The findings are hypothesis generating and require replication in larger and more diverse samples. Future studies should determine whether gaze-derived features can be measured robustly at the individual level and translated into validated adaptive difficulty algorithms that improve user outcomes. However, this study does not provide direct evidence for real-time adaptation or clinical utility.

Acknowledgments

The author used ChatGPT (OpenAI) to assist with language editing and refinement of manuscript wording during the preparation of this manuscript. ChatGPT was not used to generate study data or interpret the findings. All artificial intelligence–assisted output was reviewed, edited, and verified by the author, who takes full responsibility for the final content of the manuscript.

Funding

This study was supported by the Tokyo Professional University of Health Sciences Research Grant. The funder had no role in the study design, data collection, data analysis, interpretation of the findings, or writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: KT

XSL•FO
RenderX

Methodology: KT
Investigation: KT
Formal analysis: KT
Data curation: KT
Writing – original draft: KT
Writing – review and editing: KT

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures (Figures S1-S5) showing additional visual materials for the exergame task, area of interest setting, exploratory correlation analyses, and supportive gaze distribution analyses.
[PDF File (Adobe PDF File), 3238 KB - rehab_v13i1e81667_app1.pdf]

References

1. Dale G, Joessel A, Bavelier D, Green CS. A new look at the cognitive neuroscience of video game play. Ann N Y Acad Sci 2020 Mar;1464(1):192-203. [doi: 10.1111/nyas.14295] [Medline: 31943260]
2. Bediou B, Adams DM, Mayer RE, Tipton E, Green CS, Bavelier D. Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. Psychol Bull 2018 Jan;144(1):77-110. [doi: 10.1037/bul0000130] [Medline: 29172564]
3. Wang RY, Huang YC, Zhou JH, Cheng SJ, Yang YR. Effects of exergame-based dual-task training on executive function and dual-task performance in community-dwelling older people: a randomized-controlled trial. Games Health J 2021 Oct;10(5):347-354. [doi: 10.1089/g4h.2021.0057] [Medline: 34491113]
4. Gallou-Guyot M, Mandigout S, Marie R, Robin L, Daviet JC, Perrochon A. Feasibility and potential cognitive impact of a cognitive-motor dual-task training program using a custom exergame in older adults: a pilot study. Front Aging Neurosci 2023 Feb 2;15:1046676 [FREE Full text] [doi: 10.3389/fnagi.2023.1046676] [Medline: 36819724]
5. Attoh-Mensah E, Bel JF, El Ayoubi K, Boujut A, Ramanoël S, Perrochon A. Neurobiological effects of exergame interventions in older adults with or without a neurocognitive disorder: a systematic review. Neurobiol Aging 2025 Oct;154:47-61 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2025.06.009] [Medline: 40609365]
6. Tuena C, Borghesi F, Bruni F, Cavedoni S, Maestri S, Riva G, et al. Technology-assisted cognitive motor dual-task rehabilitation in chronic age-related conditions: systematic review. J Med Internet Res 2023 May 22;25:e44484 [FREE Full text] [doi: 10.2196/44484] [Medline: 37213200]
7. Lapierre N, Um Din N, Belmin J, Lafuente-Lafuente C. Exergame-assisted rehabilitation for preventing falls in older adults at risk: a systematic review and meta-analysis. Gerontology 2023;69(6):757-767 [FREE Full text] [doi: 10.1159/000528854] [Medline: 36580901]
8. Chen Y, Zhang Y, Guo Z, Bao D, Zhou J. Comparison between the effects of exergame intervention and traditional physical training on improving balance and fall prevention in healthy older adults: a systematic review and meta-analysis. J Neuroeng Rehabil 2021 Nov 24;18(1):164 [FREE Full text] [doi: 10.1186/s12984-021-00917-0] [Medline: 34819097]
9. Jiang J, Guo W, Wang B. Effects of exergaming on executive function of older adults: a systematic review and meta-analysis. PeerJ 2022 Apr 11;10:e13194 [FREE Full text] [doi: 10.7717/peerj.13194] [Medline: 35433124]
10. Mura G, Carta MG, Sancassiani F, Machado S, Prosperini L. Active exergames to improve cognitive functioning in neurological disabilities: a systematic review and meta-analysis. Eur J Phys Rehabil Med 2018 Jun;54(3):450-462 [FREE Full text] [doi: 10.23736/S1973-9087.17.04680-9] [Medline: 29072042]
11. Stojan R, Voelcker-Rehage C. A systematic review on the cognitive benefits and neurophysiological correlates of exergaming in healthy older adults. J Clin Med 2019 May 23;8(5):734 [FREE Full text] [doi: 10.3390/jcm8050734] [Medline: 31126052]
12. Gallou-Guyot M, Perrochon A, Marie R, Bourgeois M, Mandigout S. Measured and perceived exercise intensity during the performance of single-task, cognitive-motor dual-task, and exergame training: transversal study. JMIR Serious Games 2023 Feb 02;11:e36126 [FREE Full text] [doi: 10.2196/36126] [Medline: 36729572]
13. Béraud-Peigné N, Perrot A, Maillot P. Active video games training for older adults: comparative study of user experience, workload, pleasure, and intensity. JMIR Serious Games 2025 May 26;13:e67314 [FREE Full text] [doi: 10.2196/67314] [Medline: 40418855]
14. Dias RD, Ngo-Howard MC, Boskovski MT, Zenati MA, Yule SJ. Systematic review of measurement tools to assess surgeons' intraoperative cognitive workload. Br J Surg 2018 Apr;105(5):491-501 [FREE Full text] [doi: 10.1002/bjs.10795] [Medline: 29465749]
15. Young JQ, Van Merrienboer J, Durning S, Ten Cate O. Cognitive Load Theory: implications for medical education: AMEE guide no. 86. Med Teach 2014 May;36(5):371-384. [doi: 10.3109/0142159X.2014.889290] [Medline: 24593808]
16. Muir-Hunter SW, Wittwer JE. Dual-task testing to predict falls in community-dwelling older adults: a systematic review. Physiotherapy 2016 Mar;102(1):29-40. [doi: 10.1016/j.physio.2015.04.011] [Medline: 26390824]

XSL•FO
RenderX

17. Pashler H. Dual-task interference in simple tasks: data and theory. Psychol Bull 1994;116(2):220-244. [doi: 10.1037//0033-2909.116.2.220]
18. Bhatt M, Mahana B, Ko JH, Kolesar TA, Kanitkar A, Szturm T. Computerized dual-task testing of gait visuomotor and cognitive functions in Parkinson's disease: test-retest reliability and validity. Front Hum Neurosci 2021 Jul 15;15:706230 [FREE Full text] [doi: 10.3389/fnhum.2021.706230] [Medline: 34335213]
19. Kopiske K, Koska D, Baumann T, Maiwald C, Einhäuser W. Icy road ahead-rapid adjustments of gaze-gait interactions during perturbed naturalistic walking. J Vis 2021 Aug 02;21(8):11 [FREE Full text] [doi: 10.1167/jov.21.8.11] [Medline: 34351396]
20. Ekin M, Krejtz K, Duarte C, Duchowski AT, Krejtz I. Prediction of intrinsic and extraneous cognitive load with oculometric and biometric indicators. Sci Rep 2025 Feb 12;15(1):5213 [FREE Full text] [doi: 10.1038/s41598-025-89336-y] [Medline: 39939345]
21. Müller C, Baumann T, Einhäuser W, Kopiske K. Slipping while counting: gaze-gait interactions during perturbed walking under dual-task conditions. Exp Brain Res 2023 Mar;241(3):765-780 [FREE Full text] [doi: 10.1007/s00221-023-06560-6] [Medline: 36725725]
22. Shah SM, Sun Z, Zaman K, Hussain A, Shoaib M, Pei L. A driver gaze estimation method based on deep learning. Sensors (Basel) 2022 May 23;22(10):3959 [FREE Full text] [doi: 10.3390/s22103959] [Medline: 35632365]
23. Ahlström C, Kircher K, Nyström M, Wolfe B. Eye tracking in driver attention research-how gaze data interpretations influence what we learn. Front Neuroergon 2021 Dec 8;2:778043 [FREE Full text] [doi: 10.3389/fnrgo.2021.778043] [Medline: 38235213]
24. Ding W, Murzello Y, Cao S, Samuel S. Where to gaze during take-over: eye gaze strategy analysis of different situation awareness and hazard perception levels. Proc Hum Factors Ergon Soc Annu Meet 2023 Oct 21;67(1):2042-2047. [doi: 10.1177/21695067231193650]
25. Krebs C, Falkner M, Niklaus J, Persello L, Klöppel S, Nef T, et al. Application of eye tracking in puzzle games for adjunct cognitive markers: pilot observational study in older adults. JMIR Serious Games 2021 Mar 22;9(1):e24151 [FREE Full text] [doi: 10.2196/24151] [Medline: 33749607]
26. Nian Q, Lu W, Xu Y. Effects of object working memory load on visual search in basketball players: an eye movement study. BMC Psychol 2023 Dec 19;11(1):446 [FREE Full text] [doi: 10.1186/s40359-023-01488-6] [Medline: 38115097]
27. Alemanno M, Di Pompeo I, Marcaccio M, Canini D, Curcio G, Migliore S. From gaze to game: a systematic review of eye-tracking applications in basketball. Brain Sci 2025 Apr 20;15(4):421 [FREE Full text] [doi: 10.3390/brainsci15040421] [Medline: 40309899]
28. Oberfrank L, Kruse L, Steinicke F. Sphere Saber: a virtual reality exergame to study age-related differences in selective visual attention and information processing. In: Huang T, Sra M, Argelaguet F, Lopes P, Machuca MB, editors. SUI '23: Proceedings of the 2023 ACM Symposium on Spatial User Interaction. New York, NY: Association for Computing Machinery; 2023:1-10. [doi: 10.1145/3607822.3614534]
29. Cikajlo I, Rudolf M, Mainetti R, Borghese NA. Multi-exergames to set targets and supplement the intensified conventional balance training in patients with stroke: a randomized pilot trial. Front Psychol 2020 Apr 2;11:572 [FREE Full text] [doi: 10.3389/fpsyg.2020.00572] [Medline: 32300321]
30. Lee HI, Park S, Lim J, Chang SH, Ji JH, Lee S, et al. Influence of driver's career and secondary cognitive task on visual search behavior in driving: a dual-task paradigm. Adv Phys Educ 2015 Nov;05(04):245-254. [doi: 10.4236/ape.2015.54029]
31. Chatterjee SA, Fox EJ, Daly JJ, Rose DK, Wu SS, Christou EA, et al. Interpreting prefrontal recruitment during walking after stroke: influence of individual differences in mobility and cognitive function. Front Hum Neurosci 2019 Jun 18;13:194 [FREE Full text] [doi: 10.3389/fnhum.2019.00194] [Medline: 31316360]
32. DeFOREST BA, Cantrell GS, Schilling BK. Muscle activity in single- vs. double-leg squats. Int J Exerc Sci 2014 Oct 01;7(4):302-310 [FREE Full text] [doi: 10.70252/MXVZ7653] [Medline: 27182408]
33. Graham-Smith P, Natera A, Jarvis M. Load comparison ratio in single and double leg movements. In: Proceedings of the UKSCA's 11th Annual Conference. 2015 Presented at: UKSCA 2015; July 31-August 2, 2015; Chesford Grange, Warwickshire. [doi: 10.13140/RG.2.2.33093.88804]
34. Sato T, Shimizu K, Shiko Y, Kawasaki Y, Orita S, Inage K, et al. Effects of Nintendo Ring Fit Adventure exergame on pain and psychological factors in patients with chronic low back pain. Games Health J 2021 Jun;10(3):158-164. [doi: 10.1089/g4h.2020.0180] [Medline: 33891508]
35. Wu YS, Wang WY, Chan TC, Chiu YL, Lin HC, Chang YT, et al. Effect of the Nintendo Ring Fit adventure exergame on running completion time and psychological factors among university students engaging in distance learning during the COVID-19 pandemic: randomized controlled trial. JMIR Serious Games 2022 Mar 22;10(1):e35040. [doi: 10.2196/35040] [Medline: 35315780]
36. Takei K, Morita S, Watanabe Y, Suganami M, Inao M. Safety, feasibility, and acceptability of physiotherapy combined with strength training using active video games for older patients with musculoskeletal conditions. Disabil Rehabil Assist Technol 2024 Apr;19(3):641-647. [doi: 10.1080/17483107.2022.2115562] [Medline: 36001528]
37. Chan WL, Chan CW, Lam FM, Chan HH, Chan KC, Chan JS, et al. Feasibility, safety, and effects of a Nintendo Ring Fit Adventure™ balance and strengthening exercise program in community-dwelling older adults with a history of falls: a

XSL•FO
RenderX

feasibility randomized controlled trial. Geriatr Gerontol Int 2024 Mar;24 Suppl 1(Suppl 1):334-341. [doi: 10.1111/ggi.14771] [Medline: 38088479]
38. Tobii Pro Lab gaze filter. Tobii. 2023. URL: https://connect.tobii.com/s/article/Gaze-Filter-functions-and-effects?language=en_US#GazeFilterI-VTAttention [accessed 2024-01-31]
39. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2024. URL: https://www.R-project.org/ [accessed 2026-06-12]
40. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biom J 2008 Jun;50(3):346-363 [FREE Full text] [doi: 10.1002/bimj.200810425] [Medline: 18481363]
41. Ben-Shachar MS, Lüdecke D, Makowski D. effectsize: estimation of effect size indices and standardized parameters. J Open Source Softw 2020 Dec 23;5(56):2815. [doi: 10.21105/joss.02815]
42. Konietschke F, Placzek M, Schaarschmidt F, Hothorn LA. nparcomp: an R software package for nonparametric multiple comparisons and simultaneous confidence intervals. J Stat Softw 2015 Mar;64(9):1-17. [doi: 10.18637/jss.v064.i09]
43. Martinez-Cedillo AP, Gavrila N, Mishra A, Geangu E, Foulsham T. Cognitive load affects gaze dynamics during real-world tasks. Exp Brain Res 2025 Mar 03;243(4):82. [doi: 10.1007/s00221-025-07037-4] [Medline: 40029455]
44. Glenberg AM, Schroeder JL, Robertson DA. Averting the gaze disengages the environment and facilitates remembering. Mem Cognit 1998 Jul;26(4):651-658. [doi: 10.3758/bf03211385] [Medline: 9701957]
45. Liu JC, Li KA, Yeh SL, Chien SY. Assessing perceptual load and cognitive load by fixation-related information of eye movements. Sensors (Basel) 2022 Feb 04;22(3):1187. [doi: 10.3390/s22031187] [Medline: 35161930]
46. Walter K, Bex P. Cognitive load influences oculomotor behavior in natural scenes. Sci Rep 2021 Jun 11;11(1):12405 [FREE Full text] [doi: 10.1038/s41598-021-91845-5] [Medline: 34117336]
47. Hülsmann F, Frank C, Senna I, Ernst MO, Schack T, Botsch M. Superimposed skilled performance in a virtual mirror improves motor performance and cognitive representation of a full body motor action. Front Robot AI 2019 Jun 21;6:43 [FREE Full text] [doi: 10.3389/frobt.2019.00043] [Medline: 33501059]

Abbreviations

AOI: area of interest

Edited by S Munce; submitted 18.Aug.2025; peer-reviewed by JH Song, G Renganathan, J Fei; comments to author 19.Feb.2026; accepted 01.Jun.2026; published 23.Jun.2026.

Please cite as:
Takei K
Gaze Allocation and Performance Across Task-Demand Conditions During Squat-Based Exergaming: Pilot Study Using Eye Tracking
JMIR Rehabil Assist Technol 2026;13:e81667
URL: https://rehab.jmir.org/2026/1/e81667
doi:10.2196/81667
PMID:

©Keiichi Takei. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Designing an mHealth App for Stroke Rehabilitation in Indonesia: Mixed Methods Design Science Research Study

Sabrina Atha Shania*, SKom; Putu Wuri Handayani*, Prof Dr; Kamila Alifia Imanuddin*, MTI

Faculty of Computer Science, University of Indonesia, Jl Kampus UI, Depok, West Java, Indonesia

*all authors contributed equally

Corresponding Author:
Putu Wuri Handayani, Prof Dr
Faculty of Computer Science, University of Indonesia, Jl Kampus UI, Depok, West Java, Indonesia

Abstract

Background: Indonesia has the second-highest prevalence of stroke in Southeast Asia, with 4,918,487 people (1.7% of the population) living with this condition as of 2025. This is attributable to the living conditions in Indonesia, where limited health care workers and facilities are available to assist with stroke rehabilitation via mobile health (mHealth) interventions.

Objective: This study designed an mHealth app to facilitate stroke rehabilitation in Indonesia.

Methods: In this mixed methods study, interviews were conducted with 10 experts (doctors, therapists, and medical students) and 8 patients who have experienced a stroke to explore the need for stroke rehabilitation management apps. Using a design science research approach, we conducted 3 iterative design cycles, each comprising design, development, and evaluation, to construct an mHealth app for stroke rehabilitation. The first iteration produced a low-fidelity app prototype informed by intervention theory. The prototype, which comprised 13 features for patients and 12 for medical personnel, was evaluated through interviews with medical personnel. The second iteration produced a design for a high-fidelity app prototype based on feedback from the first iteration. The third iteration yielded an improved app based on feedback from the second iteration.

Results: This study identified app features that support patient rehabilitation, including patient training programs with progress monitoring in functional, speech, and occupational therapy, along with consultations and home visits to connect rehabilitating patients with medical personnel. Other features of the app include onboarding, gamification, users' authentication, user profile management, schedule reminders, activity targets, therapeutic progress reports, and educational materials and videos. The prototypes received a System Usability Scale score of 71.83 ("good") and average scores of 2.32 (system usefulness), 2.39 (information quality), 2.42 (interface quality), and 2.35 (overall).

Conclusions: This study provides guidance for health care facilities and regulators to improve the quality of rehabilitation services for patients with stroke in Indonesia.

(JMIR Rehabil Assist Technol 2026;13:e91464) doi:10.2196/91464

KEYWORDS

stroke; rehabilitation; mobile health; mHealth; design science research; intervention theory; Indonesia

Introduction

Background

Stroke can cause a wide range of long-term symptoms and remains a significant global health burden, requiring comprehensive and innovative rehabilitative approaches to optimize recovery outcomes [1]. The World Health Organization (WHO) [2] reports that every patient with acute stroke or cognitive dysfunction should undergo an initial functional assessment to determine rehabilitation needs and develop an individualized rehabilitation plan. The Barrow Neurological Institute [3] argues that most patients with stroke (80% - 90%) require some form of rehabilitation, whether acute, subacute, inpatient, or outpatient.

Various types of rehabilitation are available for patients with stroke, including occupational therapy, speech therapy, and physiotherapy [4]. Occupational therapists teach patients with stroke how to carry out daily living activities, such as cooking, cleaning the house, and taking care of themselves [4]. Speech therapists draw on scientific and technological knowledge to help patients communicate (touching on speech, voice, fluency, and other aspects of language) and cope with swallowing disorders [5]. However, the implementation of stroke rehabilitation faces significant and persistent challenges, including patient noncompliance and the gap between the need for post–inpatient intensive therapy and its actual availability in health care systems [6]. Globally, the cost of treating stroke in the year 2022 was estimated to reach more than US $721 billion, or about 0.66% of the world's gross domestic product [7].

XSL•FO
RenderX

Indonesia has the second-highest incidence and prevalence of stroke in Southeast Asia, with 642,943 new cases recorded and 4,918,487 people living with stroke between 2017 and 2022 [8]. In 2023, stroke ranked third among catastrophic diseases in Indonesia [9]. Stroke prevalence differs by age group, with individuals aged 75 years and older demonstrating the highest prevalence (41.3 per 1000 people) and the 15 to 54 years age group exhibiting the lowest at less than 9 per 1000 [10].

The challenges of stroke rehabilitation in developing countries, such as Indonesia, include health care factors, such as the demands of a growing health system and a lack of income for specialized training, rehabilitation facilities, and rehabilitation staff [11]. Other challenges include a lack of adequate learning opportunities and formal education programs for nurses, therapists, and even doctors [12]. There are also internal challenges on the patient side. Kayola et al [11] note that individual-level factors, such as low health literacy, can increase drug and rehabilitation costs as well as caregiver wages. Poststroke traumatic stress disorder can lead to nonadherence to treatment, increase disability rates, and increase the frequency of lesions in the right brain and brainstem [13].

Mobile health (mHealth) can provide innovative solutions to improve rehabilitation efficiency, offer more personalized services, and help health care facilities better understand patient needs [14-16]. Szeto et al [17] explain that the most beneficial apps integrate therapeutic elements such as social interaction, real-time feedback, rhythmic cues, and multisensory stimulation, as well as repetitive, task-specific, and/or goal-oriented exercises. However, user compliance with rehabilitation remains relatively low, and existing apps lack interoperability with electronic health records [17]. Song et al [18] discuss apps that specify daily rehabilitation exercises via asynchronous video recordings monitored by clinicians and offer reminder features to boost patient adherence. However, most of these studies were conducted in developed countries, including Spain, the Netherlands, the United States, and the United Kingdom [15].

Design science research (DSR) is a methodology that seeks to overcome real-world challenges through the design and evaluation of innovative artifacts [19]. DSR was used as this study's main methodological framework to ensure that the artifact, in this study named as a prototype, would be developed appropriately for users in Indonesia. Additionally, intervention theory is used as a foundation for designing more effective problem-solving systems [20]. Intervention theory holds that user behavior can be influenced through 3 main principles: the provision of valid and useful information, the freedom of the user to act based on the information obtained, and the formation of an internal commitment to the chosen action [20].

Objective

This research aimed to design an mHealth app for stroke rehabilitation management, focusing on the promotive, preventive, curative, and rehabilitative aspects. Our research question was as follows: How can an mHealth app for stroke rehabilitation in Indonesia be designed based on DSR and intervention theory? This research provides guidance for the development of health apps that meet the needs and expectations of patients with stroke in Indonesia.

Methods

Ethical Considerations

This research was carried out in collaboration with Indonesia's National Brain Center Hospital and received ethical approval from Mahar Mardjono on September 26, 2025 (reference DP.04.03/D.XXIII.9/208/2025). All respondents provided written informed consent to participate, and the data submitted were anonymous. The respondent data were used only for the purposes of this research.

Study Design

This mixed methods study used both quantitative and qualitative approaches. This research followed the 5 stages of the DSR methodology: problem identification and motivation, definition of solution objectives, design and development, demonstration, and evaluation and communication [21]. We collected data at each stage through interviews. However, for the evaluation and communication stage, we used both interviews (qualitative approach) and questionnaires (quantitative approach) to evaluate the prototypes. We triangulated the results to ensure that the prototypes aligned with the respondents' needs.

Using design science research, we conducted 3 iterative design cycles, each comprising design, development, and evaluation, to construct an mHealth app for stroke rehabilitation. The first iteration—the problem identification and motivation stage—was carried out through a literature review, application benchmarking, and interviews with patients with stroke and experts (including doctors, therapists, and medical students), who provided written consent. A thematic analysis [22] of the interview data was conducted to group the speakers' statements into problem categories. These categories allow for the simplification of open-ended data and indicate the problems' frequencies and levels of validation. Based on these findings, we proceeded to define the objectives and establish what would qualify as solutions, following the "how might we" (HMW) approach [23]. Mapping with the HMW approach offered solutions and ideas for features of the proposed app. To ensure the relevance of the app to user needs, we defined personas. Personas represented concrete pictures of user characteristics [24], making the entire mHealth solution design process more targeted and contextual. The next stage was design and development; we aimed to develop solutions in the form of low-fidelity prototypes, using the principle of the 8 golden rules as the interface design principle [21]. The last stage of the first iteration was demonstration, as we tested the prototype through interviews with doctors, therapists, and health students. Each stage of the first iteration included an evaluation and validation mechanism to ensure that the developed artifact matched the users' needs.

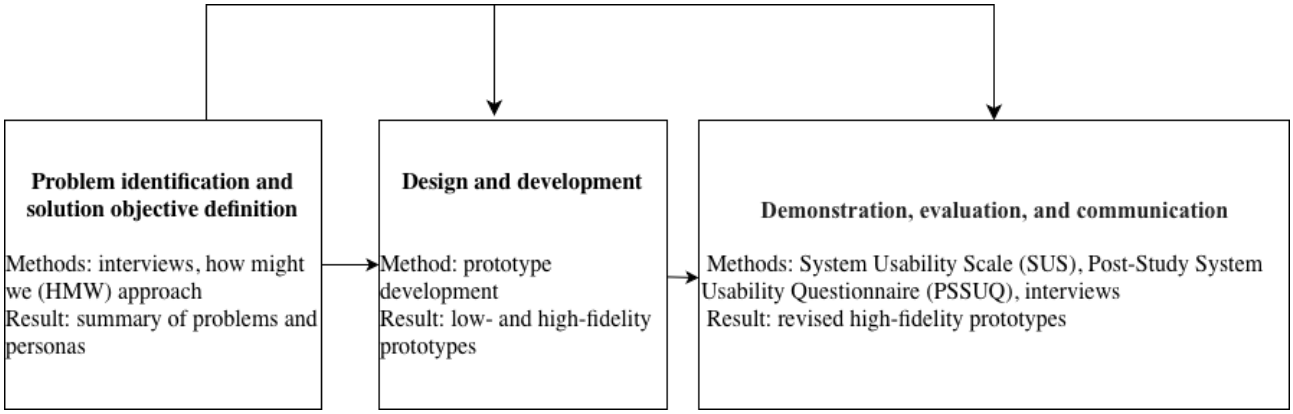
The second iteration focused on refining the prototype design based on the evaluation results from the first phase. The second iteration returned to design and development, using input in the form of feedback on the previous low-fidelity prototype [21]. This feedback guided the design of a higher-fidelity prototype in accordance with the 8 golden rules [21]. After the artifacts were repaired and determined suitable for our purposes, we proceeded to evaluation and communication [21]. This stage

XSL•FO
RenderX

included usability testing via the System Usability Scale (SUS), as well as interviews with patients and experts (doctors, therapists, and medical students). This evaluation aimed to explore user perceptions and experiences and to map obstacles and potential improvements to the interface and features [25]. This iteration yielded more targeted design recommendations and laid the foundation for the third phase.

The third and final iteration of the DSR development process began with design and development based on feedback and recommendations from the second stage. We then returned to evaluation and communication by measuring the satisfaction and effectiveness of the system [21]. For this, we used the Post-Study System Usability Questionnaire (PSSUQ) with various respondents. After ensuring that the evaluation results met the criteria, we proceeded to formulate design principles. This involved cross-iteration analysis to develop conceptual design principles based on user needs, feature responses, and solution effectiveness. SUS and PSSUQ have been widely adopted to assess the usability of digital systems [25]. The final stage was the preparation of conclusions and recommendations that summarized all findings [21]. Figure 1 describes the research flow based on DSR.

Figure 1. Research flow based on design science research.



Inclusion Criteria

We recruited patients who have experienced a stroke, experts in stroke rehabilitation, and respondents interested in stroke rehabilitation to obtain in-depth data. Patients treated at the National Brain Center Hospital were selected if they were 18 to 80 years old, had experienced a stroke, and experienced cognitive or physical impairments. Experts were selected if they were medical rehabilitation specialists, therapists, or students with a health-related major who were willing to participate (purposive sampling). Table 1 describes the respondent demographics for each DSR stage.

XSL•FO
RenderX

Table . Demographics of respondents.

DSR[a] stage, data collection method, and respondents' role	Respondents, n
Problem and solution objective identification	
Interview	
Experts	10
Patients with stroke	8
Design and development	
Prototype development	
No respondents involved	0
Demonstration, evaluation, and communication	
SUS[b] questionnaire	
Patients with stroke	14
Caregivers of patients with stroke	109
Health care workers	23
PSSUQ[c] questionnaire	
Health care workers	14
Patients with stroke	41
Caregivers of patients with stroke	73
Interview	
Health care workers	20
Patients with stroke	5

[a]DSR: design science research.

[b]SUS: System Usability Scale.

[c]PSSUQ: Post-Study System Usability Questionnaire.

Analysis Methods

The data were analyzed using 2 approaches. Qualitatively, we identified important characteristics or features required in the app design; quantitatively, we evaluated and measured the quality of the app prototypes. The analysis process began in the first iteration with semistructured interviews to identify user needs. In the next stage, we implemented a thematic analysis technique proposed by Ozuem et al [22]. This involved classifying and grouping problems while counting the number of respondents who confirmed each problem. Once categorized, the problems were mapped onto questions. The HMW framework was used as a guide to develop answers to each question [23]. We then converted the data into persona profiles that described the users' characteristics. The information architecture design was also developed at this stage to systematically organize the app content and facilitate user accessibility according to the corresponding personas. Semistructured interviews were conducted as a form of design evaluation in the second and third iterations. The evaluation was carried out using 2 evaluation tools—SUS and PSSUQ—to analyze 3 key components: system usability, information quality, and interface quality [25].

Research Instruments

Questionnaires and interviews were used during the evaluation stage. The questionnaire was conducted in the second and third iterations. In the second iteration, the questionnaire consisted of questions about demographics and the use of health apps. This section addressed patients who experienced stroke. The demographic questions touched on age, residence, occupation, number of devices, and email. Questions about the use of health apps covered experiences with mHealth related to stroke rehabilitation (such as app names and disadvantages). In the third iteration, results from the SUS questionnaire in the second iteration and the PSSUQ in the third iteration were used to test usability levels and questions about stroke rehabilitation. The SUS uses a 5-point Likert scale, where 1 indicates "strong disagreement" and 5 indicates "strong agreement," while the PSSUQ uses a different Likert scale, where 7 indicates "strong disagreement" and 1 indicates "strong agreement." Multimedia Appendix 1 further explains SUS, and Multimedia Appendix 2 details PSSUQ. Multimedia Appendix 3 presents the interview questions used for each respondent.

XSL•FO
RenderX

Results

Identifying Problems and Defining Solutions

This stage began with a literature review on the epidemiology of stroke globally and in Indonesia, as well as challenges in handling stroke rehabilitation. We also conducted interviews with 10 health workers (general practitioners, medical students, and therapists) and 8 patients who have experienced a stroke (or their caregivers as intermediaries) to understand patients' needs and challenges. Multimedia Appendix 4 summarizes the problems emerging from interviews with 18 respondents, comprising 10 experts and 8 patients. These respondents identified various problems in stroke rehabilitation, including a lack of self-rehabilitation exercises at home, low patient compliance with exercises, unclear roles and low willingness among medical personnel, the decreased ability of patients to undergo rehabilitation in the hospital due to age, limited health care facilities, the variety of conditions among patients who have experienced a stroke, and the absence of evidence on home exercises as a form of evaluation, which leaves evaluations reliant solely on patient reports, which are often subjective and not always accurate.

Problem solution mapping was used to understand these problems and find solutions. This study adopted the HMW method, in which questions were formulated based on identified problems. The research continued by exploring all possible solutions that were relevant and effective for solving the problems. The solutions discovered through the HMW method are summarized in Multimedia Appendix 5.

Subsequently, personas were constructed to understand users' needs, motivations, and goals in stroke rehabilitation. Based on the previous stages, 2 main personas with different characteristics were identified: health care workers and patients who experienced a stroke. These 2 types have different roles and needs in stroke rehabilitation management. Thus, the app was designed for these individuals.

The patient persona represents a man who had a stroke at the age of 56 years and worked as a lawyer. Currently, he is undergoing a physical rehabilitation process at home with the support of his family members. The stroke he experienced impacted his physical abilities, especially in terms of mobility and body coordination. In addition, he experiences communication disorders, which make his pronunciation unclear. Such conditions make it hard for patients to remember the frequency and types of exercise that need to be performed after discharge from the hospital, while also making it difficult to visit the hospital due to cognitive impairment. This persona has 2 basic needs in his rehabilitation process: the support and encouragement of his family or friends to maintain his spirit, and the presence of his family to help guide him through the rehabilitation process.

The persona of the medical personnel represents a female who is aged 23 years and lives in Jakarta. As a professional in physical rehabilitation, she has 2 years of experience as a physiotherapist. She has treated several patients who have experienced a stroke, but the rehabilitation process has not always proceeded smoothly. She recognizes the challenges in dealing with patients who have experienced a stroke, which involve both internal patient factors and external factors that affect the recovery process. She has 2 basic needs in carrying out her duties as a physiotherapist: the ability to effectively manage post-stroke rehabilitation, and the need to improve both her nonclinical and clinical skills to offer optimal care. Because she is motivated, she has a realistic perspective and always appreciates patients' progress. However, she has difficulty motivating patients to undergo stroke rehabilitation consistently and is unable to monitor patients' exercises at home.

Design and Development

Based on the mapping of problems and solutions in the previous stage, this stage involved translating user needs into an initial design that could be tested and evaluated. The low-fidelity (lo-fi) app prototype from the first iteration was designed to visualize the flow of user interaction and ensure that every feature and flow was user-friendly. The high-fidelity (hi-fi) prototypes from the second and third iterations were used to determine how users would interact with an app featuring visual elements and interactivity. This design also tested user flow and design consistency. Figures 2-4 present examples of the lo-fi prototype design, featuring the home screen, stroke education, and stroke rehabilitation activity targets. Figures 5-7 display examples of the hi-fi prototype design with the same content.

XSL•FO
RenderX

Figure 2. Low-fidelity home screen.



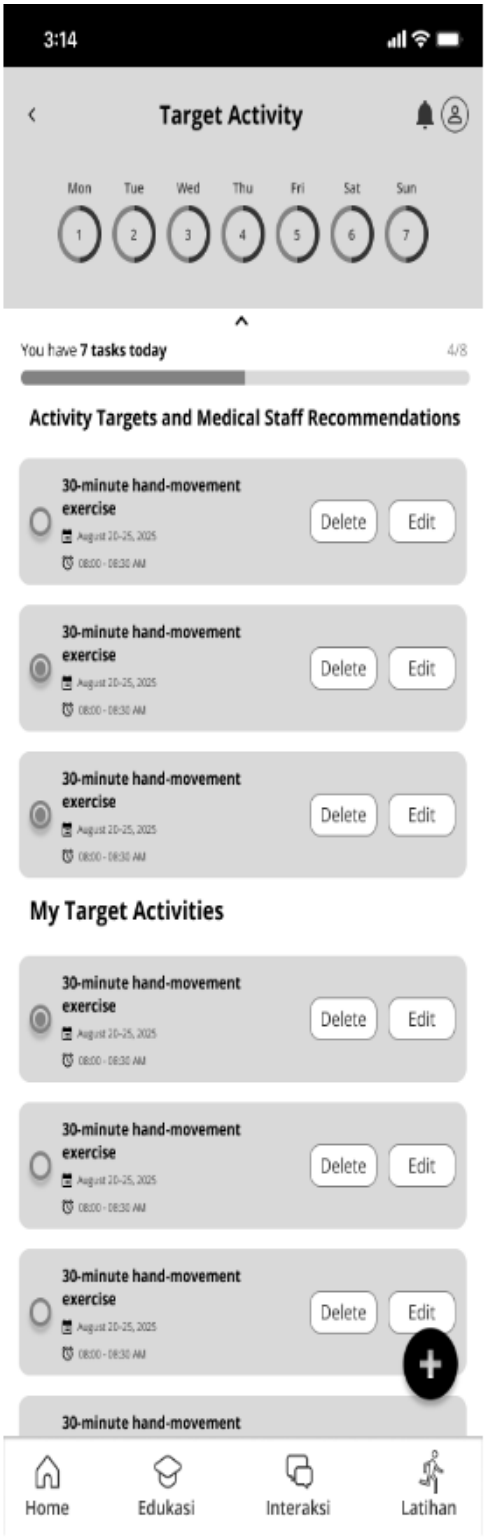
XSL•FO
RenderX

Figure 3. Low-fidelity stroke education.



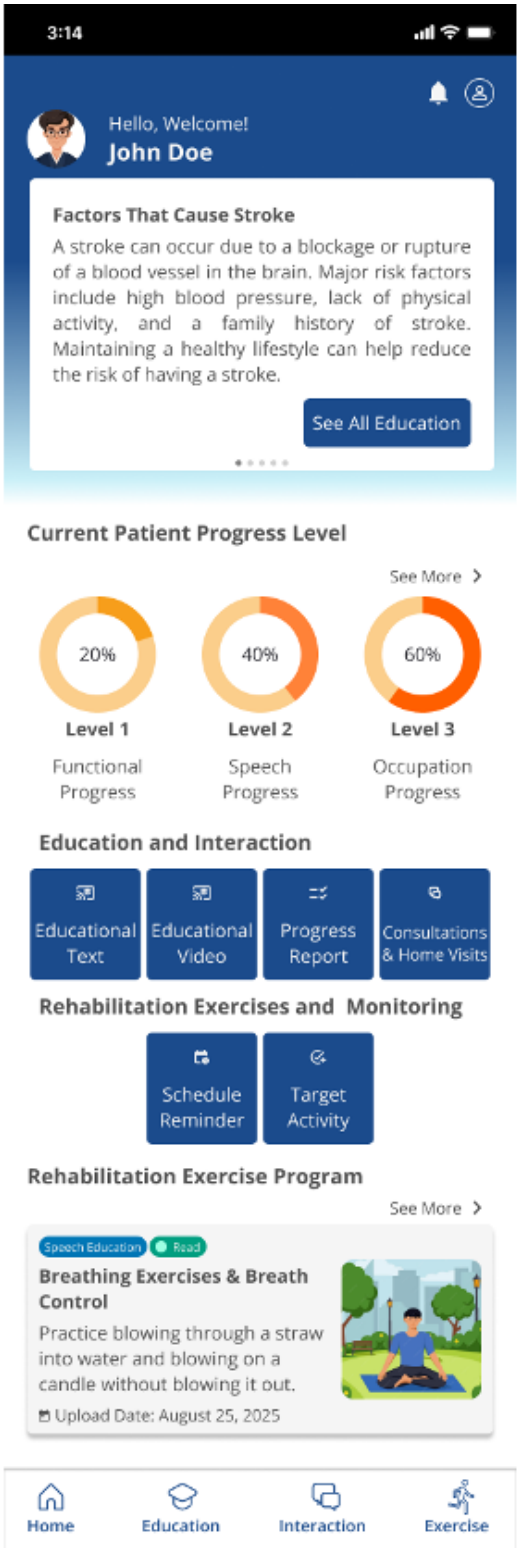
XSL•FO
RenderX

Figure 4. Low-fidelity stroke activity.



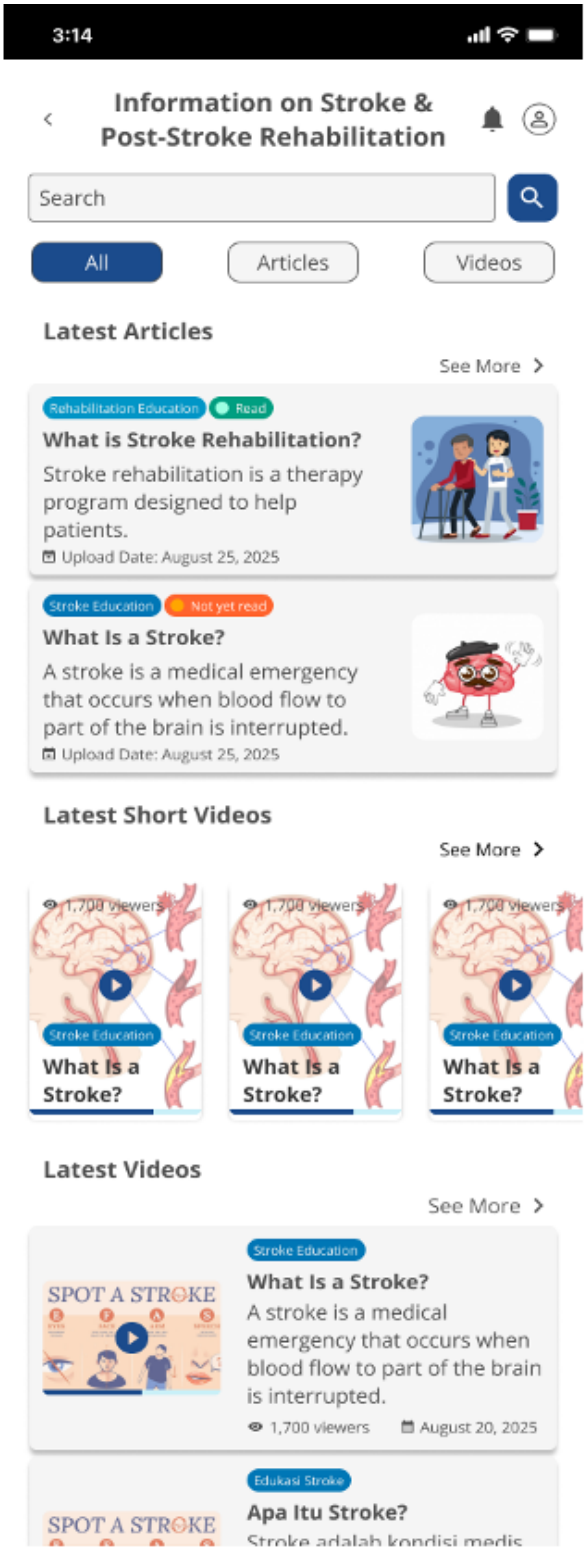
XSL•FO
RenderX

Figure 5. High-fidelity home screen.



XSL•FO
RenderX

Figure 6. High-fidelity stroke education.



XSL•FO
RenderX

Figure 7. High-fidelity stroke activity.

XSL•FO
RenderX

Demonstration, Evaluation, and Communication

In the demonstration stage, we showed interviewees the lo-fi app prototype, and they provided feedback on the design in the evaluation and communication stages. This yielded various inputs regarding the initial design features. All reviews submitted by the interviewees were validated and analyzed to determine improvement priorities. Table 2 summarizes the lo-fi prototype evaluation, which led to improvements in the hi-fi version.

Table . Summary of low-fidelity prototype evaluations.

Improved features	Improvements
Authentication	• Using a registration certificate number (STR) instead of a practice license number
User profile management features	• Adding data on patients' disease history, drug allergies, and medical records • Adding an STR number for medical personnel data • Linking to the national data system and restricting access to authorized medical personnel
Therapy results report submission feature	• Connecting to a consultation or home-visit service • Sending examination results to medical personnel in video format
Gamification (pop-up quiz)	• Adding a "Skip" button for questions
Activity target features	• Improving concision • Adding long- and short-term activity targets
Consultation and home-visit features	• Adding patient medical record data to the home-visit feature • For consulting services, removing rehabilitation progress trackers

The evaluation and communication stages were carried out using surveys and interviews to collect feedback on the design. Both quantitative (SUS) and qualitative methods were used. Evaluation data for SUS were collected from 146 respondents from November 1, 2025, to November 7, 2025. The respondents who filled out the SUS questionnaire included a balanced number of men and women, with 73 respondents aged 26 to 35 years (n=146, 50%). The average SUS score was 71.83, a "Good" result [26]. Based on these results, we interviewed 20 medical personnel and 5 patients with stroke from November 7, 2025, to November 9, 2025. Multimedia Appendix 6 summarizes the recommended improvements.

The final evaluation was carried out using the PSSUQ to measure app design quality based on the users' perceptions of ease of use and information quality. The PSSUQ was distributed from November 17, 2025, to November 21, 2025, and was returned by 128 respondents, including 69 (54%) women and 59 (46%) men. Participants aged 26 to 35 years comprised the largest group with 40 (31%) respondents, followed by the 17 to 25 group with 33 (26%) respondents. Table 3 indicates that these participants found the improved app easy to implement and use.

Table . Post-Study System Usability Questionnaire results.

Category	System usefulness	Information quality	Interface quality	Overall system
Red	2.32	2.39	2.42	2.35
Lower limit	2.12	2.19	2.23	2.16
Upper limit	2.52	2.9	2.60	2.54

The app's features were grouped into 4 categories of health service approaches, namely promotive, preventive, curative, and rehabilitative. Promotion focuses on knowledge, awareness, and user participation so that they are proactive in maintaining their health. Prevention pertains to the avoidance or early detection of health problems through routine monitoring and assessment. Curative care concerns the treatment of conditions through access to medical services. Rehabilitation focuses on restoring cognitive, emotional, and social functions through fun, therapeutic activities.

The entire list of app features mapped to each persona is listed in Multimedia Appendix 7. Promotion is supported by onboarding, video and text education features, and rehabilitation progress monitoring. Prevention is facilitated by target activity, gamification, and rehabilitation schedule reminder features. Curative care is supported by consultation, therapy results reports, and home-visit service features. Rehabilitation is facilitated by a rehabilitation training program and user profile management to tailor rehabilitation interventions to the individual circumstances of each patient. In addition, we implemented authentication features that allow each user to securely log in and access the system with a verified identity, thereby ensuring the security and personalization of patients' rehabilitation data.

XSL•FO
RenderX

All features were mapped onto intervention theory principles to ensure that the features had a strong interventional basis. Intervention theory focuses on 3 main aspects: valid and useful information; free, informed choices; and internal commitment [20]. Valid and useful information requires that the data provided by the system or intervention must be accurate, trustworthy, and relevant to the needs of the user [20]. Free, informed choice means that users have the liberty to make decisions without coercion or manipulation [20]. Internal commitment mandates that users have the internal motivation and commitment to carry out and complete intervention programs without any external pressure or coercion [20]. Feature mapping was carried out to document the theoretical basis of each feature, ensure the completeness of behavioral change components, and provide theoretical justifications for design decisions. Mapping each feature onto the aforementioned principles (Table 4) allowed us to evaluate how well each feature contributed to an effective intervention process.

XSL•FO
RenderX

Table . Feature mapping with intervention theory.

Persona, principle of intervention theory, and feature	Theoretical justification
Patients who experienced a stroke	
Valid and useful information	
Onboarding	Provides accurate initial information on the app's functionality and a thorough introduction to the app's features prior to direct use by patients who have experienced a stroke.
General educational video features related to stroke and stroke rehabilitation	Provide education in video format about stroke and stroke rehabilitation that is relevant to patients performing stroke rehabilitation.
General educational articles	Provide education in article format about stroke and stroke rehabilitation that is relevant to patients performing stroke rehabilitation.
Rehabilitation progress monitoring	Provides detailed information related to the patient's rehabilitation progress based on the condition and the latest assessment of medical personnel.
Rehabilitation training program	Provides education in the form of articles, videos, and short videos on stroke rehabilitation exercises at home.
Free, informed choice	
Pop-up quiz (gamification)	Provides freedom for patients who have experienced a stroke to fill out pop-up quizzes or skip them.
Reminders	Provide freedom for patients who have experienced a stroke to manage schedules, whether exercise schedules, consultations, or home visits.
User profile management	Provides freedom for patients who have experienced a stroke to manage their personal data.
Therapy reports	Provide freedom for patients who have experienced a stroke to submit therapy results to medical personnel in video format.
Health consultation	Provides freedom for patients who have experienced a stroke to consult with and choose available medical personnel.
Home-visit service	Provides freedom for patients who have experienced a stroke to conduct home visits and choose available medical personnel.
Internal commitment	
Activity targets	Giving patients control over setting personal goals creates a sense of ownership, so that the motivation to achieve them comes from their own intrinsic desires, not external compulsions.
Medical personnel	
Valid and useful information	
Onboarding	Provides accurate initial information on the app's functionality and a thorough introduction to the app's features before being used directly by medical personnel.
General educational video	Allows medical personnel to manage education in video format about stroke and stroke rehabilitation by ensuring that the information provided to patients remains accurate and relevant.
General educational articles	Allow medical personnel to manage education in article format on stroke and stroke rehabilitation by ensuring that the information provided to patients remains accurate and relevant.
Rehabilitation progress monitoring	Allows medical personnel to update patient rehabilitation progress data validly and accurately according to the patient's current condition.
Rehabilitation training program	Allows medical personnel to manage education in the form of articles, videos, and short videos on home stroke rehabilitation exercises by ensuring that the information provided to patients remains accurate and relevant.
Pop-up quiz (gamification)	Allows medical personnel to add pop-up quiz questions with correct answers.
Therapy results report	Allows medical personnel to provide verified and accurate evaluation results for rehabilitation progress.
Medical consultation	Allows medical personnel to serve patients who have experienced a stroke and provide accurate and verified evaluation results.
Home-visit service	Allows medical personnel to serve patients who have experienced a stroke and provide accurate and verified evaluation results.
Free, informed choice	

XSL•FO
RenderX

Persona, principle of intervention theory, and feature	Theoretical justification
User profile management	Provides freedom for medical personnel to manage their personal data in accordance with the decisions of medical personnel.
Internal commitment	
Activity target	Allows medical personnel to provide recommendations for targeted activities tailored to the patient's condition. This encourages patients to have a personal stake in completing predetermined rehabilitation activities.

Discussion

Principal Findings

Across 3 iterations involving doctors, therapists, medical students, and individuals interested in stroke rehabilitation issues, this study collected feedback to build an app based on patients' and related stakeholders' feedback. The app's main features are grouped into several interface categories, including exercise and progress monitoring programs, consultation and home-visit services, and educational content designed to support poststroke rehabilitation. These features were mapped onto the principles of intervention theory, namely valid information; free, informed choice; and internal commitment. The evaluations of the app design suggested an increase in usage and user satisfaction, as well as success in addressing the challenges and limitations identified in previous research.

In addition, this study discovered a new problem: decreased patient ability to undergo rehabilitation in hospitals due to age. This problem was overcome by adding a feature for home visits, which remains rarely addressed in stroke rehabilitation apps [17,18,27]. Another new problem in the Indonesian context concerns the mental struggles of patients, which impede rehabilitation. For example, low patient motivation to carry out stroke rehabilitation was overcome by adding activity targets to the app, thus ensuring consistency and triggering patient motivation.

In Indonesia, there is a lack of objective evidence regarding home care because evaluations rely on subjective patient reports. This problem was addressed by designing a rehabilitation progress monitoring feature and therapy result reports, integrated with patient medical records in the app; this allows patient progress to be evaluated properly, rather than relying solely on oral testimonials. Thus, several features of the app were adapted to the context in Indonesia, including home-visit features, rehabilitation training programs, gamification (pop-up quizzes), rehabilitation progress monitoring, activity targets, and therapy reports.

Theoretical Implications

Researchers have previously conducted systematic reviews or usability assessments of existing systems using various evaluation frameworks [15,17,18,27,28]. In contrast, this study designed a stroke rehabilitation management app by consulting patients and medical personnel using DSR. Thus, the feedback obtained represents the true needs of patients who experienced a stroke, including recovery processes and mechanisms underlying behavioral changes, as well as the promotive, preventive, curative, and rehabilitative aspects of health promotion. This study also expands the limited literature on stroke rehabilitation in developing countries [11].

There is a gap in the research [15,17] regarding the low adherence of users to rehabilitation procedures, which we addressed by adding activity target features and schedule reminders. This will help patients, families, and caregivers monitor patient rehabilitation in a more structured and scheduled manner. Another gap [18] is the limited variety of content for different stages of stroke and patient conditions, which we addressed by providing a range of educational content and exercises to suit individual needs. The app also provides tutorials on usage and onboarding, which were not included in the research conducted by Song et al [18]. Furthermore, previous research used relatively small samples and did not include all relevant respondents, especially patients [28]. Our app also minimizes or eliminates certain challenges of stroke rehabilitation. The app overcomes the limitations of health facilities by providing consultation and home-visit services and increasing digital literacy and compliance with rehabilitation exercises. The app can also mitigate the impact of stroke on patients by optimizing the body's functionality and preventing complications.

The importance of this study stems from the high prevalence of stroke in Indonesia and the lack of applications that support patient-centered stroke rehabilitation. Research in Canada [17], Belgium [15], and the United States [18] has discussed various mHealth apps for stroke rehabilitation, covering improved physical function, physical activity, and patient quality of life. However, these are all developed countries, and no previous work has focused on Indonesian patients with stroke. The problems identified in our interviews reflect those noted in previous works [15,17]. Furthermore, the app offers solutions such as schedule reminder features, gamification in the form of pop-up quizzes, and rehabilitation progress monitoring.

Practical Implications

For health care facilities, such as hospitals, therapy clinics, and health centers, our app can be used to monitor patients undergoing stroke rehabilitation or to help prevent recurrent strokes. For medical personnel, the app can help carry out rehabilitation services both online (through messages or voice media) and offline (by visiting the patient's home). This would make rehabilitation services more efficient and help evaluate the rehabilitation process based on patient medical data. Medical personnel can also provide education and training on stroke rehabilitation at home. Zeng et al [29] noted that improved health literacy among stroke survivors and caregivers improves self-management for survivors.

XSL•FO
RenderX

For patients, families, and caregivers, the app can provide educational information about stroke rehabilitation, including exercises that can be performed at home easily and affordably. Stroke apps should involve online communities to support survivors and caregivers [16,30]. Moreover, the app facilitates consultation services and home visits from medical personnel. There are also schedule reminders, activity targets, and progress-tracking features for patients, families, and caregivers to monitor speech, occupation, and function.

For mHealth developers, this app is built to serve the needs and expectations of users in Indonesia. Likewise, for the government or policymakers, it can be connected to hospitals and official services. The app can be used to collect data on patient activity patterns after stroke and to provide evidence of its effectiveness in supporting stroke rehabilitation. The government should consider the value of medical personnel being able to access educational and training information for patients, families, and caregivers.

Limitations

Despite these strengths, the study has limitations, such as limited digital literacy among participants, which made it difficult for them to understand certain features or use the app independently. Therefore, training on using the system is needed. Other difficulties included technical limitations when using the application, such as internet connection problems, device availability, operating system compatibility, and device storage capacity.

Digital apps can be difficult for older adults to use. Most of our participants were young adult caregivers of patients with stroke, which explains the lack of older participants. It was difficult to accommodate all users due to the limited accessibility of technology in remote areas and the age of patients who experienced a stroke. Although older patients may have difficulty using the app independently, caregivers, families, or online communities can help them use features that present challenges, ensuring that the collected data represent the needs of all age groups.

Further research is recommended to conduct longitudinal analyses measuring long-term adherence, app usage sustainability, and the impact on rehabilitation outcomes among patients who have experienced a stroke. Researchers could also adopt supportive technologies, such as wearable devices, to make the app accessible to users with limitations.

Acknowledgments

The authors confirm that no AI tools were used in the preparation of this manuscript.

Funding

This study was supported by the Faculty of Computer Science University of Indonesia for the internal research year 2026.

Data Availability

The data generated and analyzed during this study are not publicly available due to a lack of authority to share the data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
System Usability Scale (SUS) questionnaire questions.
[DOCX File, 14 KB - rehab_v13i1e91464_app1.docx]

Multimedia Appendix 2
Post-Study System Usability Questionnaire (PSSUQ) questionnaire questions.
[DOCX File, 16 KB - rehab_v13i1e91464_app2.docx]

Multimedia Appendix 3
Interview questions.
[DOCX File, 20 KB - rehab_v13i1e91464_app3.docx]

Multimedia Appendix 4
Summary of problems.
[DOCX File, 15 KB - rehab_v13i1e91464_app4.docx]

Multimedia Appendix 5
Problem solution mapping with the “how might we” technique.
[DOCX File, 16 KB - rehab_v13i1e91464_app5.docx]

XSL•FO
RenderX

Multimedia Appendix 6
Summary of high-fidelity prototype improvements.
[DOCX File, 17 KB - rehab_v13i1e91464_app6.docx]

Multimedia Appendix 7
Feature summaries.
[DOCX File, 34 KB - rehab_v13i1e91464_app7.docx]

References

1. Li X, He Y, Wang D, Rezaei MJ. Stroke rehabilitation: from diagnosis to therapy. Front Neurol 2024;15:1402729. [doi: 10.3389/fneur.2024.1402729] [Medline: 39193145]
2. Integrating stroke services in health-care systems: a practical approach. : World Health Organization; 2025 URL: https://iris.who.int/server/api/core/bitstreams/66ab26d4-ec9b-4b72-a9f6-17832d32ca34/content [accessed 2026-06-25]
3. Stroke Rehabilitation. Barrow Neurological Institute. URL: https://www.barrowneuro.org/treatment/stroke-rehabilitation/ [accessed 2026-06-25]
4. Shahid J, Kashif A, Shahid MK. A comprehensive review of physical therapy interventions for stroke rehabilitation: impairment-based approaches and functional goals. Brain Sci 2023 Apr 25;13(5):717. [doi: 10.3390/brainsci13050717] [Medline: 37239189]
5. Laili ZR, Prasetyo J, Ellina AD. Analysis of the effect of speech therapy services on post-stroke patients' quality of life through the role of family in medical rehabilitation. Indones J Glob Health Res 2025;7(3):343-350. [doi: 10.37287/ijghr.v7i3.6093]
6. Vadas D, Prest K, Turk A, Tierney S. Understanding the facilitators and barriers of stroke survivors' adherence to recovery-oriented self-practice: a thematic synthesis. Disabil Rehabil 2022 Nov;44(22):6608-6619. [doi: 10.1080/09638288.2021.1968512] [Medline: 34460334]
7. Feigin VL, Brainin M, Norrving B, et al. World Stroke Organization (WSO): global stroke fact sheet 2022. Int J Stroke 2022 Jan;17(1):18-29. [doi: 10.1177/17474930211065917] [Medline: 34986727]
8. Zainuddin AA, Kadir RRA, Kuswanto H, et al. The major risk factor of stroke across Indonesia; a nationwide geospatial analysis of universal health coverage program. Arch Public Health 2025 Jul 1;83(1):169. [doi: 10.1186/s13690-025-01613-4] [Medline: 40598345]
9. Profil Kesehatan Indonesia 2023 [Report in Indonesian]. : Kementerian Kesehatan Republik Indonesia; 2024 URL: https://kemkes.go.id/id/profil-kesehatan-indonesia-2023 [accessed 2026-06-25]
10. Survei Kesehatan Indonesia (SKI) 2023 Dalam Angka [Report in Indonesian]. : Badan Kebijakan Pembangunan Kesehatan (BKPK), Kementerian Kesehatan Republik Indonesia; 2024 URL: https://www.badankebijakan.kemkes.go.id/ski-2023-dalam-angka/ [accessed 2026-06-25]
11. Kayola G, Mataa MM, Asukile M, et al. Stroke rehabilitation in low- and middle-income countries: challenges and opportunities. Am J Phys Med Rehabil 2023 Feb 1;102(2S Suppl 1):S24-S32. [doi: 10.1097/PHM.0000000000002128] [Medline: 36634327]
12. Morone G, Pichiorri F. Post-stroke rehabilitation: challenges and new perspectives. J Clin Med 2023 Jan 10;12(2):550. [doi: 10.3390/jcm12020550] [Medline: 36675486]
13. Devereux N, Berns AM. Evaluation & treatment of psychological effects of stroke. Dela J Public Health 2023 Aug;9(3):62-69. [doi: 10.32481/djph.2023.08.011] [Medline: 37701469]
14. Bocean CG, Vărzaru AA. Health status in the era of digital transformation and sustainable economic development. BMC Health Serv Res 2025 Mar 5;25(1):343. [doi: 10.1186/s12913-025-12498-y] [Medline: 40045359]
15. Rintala A, Kossi O, Bonnechère B, Evers L, Printemps E, Feys P. Mobile health applications for improving physical function, physical activity, and quality of life in stroke survivors: a systematic review. Disabil Rehabil 2023 Dec;45(24):4001-4015. [doi: 10.1080/09638288.2022.2140844] [Medline: 36325613]
16. Handayani PW, Alifia K, Sutanto J, et al. Online community support for stroke survivors and caregivers: scoping review. J Med Internet Res 2026 Apr 29;28:e71190. [doi: 10.2196/71190] [Medline: 42054545]
17. Szeto SG, Wan H, Alavinia M, Dukelow S, MacNeill H. Effect of mobile application types on stroke rehabilitation: a systematic review. J Neuroeng Rehabil 2023 Jan 24;20(1):12. [doi: 10.1186/s12984-023-01124-9] [Medline: 36694257]
18. Song AJ, Lugo L, Muccini J, Mlynash M, Lansberg MG. EngageHealth: a mobile device application designed to deliver stroke rehabilitation exercises using asynchronous video recordings. Front Stroke 2024;3:1418298. [doi: 10.3389/fstro.2024.1418298] [Medline: 41542231]
19. Jat AS, Grønli TM, Ghinea G. Navigating design science research in mHealth applications: a guide to best practices. IEEE Trans Eng Manage 2024;71:14472-14484. [doi: 10.1109/TEM.2024.3450178]
20. Nguyen A, Tuunanen T, Gardner L, Sheridan D. Design principles for learning analytics information systems in higher education. Eur J Inf Syst 2021 Sep 3;30(5):541-568. [doi: 10.1080/0960085X.2020.1816144]

XSL•FO
RenderX

21. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inf Syst 2007 Dec;24(3):45-77. [doi: 10.2753/MIS0742-1222240302]
22. Ozuem W, Willis M, Howell K. Thematic analysis without paradox: sensemaking and context. Qual Mark Res Int J 2022 Jan 19;25(1):143-157. [doi: 10.1108/QMR-07-2021-0092]
23. Laine J, Zimmer MP, Minkkinen M, Salmela H, Mäntymäki M. Clustering design science research based on the nature of the designed artifact. In: Papagiannidis S, Alamanos E, Gupta S, Dwivedi YK, Mäntymäki M, Pappas IO, editors. The Role of Digital Technologies in Shaping the Post-Pandemic World: Springer; 2022. [doi: 10.1007/978-3-031-15342-6_20]
24. Pérez-Montoro M, Codina L. Chapter 3—designing user experience. In: Navigation Design and SEO for Content-Intensive Websites: Chandos Publishing; 2017:65-84. [doi: 10.1016/B978-0-08-100676-4.00003-1]
25. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
26. Nicora G, Santangelo G, Marzero G, Dagliati A, Aprile I, Quaglini S. Exploring usability standards in robotic rehabilitation: a meta-analysis of system usability scale scores. IEEE Int Conf Rehabil Robot 2025 May;2025:476-482. [doi: 10.1109/ICORR66766.2025.11063026] [Medline: 40644156]
27. Chien SY. A usability evaluation framework for a mobile application in supporting home-based rehabilitation for stroke patients: a qualitative study. Digit Health 2025;11:20552076251340183. [doi: 10.1177/20552076251340183] [Medline: 42186542]
28. Rätz R, Ratschat AL, Cividanes-Garcia N, Ribbers GM, Marchal-Crespo L. Designing for usability: development and evaluation of a portable minimally-actuated haptic hand and forearm trainer for unsupervised stroke rehabilitation. Front Neurorobot 2024;18:1351700. [doi: 10.3389/fnbot.2024.1351700] [Medline: 38638360]
29. Zeng M, Liu Y, He Y, Huang W. Relationship between stroke knowledge, health information literacy, and health self-management among patients with stroke: multicenter cross-sectional study. JMIR Med Inform 2025 Jun 23;13:e63956. [doi: 10.2196/63956] [Medline: 40549967]
30. Hestetun-Mandrup AM, Hamre C, Lund A, Martinsen ACT, He HG, Pikkarainen M. Exploring the influence of digitalization on multidisciplinary poststroke rehabilitation practice: qualitative study. JMIR Rehabil Assist Technol 2026 Feb 17;13:e77753. [doi: 10.2196/77753] [Medline: 41701937]

Abbreviations

DSR: design science research
hi-fi: high-fidelity
HMW: how might we
lo-fi: low-fidelity
mHealth: mobile health
PSSUQ: Post-Study System Usability Questionnaire
SUS: System Usability Scale
WHO: World Health Organization

Edited by A Mastropietro; submitted 14.Jan.2026; peer-reviewed by J Zlamal, L Herdiman; revised version received 10.Jun.2026; accepted 16.Jun.2026; published 23.Jul.2026.

Please cite as:
Shania SA, Handayani PW, Imanuddin KA
Designing an mHealth App for Stroke Rehabilitation in Indonesia: Mixed Methods Design Science Research Study
JMIR Rehabil Assist Technol 2026;13:e91464
URL: https://rehab.jmir.org/2026/1/e91464
doi:10.2196/91464

© Sabrina Atha Shania, Putu Wuri Handayani, Kamila Alifia Imanuddin. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Deep Learning–Based Automated Clinical Gait Assessment From Kinematic Data in People With Stroke: Development and Validation Study

Zhexuan Gu[1*], MSc; Jiaqi Li[2,3*], MSc; Wenbin Zhou[4], MSc; Kenneth N K Fong[2], PhD; Lin Wang[3], PhD; Eng Keong Lua[5], PhD; Patrick W H Kwong[2], PhD; Yancheng Yuan[6], PhD

[1]Department of Data Science & Artificial Intelligence, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, KLN, Hong Kong, China (Hong Kong)

[3]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

[4]Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong, China (Hong Kong)

[5]Computer Laboratory, University of Cambridge, Cambridge, England, United Kingdom

[6]Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Patrick W H Kwong, PhD
Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, KLN, Hong Kong, China (Hong Kong)

Abstract

Background: Gait impairment is a prevalent sequela of stroke. Although observational gait analysis remains a standard clinical practice for assessing neuromotor impairments in people with stroke, it is prone to subjective bias. Consequently, objective assessment is required to inform effective rehabilitation protocols.

Objective: This study aimed to validate a deep learning framework for automating clinical gait assessment using kinematic data.

Methods: This study was conducted in a university-affiliated gait analysis laboratory. Kinematic data were collected from 51 individuals with hemiparetic stroke and 18 healthy controls using wearable inertial measurement units, and video recordings were obtained for the Wisconsin Gait Scale (WGS) scoring. WGS scores, rated by 2 experienced physiotherapists, were used to establish the expert reference standard. A statistical graph convolutional network (STAT-GCN), incorporating a STAT-Attention Head module, was developed to predict WGS items 2 to 14 from kinematic data. Model evaluation was performed using a stratified, participant-level 5-fold cross-validation protocol, with all gait cycles from the same participant kept within the same fold. Model performance was assessed at the participant level using exact accuracy, mean absolute error, and weighted κ metrics. A descriptive comparison involving 4 final-year physiotherapy students was also conducted.

Results: Among participants with stroke, STAT-GCN achieved an overall exact accuracy of 75.42% for WGS items 2 to 14, with a mean absolute error of 0.290 score levels. Across the participant-level 5-fold cross-validation test sets, STAT-GCN achieved a mean exact accuracy of 75.64% (SD 6.58%), whereas the 4 student evaluators achieved mean accuracies ranging from 56.19% (SD 7.03%) to 67.09% (SD 4.03%).

Conclusions: The proposed STAT-GCN framework demonstrated reasonable preliminary performance for automated WGS item prediction in individuals with stroke. The results suggest that kinematic data–driven models may support more standardized gait assessment, but performance should be interpreted cautiously given the limited sample size, class imbalance across WGS score items, and absence of external validation. Further studies with larger and more clinically diverse stroke cohorts are needed before clinical implementation.

(JMIR Rehabil Assist Technol 2026;13:e94031) doi:10.2196/94031

KEYWORDS

deep learning; gait; graph convolution network; stroke; Wisconsin Gait Scale

XSL•FO
RenderX

Introduction

Stroke is a principal contributor to global mortality and long-term disability [1,2]. It constitutes a major public health challenge, imposing substantial societal and health care burdens worldwide. The complications of stroke are debilitating and commonly include muscle weakness, spasticity, and limb apraxia, which contribute to gait impairments [3,4]. These impairments present as reduced cadence and speed, resulting in an abnormal gait pattern. These impairments may also lead to difficulties in maintaining balance [5-7]. Consequently, gait impairments affect the level of community integration in patients with stroke [8].

Observational gait analysis remains a conventional clinical practice for assessing neuromotor impairments in people with stroke; however, this methodology is prone to subjective bias in clinical assessments [9]. Thus, for greater systematicity, standardized gait analysis tools such as the Wisconsin Gait Scale (WGS) have been adopted to evaluate videotaped gait patterns in individuals with stroke [10]. The WGS [11] is a widely adopted assessment tool for evaluating gait abnormalities. It comprises 14 items, of which 13 are related to gait performance and gait patterns and are rated through observation of gait. The WGS is clinically efficient, requiring minimal instrumentation, and has demonstrated high intrarater reliability (intraclass correlation coefficient [ICC]=0.83) in poststroke validation studies [12]. However, manual scoring by a rehabilitation specialist is not only time-consuming but also subject to bias in subjective clinical judgment [13]. Furthermore, rating many WGS items requires the observation of subtle changes in joint angles, such as the rotation angle of a specific joint. The visual identification of these subtle changes is highly challenging.

AI has been increasingly used to automate and reduce subjective bias in clinical assessment [14,15]. A multimodal graph fusion network was proposed in a previous study [16] to detect the freezing of gait, a disabling symptom of Parkinson disease. The graph fusion network comprised 4 modalities, each corresponding to a distinct data source. Graph fusion networks have also been used successfully to determine the severity of parkinsonism [17,18]. Similarly, a study [19] successfully distinguished Parkinson disease cases from controls using a novel model, namely a multiview graph convolutional network (GCN). Human motion analysis represents an evolving computational paradigm where skeleton-based kinematic time-series data serve as critical inputs for emerging action recognition algorithms [18,20-22]. GCN-based models have demonstrated state-of-the-art performance on established motion analysis benchmarks, achieving >93% accuracy on the NTU RGB+D dataset [23]. However, translational applications in clinical gait assessment remain limited [24], primarily due to the fundamental disparity between action classification and quantitative clinical evaluation. Evaluating an action in clinical settings requires a more nuanced level of analysis than simple recognition, as it involves complex assessments of quality, precision, and clinical relevance [25]. Standard deep learning architectures, such as convolutional neural networks and recurrent neural networks, often struggle with this level of granularity because they treat skeletal data as flat sequences or grids, ignoring the intrinsic biomechanical connectivity between joints. This loss of topological information makes it difficult to quantify the subtle deviations in interjoint coordination that characterize pathological gait.

To address this gap, we propose a novel statistical graph convolutional network (STAT-GCN) for the automation of WGS scoring. By using gait kinematic data collected from inertial measurement units (IMUs) [26], this model enables the representation of joint angular trajectories as spatiotemporal features on a graph structure. By preserving the topological structure of lower-limb joints, this approach is intended to improve the model's ability to learn clinically relevant gait features for automated WGS item scoring.

Methods

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of The Hong Kong Polytechnic University (approval number HSEAR20210715003). All participants provided written informed consent. Participants' privacy and confidentiality were maintained throughout the study. All collected data were deidentified and stored securely, and only authorized members of the research team had access to the data. Participants received reimbursement for transportation expenses associated with study participation.

Participants

This study enrolled individuals with stroke according to the following criteria: (1) aged 18 to 80 years, (2) diagnosis of ischemic or hemorrhagic stroke confirmed by magnetic resonance imaging or computed tomography more than 6 months before assessment, (3) able to walk 10 m independently or with assistance, and (4) capable of providing informed consent. Individuals were excluded if they had any medical conditions that hindered their ability to participate in the assessment. In addition, 18 healthy individuals were recruited according to the following criteria: (1) aged 18 to 80 years, (2) no history of stroke or other diseases that may affect gait patterns, (3) able to walk independently, and (4) capable of providing informed consent. From March to August 2024, eligible participants were recruited through convenience sampling with assistance from local self-help groups.

Because no consensus has been established on sample size calculations in studies focused on developing deep learning models for various clinical applications [24], we referred to our previous study, which included 40 participants. In addition, an AI model using gait kinematic data achieved an accuracy of 84% in predicting the outcomes of clinical assessments of functional mobility conducted by physiotherapists [14]. Based on the sample sizes of these previous studies, for the development of a deep learning model, our trial aimed to recruit 51 participants. The study protocol was conducted in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1).

XSL•FO
RenderX

Outcome Measures

Gait Assessment and Clinical Measures

In this study, gait kinematics were measured in participants with stroke. Gait analysis and kinematic assessment were conducted using the MTw Awinda system (Xsens Technologies B.V.). This system consists of an IMU containing a 3D gyroscope, 3D accelerometer, and 3D magnetometer. Data were sampled at a frequency of 60 Hz. According to the official guidelines, body motion was captured by attaching 17 wireless sensors to the head, shoulders, sternum, upper arms, forearms, hands, pelvis, upper legs, lower legs, and feet of the individual. The IMUs were secured using MTw Velcro Body Straps and tape. Data were transmitted between the wireless tracker and the Awinda Station using the patented Awinda radio protocol, which was specially developed to ensure accurate 3D motion tracking even with unstable data transmission [27], and were processed using Xsens MVN 2024 software for gait analysis and kinematic assessment.

This study also used 2 distinct outcome measures:

1. WGS served as the primary outcome measure and our modeling target.
2. The Motor Subscale of the Fugl-Meyer Assessment of the Lower Extremity (FMA-LE) was used as a supplementary assessment to provide a broader characterization of participants' motor impairment.

WGS

The WGS is a gait assessment tool specifically designed for individuals with hemiplegia (Multimedia Appendix 1). The scale comprises 14 items and is designed to measure clinically relevant gait patterns during the stance phase, toe-off, swing phase, and heel strike of the gait cycle. Rodriguez et al [28] first introduced the WGS using videotaped gait recordings obtained in the frontal and sagittal planes. Each item is rated on an ordinal scale of either 1 to 3 or 1 to 4. The total score ranges from 14 to 45, with higher scores indicating greater gait impairment. As the primary outcome measure, WGS scores were assigned based on participants' gait. The ground truth was established based on consensus WGS scores agreed upon by 2 Hong Kong–registered physiotherapists, each with more than 10 years of clinical experience. Additionally, 4 final-year physiotherapy students provided WGS ratings to represent the perspectives of inexperienced raters.

FMA-LE

As a supplementary measure to characterize participants, the FMA-LE was used. Although the full FMA-LE evaluates motor function, sensation, range of motion, and pain [29], the motor subscale specifically is widely used to assess poststroke lower extremity motor impairment. This subscale evaluates reflex activity, movement in and out of synergy, as well as coordination and speed. It comprises 17 items, each rated on a 3-point ordinal scale (0‐2), summing to a maximum total score of 34 (indicating normal function) [30]. It demonstrates excellent intrarater reliability (ICC=0.99) and interrater reliability (ICC=0.91) [31].

Study Procedures

Prior to kinematic analysis using the Xsens MVN 2024 software, participants with stroke underwent a standardized assessment using the motor subscale of the FMA-LE [32]. Subsequently, following the Xsens user manual [33], 17 sensors were attached to the corresponding body segments of each participant for motion capture. Participants were instructed to walk back and forth in a straight line at a self-selected, comfortable speed over a flat 10 m surface. Three trials were recorded, ensuring that participants completed at least 5 full gait cycles in each direction. All walking trials were videotaped in both the sagittal and frontal planes. Finally, each trial was reviewed by 2 experienced physiotherapists to generate a consensus WGS score. The gait kinematic data exported from the software were then used to develop a deep learning model, namely, the STAT-GCN model, to predict WGS scores for participants with stroke. Item 1 of the WGS (ie, use of a walking aid) was excluded from the analysis because this item is easier to rate through direct observation.

Model Training and Testing

Cross-Validation and Data Partitioning

To ensure a rigorous evaluation of the proposed model, we implemented a stratified, participant-level 5-fold cross-validation protocol across the entire cohort (51 individuals with stroke and 18 healthy controls). The dataset partitioning was executed strictly at the participant level; that is, all gait cycles originating from a single participant were assigned exclusively to either the training set D_{train} or the testing set D_{test}. This participant-isolated strategy guarantees that the model was tested solely on unseen individuals, thereby eliminating the risk of intraparticipant data leakage.

The cross-validation folds were stratified by clinical groups to maintain an approximately balanced proportion of patients and controls across the folds. During the inference phase, a majority-voting mechanism was applied to aggregate cycle-level predictions into a definitive participant-level clinical score. The most frequent prediction across all available gait cycles for a given participant was designated as the final predicted score for each WGS item.

Data augmentation was applied only to the training gait cycles within each cross-validation fold [34]. Three time-series augmentation strategies were used to simulate natural gait variability and sensor measurement noise: jittering, scaling, and time warping. Jittering added small Gaussian noise to the original gait trajectories, scaling adjusted feature amplitudes to represent variations in stride intensity, and time warping introduced smooth temporal distortions to simulate cadence fluctuations within a gait cycle. Through these augmentation procedures, the number of training gait cycles increased from 979 to 10,152. No augmented samples derived from test participants were included in the training set, and test set performance was evaluated using nonaugmented gait cycles from unseen participants.

To reduce the risk of overfitting, the STAT-GCN was designed as a compact model. It consists of 3 spatiotemporal graph convolutional blocks followed by a statistical head, with channel

XSL•FO
RenderX

widths of 24, 48, and 48 across the 3 blocks. The STAT-Attention module was implemented as a lightweight feature-scaling mechanism rather than a dense pairwise self-attention layer. The same model configuration and hyperparameter settings were applied across all cross-validation folds. The model was trained with a batch size of 16 for 40 epochs. The initial learning rate was set to 1×10^{-3}, with a step decay factor of 0.8, and the weight decay was set to 0.1 for L2 regularization. No fold-specific hyperparameter tuning or test set–based model selection was performed.

Backbone Network

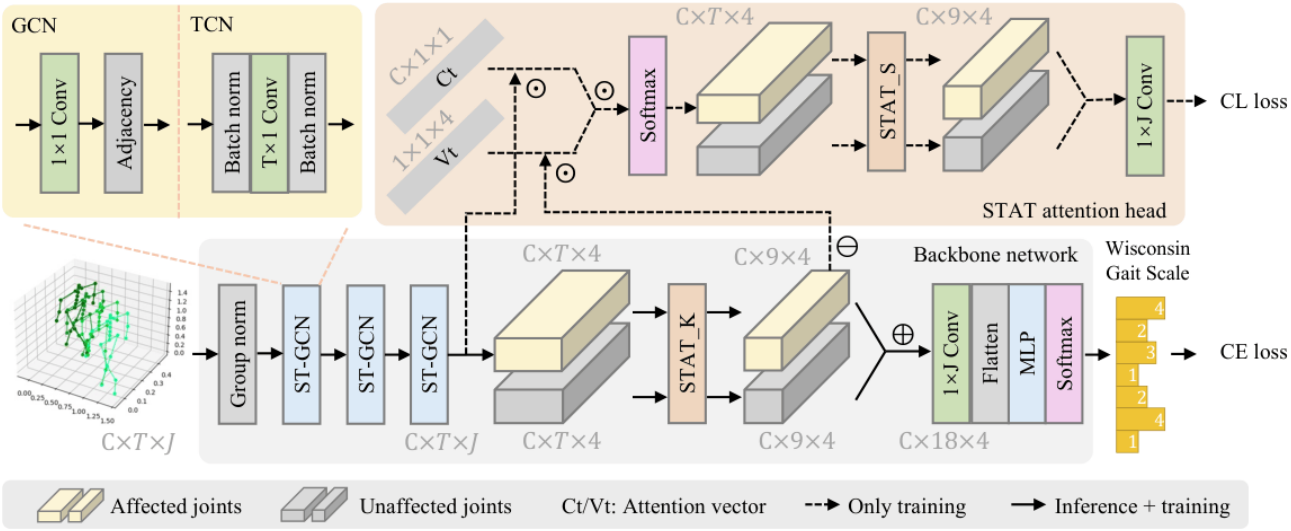
For our model, a batch of skeleton data, $D \in R^{(B\times C\times T\times J)}$, is used as input, where B is the batch size, C represents the number of raw features, T represents the number of frames in 1 gait cycle, and J is the number of joints. The backbone network is adapted from previous studies [18,35], with an additional statistics module.

The role of a spatio-temporal graph convolutional network (ST-GCN) block is 3-fold. It embeds low-level features into a high-dimensional hidden space, facilitates the exchange of information between joints, and aggregates features along the temporal dimension for each joint. By stacking multiple ST-GCN blocks, the model can gradually expand its receptive field and learn meaningful features.

STAT-GCN

To address the inadequate performance of prior data-driven approaches due to limited clinical data, we propose STAT-GCN (Figure 1). This model processes input skeleton data derived from motion capture to predict WGS scores. The inference architecture consists of a backbone network built from multiple ST-GCN blocks followed by a statistics module. To enhance feature discrimination, we introduce a STAT-Attention Head (SA Head) exclusively during training. This module uses contrastive loss and attention mechanisms to explicitly extract critical statistical features and highlight significant joints. Furthermore, by leveraging clinical labels regarding the affected side in participants with stroke, we segregate joint features into "affected" and "unaffected" categories to minimize misinterpretation. Consequently, STAT-GCN generates highly representative features tailored to each patient's specific mobility characteristics.

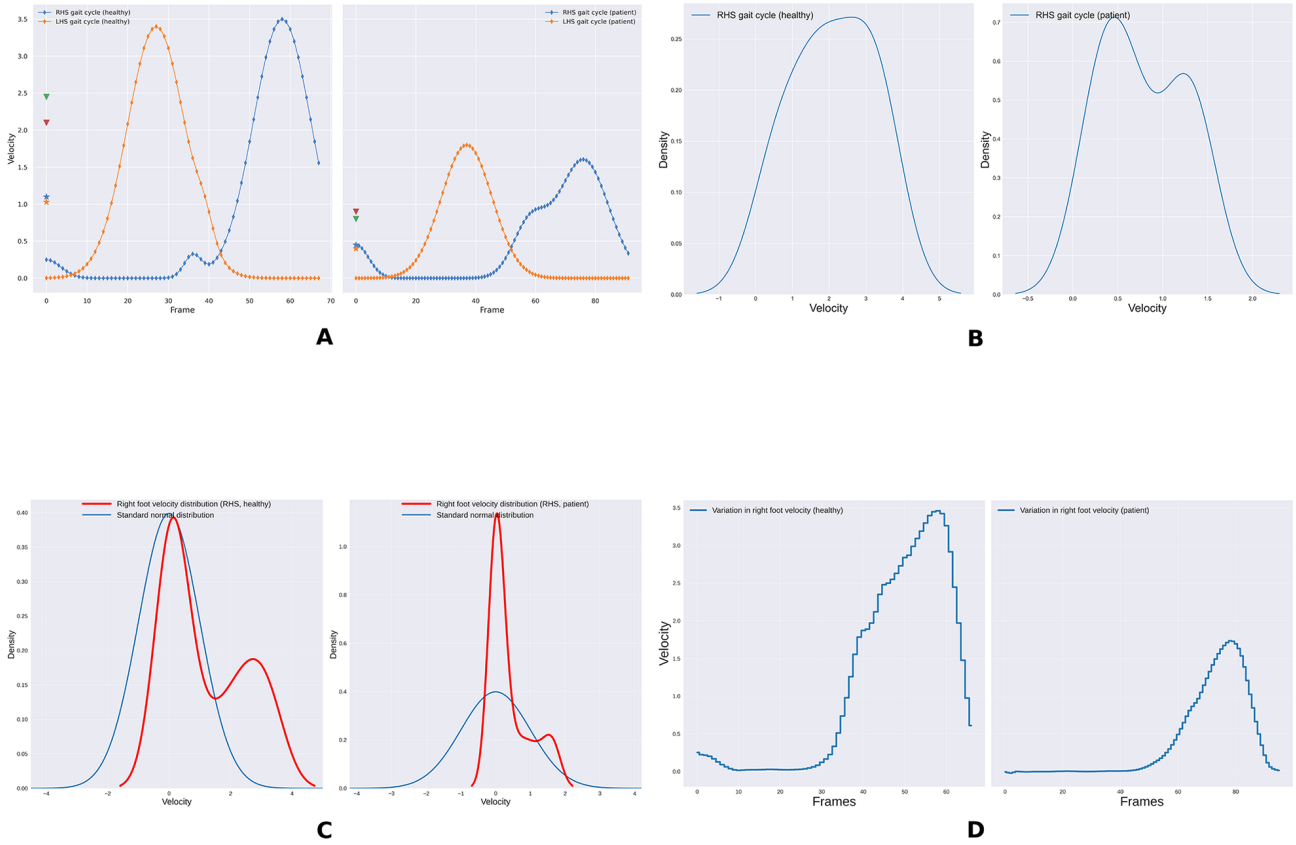
Figure 1. Proposed STAT-GCN architecture. CE: cross-entropy loss; CL: contrastive loss; Ct: channel attention vector; MLP: multilayer perceptron; ST-GCN: spatio-temporal graph convolutional network; STAT-GCN: statistical graph convolutional network; TCN: temporal convolutional network; Vt: vertex attention vector (graph vertex=joint).



The statistics module summarized gait trajectories using commonly used statistical measures, including range, maximum, mean, median, SD, and the 25th and 75th percentiles (IQR), as previously described [36]. These statistical descriptors were used to provide cycle-level summaries that were less dependent on the starting point of the gait cycle. As illustrated in Figure 2, gait cycles initiated from different events, such as right heel strike and left heel strike, may have different temporal representations despite reflecting comparable movement patterns.

XSL•FO
RenderX

Figure 2. Gait pattern visualizations based on velocity. (A) Comparison of right foot velocity across gait cycles and health conditions. The symbol ∇ indicates the first and third quartiles of the corresponding pattern, and the symbol indicates the median and mean. (B) Kernel density estimate of the combined left and right velocities during the right-heel-strike gait cycle for both healthy individuals and patients with stroke. (C) Comparison between the right foot velocity distribution and the standard normal distribution. (D) Variations in right foot velocity during the right-heel-strike gait cycle for both healthy individuals and patients with stroke.



To further characterize gait trajectory patterns, skewness, kurtosis, and total variation were incorporated into the statistics module. Skewness was used to summarize distributional asymmetry, kurtosis to describe the peakedness of the distribution, and total variation to depict joints' ranges of motion across the gait cycle. These additional descriptors were intended to capture complementary characteristics of poststroke gait patterns beyond the standard statistical measures. The captured IMU data were categorized into joint position and joint angle modalities, and separate models were trained for each modality. The output probabilities from both modalities were then combined to generate the final prediction.

Contrastive Learning

In the subsequent phase of the SA Head module, the global prototype for label *k* is refined by incorporating the prototypes of true-positive (TP) samples in *D* through an exponential moving average. Let DTPk be the set of TP samples for label *k* in a batch with size nTPk. The exponential moving average operation can be defined as follows:

(1)Pk=(1−α)·1nTPk∑i∈ DTPkFi+α·Pk

where F_i is the prototype of the *i*th sample, and P_k is the global prototype for label *k*. P_k will converge to a consistent representation after training. For the false-negative (FN) and false-positive (FP) samples of label *k*, only a local cluster at the center is computed. Similarly, for the 2 types of samples, the cluster at the center is calculated:

(2)μFPk=1nFPk∑i∈ DFPkFi, μFNk=1nFNk∑i∈ DFNkFi

The FN and FP samples, which are often referred to as ambiguous samples, are calibrated by increasing the distance between the TP prototype and the FP samples, while decreasing the distance between the TP prototype and the FN samples. Given a confident TP sample F_i, the compensation term for the FN samples is as follows:

(3)φi={1−dis (Fi,μFNk)if i∈ ITPk∧nFNk>00otherwise

The penalty term for the FP samples is as follows:

(4)βi={1+dis (Fi,μFPk),if i∈ ITPk∧nFPk>00otherwise

The cosine similarity function is used to calculate the distance between the TP prototype and the FP samples. The cosine distance is maximized to ensure a shorter distance between the center of the FN cluster and the confident TP samples, thus minimizing the compensation term. Moreover, the penalty term should be minimized. Finally, the contrastive learning loss for the confident sample F_i is defined as follows:

[illegible]

Here, p_{ik} is the probability of the sample F_i for label *k*, and τ is the temperature. The final training objective is to simultaneously minimize the cross-entropy and contrastive learning loss.

Model Interpretation Analysis

To quantitatively evaluate the faithfulness of the STAT-Attention module, we conducted an attention-guided

XSL•FO
RenderX

occlusion analysis using the trained participant-level 5-fold cross-validation models. For each WGS item, attention weights were aggregated into a 2×4 grid representing 2 modalities, joint position and joint angle, and 4 lower-limb joint groups, including upper leg, lower leg, foot, and toe. For each test participant, these modality-joint cells were ranked according to their attention weights. Single-cell occlusion was then performed by masking the input features corresponding to the selected cell. Four occlusion conditions were examined: the highest-attention cell, the second-highest-attention cell, the lowest-attention cell, and the average across the remaining cells. Faithfulness was evaluated by calculating the decrease in true-class probability after occlusion. A larger decrease indicated that the occluded feature group was more important to the model's prediction.

Statistical Analysis

Model performance was evaluated at the participant level. Accuracy was defined as the proportion of predicted WGS item scores that exactly matched the expert reference scores. Because WGS item scores are ordinal, mean absolute error (MAE), linear weighted κ, and quadratic weighted κ were additionally calculated to account for the magnitude of disagreement between predicted and reference scores. Item-level performance was reported for WGS items 2 to 14, and overall performance was calculated by pooling all participant-level item predictions across these items.

To provide a robust estimate of model generalizability, performance metrics were summarized across the participant-level 5-fold cross-validation folds as the mean (SD). The distribution of expert reference scores for each WGS item was also reported to characterize potential class imbalance. To contextualize model performance against expert-level agreement, interrater reliability between the 2 experienced physiotherapists was assessed using their independent ratings. ICCs were calculated using a 2-way random-effects model with absolute agreement and single measures (ICC [2,1]) for each WGS item, the total WGS score, and the summed score of WGS items 2 to 14.

Data Availability

The raw clinical gait data associated with this study are not publicly available because they contain participant-level movement recordings and are subject to ethical and institutional data-sharing restrictions. Deidentified processed data may be made available from the corresponding author upon reasonable request and subject to institutional approval.

Results

This study included 51 patients with stroke and 18 healthy individuals. The mean age was 60.3 (SD 10.2) years in healthy individuals and 62.7 (SD 9.5) years in patients with stroke. The poststroke duration was 9.0 (SD 7.2) years. In this sample, 21 patients with stroke and 9 healthy individuals were men. The median FMA-LE score was 24 (IQR 21-29), and the median WGS score was 22 (IQR 18-26; Table 1).

Figure 3 summarizes the experimental workflow and results. Using instrumented gait data as input, we achieved higher accuracy in WGS ratings compared to 4 final-year physiotherapy students. As presented in Table 2, the STAT-Net model was descriptively compared with 4 final-year physiotherapy students across the participant-level 5-fold cross-validation test sets. The STAT-Net model achieved a mean exact accuracy of 75.64% (SD 6.58%), whereas the 4 student evaluators achieved mean accuracies of 67.09% (SD 4.03%), 65.90% (SD 6.50%), 56.19% (SD 7.03%), and 58.72% (SD 4.22%), respectively. These findings suggest that the model showed higher exact agreement with the expert reference scores than the student evaluators in this novice-rater comparison.

Table 3 presents the item-level prediction performance of STAT-Net among participants with stroke. Across WGS items 2 to 14, the model achieved an overall exact accuracy of 75.42%, with an MAE of 0.290 score levels. The overall linear and quadratic weighted κ values were 0.625 and 0.660, respectively, indicating moderate ordinal agreement with the expert reference scores.

XSL•FO
RenderX

Table . Characteristics of the participants (N=69).

Characteristics	Stroke	Healthy
Gender (male:female)	24:27	8:10
Side of hemiplegia (right:left)	23:28	N/A[a]
Type of stroke (ischemic:hemorrhagic)	34:17	N/A
Age (y), mean (SD)	62.74 (9.5)	60.3 (10.2)
Height (m), mean (SD)	1.63 (0.09)	1.66 (0.09)
Weight (kg), mean (SD)	65.77 (12.12)	64.10 (12.98)
BMI (kg/m^2), mean (SD)	23.76 (4.03)	23.11 (3.20)
Poststroke duration (y), mean (SD)	9.0 (7.2)	N/A
FMA-LE[b], median (IQR)	24 (21-29)	N/A
WGS[c], median (IQR)	22 (18-26)	N/A

[a]Not applicable.

[b]FMA-LE: Fugl-Meyer Assessment for the Lower Extremity.

[c]WGS: Wisconsin Gait Scale.

Figure 3. Overview of the gait assessment workflow and performance comparison. CV: cross-validation; IMU: inertial measurement units; MVN: inertial motion capture system; STAT-GCN: statistical graph convolutional network; WGS: Wisconsin Gait Scale.

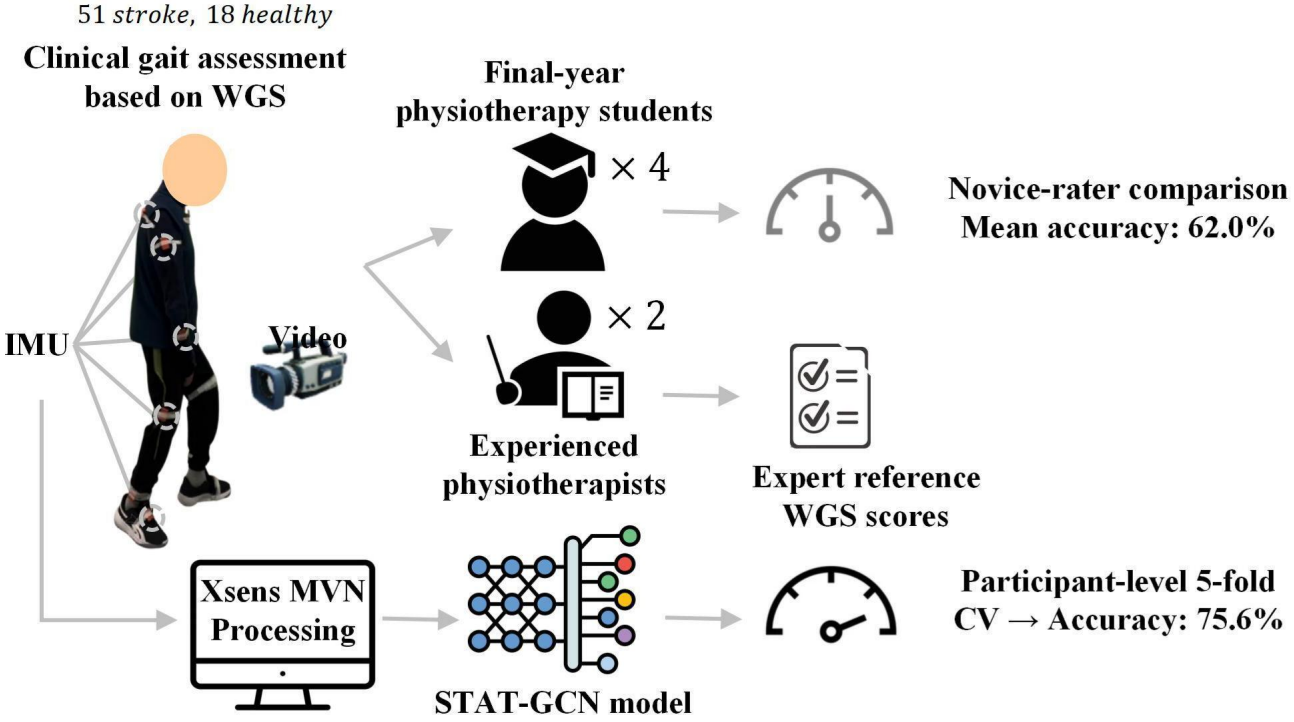


Table . Comparison of Wisconsin Gait Scale (WGS) item classification accuracy (%) between the STAT-Net model and 4 final-year physiotherapy students across the participant-level 5-fold cross-validation test sets.

Method	Fold 0	Fold 1	Fold 2	Fold 3	Fold 4	5-fold mean (SD)
STAT-Net	73.08	82.05	67.52	70.94	84.62	75.64 (6.58)
Evaluator A	65.38	66.67	60.68	71.79	70.94	67.09 (4.03)
Evaluator B	65.38	70.94	55.56	63.25	74.36	65.90 (6.50)
Evaluator C	48.46	55.56	48.72	65.81	62.39	56.19 (7.03)
Evaluator D	57.69	57.26	52.14	62.39	64.10	58.72 (4.22)

XSL•FO
RenderX

Table . Item-level prediction performance of the model among patients with stroke (n=51).

WGS[a] item	Accuracy (%)	MAE[b]	Linear weighted κ	Quadratic weighted κ
Item 2	76.47	0.235	0.673	0.749
Item 3	80.39	0.275	0.486	0.357
Item 4	68.63	0.392	0.493	0.472
Item 5	76.47	0.294	0.534	0.570
Item 6	78.43	0.235	0.638	0.726
Item 7	72.55	0.333	0.529	0.527
Item 8	64.71	0.392	0.377	0.387
Item 9	88.24	0.118	0.805	0.860
Item 10	86.27	0.137	0.717	0.740
Item 11	66.67	0.392	0.617	0.710
Item 12	70.59	0.314	0.443	0.513
Item 13	70.59	0.431	0.419	0.320
Item 14	80.39	0.216	0.759	0.812
Overall (N=663)	75.42	0.290	0.625	0.660

[a]WGS: Wisconsin Gait Scale.

[b]MAE: mean absolute error.

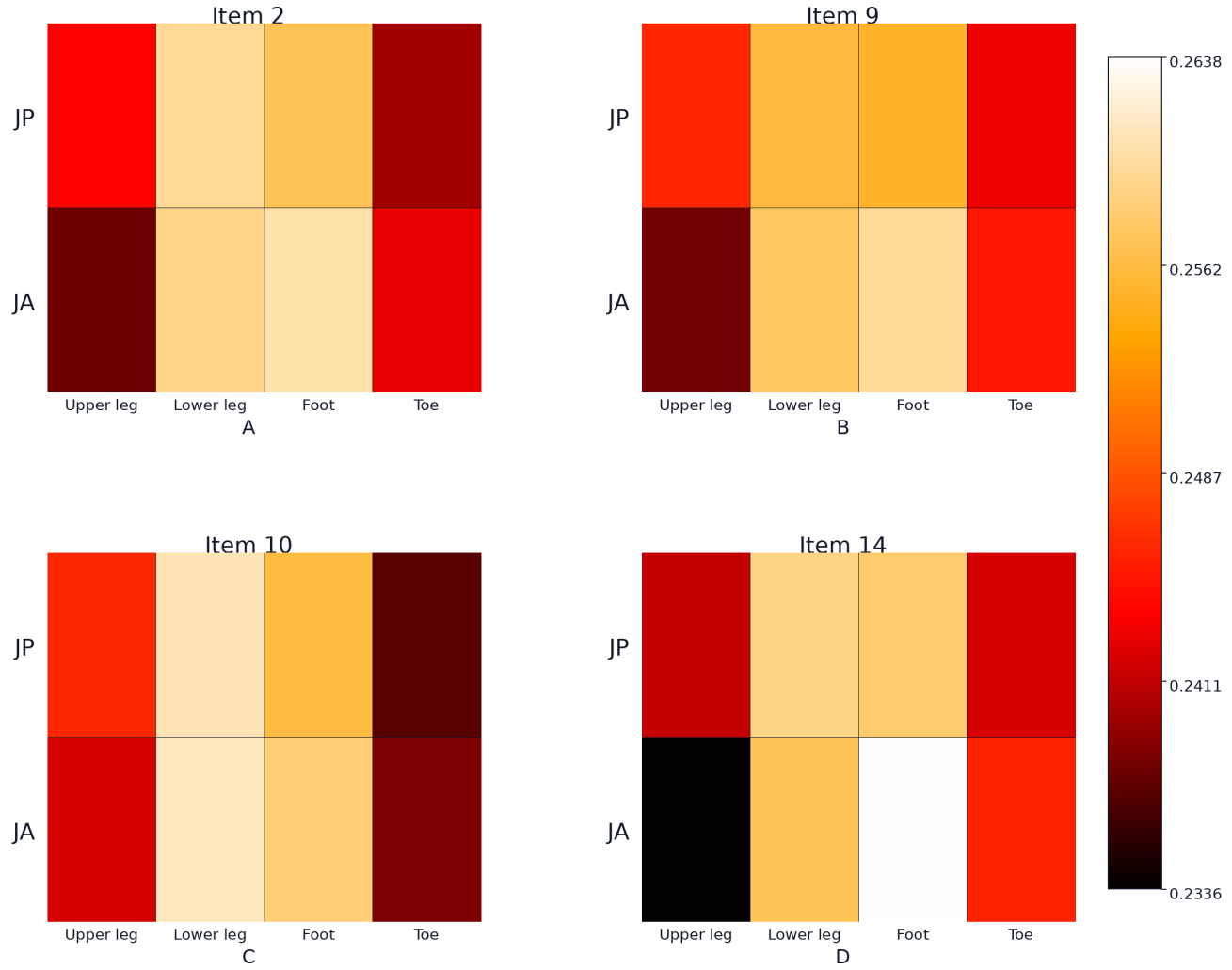
Item-level performance varied across WGS items. Higher agreement was observed for items 9, 10, and 14, whereas items 8 and 13 showed relatively lower weighted agreement. The distribution of expert reference scores for WGS items 2 to 14 is presented in Table S1 of Multimedia Appendix 2. The score distribution was imbalanced, with lower score categories more frequently represented and severe score categories underrepresented. Score 4 was observed only for item 11 and was absent from the remaining items. The 2 experienced physiotherapists demonstrated high interrater reliability, with ICC (2,1) values of 0.981 for the total WGS score and 0.975 for the summed score of WGS items 2 to 14. Item-level ICC (2,1) values ranged from 0.829 to 0.944, as shown in Table S2 of Multimedia Appendix 2.

The training and testing loss curves for representative WGS items are shown in Figure S1 in Multimedia Appendix 2. The training loss generally decreased across epochs, whereas the testing loss remained relatively stable without clear late-stage divergence, providing supporting evidence for stable model convergence.

To provide an exploratory interpretation of the learned attention mechanism, we performed qualitative attention-weighted visualizations and a post hoc feature occlusion analysis. As illustrated in Figure 4, the attention heatmaps show the normalized weights assigned to different lower-limb feature groups across representative WGS items. In each subplot, the rows represent joint-related feature types, and the columns represent lower-limb regions, including the upper leg, lower leg, foot, and toe. Darker pixels indicate lower-attention weights, whereas lighter pixels indicate higher-attention weights. The attention maps show item-specific weighting patterns across lower-limb regions. For example, the model assigned relatively higher weights to foot-related features for item 14, which is consistent with the clinical relevance of foot position during initial contact. These visualizations suggest that the model assigns different levels of priority to lower-limb features across WGS items.

XSL•FO
RenderX

Figure 4. STAT attention heatmaps for selected Wisconsin Gait Scale items. (A-D) Post-Softmax attention heatmaps for WGS Items 2, 9, 10, and 14, respectively. The heatmaps visualize how the model distributes attention across lower-limb joints in the joint-angle (JA) and joint-position (JP) modalities, where warmer colors indicate higher normalized attention weights and suggest the regions that contribute more strongly to the corresponding item prediction.



To further evaluate whether the learned attention weights reflected functionally important input features, an attention-guided occlusion analysis was performed for representative well-performing WGS items. As shown in Table 4 and Figure 5, occluding the highest-attention feature group generally produced larger decreases in true-class probability than occluding the lowest-attention feature group or the remaining feature groups. This pattern was most evident for items 2, 9, and 14. For item 2, top-1 occlusion produced the largest decrease in true-class probability (Δp_true=0.0792), compared with top-2nd (0.0381), bottom-1 (0.0015), and rest-avg (0.0178). Similar patterns were observed for item 9 and item 14. Item 10 showed a less distinct pattern, suggesting that attention faithfulness may vary across WGS items.

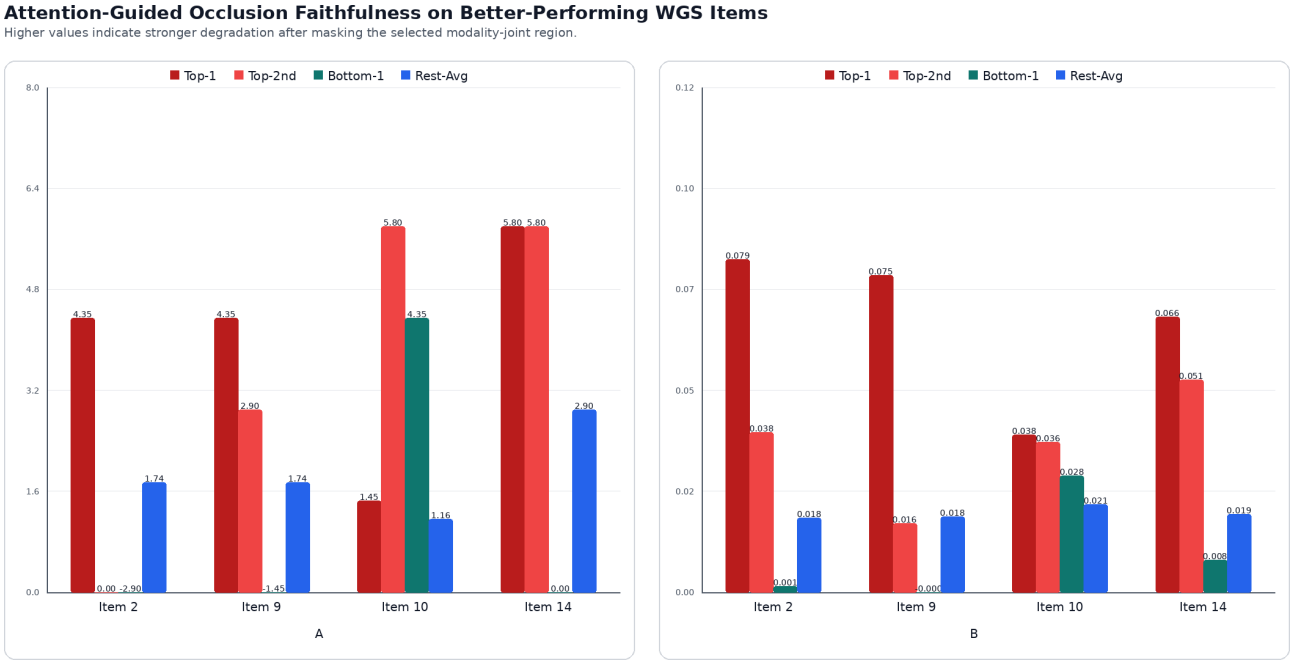
Table . Single-cell attention-guided occlusion faithfulness on selected Wisconsin Gait Scale (WGS) items[a].

Item	Top-1 Δp_true	Top-2nd Δp_true	Bottom-1 Δp_true	Rest-avg Δp_true
Item 2	0.0792	0.0381	0.0015	0.0178
Item 9	0.0754	0.0164	−0.0001	0.0180
Item 10	0.0375	0.0358	0.0277	0.0210
Item 14	0.0655	0.0505	0.0077	0.0186

[a]Δp_true indicates the decrease in the predicted probability of the true class after occlusion. Top-1 and Top-2nd refer to the modality-joint cells with the highest and second-highest-attention weights, respectively. Bottom-1 refers to the cell with the lowest-attention weight. Rest-avg represents the average occlusion effect across the remaining cells. Larger positive values indicate a greater reduction in true-class probability after masking and therefore suggest a greater functional contribution of the corresponding feature group to the model prediction.

XSL•FO
RenderX

Figure 5. Attention-guided occlusion analysis for selected Wisconsin Gait Scale (WGS) items. (A) Participant-level accuracy drop (%); drop relative to baseline 5-fold test accuracy. (B) True-class confidence drop (decrease in participant-level true-class probability after occlusion).



Discussion

Principal Findings

In this study, we developed and evaluated STAT-GCN, a deep learning framework for automated WGS item prediction using IMU-derived kinematic data. Using a stratified, participant-level 5-fold cross-validation protocol, all gait cycles from the same participant were kept within the same fold to reduce the risk of intraparticipant information leakage. The findings suggest that kinematic data-driven models may provide preliminary support for more standardized WGS item scoring; however, they should be interpreted cautiously, given the limited sample size and absence of external validation. Previous work on digital WGS assessment has also suggested that technology-assisted approaches may help reduce subjectivity in observational gait evaluation [37].

The comparison with 4 final-year physiotherapy students was interpreted as a descriptive novice-rater comparison. In this dataset, the model showed higher agreement with the expert reference scores than did the novice raters. However, this comparison should be interpreted cautiously, as the students represent a novice-rater benchmark rather than an expert clinical standard. The high interrater reliability between the 2 experienced physiotherapists indicates that expert-level WGS assessment remains more consistent than that of the current automated model. Therefore, the findings should not be interpreted as evidence that STAT-GCN has achieved expert-level clinical performance. Instead, the model may provide preliminary support for more standardized WGS item scoring, particularly when expert raters are unavailable or when repeated assessments are required.

Item-level performance varied across WGS items, suggesting that some gait components were more difficult for the model to predict than others. Relatively low agreement was observed for items 8 and 13, both of which involve rotational components. Item 8 assesses external rotation during the initial swing phase of the affected leg, whereas item 13 assesses pelvic rotation during terminal swing. These transverse-plane movements may be challenging because they are difficult to estimate visually and may also be sensitive to sensor placement and movement variability. In addition, item 11, which assesses knee flexion from toe-off to mid-swing by comparing the affected and unaffected limbs, may require more explicit bilateral and phase-specific information than the current model input provides. Future model development may benefit from incorporating clinically informed features, such as inter-limb ratios, phase-specific joint trajectories, or additional transverse-plane kinematic descriptors.

The distribution of expert reference scores also indicated class imbalance across WGS items. Lower score categories were more frequently represented, whereas severe score categories were underrepresented. This imbalance may have affected both model learning and performance interpretation, because the model had fewer examples from which to learn severe gait deviations. Similar concerns have been reported in other gait-related machine learning studies, where class imbalance and limited representation of severe cases were found to influence model performance and stability, highlighting the importance of balanced datasets and appropriate augmentation strategies [38]. It may also explain why exact accuracy alone could provide an incomplete picture of model performance. For this reason, ordinal metrics such as MAE and weighted κ were included to better account for the magnitude of disagreement between predicted and expert reference scores. Future studies should include larger and more clinically diverse samples and may consider class-balanced training strategies, cost-sensitive learning, or carefully validated data augmentation methods.

The attention-guided occlusion analysis provided partial quantitative support for the faithfulness of the STAT-Attention

XSL•FO
RenderX

module. In several representative WGS items, masking high-attention modality-joint regions produced larger reductions in true-class probability than masking low-attention regions, suggesting that the model relied more heavily on features assigned higher attention weights. However, this pattern was item-dependent and should not be interpreted as complete evidence of clinical explainability. In addition, aggregating attention weights into a 2×4 modality-joint grid may obscure more detailed joint-specific and time-specific information. Therefore, the attention maps should be considered supportive indicators of feature relevance rather than definitive explanations of the model's decision-making process. Future studies should compare this approach with established explainability methods and further examine whether the highlighted features correspond to specific clinically defined WGS gait deviations.

Although automated WGS item prediction may help reduce subjectivity and improve the standardization of repeated gait assessments, the present study did not directly examine whether model-predicted WGS scores improve rehabilitation decision-making, treatment planning, or clinical outcomes. Therefore, the translational value of the proposed framework remains preliminary. Future studies should investigate whether automated item-level WGS predictions can be linked to individualized exercise prescription, longitudinal treatment response, or clinically meaningful functional recovery.

Limitations

This study has several limitations. First, although participant-level cross-validation and training-only data augmentation were used to reduce the risk of overly optimistic performance estimates, the sample size remained limited for deep learning model development, and no external validation dataset was available. The limited sample size and a relatively low representation of participants with severe gait impairment may have affected model stability, particularly for higher WGS score categories, and may limit generalizability to broader populations of individuals with stroke and more diverse gait impairments. Independent validation across different clinical sites, devices, and participant populations is required before clinical implementation. Second, the model relied primarily on kinematic data. Additional modalities, such as plantar pressure, force-plate measurements, or video-based movement features, may improve the prediction of items related to weight shifting, limb comparison, and transverse-plane rotation. However, the use of additional equipment may reduce clinical feasibility, and future work should balance model performance with practical usability.

Conclusions

STAT-GCN demonstrated reasonable preliminary performance for automated WGS item prediction in individuals with stroke using IMU-derived kinematic data. By combining a graph-based representation of lower-limb kinematics with participant-level validation, this study provides initial evidence that data-driven models may support more standardized observational gait assessment. However, further studies using larger and more clinically diverse stroke cohorts, additional data modalities, and independent external validation are needed to determine the clinical utility and generalizability of this approach.

Acknowledgments

Generative AI tools were used only for language editing and polishing. All scientific content, data analyses, interpretations, and references were produced and verified by the authors, who take full responsibility for the accuracy and integrity of the manuscript.

Funding

This work was supported in part by The Hong Kong Polytechnic (grant P0036617) and the Hong Kong Research Grants Council (grant T42-513/24 R).

Authors' Contributions

ZG, JL, PWHK, and YY contributed substantially to the conception and design of the study. ZG, JL, and WZ were responsible for data collection, analysis, and interpretation. All authors participated in drafting the manuscript. PWHK, LW, KNKF, and EKL critically revised the manuscript for important intellectual content. All authors contributed equally to the final version, read, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Wisconsin Gait Scale full items.
[DOCX File, 388 KB - rehab_v13i1e94031_app1.docx]

Multimedia Appendix 2
Wisconsin Gait Scale score distributions and interrater reliability.
[DOCX File, 67 KB - rehab_v13i1e94031_app2.docx]

XSL•FO
RenderX

Checklist 1
STROBE checklist.
[DOC File, 84 KB - rehab_v13i1e94031_app3.doc]

References

1. Wafa HA, Wolfe CDA, Emmett E, Roth GA, Johnson CO, Wang Y. Burden of stroke in Europe: thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years. Stroke 2020 Aug;51(8):2418-2427. [doi: 10.1161/STROKEAHA.120.029606] [Medline: 32646325]
2. Pu L, Wang L, Zhang R, Zhao T, Jiang Y, Han L. Projected global trends in ischemic stroke incidence, deaths and disability-adjusted life years from 2020 to 2030. Stroke 2023 May;54(5):1330-1339. [doi: 10.1161/STROKEAHA.122.040073] [Medline: 37094034]
3. Heshmatollah A, Darweesh SKL, Dommershuijsen LJ, Koudstaal PJ, Ikram MA, Ikram MK. Quantitative gait impairments in patients with stroke or transient ischemic attack: a population-based approach. Stroke 2020 Aug;51(8):2464-2471. [doi: 10.1161/STROKEAHA.120.029829] [Medline: 32654631]
4. Pan JW, Sidarta A, Wu TL, et al. Unraveling stroke gait deviations with movement analytics, more than meets the eye: a case control study. Front Neurosci 2024;18:1425183. [doi: 10.3389/fnins.2024.1425183] [Medline: 39104608]
5. Lattouf NA, Tomb R, Assi A, Maynard L, Mesure S. Eccentric training effects for patients with post-stroke hemiparesis on strength and speed gait: a randomized controlled trial. NeuroRehabilitation 2021;48(4):513-522. [doi: 10.3233/NRE-201601] [Medline: 33967063]
6. Wang Y, Mukaino M, Ohtsuka K, et al. Gait characteristics of post-stroke hemiparetic patients with different walking speeds. Int J Rehabil Res 2020 Mar;43(1):69-75. [doi: 10.1097/MRR.0000000000000391] [Medline: 31855899]
7. Kwong PW, Ng SS, Liu TW, Chung RC, Ng GY. Effect of leg selection on the Berg Balance Scale scores of hemiparetic stroke survivors: a cross-sectional study. Arch Phys Med Rehabil 2016 Apr;97(4):545-551. [doi: 10.1016/j.apmr.2015.11.017] [Medline: 26707457]
8. Kwong PWH, Ng SSM, Chung RCK, Ng GYF. A structural equation model of the relationship between muscle strength, balance performance, walking endurance and community integration in stroke survivors. PLoS One 2017;12(10):e0185807. [doi: 10.1371/journal.pone.0185807] [Medline: 29049293]
9. Li J, Kwong PWH, Li H, Tham T, Fong KNK, Wang L. Validating gait profile score for post-stroke gait: correlation with observational scales and clinical outcomes. Physiother Pract Res 2025 Aug;46(2):138-146. [doi: 10.1177/22130683251342181]
10. Kerr A, Rowe P, Clark A, et al. Biomechanical correlates for recovering walking speed following a stroke. The potential of tibia to vertical angle as a therapy target. Gait Posture 2020 Feb;76:162-167. [doi: 10.1016/j.gaitpost.2019.12.009] [Medline: 31862664]
11. Guzik A, Drużbicki M, Wolan-Nieroda A, Przysada G, Kwolek A. The Wisconsin Gait Scale—the minimal clinically important difference. Gait Posture 2019 Feb;68:453-457. [doi: 10.1016/j.gaitpost.2018.12.036] [Medline: 30599331]
12. Yaliman A, Kesiktas N, Ozkaya M, Eskiyurt N, Erkan O, Yilmaz E. Evaluation of intrarater and interrater reliability of the Wisconsin Gait Scale with using the video taped stroke patients in a Turkish sample. NeuroRehabilitation 2014;34(2):253-258. [doi: 10.3233/NRE-131033] [Medline: 24401828]
13. Pandis N. Introduction to observational studies: part 2. Am J Orthod Dentofacial Orthop 2014 Feb;145(2):268-269. [doi: 10.1016/j.ajodo.2013.11.002] [Medline: 24485743]
14. Li J, Kwong PWH, Lua EK, Chan MYL, Choo A, Donnelly CJW. Development of a convolutional neural network (CNN) based assessment exercise recommendation system for individuals with chronic stroke: a feasibility study. Top Stroke Rehabil 2023 Dec;30(8):786-795. [doi: 10.1080/10749357.2022.2127669] [Medline: 36189968]
15. Bivard A, Churilov L, Parsons M. Artificial intelligence for decision support in acute stroke—current roles and potential. Nat Rev Neurol 2020 Oct;16(10):575-585. [doi: 10.1038/s41582-020-0390-y] [Medline: 32839584]
16. Hu K, Wang Z, Martens KAE, et al. Graph fusion network-based multimodal learning for freezing of gait detection. IEEE Trans Neural Netw Learning Syst 2021;34(3):1588-1600. [doi: 10.1109/TNNLS.2021.3105602]
17. Sabo A, Mehdizadeh S, Iaboni A, Taati B. Estimating parkinsonism severity in natural gait videos of older adults with dementia. IEEE J Biomed Health Inform 2022 May;26(5):2288-2298. [doi: 10.1109/JBHI.2022.3144917] [Medline: 35077373]
18. Yan S, Xiong Y, Lin D. Spatial temporal graph convolutional networks for skeleton-based action recognition. Proc AAAI Conf Artif Intell 2018;32(1). [doi: 10.1609/aaai.v32i1.12328]
19. Zhang X, He L, Chen K, Luo Y, Zhou J, Wang F. Multi-view graph convolutional network and its applications on neuroimage analysis for Parkinson's disease. AMIA Annu Symp Proc 2018;2018:1147-1156. [Medline: 30815157]
20. Li M, Chen S, Chen X, Zhang Y, Wang Y, Tian Q. Actional-structural graph convolutional networks for skeleton-based action recognition. Presented at: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 15-20, 2019. [doi: 10.1109/CVPR.2019.00371]
21. Zhang X, Xu C, Tao D. Context aware graph convolution for skeleton-based action recognition. Presented at: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 14-19, 2020. [doi: 10.1109/CVPR42600.2020.01434]

XSL•FO
RenderX

22. Huang L, Huang Y, Ouyang W, Wang L. Part-level graph convolutional network for skeleton-based action recognition. Proc AAAI Conf Artif Intell 2020;34(7):11045-11052. [doi: 10.1609/aaai.v34i07.6759]
23. Shahroudy A, Liu J, Ng TT, Wang G. NTU RGB+D: a large scale dataset for 3D human activity analysis. Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016. [doi: 10.1109/CVPR.2016.115]
24. Scheper MC, van Velzen M, van Meeteren NLU. Towards responsible use of artificial intelligence in daily practice: what do physiotherapists need to know, consider and do? J Physiother 2024 Apr;70(2):81-84. [doi: 10.1016/j.jphys.2023.07.012] [Medline: 38036398]
25. Yu BXB, Liu Y, Zhang X, Chen G, Chan KCC. EGCN: an ensemble-based learning framework for exploring effective skeleton-based rehabilitation exercise assessment. Presented at: Thirty-First International Joint Conference on Artificial Intelligence {IJCAI-22}; Jul 23-29, 2022. [doi: 10.24963/ijcai.2022/511]
26. Li J, Kwong PW, Lin W, Fong KN, Wu W, Sidarta A. Assessment of ambulation functions through kinematic analysis in individuals with stroke: a systematic review. Eur J Phys Rehabil Med 2025 Feb;61(1):28-40. [doi: 10.23736/S1973-9087.25.08767-2] [Medline: 40008910]
27. Myn U, Link M, Awinda M. Xsens MVN user manual. : Xsens Motion Technologies BV; 2015 URL: https://www.xsens.com/hubfs/Downloads/usermanual/MVN_User_Manual.pdf [accessed 2026-07-21]
28. Rodriquez AA, Black PO, Kile KA, et al. Gait training efficacy using a home-based practice model in chronic hemiplegia. Arch Phys Med Rehabil 1996 Aug;77(8):801-805. [doi: 10.1016/s0003-9993(96)90260-9] [Medline: 8702375]
29. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. Scand J Rehabil Med 1975;7(1):13-31. [Medline: 1135616]
30. Hernández ED, Forero SM, Galeano CP, Barbosa NE, Sunnerhagen KS, Alt Murphy M. Intra- and inter-rater reliability of Fugl-Meyer Assessment of Lower Extremity early after stroke. Braz J Phys Ther 2021;25(6):709-718. [doi: 10.1016/j.bjpt.2020.12.002] [Medline: 33358073]
31. Sullivan KJ, Tilson JK, Cen SY, et al. Fugl-Meyer assessment of sensorimotor function after stroke: standardized training procedure for clinical practice and clinical trials. Stroke 2011 Feb;42(2):427-432. [doi: 10.1161/STROKEAHA.110.592766] [Medline: 21164120]
32. Twitchell TE. The restoration of motor function following hemiplegia in man. Brain 1951 Dec;74(4):443-480. [doi: 10.1093/brain/74.4.443] [Medline: 14895765]
33. Sensor placement in Xsens Awinda system. Xsens. 2023. URL: https://base.movella.com/s/article/Sensor-Placement-in-Xsens-Awinda-System?language=en_US [accessed 2026-07-21]
34. Um TT, Pfister FMJ, Pichler D, et al. Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks. Presented at: ICMI '17:The 19th ACM International Conference on Multimodal Interaction; Nov 13-17, 2017. [doi: 10.1145/3136755.3136817]
35. Zhou H, Liu Q, Wang Y. Learning discriminative representations for skeleton based action recognition. Presented at: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 17-24, 2023. [doi: 10.1109/CVPR52729.2023.01022]
36. Dong W, Yuan T, Yang K, Li C, Zhang S. Autoencoder regularized network for driving style representation learning. Presented at: Twenty-Sixth International Joint Conference on Artificial Intelligence; Aug 19-25, 2017. [doi: 10.24963/ijcai.2017/222]
37. Guzik A, Wolan-Nieroda A, Drużbicki M. Assessment of agreement between a new application to compute the Wisconsin gait score and 3-dimensional gait analysis, and reliability of the application in stroke patients. Front Hum Neurosci 2022;16:775261. [doi: 10.3389/fnhum.2022.775261] [Medline: 35185497]
38. Trabassi D, Castiglia SF, Bini F, et al. Optimizing rare disease gait classification through data balancing and generative AI: insights from hereditary cerebellar ataxia. Sensors (Basel) 2024 Jun 3;24(11):3613. [doi: 10.3390/s24113613] [Medline: 38894404]

Abbreviations

FMA-LE: Fugl-Meyer Assessment of the Lower Extremity
FN: false-negative
FP: false-positive
GCN: graph convolutional network
ICC: intraclass correlation coefficient
IMU: inertial measurement unit
MAE: mean absolute error
ST-GCN: spatio-temporal graph convolutional network
STAT-GCN: statistical graph convolutional network
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TP: true-positive

XSL•FO
RenderX

WGS: Wisconsin Gait Scale

Edited by S Munce; submitted 24.Feb.2026; peer-reviewed by D Trabassi, L Zhou; revised version received 30.Jun.2026; accepted 14.Jul.2026; published 19.Aug.2026.

Please cite as:
Gu Z, Li J, Zhou W, Fong KNK, Wang L, Lua EK, Kwong PWH, Yuan Y
Deep Learning–Based Automated Clinical Gait Assessment From Kinematic Data in People With Stroke: Development and Validation Study
JMIR Rehabil Assist Technol 2026;13:e94031
URL: https://rehab.jmir.org/2026/1/e94031
doi:10.2196/94031

© Zhexuan Gu, Jiaqi Li, Wenbin Zhou, Kenneth N K Fong, Lin Wang, Eng Keong Lua, Patrick W H Kwong, Yancheng Yuan. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of High-Intensity Laser Therapy Combined With Plantar Fascia Stretching Exercises in the Treatment of Plantar Fasciitis: Randomized, Double-Blind, Sham-Controlled Trial

Nantaporn Jitpimolmard, MSc, MD; Phonlawat Ouemphancharoen, MD; Preeda Arayawichanon, MD

Department of Rehabilitation Medicine, Faculty of Medicine, Khon Kaen University, 123 Mittraphap Road, Mueang, Khon Kaen, Thailand

Corresponding Author:
Nantaporn Jitpimolmard, MSc, MD
Department of Rehabilitation Medicine, Faculty of Medicine, Khon Kaen University, 123 Mittraphap Road, Mueang, Khon Kaen, Thailand

Abstract

Background: Plantar fasciitis causes heel pain and functional limitations; conservative treatment typically includes plantar fascia and calf stretching. High-intensity laser therapy (HILT) offers deeper photobiomodulation and potential tissue-healing benefits, but robust evidence of added clinical benefit remains limited.

Objective: This study aims to evaluate the efficacy of HILT combined with plantar fascia stretching exercises compared with a sham control for the treatment of plantar fasciitis.

Methods: This study was designed as a randomized, double-blind, sham-controlled trial conducted at the outpatient clinic of a university hospital. Participants were randomly allocated into 2 groups: the HILT group and the sham treatment group. Both groups received 9 treatment sessions over 3 weeks. The HILT group received active laser therapy, while the sham group received identical treatment without laser emission. The HILT used a wavelength of 1064 nm in continuous mode, with a power output of 12 W applied for 250 seconds, delivering an energy density of 120 J/cm² applied to a 25-cm² area for a total energy of 3000 J. In addition to the assigned interventions, all participants performed a standardized self-stretching exercise program targeting the plantar fascia and Achilles tendon throughout the study period. The primary outcome was pain intensity measured using a visual analog scale. Secondary outcomes included ultrasonographic measurement of plantar fascia thickness (PFT) and the subjective foot and ankle ability measure (FAAM) score, recorded before and immediately after the intervention.

Results: A total of 34 patients diagnosed with unilateral plantar fasciitis were enrolled in this study. Based on intragroup comparison, both groups demonstrated statistically significant improvements in all outcomes compared with baseline (P<.001). However, no significant differences were found between the 2 groups across all outcomes. The mean difference in pain reduction, measured by the visual analog scale, was –35.3 (95% CI –45.3 to –25.0) mm in the HILT group and –30.4 (95% CI –46.3 to –14.4) mm in the sham group (–5.0 mm, 95% CI –14.3 to 4.3; P=.59). Similarly, reductions in PFT and improvements in FAAM scores showed no significant differences between groups (mean difference –0.02 mm, 95% CI –0.2 to 0.1; P=.90 and 5.6 points, 95% CI –1.1 to 12.4; P=.40, respectively).

Conclusions: There was no additional clinical effectiveness of HILT on pain reduction, decreased PFT, or increased FAAM scores compared with sham laser when combined with standard stretching exercises for plantar fasciitis and the Achilles tendon.

Trial Registration: Thai Clinical Trials Registry TCTR20200724001; https://www.thaiclinicaltrials.org/show/TCTR20200724001

(JMIR Rehabil Assist Technol 2026;13:e77419) doi:10.2196/77419

KEYWORDS

plantar fasciitis; high-intensity laser therapy; lasers; pain measurement; ultrasonography; fascia; foot; ankle

Introduction

Plantar fasciitis is a common and often disabling condition characterized by pain at the plantar aspect of the heel, which can substantially limit daily activities [1]. Risk factors include increasing age, elevated body mass index, prolonged standing, and foot posture abnormalities that increase stress on the plantar fascia [2]. First-line management is conservative and heterogeneous, encompassing manual therapy, night splints, taping, orthoses, and exercise programs; among these, stretching of the plantar fascia and associated calf musculature is widely recommended and frequently used in clinical practice [3].

For this study, passive stretching was chosen as the standardized backbone intervention. Mechanistically, stretching directly targets gastrocnemius–soleus tightness and plantar fascia strain, two important contributors to symptom persistence, and has been shown in clinical studies to reduce pain and improve

XSL•FO
RenderX

function in the short to medium term [4]. Practically, passive stretching is easily standardized, low cost, self-administered, well tolerated, and associated with good patient adherence compared with more resource-intensive or operator-dependent modalities (eg, manual therapy or supervised exercise programs). Using a single, well-established conservative treatment as the common background therapy therefore reduces between-participant variability and more clearly isolates any incremental effect of the adjunctive intervention under study.

Adjunctive physical modalities—including extracorporeal shockwave therapy, ultrasound, and low-intensity laser therapy—are commonly used to enhance recovery [5]. Laser therapy delivers photons that may stimulate cellular processes involved in tissue repair, including increased adenosine triphosphate production, enhanced collagen synthesis, and improved microcirculation [6-8]. Systematic reviews indicate that low-intensity laser therapy can reduce pain in plantar fasciitis [5,9-12]. In theory, high-intensity laser therapy (HILT) has greater tissue penetration and may produce stronger collagen stimulation, greater blood flow effects, and more pronounced anti-inflammatory responses than low-intensity devices [13,14]. However, studies comparing HILT with low-intensity laser or sham treatment have reported conflicting results [15,16], and recent meta-analyses have not comprehensively evaluated HILT for plantar fasciitis.

Prior HILT studies have often been limited by methodological weaknesses, inconsistent control treatments, and a scarcity of objective outcomes such as ultrasound-measured plantar fascia thickness (PFT) or functionally relevant measures. To address these gaps, this randomized, sham-controlled trial evaluated the efficacy of HILT added to a standardized passive stretching program compared with sham laser plus the same conservative regimen. We hypothesized that adjunctive HILT would produce greater improvements in pain, functional capacity, and PFT compared with conservative care using sham laser.

Methods

Study Design

This study was a prospective, randomized, double-blinded (participant and assessor), sham-controlled trial conducted in an outpatient rehabilitation setting at Srinagarind Hospital, Khon Kaen University.

Ethical Considerations

The study protocol was approved by the Khon Kaen University Ethics Committee (HE621318). Written informed consent was obtained from all participants prior to enrollment. Participant privacy and confidentiality were strictly maintained. All data were deidentified prior to analysis, and personal identifiers were removed. The dataset was stored securely and was accessible only to authorized members of the research team. Participants received compensation of 900 for their participation. Based on an exchange rate of US $1 to 31, this is approximately US $29 per participant. Compensation was provided regardless of study completion.

The study adhered to the CONSORT (Consolidated Standards of Reporting Trials) guidelines.

Participants

A total of 36 participants diagnosed with unilateral plantar fasciitis were recruited between January 2020 and September 2020.

Inclusion Criteria

Participants were eligible for inclusion if they met all of the following criteria:

- Adults aged 18 years or older
- Foot pain persisting for a minimum of 6 weeks
- Diagnosis of plantar fasciitis confirmed by a rehabilitation medicine physician based on clinical history, characterized by pain primarily felt during the first steps after waking, which worsens with prolonged weight-bearing activities throughout the day [17]
- Tenderness at the medial calcaneal tubercle upon palpation [17]

Exclusion Criteria

Participants were excluded from the study if they met any of the following criteria:

- History of foot surgery
- Chronic inflammatory arthritis (eg, rheumatoid arthritis, gout, systemic lupus erythematosus)
- Conditions affecting sensory perception (eg, diabetes mellitus with peripheral neuropathy)
- History of corticosteroid injection to the plantar fascia within the preceding 6 months
- Neuropathic pain in the foot, or evidence of intrinsic foot muscle weakness
- Pregnancy

Baseline Assessment

At baseline, the following data were collected: age, gender, body mass index, side with the most severe pain, and Foot Posture Index (FPI-6) score.

Randomization and Blinding

Participants were randomly assigned to either the HILT group or the sham laser (control) group using a permuted block randomization method with a block size of 4. The randomization sequence was computer generated and concealed within sequentially numbered, opaque, sealed envelopes. Envelopes were opened by the treating therapist immediately prior to the first treatment session. Participants were blinded to their treatment allocation. The assessor performing outcome measurements was also blinded to treatment allocation.

Intervention

Overview

All participants received 9 treatment sessions administered every 3 days. Both groups received a standardized home exercise program consisting of plantar fascia and Achilles tendon stretching exercises. Participants were instructed to perform the following stretch while seated or supine: with a towel or strap around the forefoot, the knee was extended and the forefoot pulled toward the body, maintaining dorsiflexion for 30 seconds. The exercise was performed for 10 repetitions per session, once

XSL•FO
RenderX

daily. Adherence to the home exercise program was monitored using patient logs.

HILT Group

HILT was administered by a trained medical doctor using the BTL-6000 high-intensity laser device (BTL Industries) operating at a wavelength of 1064 nm. The treatment parameters were as follows: continuous mode, power output of 12 W applied for 250 seconds, energy density of 120 J/cm² delivered over a 25-cm² area, resulting in a total energy of 3000 J [15,16]. The participant was positioned prone with the foot in a neutral position. The laser was applied using a 10-mm pen applicator and swept over the calcaneal insertion and along the medial border of the plantar fascia.

Sham Laser Group

The sham treatment mimicked the HILT procedure, including the visual and auditory cues of the laser device, but without actual laser energy delivery. The same treatment protocol as the HILT group was followed in terms of positioning, applicator movement, and device sounds.

Medication and High Plantar Fascial Loading Activity

Participants were instructed to maintain their usual pain medication regimen, which was limited to 500 mg of paracetamol, and not to exceed 3000 mg per day. They were also advised to strictly refrain from strenuous exercises that involve high plantar fascial loading. Crucially, the use of anti-inflammatory drugs was strictly prohibited during the study per the protocol. The number of pain medication pills taken was recorded throughout the study period.

Outcome Measures

Outcomes were assessed at baseline and immediately postintervention by a blinded assessor. Pain intensity was measured using a 100-mm visual analog scale (VAS). Participants were asked to recall the pain experienced during the first 3 steps taken in the morning. Scores were assessed immediately preintervention and again via telephone 1 day after the final session to capture pain during participants' first steps in the morning following the treatment period.

PFT was measured by ultrasound using a LOGIQ E Premium system (GE Healthcare) with a 5 MHz linear probe at the calcaneal insertion. Participants were positioned prone with feet in a neutral position, and the feet were placed at the end of the examination bed. The ultrasound probe was placed perpendicular to the plantar fascia on the plantar surface, and the thickness of the thickest portion was measured [18]. The reported value is the mean of 3 measurements.

Foot and ankle function was assessed using the activities of daily living subscale of the Thai version of the foot and ankle ability measure (FAAM). The FAAM subjective form evaluates foot and ankle function for daily activities and sports; however, only the activities of daily living subscale was used in this study. The reliability of the FAAM was examined by Arunakul et al [19], who reported strong internal consistency.

Data Analysis

All analyses were conducted using Python (version 3.10), executed in Google Colaboratory. Continuous variables are presented as mean (SD) or as median (IQR) for nonnormally distributed data, while categorical variables are presented as frequency (percentage). Normality of continuous data was assessed using the Shapiro-Wilk test.

The primary analysis followed the intention-to-treat (ITT) principle. Missing outcome data (5%) were addressed using multiple imputation by chained equations with 20 imputations. The imputation model included baseline outcome values, treatment group, and relevant covariates. Each imputed dataset was analyzed separately, and the results were pooled using Rubin's rules.

Between-group comparisons of continuous posttreatment outcomes (primary and secondary) were conducted using independent *t* tests when distributional assumptions were met. For nonnormally distributed outcomes, between-group comparisons used the Mann-Whitney *U* test on observed data, with additional sensitivity analyses performed on the imputed datasets. Within-group changes were evaluated using paired *t* tests or the Wilcoxon signed rank test, as appropriate.

A per-protocol analysis including participants who completed at least 7 treatment sessions was conducted as a sensitivity analysis. Effect sizes (Cohen *d*) and 95% CIs were reported for parametric comparisons. All statistical tests were 2-tailed, and a *P* value <.05 was considered statistically significant.

Sample Size Calculation

Based on prior studies [15], a sample size of 16 participants per group was required to detect a mean difference of 15 (SD 15) mm on the VAS, with 80% power and a 2-sided α of .05. To allow for an anticipated 10% dropout rate, 36 participants (18 per group) were recruited.

Withdrawal and Discontinuation Criteria

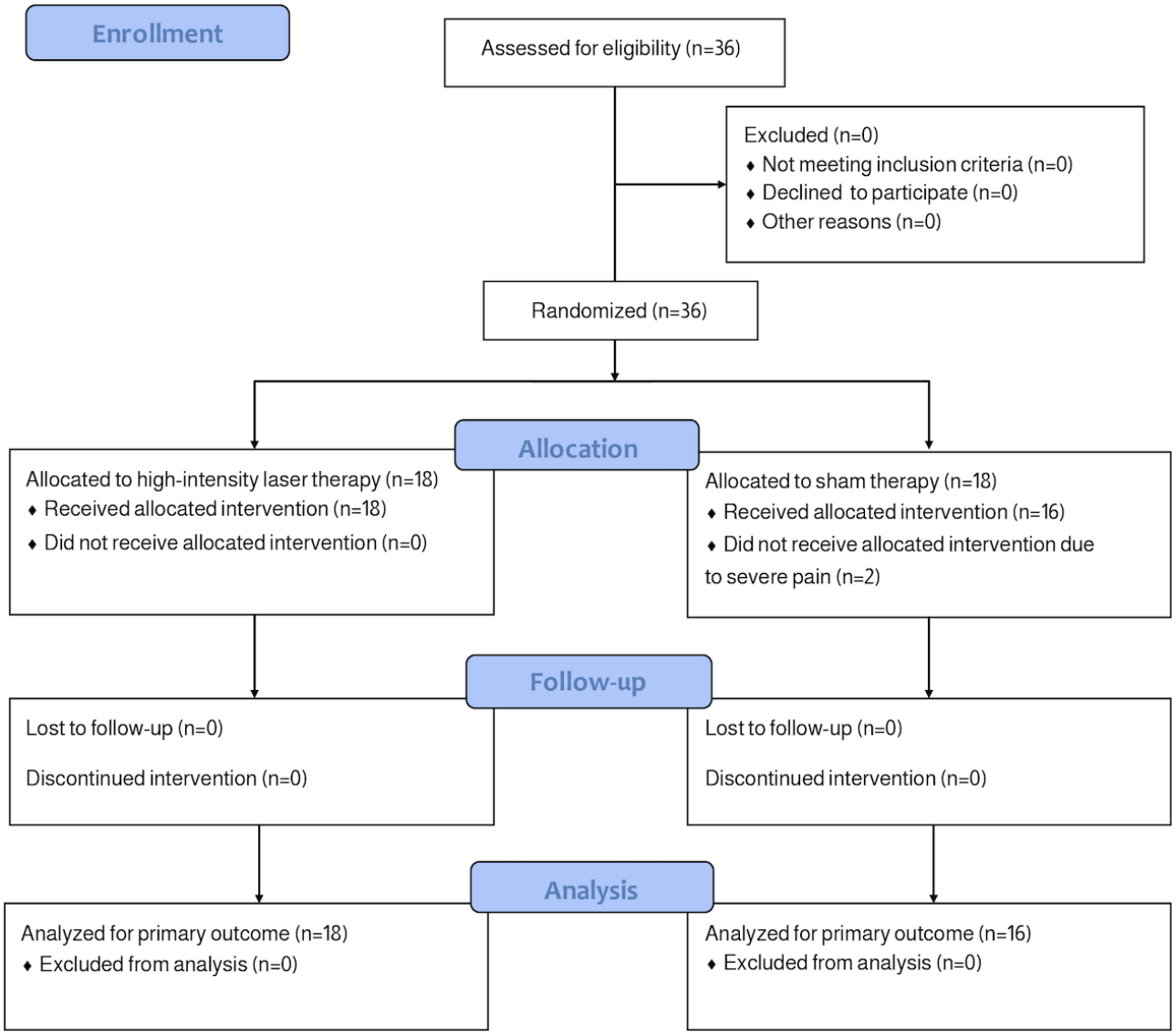
Participants could withdraw from the study at any time without affecting their care. Discontinuation criteria included withdrawal of consent or adverse effects (eg, burns, eye irritation).

Results

A total of 36 participants were enrolled, with 34 participants completing the study. Two participants withdrew due to severe pain and opted for alternative treatments (Figure 1).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram outlining the identification, enrollment, and allocation of participants to the 2 groups: high-intensity laser therapy group and sham group.



The HILT group included 18 participants (n=2 men, n=16 women; mean age 46, SD 11.4 years), while the sham laser group included 16 participants (n=1 man, n=15 women; mean age 49, SD 13.0 years). The mean body mass index was 23.3 (SD 2.9) kg/m² in the HILT group and 26.0 (SD 2.8) kg/m² in the sham laser group (*P*=.20). The FPI-6 mean scores were –0.4 (SD 1.1) in the HILT group and –0.6 (SD 1.1) in the sham group, indicating a slightly supinated foot posture [20]. Moreover, no patients in our study had markedly supinated or pronated feet. There was no significant difference in the duration of foot pain, with a mean of 14.7 (SD 4.2) weeks in the HILT group and 14.8 (SD 4.4) weeks in the sham group (Table 1).

XSL•FO
RenderX

Table . Demographic characteristics of the 2 study groups.

Characteristics	High-intensity laser therapy group (n=18)	Sham group (n=16)	*P* value
Age (y), mean (SD)	46 (11.4)	49 (13.0)	.52
Gender, n (%)			.61
Male	2 (11.1)	1 (6.3)	
Female	16 (88.9)	15 (93.7)	
FPI-6[a] (points), mean (SD)	–0.4 (1.1)	–0.6 (1.1)	.51
BMI (kg/m²), mean (SD)	23.3 (2.9)	26.0 (2.8)	.21
Duration of foot pain (wk), mean (SD)	14.7 (4.2)	14.8 (4.4)	>.99

[a]FPI-6: Foot Posture Index.

The ITT analysis (n=36; pooled across 20 imputations) showed the following posttreatment results. For pain (VAS), the mean posttreatment score was 23.0 (SD 17.5) mm in the HILT group versus 26.8 (SD 20.7) mm in the sham group. The between-group mean difference (laser and sham groups) in posttreatment VAS was –5.0 mm (95% CI –14.3 to 4.3; *P*=.59). Cohen *d* for the between-group difference at posttreatment (pooled SD) was small (d≈–0.20).

Mean PFT decreased from 4.8 (SD 1.4) mm to 4.1 (SD 1.2) mm (mean change –0.7 mm, 95% CI –1.0 to −0.5; *P*<.001) in the HILT group. For PFT, the median decreased from 5.3 (IQR 4.5 - 5.4) mm to 4.5 (IQR 3.7 - 4.7) mm (median change –0.7 mm, 95% CI –1.0 to −0.5; *P*<.001) in the sham group. The between-group difference in change was –0.02 mm (95% CI –0.2 to 0.1; *P*=.90). Because values in the sham group were nonnormally distributed, a nonparametric test (Mann-Whitney *U* test) was used for between-group comparisons of this outcome, and Cohen *d* was not reported for this comparison.

For FAAM, the mean posttreatment scores were 82.3 (SD 12.8) in the HILT group and 82.8 (SD 17.7) in the sham group. Within-group improvements were significant in both groups. The mean change was 18.7 (95% CI 10.4 to 27.1; *P*<.001) in the HILT group and 14.5 (95% CI 3.2 to 25.8; *P*=.02) in the sham group. The between-group mean difference in change was 5.6 (95% CI –1.1 to 12.4; *P*=.40). Cohen *d* for the between-group posttreatment comparison was very small (d≈0.03; Table 2).

Table . Comparison of pain intensity measured by the 100-mm VAS, plantar fascial thickness, and FAAM score.

Outcomes	High-intensity laser therapy group (n=18)				Sham group (n=16)				High-intensity laser therapy versus sham	
	Before, mean (SD)	After, mean (SD)	Mean difference (95% CI)	*P* value	Before, mean (SD)	After, mean (SD)	Mean difference (95% CI)	*P* value	Between-group mean difference (95% CI)	*P* value
VAS[a] (mm)	58.3 (19.1)	23.0 (17.5)	–35.3[b] (–45.3 to –25.0)	<.001	57.1 (22.2)	26.8 (20.7)	–30.4[b] (–46.3 to –14.4)	.001	–5.0 (–14.3 to 4.3)[c]	.59
Plantar fascia thickness (mm)	4.8 (1.4)	4.1 (1.2)	–0.7[b] (–1.0 to –0.5)	<.001	5.3 (4.5 to 5.4)[d]	4.5 (3.7 to 4.7)[d]	–0.7[e] (–1.0 to –0.5)	<.001	–0.02 (–0.2 to 0.1)[c]	.90
FAAM[f] (point)	63.6 (13.6)	82.3 (12.8)	18.7[b] (10.4 to 27.1)	<.001	68.3 (23.8)	82.8 (17.7)	14.5[b] (3.2 to 25.8)	.02	5.62 (–1.1 to 12.4)[c]	.40

[a]VAS: visual analog scale.

[b]Paired *t* test was used.

[c]Independent *t* test was used.

[d]Median (IQR).

[e]Wilcoxon signed rank test was used.

[f]FAAM: foot and ankle ability measure.

The per-protocol analysis (n=34 completers) produced results consistent with the ITT analysis; between-group differences in change scores remained small and nonsignificant for VAS, PFT, and FAAM, supporting the robustness of the findings. Among participants with plantar fasciitis who had a history of strenuous weight-bearing exercise, 1 participant (a runner) in the HILT

XSL•FO
RenderX

group was advised to refrain from such activity during the treatment period. Furthermore, 2 participants (1 from each group) reported taking paracetamol at a dosage of 2000 mg per day for 2 days. No adverse events (eg, increased pain, skin burns) were reported in the HILT group. Two participants in the sham group withdrew owing to pain intolerance. Adherence to the exercise protocol did not differ significantly between groups: there were a mean 15.1 (SD 3.8) sessions in the HILT group versus 17.5 (SD 3.7) in the sham group (P=.06).

Discussion

This randomized controlled trial assessed the effectiveness of HILT combined with conventional therapy (stretching exercises) compared with sham laser plus conventional therapy in patients with plantar fasciitis. A key strength of our methodology, which distinguishes it from many prior trials, was the inclusion of objective structural outcomes (ultrasound) and validated quality-of-life assessment (FAAM), coupled with a rigorous sham control. Both treatment arms showed significant improvements in pain scores, a reduction in PFT, and an increase in foot and ankle function following treatment. However, HILT did not demonstrate superiority over the sham intervention.

This primary finding—that HILT offered no clear added benefit—aligns with several previous investigations. For instance, Tkocz et al [21], studying chronic heel pain (including plantar fasciitis and calcaneal spurs) using a HILT device at a lower power (7 W) and high total energy, similarly found no clear advantage of HILT over sham combined with ultrasound. Furthermore, Naruseviciute and Kubilius [16] also reported no superiority of HILT versus low-intensity laser therapy regarding either pain or fascial thickness. Conversely, Akkurt et al [22] observed improved outcomes when HILT was combined with insoles; however, the absence of a true sham laser control in that study makes it impossible to definitively exclude a significant placebo contribution.

The reasons for this consistent lack of incremental benefit are likely multifactorial. High-intensity laser protocols vary widely across trials concerning power, energy per session, mode, number of sessions, and application technique [15,16,21,22]. This substantial heterogeneity may partly explain discrepant results in the literature, as these parameters dictate whether the tissue effects are primarily photobiomodulatory or thermal. Nevertheless, the majority of trials to date do not show a consistent, clinically meaningful incremental effect of HILT over standard conservative care. Future research must prioritize robust dose-finding studies, provide complete dosimetry and application details, and include mechanistic measures to determine whether any specific parameter set reliably improves outcomes in plantar fasciitis before routine adoption.

Beyond methodological variations, anatomical considerations may temper the potential benefit of HILT. The plantar fascia lies only a few millimeters below the skin, varying by location and individual subcutaneous thickness. Consequently, the deeper penetration advantage often ascribed to HILT may be less relevant for plantar fasciitis compared with deeper targets (eg, knee osteoarthritis), where reaching intra-articular structures is more pertinent [23]. This superficial anatomical location of the target may inherently limit the potential incremental benefit of HILT in this disease.

Given that all 3 primary outcomes improved significantly in both study arms, it strongly suggests that the standardized stretching program was the major driver of clinical benefit. Plantar fascia and calf stretching target the gastrocnemius–soleus complex and the fascia because tightness increases tensile strain on the arch [4]. Mechanistically, regular stretching likely contributed to the reduction in pain and functional gains by increasing ankle dorsiflexion, lowering passive tension transmitted to the fascial insertion, and improving tissue viscoelasticity.

The consistent improvements observed in our objective and patient-centric measures validate the effectiveness of this conservative approach. The reduction in PFT, quantified objectively via ultrasound, confirms that the therapy successfully induced a positive structural change in the affected tissue, moving beyond mere symptomatic relief. PFT is considered the anatomical hallmark of chronic plantar fasciitis (reflecting chronic microtearing and disorganized collagen repair) [18]. Similarly, the significant increase in scores on the FAAM further substantiates the clinical relevance of the treatment. Since the FAAM is a validated patient-reported outcome measure focusing specifically on function and participation in daily living and sports activities [19], the documented improvements indicate that the structural and pain changes successfully translated directly into better functional capacity for participants.

An alternative explanation for the null between-group difference lies in statistics. The study was powered to detect a 15-mm VAS difference, yet the observed between-group difference was statistically small (only –5.0 mm). This small effect size combined with substantial variability reduced the power to find a statistically significant difference. Additionally, because many participants improved substantially with stretching (and some used short courses of acetaminophen), a ceiling effect is possible, leaving little room for additional measurable improvement attributable solely to HILT. Reporting responder rates, change-score distributions, and conducting larger trials or trials enrolling participants with higher baseline pain would help determine whether a clinically meaningful subgroup benefit exists.

Most participants in our sample had near-normal alignment with slight supination, and baseline alignment was balanced between the 2 groups, so imbalance in foot posture is unlikely to account for the null between-group findings. Nonetheless, because plantar fascia loading differs across foot types and many patients with plantar fasciopathy present with pes planus or overpronation [2], the sample's alignment profile may limit external validity. Future trials should objectively quantify foot posture (eg, FPI-6, navicular drop, calcaneal eversion, or dynamic pressure measures), consider stratified randomization or prespecified subgroup analyses by alignment, and, where appropriate, tailor background conservative care to alignment to test whether foot posture modifies the effect of adjunctive treatments such as HILT.

In terms of safety and adherence, a small number of participants used brief acetaminophen therapy, and 1 HILT participant was

XSL•FO
RenderX

advised to avoid strenuous running during treatment. This activity modification is an independent potential confounder. Importantly, however, no adverse events were reported among HILT recipients, supporting short-term tolerability in our sample. Although 2 participants in the sham group withdrew because of pain intolerance, both ITT and per-protocol sensitivity analyses produced similar results, supporting the robustness of our conclusions.

Finally, this study assessed outcomes only immediately postintervention, meaning the durability of effects remains unknown. Furthermore, the lack of standardized HILT protocols for plantar fasciitis complicates cross-trial comparison. Future research must prioritize longer follow-up, dose-finding studies, standardized reporting of dosimetry, and adequately powered trials (including stratified or targeted enrollment) to identify the specific patient subgroups most likely to benefit from HILT.

This study provides high-quality evidence that HILT does not offer additional benefits over standard stretching exercises in the treatment of plantar fasciitis. Given the statistically similar outcomes between the HILT and sham laser groups, our findings reinforce the effectiveness of conservative treatment approaches as the primary management strategy for plantar fasciitis.

Acknowledgments

The authors would like to express their sincere gratitude for the financial support provided, which made this study possible.

Funding

This research was supported by the Research and Innovation Division, Faculty of Medicine, Khon Kaen University, Thailand.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH (V 1.6.1) checklist.
[PDF File, 1162 KB - rehab_v13i1e77419_app1.pdf]

References

1. Scher DL, Belmont PJ, Bear R, Mountcastle SB, Orr JD, Owens BD. The incidence of plantar fasciitis in the United States military. J Bone Joint Surg Am 2009 Dec;91(12):2867-2872. [doi: 10.2106/JBJS.I.00257] [Medline: 19952249]
2. Johannsen FE, Rydahl JP, Jacobsen AS, Brahe CCH, Magnusson PS. Foot posture and ankle dorsiflexion as risk factors for developing Achilles tendinopathy and plantar fasciitis: a case-control study. Foot Ankle Int 2024 Dec;45(12):1380-1389. [doi: 10.1177/10711007241281289] [Medline: 39422991]
3. Koc TA, Bise CG, Neville C, Carreira D, Martin RL, McDonough CM. Heel pain - plantar fasciitis: revision 2023. J Orthop Sports Phys Ther 2023 Dec;53(12):CPG1-CPG39. [doi: 10.2519/jospt.2023.0303] [Medline: 38037331]
4. DiGiovanni BF, Nawoczenski DA, Lintal ME, et al. Tissue-specific plantar fascia-stretching exercise enhances outcomes in patients with chronic heel pain. a prospective, randomized study. J Bone Joint Surg Am 2003 Jul;85(7):1270-1277. [doi: 10.2106/00004623-200307000-00013] [Medline: 12851352]
5. Guimarães JDS, Arcanjo FL, Leporace G, et al. Effects of therapeutic interventions on pain due to plantar fasciitis: a systematic review and meta-analysis. Clin Rehabil 2023 Jun;37(6):727-746. [doi: 10.1177/02692155221143865] [Medline: 36571559]
6. Kujawa J, Zavodnik L, Zavodnik I, Buko V, Lapshyna A, Bryszewska M. Effect of low-intensity (3.75-25 J/cm2) near-infrared (810 nm) laser radiation on red blood cell ATPase activities and membrane structure. J Clin Laser Med Surg 2004 Apr;22(2):111-117. [doi: 10.1089/104454704774076163] [Medline: 15165385]
7. De Oliveira MF, Johnson DS, Demchak T, Tomazoni SS, Leal-junior EC. Low-intensity LASER and LED (photobiomodulation therapy) for pain control of the most common musculoskeletal conditions. Eur J Phys Rehabil Med 2021;58(2):282-289. [doi: 10.23736/S1973-9087.21.07236-1]
8. de Oliveira VLC, Silva JA Jr, Serra AJ, et al. Photobiomodulation therapy in the modulation of inflammatory mediators and bradykinin receptors in an experimental model of acute osteoarthritis. Lasers Med Sci 2017 Jan;32(1):87-94. [doi: 10.1007/s10103-016-2089-2] [Medline: 27726041]
9. Dos Santos SA, Sampaio LM, Caires JR, et al. Parameters and effects of photobiomodulation in plantar fasciitis: a meta-analysis and systematic review. Photobiomodul Photomed Laser Surg 2019 Jun;37(6):327-335. [doi: 10.1089/photob.2018.4588] [Medline: 31107161]
10. Guimarães JS, Arcanjo FL, Leporace G, et al. Effect of low-level laser therapy on pain and disability in patients with plantar fasciitis: a systematic review and meta-analysis. Musculoskelet Sci Pract 2022 Feb;57:102478. [doi: 10.1016/j.msksp.2021.102478] [Medline: 34847470]

XSL•FO
RenderX

11. Naterstad IF, Joensen J, Bjordal JM, Couppé C, Lopes-Martins RAB, Stausholm MB. Efficacy of low-level laser therapy in patients with lower extremity tendinopathy or plantar fasciitis: systematic review and meta-analysis of randomised controlled trials. BMJ Open 2022 Sep 28;12(9):e059479. [doi: 10.1136/bmjopen-2021-059479] [Medline: 36171024]
12. Wang W, Jiang W, Tang C, Zhang X, Xiang J. Clinical efficacy of low-level laser therapy in plantar fasciitis: a systematic review and meta-analysis. Medicine (Baltimore) 2019 Jan;98(3):e14088. [doi: 10.1097/MD.0000000000014088] [Medline: 30653125]
13. Kaub L, Schmitz C. Comparison of the penetration depth of 905 nm and 1064 nm laser light in surface layers of biological tissue ex vivo. Biomedicines 2023 May 4;11(5):1355. [doi: 10.3390/biomedicines11051355] [Medline: 37239026]
14. Astri SW, Murdhana N, Nusdwinuringtyas N, Kekalih A, Sunarjo P, Soewito F. The comparison of the low-level laser therapy and high intensity laser therapy on pain and functional ability in knee osteoarthritis. J Indon Med Assoc 2022;72(6):275-283. [doi: 10.47830/jinma-vol.72.6-2022-826]
15. Ordahan B, Karahan AY, Kaydok E. The effect of high-intensity versus low-level laser therapy in the management of plantar fasciitis: a randomized clinical trial. Lasers Med Sci 2018 Aug;33(6):1363-1369. [doi: 10.1007/s10103-018-2497-6] [Medline: 29627888]
16. Naruseviciute D, Kubilius R. The effect of high-intensity versus low-level laser therapy in the management of plantar fasciitis: randomized participant blind controlled trial. Clin Rehabil 2020 Aug;34(8):1072-1082. [doi: 10.1177/0269215520929073] [Medline: 32513018]
17. Goff JD, Crawford R. Diagnosis and treatment of plantar fasciitis. Am Fam Physician 2021 Sep 15;84(6) [FREE Full text]
18. Manske RC, Wolfe C, Page P, Voight M. Diagnostic musculoskeletal ultrasound in the evaluation of the plantar fascia. Int J Sports Phys Ther 2025;20(7):1091-1096. [doi: 10.26603/001c.141177] [Medline: 40620402]
19. Arunakul M, Arunakul P, Suesiritumrong C, Angthong C, Chernchujit B. Validity and reliability of thai version of the foot and ankle ability measure (FAAM) subjective form. J Med Assoc Thai 2015 Jun;98(6):561-567. [Medline: 26219160]
20. Kyung MG, Cho YJ, Lee JH, Shin MS, Park JH, Lee DY. Reliability and radiographic correlation of the foot posture index-6: a multi-rater analysis in symptomatic and asymptomatic individuals. Diagnostics (Basel) 2025 May 12;15(10):1214. [doi: 10.3390/diagnostics15101214] [Medline: 40428206]
21. Tkocz P, Matusz T, Kosowski Ł, et al. A randomised-controlled clinical study examining the effect of high-intensity laser therapy (HILT) on the management of painful calcaneal spur with plantar fasciitis. J Clin Med 2021 Oct 23;10(21):4891. [doi: 10.3390/jcm10214891] [Medline: 34768411]
22. Akkurt F, Akkurt HE, Yılmaz H, Olgun Y, Sen Z. Efficacy of high-intensity laser therapy and silicone insole in plantar fasciitis. Int J Phys Med Rehabil 2018;06(05). [doi: 10.4172/2329-9096.1000484]
23. Wyszyńska J, Bal-Bocheńska M. Efficacy of high-intensity laser therapy in treating knee osteoarthritis: a first systematic review. Photomed Laser Surg 2018 Jul;36(7):343-353. [doi: 10.1089/pho.2017.4425] [Medline: 29688827]

Abbreviations

FAAM: foot and ankle ability measure
FPI-6: Foot Posture Index
HILT: high-intensity laser therapy
ITT: intention-to-treat
PFT: plantar fascia thickness
VAS: visual analog scale

Edited by K Kokorelias; submitted 13.May.2025; peer-reviewed by P Areeudomwong, P Sakulsriprasert, SP Madani, W Namwong; revised version received 22.Dec.2025; accepted 22.Jan.2026; published 20.Feb.2026.

Please cite as:
Jitpimolmard N, Ouemphancharoen P, Arayawichanon P
Efficacy of High-Intensity Laser Therapy Combined With Plantar Fascia Stretching Exercises in the Treatment of Plantar Fasciitis: Randomized, Double-Blind, Sham-Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e77419
URL: https://rehab.jmir.org/2026/1/e77419
doi:10.2196/77419

© Nantaporn Jitpimolmard, Phonlawat Ouemphancharoen, Preeda Arayawichanon. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology,

XSL•FO
RenderX

is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Validity and Reliability of an Immersive Virtual Reality System for Multidimensional Assessment of Cervical Sensorimotor Control: Cross-Sectional Study

Ya-Lan Chiu[1,2], PT, PhD; Pei-Yun Lee[1,2], PT, PhD; Pooi-Ling Lim[1], PT, MS; Kai-Chia Cheng[3,4], PT, PhD; Zong-Xian Yin[5], PhD; Yi-Ju Tsai[1,2,3], PT, PhD

[1]Department of Physical Therapy, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[2]Physical Therapy Center, National Cheng Kung University Hospital, Tainan, Taiwan

[3]Institute of Allied Health Sciences, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[4]National Center for Geriatrics and Welfare Research, National Health Research Institute, Yunlin, Taiwan

[5]Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan

Corresponding Author:
Yi-Ju Tsai, PT, PhD
Institute of Allied Health Sciences
College of Medicine
National Cheng Kung University
No.1, University Road,
Tainan, 701
Taiwan
Phone: 886 62353535 ext 5021
Email: yijutsai@mail.ncku.edu.tw

Abstract

Background: Cervical sensorimotor control (SMC) is often disrupted in individuals with chronic neck pain, contributing to persistent symptoms and functional limitations. Traditional cervical SMC assessments are limited by complex setups, single-domain testing, and examiner dependency. Virtual reality (VR) technology offers a promising platform for multidimensional, standardized, and user-friendly assessment.

Objective: This study aimed to develop and evaluate the validity and reliability of a VR-based system for assessing cervical SMC in healthy adults.

Methods: A cross-sectional observational study was conducted in 30 healthy adults (aged 18-60 years). The custom-developed VR system (HP Reverb G2 Omnicept Edition, Unity engine) incorporated 5 SMC tests: cervical range of motion (ROM), joint position error (JPE), head-tilt response, figure of eight (FOE), and postural sway (PS). Test-retest reliability was assessed across 2 sessions, separated by 1 week, using intraclass correlation coefficients (ICCs). Concurrent validity was examined by comparing VR-based measures with gold standard optical motion capture or established clinical tools using Pearson correlation coefficients.

Results: The VR system demonstrated good to excellent test-retest reliability across most outcome measures. The ICC for cervical ROM ranged from 0.851 to 0.968 across movement directions. The ICC for JPE in each direction ranged from 0.813 to 0.827. The ICC for the FOE test's deviation frequency and task duration were 0.810 and 0.913, respectively. The ICC for the head-tilt response was 0.742 and ranged from 0.720 to 0.843 for PS under both visual conditions. The VR-based assessments for ROM, JPE, FOE, and PS showed strong correlations with reference measures (r=0.723-0.980), supporting concurrent validity.

Conclusions: This VR-based assessment system provides a valid, reliable, and user-friendly multidimensional evaluation of cervical SMC. It offers a standardized, integrated, and clinically feasible alternative to conventional assessments, with potential applications in both clinical diagnostics and rehabilitation monitoring.

Trial Registration: ClinicalTrials.gov NCT06474130; https://clinicaltrials.gov/study/NCT06474130

(JMIR Rehabil Assist Technol 2026;13:e88498) doi:10.2196/88498

KEYWORDS

cervical spine; neck pain; proprioception; reliability; sensorimotor control; validity; virtual reality

XSL•FO
RenderX

Introduction

Neck pain is a common global health concern, with epidemiological studies suggesting that approximately 30% to 50% of adults experience it at least once in their lifetime [1,2]. Although many acute episodes resolve spontaneously, a considerable proportion of individuals develop chronic neck pain (CNP), typically defined as pain persisting for more than 3 months [3]. Recent estimates suggest that more than 200 million people worldwide are affected by CNP, making it one of the leading causes of musculoskeletal disability globally. The prevalence of CNP is expected to rise sharply in the coming decades due to population aging and sedentary lifestyles [2,4]. Beyond physical discomfort, CNP is associated with psychological distress, decreased work productivity, and substantial long-term health care expenditures [5,6]. Moreover, evidence suggests that individuals with CNP frequently exhibit various sensorimotor control (SMC) impairments, including impaired proprioception, delayed neuromuscular responses, reduced joint mobility, altered eye-head coordination, and increased movement variability and postural sway (PS), which may complicate rehabilitation efforts and contribute to the persistence of symptoms [7-11].

Accurate assessment of cervical SMC is important for identifying underlying deficits associated with CNP and tailoring rehabilitation strategies. Previously, assessment methods targeting distinct functional components of cervical SMC were established and used in both clinical and research settings [7,12,13]. However, they often require multiple instruments, experienced examiners, and controlled environments. For example, cervical range of motion (ROM) assessments vary in complexity, from the use of basic tools such as goniometers and inclinometers to more advanced systems including optical motion capture and inertial measurement units (IMUs) [14]. The joint position error (JPE) test for cervical proprioception typically requires a head-mounted laser pointer and a target board to quantify repositioning error [9,10,15]. PS, which reflects whole-body balance control, is usually measured using force platforms or stabilometric systems [16]. Their fragmented nature and susceptibility to human error limit their clinical scalability. Therefore, there is an increasing need for more accessible, integrated, and standardized approaches to comprehensively assess cervical SMC across its multiple dimensions.

Virtual reality (VR) technology enables immersive, interactive, and quantifiable testing environments that integrate multiple SMC domains [17-20]. Head-mounted VR systems with embedded motion sensors enable real-time tracking of head kinematics, providing standardized and automated feedback [18,21]. Recent research has increasingly explored the potential of VR as a tool for assessing cervical SMC in individuals with CNP, although the studies are still limited. Few studies have validated VR-based measures of cervical motion and proprioception, showing strong agreement with gold-standard systems [10,22-24]. The VR-based assessment of the subjective visual vertical or head-tilt response (HTR) has also been shown to be reliable for screening otolithic dysfunctions and measuring perceptual orientation [25-27]. These findings suggest that VR may offer a compelling, multidimensional platform for the precise and engaging assessment of SMC in both clinical and research settings. However, current studies still focus on a single domain of SMC and lack integrated evaluation across multiple sensorimotor components. To address this gap, this study developed a multidimensional immersive VR-based cervical SMC assessment system that integrates mobility, proprioception, coordination, verticality perception, and postural stability within a single protocol. The study aimed to (1) evaluate the test-retest reliability of the VR system and (2) examine its concurrent validity compared with established reference methods. We hypothesized that the VR-based assessments would demonstrate good to excellent reliability (intraclass correlation coefficient [ICC] ≥0.75) and strong correlations (r≥0.70) with gold-standard measures.

Methods

Study Design

This cross-sectional observational study was conducted from October 2024 to May 2025 in a university laboratory.

Participants

A total of 30 healthy adults (15 male and 15 female; mean age 20.43, SD 1.01 years; mean height 166.08, SD 8.57 cm; mean weight 59.38, SD 12.49 kg) participated in the study. Exclusion criteria included (1) current spinal pain and spinal deformities, (2) history of spinal fracture or surgery, (3) neurological or musculoskeletal conditions that may affect performance, (4) diseases causing balance dysfunction such as central or peripheral nervous system disorders, or vestibular system dysfunction, (5) metabolic diseases, (6) psychological conditions that may affect questionnaire responses, and (7) visual impairments requiring correction, if participants were unable to wear contact lenses during VR testing.

Power analysis (α=.05; power=0.90) indicated that a minimum of 24 samples was required for reliability (expected ICC=0.75) and 14 for validity (expected r=0.70), ensuring adequate power for all analyses [28,29]. The sample size was estimated using G*power software (version 3.1.9.6; Heinrich Heine University Düsseldorf).

Instrumentation

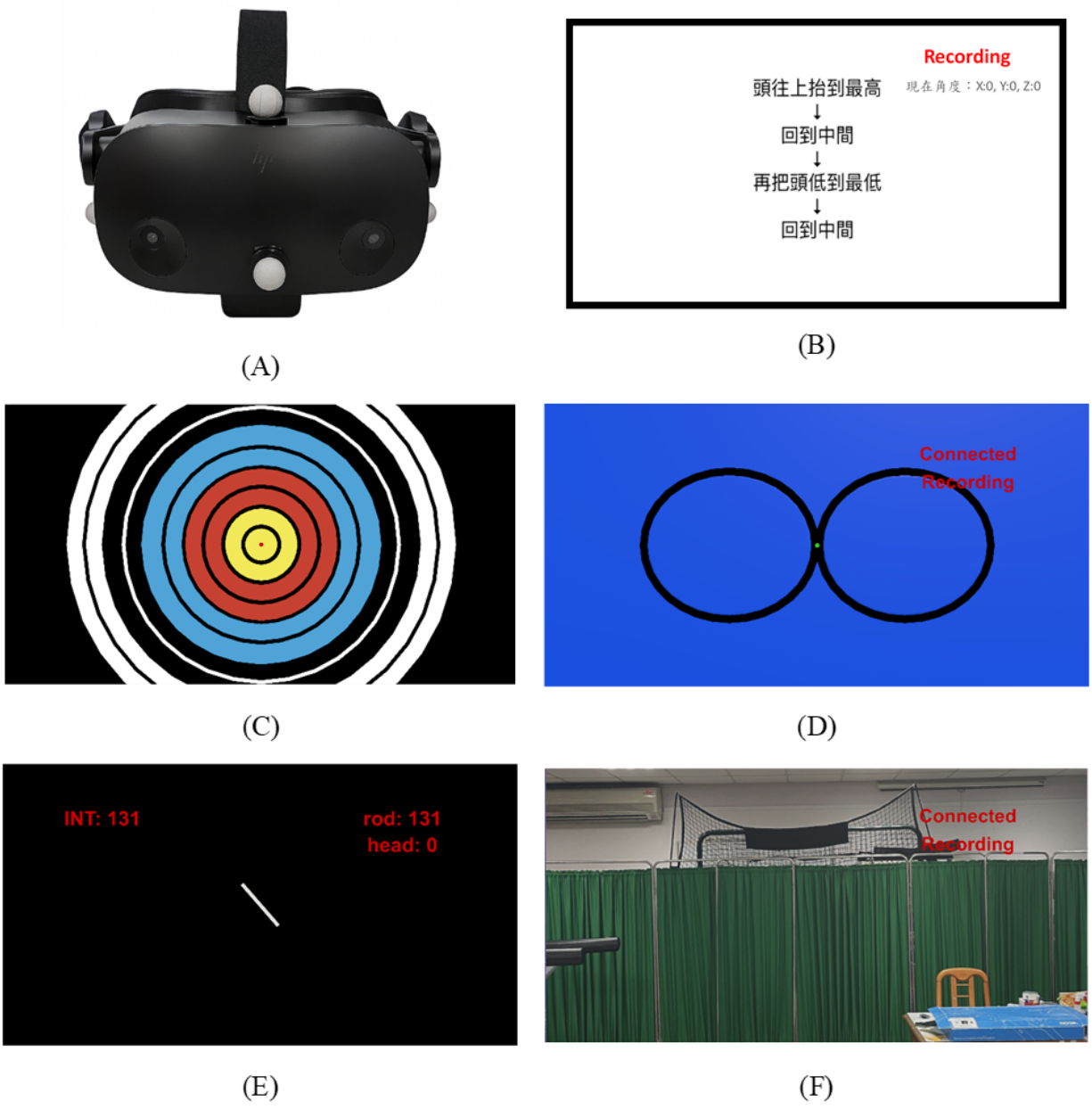
VR-Based System

A custom VR-based cervical SMC assessment system was developed using the HP Reverb G2 Omnicept Edition headset (HP Inc; Figure 1A), integrated with a custom-built virtual environment created using Unity 3D (Unity Technologies). The VR headset is equipped with embedded IMUs (gyroscope, accelerometer, and magnetometer) operating at a sampling rate of 512 Hz, capturing 3D inertial signals to enable precise tracking of head kinematics. The system incorporated 5 virtual test modules simulating standardized SMC assessments: (1) ROM tests to determine maximal active motion range, (2) JPE test to quantify cervical proprioceptive accuracy, (3) figure of eight (FOE) test to assess visual-head dynamic coordination, (4) HTR to assess verticality perception, and (5) PS to assess postural stability in standing postures. All modules provided

XSL•FO
RenderX

real-time visual feedback and standardized instructions. All visual targets and feedback across modules were presented as virtual images within the VR environment at a fixed virtual distance. No real-world targets or pass-through vision were used.

Figure 1. Virtual reality system used for cervical sensorimotor control assessment. (A) HP Reverb G2 Omnicept Edition head-mounted display and the reflective markers attached for the validity study. (B) Virtual environment designed for cervical range of motion assessment. (C) Virtual environment designed for joint position error assessment. (D) Virtual environment designed for figure-of-eight assessment. (E) Virtual environment designed for head-tilt response. (F) Virtual environment designed for postural sway assessment.



Reference Systems

To evaluate concurrent validity, the VR-derived outcomes were compared with those obtained from a motion analysis system or a clinical observation test. An optical motion capture system, including 10 cameras (VICON T10, Vicon Motion Systems Ltd) with a 100 Hz sampling rate, was used as the gold standard for ROM, JPE, and PS measures. A total of 4 reflective markers with 14 mm spheres were attached onto the front, top, left, and right sides of the VR headset to define a local coordinate system of the head (Figure 1A). Trajectory data were first digitized using the Nexus 2.10 software (Vicon Motion Systems Ltd) and subsequently processed using a custom MATLAB script (R2023a, MathWorks Inc).

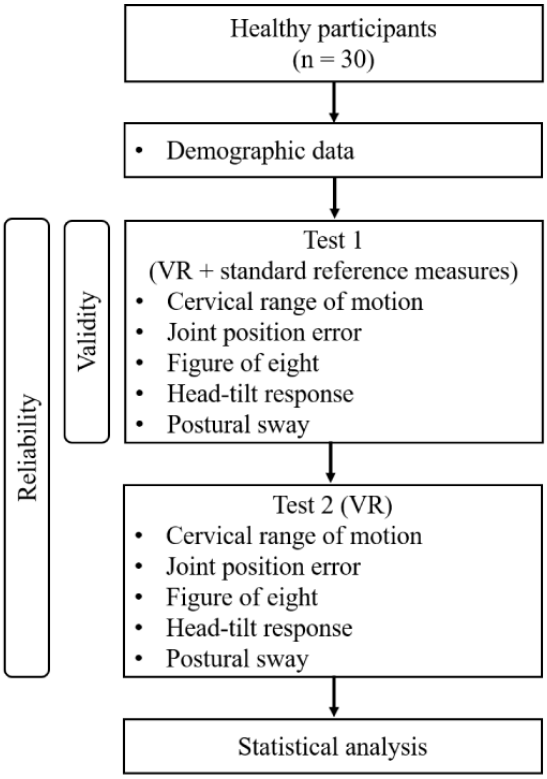
Experimental Procedures

Participants completed 2 identical VR assessment sessions, separated by 1 week, to assess test-retest reliability. To examine validity, VR and reference measures were collected concurrently within the same first session (Figure 2). All participants wore the VR headset in a comfortable and stable position throughout the assessment. For most tests, they were seated in a

XSL•FO
RenderX

back-supported chair with their feet flat on the ground and their hips/knees at approximately 90 degrees. The only exception was the PS assessment, which was performed in a quiet standing position with feet together. During all seated tasks, participants were instructed to maintain their trunk in contact with the back support and to avoid any trunk or shoulder movement during all head motion tasks. The examiner visually monitored participant performance throughout testing, and trials were repeated if obvious compensatory trunk or upper body movements were observed.

Figure 2. Study design and overall experimental procedures. VR: virtual reality.



To ensure consistency in orientation-based measurements, a neutral head position was established through a standardized calibration procedure. Participants were instructed to adopt a comfortable, upright head posture while looking straight ahead at a fixed visual reference in the virtual environment. Head orientation data were recorded over a brief 1-second static period, and the average orientation during this period was defined as the neutral reference. This calibration procedure was repeated before each trial to minimize the influence of sensor drift or initial positioning offsets. The same standardized instructions and procedures were applied across sessions to ensure consistency between test and retest measurements.

Participants completed a standardized familiarization phase prior to formal data collection. Each participant performed one practice trial for each task. Only for the HTR task, 2 practice trials were provided due to its greater complexity and to ensure adequate understanding. No performance feedback was provided beyond standardized instructions. A 1-minute rest period was provided between tests to minimize fatigue. The same testing procedures and environmental conditions were maintained across sessions to ensure consistency. All assessments were administered by a single trained examiner who provided standardized instructions using a predefined script. The same examiner conducted both test sessions for each participant to ensure consistency across sessions. Given the automated nature of the VR system, examiner involvement was limited to task instruction and safety supervision, with no manual data recording or subjective scoring involved. The examiner was blinded to the results of the previous session.

Five Cervical SMC Assessments

ROM Assessment

Cervical ROM tests were assessed to evaluate movement capacity in 6 directions: flexion, extension, lateral bending, and rotation to both the left and right sides [30]. Participants followed visual cues in the VR headset to move their head to the maximal range in each direction and then return to neutral (Figure 1B). Each movement was performed 3 times, with rest between directions to prevent fatigue.

For VR data analysis, head orientation was continuously tracked using the headset tracking system and expressed as an Euler angle sequence relative to the calibrated neutral position. Orientation data were derived from quaternion-based signals integrated from inertial and positional tracking and converted within Unity (transform.eulerAngles). Data were then exported and processed offline using MATLAB (MathWorks), where angular discontinuities were corrected, and values were expressed relative to the initial position. No additional filtering was applied. Peak angular displacement was identified for each trial, and the mean of the 3 trials was used as the final ROM value [31,32].

For concurrent validation, cervical ROM was simultaneously captured using the Vicon optical motion system. Euler angles

XSL•FO
RenderX

were computed from the head orientation relative to the global coordinate system, and peak angular displacements in each direction were identified using the same criteria as those applied to the VR-based data.

JPE Assessment

The JPE test evaluated cervicocephalic proprioception by measuring head repositioning accuracy in the absence of visual feedback [33,34]. In the VR headset, a virtual target (50 cm in diameter with concentric rings at 5 cm intervals) was displayed 1 meter in front of the participant (Figure 1C). Participants were instructed to align a red dot representing the head orientation with the target center, memorize the starting position, close their eyes, rotate their head in 1 of 4 directions (flexion, extension, right rotation, or left rotation), and attempt to return to the starting point. Three trials were completed per direction, and mean values were used for analysis [12,14].

For VR data analysis, head orientation data were continuously tracked by the embedded IMUs. The orientation data were transformed into a forward unit vector to derive 2-dimensional coordinates of the red dot. No additional filtering was applied. The difference between the starting and repositioned coordinates was converted into an angular deviation, representing repositioning error.

For concurrent validation, head orientation during the JPE test was simultaneously recorded by the Vicon motion capture system. JPE values were calculated as the angular difference between the head's initial orientation and return orientation, following the same computational criteria used for the VR data.

FOE Assessment

The FOE test assessed sensorimotor coordination, as well as the smoothness and accuracy of head-controlled tracking [35,36]. In the VR environment, a horizontal FOE path (2 cm wide) appeared 1 meter in front of the participant, with a red dot pointed to the center of the visual field serving as the controllable cursor (Figure 1D). Participants traced the figure continuously using head movements at a self-selected comfortable speed. Each participant performed 3 trials, and mean values were used for analysis.

For VR data analysis, the system recorded the cursor trajectory in a 2-dimensional X-Y plane. Two primary metrics were derived, including the completion duration and the number of deviations. Completion duration was the total time required to complete one full tracing path. The number of deviations was the total number of instances when the red dot moved outside the designated path for at least 1 consecutive frame. Continuous frames outside the boundary were counted as a single deviation until the red dot reentered the path.

For clinical validation, participants performed an equivalent head-controlled tracing task using a laser pointer mounted on the top of the head to trace a physical FOE pattern (2 cm wide, size 85 × 38 cm) displayed on a vertical board positioned 1 meter in front, replicating the VR environment conditions. Outcome measures, including completion duration and number of deviations, were manually recorded by an assessor using a stopwatch and visual inspection of the recorded videos.

HTR Assessment

The HTR test evaluated verticality perception and cervical proprioceptive control, reflecting visual dependency in spatial alignment. In the VR environment, a 15 cm white rod was displayed against a black background and randomly tilted between 45° and 135°, in 1° increments (Figure 1E). Participants were instructed to laterally tilt their heads to align the rod to the perceived vertical, after which the system automatically recorded the final orientation. A total of 20 trials were performed, and the mean angular deviation across trials was calculated as the outcome measure.

For VR data analysis, head orientation was continuously tracked via embedded IMUs. The system calculated the angular deviation between the participant-adjusted rod and the true vertical reference to quantify perceived alignment error. Mean absolute deviation across all trials represented the final HTR score.

The HTR test was excluded from the validity analysis since it primarily measures perceptual-cognitive processing rather than biomechanical movement. It reflects verticality perception and multisensory integration based on internal psychophysical judgments, and has previously demonstrated acceptable reliability and validity in computerized or virtual platforms [25,26].

PS Assessment

The PS test assessed whole-body balance control and postural stability under eyes-open and eyes-closed conditions [16,27,37]. Participants stood upright with their feet together and arms at their sides for 20 seconds while viewing a static virtual scene that replicated their natural environment (Figure 1F). Three trials were performed per condition, and mean values were used for analysis.

For VR data analysis, head motion was continuously recorded by the IMUs of the headset, which measured linear acceleration in the anteroposterior and mediolateral directions as an indirect measure of PS [38,39]. Acceleration data were processed using standard signal processing procedures, including unit conversion, low-pass filtering (6 Hz), and smoothing to reduce high-frequency noise. Sway trajectories were reconstructed using signal vector magnitude and direction cosine transformations [39]. Based on an estimated head height [40], linear displacements in the anteroposterior and mediolateral directions were derived. Four sway parameters were obtained to quantify balance performance, including maximum anteroposterior and mediolateral displacements, total path length (TPL), and sway area (SA) [40-42].

For concurrent validation, Vicon motion capture data were used to estimate head center position from the midpoint between reflective markers, serving as a proxy for the body's center of mass. The X and Y coordinates of this midpoint represented mediolateral and anteroposterior displacements, respectively. The same sway parameters, including maximum anteroposterior and mediolateral displacement, TPL, and SA, were computed to enable cross-system comparison of postural stability performance.

XSL•FO
RenderX

Statistical Analysis

Test-retest reliability of the VR-based system was examined using the ICC with the 2-way random-effects model ($ICC_{2,1}$). Reliability was interpreted as poor (<0.50), moderate (0.50-0.75), good (0.75-0.90), and excellent (≥0.90) [43]. The standard error of measurement (SEM) was calculated to estimate score precision, and the minimal detectable change (MDC) was derived to identify the smallest change exceeding measurement error. Agreement between sessions was further evaluated using Bland-Altman plots, illustrating the mean difference and 95% limits of agreement (mean ± 1.96 × SD) [44]. Concurrent validity was assessed using Pearson correlation coefficients (*r*) between VR-based outcomes and corresponding measures from clinical tests or the Vicon motion capture system. Correlation strength was interpreted as poor (<0.30), fair (0.30-0.49), moderate (0.50-0.69), strong (0.70-0.89), or excellent (≥0.90) [45]. All statistical analyses were performed using IBM SPSS Statistics 25 (IBM Corp). In line with recommendations for measurement studies, ICC and Pearson correlation coefficients were interpreted as measures of effect size, and emphasis was placed on the magnitude and precision of estimates (eg, 95% CI, SEM, and MDC) rather than statistical significance testing.

Ethical Considerations

The study was prospectively registered on ClinicalTrials.gov (NCT06474130). Ethics approval has been received from the institutional review board of the National Cheng Kung University Hospital (B-ER-113-126). Written informed consents were obtained from all participants before the experiment. Participant data were handled with strict confidentiality. All identifiable information was removed prior to analysis, and the data were stored securely with access limited to members of the research team. No financial compensation was provided to participants for their involvement in this study.

Results

Test-Retest Reliability

The VR-based assessment system demonstrated good to excellent test-retest reliability across most measures (Table 1). Cervical ROM showed good to excellent reliability for all 6 directions (ICC=0.851-0.967), with low SEMs and narrow CIs, indicating stable and precise measurements. The results of JPE showed good reliability (ICC=0.813-0.827) and small MDC values (around 1.11°-1.27°), confirming the system's ability to detect meaningful changes in repositioning accuracy. The results of FOE demonstrated good to excellent reliability, with ICCs of 0.81 for deviation frequency and 0.91 for completion duration, reflecting consistent performance across sessions. The HTR demonstrated reliability approaching the good range, with an ICC of 0.742. For PS, reliability varied by visual condition. Under eyes-open conditions, the ICC values ranged from 0.720 to 0.772, while eyes-closed conditions produced higher consistency (ICC=0.759-0.843). Overall, the VR-based system provided stable, repeatable measurements across all tests, particularly for dynamic and proprioceptive tasks.

XSL•FO
RenderX

Table 1. Test-retest reliability of virtual reality–based cervical sensorimotor control assessments.

	Test 1, mean (SD)	Test 2, mean (SD)	SEM[a]	MDC[b]	ICC[c] 95% CI
Cervical range of motion (degree)					
Flexion	57.65 (11.26)	57.30 (11.18)	2.07	5.73	0.966 (0.929-0.984)
Extension	61.77 (12.96)	63.39 (13.16)	3.50	9.71	0.928 (0.850-0.966)
Left SB[d]	40.00 (7.31)	39.78 (6.49)	1.53	4.24	0.951 (0.897-0.977)
Right SB	44.47 (8.98)	44.9 3 (8.48)	1.56	4.33	0.968 (0.934-0.985)
Left ROT[e]	72.93 (7.88)	74.53 (7.59)	2.81	7.79	0.868 (0.723-0.937)
Right ROT	76.71 (7.72)	77.17 (6.58)	2.77	7.67	0.851 (0.687-0.929)
Joint position error (degree)					
Extension	2.93 (1.00)	2.88 (0.90)	0.41	1.13	0.816 (0.611-0.912)
Flexion	2.66 (0.98)	2.57 (0.98)	0.41	1.13	0.827 (0.637-0.917)
Left ROT	3.18 (0.96)	3.28 (0.92)	0.40	1.11	0.820 (0.624-0.914)
Right ROT	3.03 (0.92)	3.31 (1.19)	0.46	1.27	0.813 (0.608-0.911)
Figure of eight					
Deviation (times)	21.34 (4.86)	20.58 (4.72)	2.09	5.79	0.810 (0.609-0.908)
Duration (seconds)	32.82 (12.36)	30.17 (10.40)	3.37	9.34	0.913 (0.803-0.960)
Head-tilt response (degree)	3.81 (1.63)	3.79 (1.41)	0.77	2.15	0.742 (0.293-0.782)
Postural sway: eyes-open condition					
AP[f] (cm)	4.93 (2.23)	5.01 (2.20)	1.06	2.93	0.772 (0.502-0.895)
ML[g] (cm)	2.24 (1.19)	2.05 (1.17)	0.59	1.64	0.749 (0.484-0.879)
TPL[h] (cm)	41.98 (13.17)	41.32 (15.68)	7.51	20.82	0.731 (0.414-0.876)
SA[i] (cm^2)	9.07 (7.90)	8.85 (10.53)	4.93	13.65	0.720 (0.387-0.871)
Postural sway: eyes-closed condition					
AP (cm)	6.40 (2.54)	6.26 (2.36)	1.13	3.14	0.787 (0.536-0.902)
ML (cm)	2.87 (1.39)	2.84 (1.42)	0.69	1.91	0.759 (0.475-0.889)
TPL (cm)	51.27 (19.93)	51.66 (22.61)	8.44	23.41	0.843 (0.658-0.927)
SA (cm^2)	13.06 (12.34)	13.07 (12.74)	5.54	15.35	0.805 (0.575-0.910)

[a]SEM: standard error of measurement.

[b]MDC: minimal detectable change.

[c]ICC: intraclass correlation coefficient.

[d]SB: side bending.

[e]ROT: rotation.

[f]AP: anteroposterior direction displacement.

[g]ML: mediolateral direction displacement.

[h]TPL: total path length.

[i]SA: sway area.

Bland-Altman plots were generated for each sensorimotor domain to visualize agreement between sessions. The plots revealed acceptable limits of agreement across all tests, further confirming the consistency of the VR-based assessments (Multimedia Appendix 1).

Concurrent Validity

Table 2 summarizes the concurrent validity results comparing VR-based cervical SMC assessments with reference measures from the motion analysis system or established clinical tests. Overall, VR outcomes showed strong to excellent correlations with their gold-standard counterparts, confirming concurrent validity across the domains for ROM, JPE, FOE, and PS. For cervical ROM, excellent correlations were observed between VR and Vicon data across 6 movement directions (r=0.892-0.980), indicating highly comparable ROM outputs. The results of JPE demonstrated strong to excellent correlations with Vicon-based angular deviations (r=0.869-0.953),

XSL•FO
RenderX

confirming accurate detection of proprioceptive errors. The FOE test demonstrated an excellent correlation for deviation frequency (r=0.923) and a strong correlation for task completion duration (r=0.723) compared with corresponding clinical reference measures, supporting its validity in assessing dynamic head coordination. For PS, VR-derived sway metrics showed strong correlations with Vicon-based measures under both visual conditions, with the highest agreement observed for TPL (r=0.895-0.917) and SA (r=0.908-0.962). These findings collectively demonstrate that the VR-based system provides accurate and valid measures of cervical motion, proprioception, coordination, and postural stability.

Table 2. Concurrent validity of virtual reality (VR)-based cervical sensorimotor control assessments.

	VR, mean (SD)	Standard reference, mean (SD)	*r*
Cervical range of motion (degree)			
Flexion	49.82 (12.97)	54.77 (13.20)	0.892
Extension	58.89 (12.57)	53.82 (12.53)	0.917
Left SB[a]	39.31 (6.85)	39.14 (7.40)	0.958
Right SB	42.85 (9.21)	42.31 (7.61)	0.927
Left ROT[b]	72.96 (8.53)	71.50 (8.50)	0.980
Right ROT	75.33 (7.27)	76.12 (7.54)	0.946
Joint position error (degree)			
Extension	4.19 (1.53)	3.52 (1.55)	0.882
Flexion	4.28 (2.72)	3.56 (2.66)	0.953
Left ROT	3.83 (1.91)	3.32 (1.61)	0.869
Right ROT	3.50 (1.24)	2.79 (1.23)	0.920
Figure of eight			
Deviations (times)	21.83 (4.63)	32.43 (7.44)	0.923
Durations (seconds)	30.94 (8.27)	31.13 (8.67)	0.723
Postural sway: eyes-open condition			
AP[c] (cm)	5.03 (2.41)	2.41 (1.01)	0.801
ML[d] (cm)	2.56 (1.25)	1.04 (0.47)	0.871
TPL[e] (cm)	41.45 (14.41)	14.06 (5.03)	0.895
SA[f] (cm^2)	9.48 (9.49)	1.82 (1.79)	0.962
Postural sway: eyes-closed condition			
AP (cm)	6.70 (3.22)	3.84 (2.15)	0.825
ML (cm)	2.69 (1.27)	1.21 (0.59)	0.851
TPL (cm)	49.64 (20.49)	20.67 (10.25)	0.917
SA (cm^2)	12.81 (13.40)	3.54 (5.11)	0.908

[a]SB: side bending.

[b]ROT: rotation.

[c]AP: anteroposterior direction displacement.

[d]ML: mediolateral direction displacement.

[e]TPL: total path length.

[f]SA: sway area.

Discussion

Principal Findings

This study demonstrated that a custom-developed VR-based system provided reliable assessments of cervical SMC across all domains (ROM, JPE, FOE, HTR, and PS) and valid measurements for ROM, JPE, FOE, HTR, and PS. The shows that this VR-based system exhibited good to excellent test-retest reliability. The consistency of repeated measurements was further supported by Bland-Altman plot analyses. Concurrent validity revealed strong correlations with gold-standard measures for ROM, JPE, FOE, and PS, confirming the measurement accuracy of this system in assessing cervical SMC.

XSL•FO
RenderX

These results are consistent with previous studies supporting the reliability and validity of VR-based approaches for assessing cervical ROM and proprioception [24,46-50]. Prior work has demonstrated strong agreement between VR-derived and optical or goniometric measurements, as well as good to excellent reliability using IMU-based or VR-based systems [24,46-52]. A recent systematic review also confirmed that VR-based ROM assessments effectively differentiate between individuals with and without CNP [24,46,49]. In this study, the neutral head position was standardized using a calibration procedure; however, slight variations in self-selected posture may still influence absolute ROM values. Repeated calibration before each trial and consistent procedures across sessions likely minimized systematic bias and contributed to the high test-retest reliability observed.

Compared with previous VR-based cervical assessment systems, this study offers several methodological and conceptual advancements. Most previous studies have focused on single-domain assessments [53], whereas the current system integrates multiple domains of cervical SMC, including mobility, proprioception, coordination, verticality perception, and postural stability, within a single immersive platform. In addition, several tasks were specifically designed for VR-based interaction and automated data capture, enabling continuous kinematic tracking without reliance on manual scoring. Furthermore, this study systematically evaluated both test-retest reliability and concurrent validity across all domains within the same cohort, providing a more comprehensive validation framework.

This study also presents an immersive VR adaptation of the FOE head-tracing task, enabling quantitative assessment of coordination through metrics, such as completion time and deviation frequency. Compared with traditional head-mounted laser-pointer methods [36,47], the VR implementation allows standardized visual presentation, continuous trajectory tracking, and automated outcome computation, thereby reducing examiner dependency and improving measurement precision. Furthermore, the VR-based format allows controlled manipulation of visual feedback and task complexity, offering new opportunities to quantify SMC across varying sensory conditions. Similarly, the HTR task demonstrated acceptable reliability, consistent with previous studies on VR-based assessments of verticality perception [7,14,25,54-56].

The VR-based PS assessment exhibited good reliability and strong correlations with reference measures under both visual conditions, consistent with previous findings [57,58]. Previous studies have also reported moderate to strong agreement between VR-based and force plate-based PS measures, supporting the use of VR systems as portable and cost-effective tools for clinical balance assessment [59]. In addition, VR-based sway metrics have been shown to differentiate between healthy individuals and clinical populations across various applications [60-62]. However, systematic differences in absolute magnitude were observed, with VR values consistently higher than those obtained from the reference system. These discrepancies may be attributed to fundamental differences in measurement principles, as the VR-based sway metrics were reconstructed from head-mounted IMU signals, whereas the reference system directly quantified head displacement using marker-based motion capture. In addition, IMU-based approaches are inherently more sensitive to high-frequency motion components and may accumulate small fluctuations in orientation signals, leading to overestimation of trajectory-based metrics such as TPL and SA. Similar patterns have been reported in previous studies, where strong correlations were observed despite differences in absolute values [39,40]. These findings suggest that a VR-based system is suitable for relative comparisons and longitudinal monitoring, although direct interchangeability with reference systems should be interpreted with caution.

Slightly lower reliability observed in the HTR and certain PS parameters is consistent with previous studies and likely reflects the intrinsic variability of perceptual, proprioceptive, and balance-related tasks rather than limitations of the VR system itself [47,49]. Such variability, influenced by sensory integration, attentional demand, and postural control strategies across sessions, may reduce test-retest stability even under standardized conditions. Nevertheless, the observed reliability levels remain within acceptable ranges for complex sensorimotor assessments.

From a clinical measurement perspective, the relatively small SEM and MDC values observed across most parameters indicate good measurement precision and support the ability to detect meaningful changes at the individual level. MDC values provide a practical threshold for distinguishing true change from measurement error, suggesting that relatively small improvements can be confidently interpreted as real changes rather than measurement variability. These findings support the use of the VR-based system for monitoring rehabilitation progress and evaluating intervention outcomes. Taken together, the present findings support the notion that the proposed VR system provides an integrated and clinically feasible platform for the multidimensional assessment of cervical SMC. Furthermore, the high repeatability observed in healthy participants may also serve as a normative reference for cervical SMC performance. Such baseline data could facilitate future comparisons with individuals presenting with cervical symptoms or pathologies, thereby supporting the clinical applicability of the VR-based system for assessment and rehabilitation.

Limitations

Despite the valuable contributions of this study, several limitations should be acknowledged. The participants in this study were healthy young adults; therefore, the results may not generalize to older populations or those with CNP. Future research should validate the system in diverse age groups and clinical populations, and further examine its predictive validity and responsiveness to intervention-induced changes. Although participants were instructed to minimize trunk movement and were visually monitored during testing, compensatory cervicothoracic or trunk motion cannot be completely excluded, particularly during lateral bending tasks. As the current system relies on head-mounted tracking without additional trunk sensors, the measured angles may partially reflect combined head and trunk motion. Future studies incorporating objective trunk motion monitoring (eg, thoracic IMUs, motion capture markers, or head-to-trunk relative angle calculations) are warranted to further isolate cervical motion and improve measurement specificity. Despite this limitation, the high

XSL•FO
RenderX

test-retest reliability observed in this study suggests that the measurement approach was consistent across sessions. Visual requirements may pose a practical constraint during the VR-based assessment. Participants who usually wore glasses were required to use contact lenses to ensure compatibility with the VR headset, which may have excluded individuals with visual impairments or low tolerance for immersive VR environments. This limitation may introduce potential bias and affect both the accessibility and ecological validity of the system in real-world clinical applications. Although familiarization trials were provided prior to formal testing, the novelty of the VR environment may have introduced learning effects or interindividual variability in task adaptation, potentially influencing performance consistency across trials and participants. Moreover, factors such as headset fit, visual delay, or cybersickness could affect measurements, though none were reported in this study. Continued optimization of head-tracking accuracy and algorithm calibration may further enhance the system's sensitivity to low-amplitude motion and subtle postural adjustments.

Conclusion

This study established the reliability and validity of a novel VR-based assessment system for the multidimensional assessment of cervical SMC. A single session provides data across 4 major SMC domains and the immersive environment sustains participant motivation. The system demonstrated good to excellent reliability across all domains and strong concurrent validity for ROM, JPE, FOE coordination, and PS. Its standardized, immersive, and user-friendly design supports efficient and objective evaluation of cervical sensorimotor function with reduced operator bias and higher reproducibility. Future studies should extend validation to clinical populations and evaluate responsiveness to therapeutic interventions. This VR-based platform uses commercially available hardware and portable software that is adaptable for both clinical and tele-assessment contexts and provides a robust foundation for establishing normative reference values and advancing digital assessment and management of CNP in both clinical and telehealth settings.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This work was supported by the National Science and Technology Council, Taiwan (NSTC 113-2410-H-006-098).

Authors' Contributions

Conceptualization: YLC (lead), YJT (equal)
Data curation: PYL
Formal analysis: PLL (lead), KCC (supporting)
Funding acquisition: YJT
Investigation: ZXY
Methodology: KCC
Project administration: YLC (lead), PYL (equal), YJT (supporting)
Resources: ZXY
Supervision: YJT
Validation: PYL
Visualization: YLC (lead), YJT (supporting)
Writing—original draft: YLC (lead), YJT (supporting)
Writing—review and editing: YJT (lead), YLC (supporting), PYL (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Bland-Altman plots for agreement between repeated virtual reality–based assessments.
[PDF File (Adobe PDF File), 329 KB - rehab_v13i1e88498_app1.pdf]

References

1. Hoy D, Protani M, De R, Buchbinder R. The epidemiology of neck pain. Best Pract Res Clin Rheumatol 2010;24(6):783-792. [doi: 10.1016/j.berh.2011.01.019] [Medline: 21665126]
2. Li Y, Zhang S, Shu P. Global burden of neck pain in 204 countries from 1990-2019. Arch Med Sci 2023;19(6):1811-1821. [doi: 10.5114/aoms/170962] [Medline: 38058733]
3. Cohen SP. Epidemiology, diagnosis, and treatment of neck pain. Mayo Clin Proc 2015;90(2):284-299. [doi: 10.1016/j.mayocp.2014.09.008] [Medline: 25659245]

XSL•FO
RenderX

4. GBD 2021 Neck Pain Collaborators. Global, regional, and national burden of neck pain, 1990-2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol 2024;6(3):e142-e155 [FREE Full text] [doi: 10.1016/S2665-9913(23)00321-1] [Medline: 38383088]
5. Fejer R, Kyvik KO, Hartvigsen J. The prevalence of neck pain in the world population: a systematic critical review of the literature. Eur Spine J 2006;15(6):834-848 [FREE Full text] [doi: 10.1007/s00586-004-0864-4] [Medline: 15999284]
6. Dieleman JL, Baral R, Birger M, Bui AL, Bulchis A, Chapin A, et al. US spending on personal health care and public health, 1996-2013. JAMA 2016;316(24):2627-2646 [FREE Full text] [doi: 10.1001/jama.2016.16885] [Medline: 28027366]
7. Treleaven J. Sensorimotor disturbances in neck disorders affecting postural stability, head and eye movement control. Man Ther 2008;13(1):2-11. [doi: 10.1016/j.math.2007.06.003] [Medline: 17702636]
8. de Vries J, Ischebeck BK, Voogt LP, van der Geest JN, Janssen M, Frens MA, et al. Joint position sense error in people with neck pain: a systematic review. Man Ther 2015;20(6):736-744. [doi: 10.1016/j.math.2015.04.015] [Medline: 25983238]
9. de Zoete RM, Osmotherly PG, Rivett DA, Farrell SF, Snodgrass SJ. Sensorimotor control in individuals with idiopathic neck pain and healthy individuals: a systematic review and meta-analysis. Arch Phys Med Rehabil 2017;98(6):1257-1271. [doi: 10.1016/j.apmr.2016.09.121] [Medline: 27771360]
10. Särkilahti N, Hirvonen M, Lavapuro J, Takatalo J, Löyttyniemi E, Tenovuo O. Sensorimotor tests in patients with neck pain and its associated disorders: a systematic review and meta-analysis. Sci Rep 2024;14(1):12764 [FREE Full text] [doi: 10.1038/s41598-024-63545-3] [Medline: 38834665]
11. Hodges PW, Tucker K. Moving differently in pain: a new theory to explain the adaptation to pain. Pain 2011;152(3 Suppl):S90-S98. [doi: 10.1016/j.pain.2010.10.020] [Medline: 21087823]
12. Revel M, Andre-Deshays C, Minguet M. Cervicocephalic kinesthetic sensibility in patients with cervical pain. Arch Phys Med Rehabil 1991;72(5):288-291. [Medline: 2009044]
13. Sarig Bahat H, Takasaki H, Chen X, Bet-Or Y, Treleaven J. Cervical kinematic training with and without interactive VR training for chronic neck pain: a randomized clinical trial. Man Ther 2015;20(1):68-78. [doi: 10.1016/j.math.2014.06.008] [Medline: 25066503]
14. Kristjansson E, Treleaven J. Sensorimotor function and dizziness in neck pain: implications for assessment and management. J Orthop Sports Phys Ther 2009;39(5):364-377. [doi: 10.2519/jospt.2009.2834] [Medline: 19411769]
15. Peng B, Yang L, Li Y, Liu T, Liu Y. Cervical proprioception impairment in neck pain-pathophysiology, clinical evaluation, and management: a narrative review. Pain Ther 2021;10(1):143-164 [FREE Full text] [doi: 10.1007/s40122-020-00230-z] [Medline: 33464539]
16. Ruhe A, Fejer R, Walker B. The test-retest reliability of centre of pressure measures in bipedal static task conditions--a systematic review of the literature. Gait Posture 2010;32(4):436-445. [doi: 10.1016/j.gaitpost.2010.09.012] [Medline: 20947353]
17. Mirelman A, Bonato P, Deutsch JE. Effects of training with a robot-virtual reality system compared with a robot alone on the gait of individuals after stroke. Stroke 2009;40(1):169-174. [doi: 10.1161/STROKEAHA.108.516328] [Medline: 18988916]
18. Niehorster DC, Li L, Lappe M. The accuracy and precision of position and orientation tracking in the HTC Vive virtual reality system for scientific research. i-Perception 2017;8(3):2041669517708205 [FREE Full text] [doi: 10.1177/2041669517708205] [Medline: 28567271]
19. Kim A, Darakjian N, Finley JM. Walking in fully immersive virtual environments: an evaluation of potential adverse effects in older adults and individuals with Parkinson's disease. J Neuroeng Rehabil 2017;14(1):16 [FREE Full text] [doi: 10.1186/s12984-017-0225-2] [Medline: 28222783]
20. Burdea GC, Coiffet P. Virtual Reality Technology. New Jersey: John Wiley & Sons; 2003.
21. Sarig Bahat H, Sprecher E, Sela I, Treleaven J. Neck motion kinematics: an inter-tester reliability study using an interactive neck VR assessment in asymptomatic individuals. Eur Spine J 2016;25(7):2139-2148. [doi: 10.1007/s00586-016-4388-5] [Medline: 26831536]
22. Chen KB, Xu X, Lin J, Radwin RG. Evaluation of older driver head functional range of motion using portable immersive virtual reality. Exp Gerontol 2015;70:150-156. [doi: 10.1016/j.exger.2015.08.010] [Medline: 26315289]
23. Forsberg K, Jirlén J, Jacobson I, Röijezon U. Concurrent validity of cervical movement tests using VR technology-taking the lab to the clinic. Sensors (Basel) 2023;23(24):9864 [FREE Full text] [doi: 10.3390/s23249864] [Medline: 38139710]
24. Forsberg K, Jirlén J, Jacobson I, Röijezon U. Cervical sensorimotor function tests using a VR headset-an evaluation of concurrent validity. Sensors (Basel) 2024;24(17):5811 [FREE Full text] [doi: 10.3390/s24175811] [Medline: 39275722]
25. Wang CH, Winnick AA, Ko YH, Wang Z, Chang TP. Test-retest reliability of subjective visual vertical measurements with lateral head tilt in virtual reality goggles. Tzu Chi Med J 2021;33(3):294-300 [FREE Full text] [doi: 10.4103/tcmj.tcmj_207_20] [Medline: 34386369]
26. Ardıç FN, Metin U, Gökcan BE. Subjective visual vertical test with the 3D virtual reality system: effective factors and cybersickness. Acta Otolaryngol 2023;143(7):570-575. [doi: 10.1080/00016489.2023.2238769] [Medline: 37493360]
27. Umibe A, Fushiki H, Tsunoda R, Kuroda T, Kuroda K, Tanaka Y. Development of a subjective visual vertical test system using a smartphone with virtual reality goggles for screening of otolithic dysfunction: observational study. JMIR Form Res 2024;8:e53642 [FREE Full text] [doi: 10.2196/53642] [Medline: 38833295]

XSL•FO
RenderX

28. Bujang MA, Baharum N. Sample size guideline for correlation analysis. World J Soc Sci Res 2016;3(1):37-46. [doi: 10.22158/wjssr.v3n1p37]
29. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Arch Orofac Sci 2017;12(1):1-11 [FREE Full text] [doi: 10.21315/mjms2018.25.1.9]
30. Youdas JW, Carey JR, Garrett TR. Reliability of measurements of cervical spine range of motion--comparison of three methods. Phys Ther 1991;71(2):98-104; discussion 105. [doi: 10.1093/ptj/71.2.98] [Medline: 1989013]
31. Cutti AG, Paolini G, Troncossi M, Cappello A, Davalli A. Soft tissue artefact assessment in humeral axial rotation. Gait Posture 2005;21(3):341-349. [doi: 10.1016/j.gaitpost.2004.04.001] [Medline: 15760751]
32. Sarig-Bahat H, Weiss PL, Laufer Y. Cervical motion assessment using virtual reality. Spine (Phila Pa 1976) 2009;34(10):1018-1024. [doi: 10.1097/BRS.0b013e31819b3254] [Medline: 19404177]
33. Treleaven J, Jull G, Sterling M. Dizziness and unsteadiness following whiplash injury: characteristic features and relationship with cervical joint position error. J Rehabil Med 2003;35(1):36-43 [FREE Full text] [doi: 10.1080/16501970306109] [Medline: 12610847]
34. Röijezon U, Clark NC, Treleaven J. Proprioception in musculoskeletal rehabilitation. Part 1: basic science and principles of assessment and clinical interventions. Man Ther 2015;20(3):368-377. [doi: 10.1016/j.math.2015.01.008] [Medline: 25703454]
35. Werner IM, Ernst MJ, Treleaven J, Crawford RJ. Intra and interrater reliability and clinical feasibility of a simple measure of cervical movement sense in patients with neck pain. BMC Musculoskelet Disord 2018;19(1):358 [FREE Full text] [doi: 10.1186/s12891-018-2287-0] [Medline: 30290759]
36. Ernst MJ, Williams L, Werner IM, Crawford RJ, Treleaven J. Clinical assessment of cervical movement sense in those with neck pain compared to asymptomatic individuals. Musculoskelet Sci Pract 2019;43:64-69. [doi: 10.1016/j.msksp.2019.06.006] [Medline: 31277033]
37. Schmitz R, Arnold B. Intertester and intratester reliability of a dynamic balance protocol using the biodex stability system. J Sport Rehabil 1998;7(2):95-101. [doi: 10.1123/jsr.7.2.95]
38. Akizuki H, Uno A, Arai K, Morioka S, Ohyama S, Nishiike S, et al. Effects of immersion in virtual reality on postural control. Neurosci Lett 2005;379(1):23-26. [doi: 10.1016/j.neulet.2004.12.041] [Medline: 15814192]
39. Duarte M, Freitas SM. Revisão sobre posturografia baseada em plataforma de força para avaliação do equilíbrio. Rev Bras Fisioter 2010;14(3):183-192. [doi: 10.1590/S1413-35552010000300003]
40. Mancini M, King L, Salarian A, Holmstrom L, McNames J, Horak FB. Mobility lab to assess balance and gait with synchronized body-worn sensors. J Bioeng Biomed Sci 2011;Suppl 1:007. [doi: 10.4172/2155-9538.S1-007] [Medline: 24955286]
41. Gao J, Hu J, Buckley T, White K, Hass C. Shannon and renyi entropies to classify effects of mild traumatic brain injury on postural sway. PLoS One 2011;6(9):e24446 [FREE Full text] [doi: 10.1371/journal.pone.0024446] [Medline: 21931720]
42. Carpenter MG, Frank JS, Silcher CP. Surface height effects on postural control: a hypothesis for a stiffness strategy for stance. J Vestib Res 1999;9(4):277-286. [doi: 10.3233/ves-1999-9405]
43. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
44. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327(8476):307-310. [doi: 10.1016/S0140-6736(86)90837-8]
45. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Milton Park: Routledge; 2013.
46. Luc A, Tamer S, Hage R, Detrembleur C, Pitance L. Do the kinematics and sensorimotor control of people with chronic non-specific neck pain differ from those of healthy individuals when assessed in an immersive virtual reality environment? A systematic review. Phys Ther Rev 2022;27(6):430-443. [doi: 10.1080/10833196.2022.2143211]
47. Röijezon U, Jull G, Blandford C, Daniels A, Michaelson P, Karvelis P, et al. Proprioceptive disturbance in chronic neck pain: discriminate validity and reliability of performance of the clinical cervical movement sense test. Front Pain Res (Lausanne) 2022;3:908414 [FREE Full text] [doi: 10.3389/fpain.2022.908414] [Medline: 35875476]
48. Paul A, Connolly J, Condell J, Gardiner P. Assessing the accuracy and reliability of a low-cost virtual reality headset versus a goniometer for measuring cervical range of motion. : IEEE; 2023 Presented at: 34th Irish Signals and Systems Conference (ISSC); 2023 June 13-14; Dublin, Ireland. [doi: 10.1109/issc59246.2023.10162076]
49. Koumantakis GA, Gkouma S, Floka C, Tatsios PI, Moutzouri M, Sakellari V. Reliability and validity of the KFORCE Sens inertial sensor for measuring cervical spine proprioception in patients with non-specific chronic neck pain. Brain Sci 2024;14(12):1165 [FREE Full text] [doi: 10.3390/brainsci14121165] [Medline: 39766364]
50. Santos-Paz J, Sánchez-Picot Á, Rojo A, Martín-Pintado-Zugasti A, Otero A, Garcia-Carmona R. A novel virtual reality application for autonomous assessment of cervical range of motion: development and reliability study. PeerJ 2022;10:e14031 [FREE Full text] [doi: 10.7717/peerj.14031] [Medline: 36124134]
51. Carley P, Favolise M, Moses M, Heumiller S, Rocchio F, Rubin A. Reliability comparisons of cervical active range of motion measuring methods: universal goniometer versus virtual reality. Phys Med Rehabil Int 2022;9(1):1195. [doi: 10.26420/physmedrehabilint.2022.1195]

XSL•FO
RenderX

52. Chang KV, Wu WT, Chen MC, Chiu YC, Han DS, Chen CC. Smartphone application with virtual reality goggles for the reliable and valid measurement of active craniocervical range of motion. Diagnostics (Basel) 2019;9(3):71 [FREE Full text] [doi: 10.3390/diagnostics9030071] [Medline: 31295869]
53. Santos-Paz JA, García-Carmona R, Sánchez-Picot Á, Otero A. Comparative evaluation of virtual reality setups for assessing cervical range of motion: a reliability study. Virtual Real 2025;30(1). [doi: 10.1007/s10055-025-01299-7]
54. Cheng Y, Zhang Y, Chen F, Chen Z, Xue T, Zhang Q, et al. Virtual reality for the measurement of SVV and SVH during static head tilt in healthy adults: a novel vestibular test. Acta Otolaryngol 2022;142(3-4):291-297. [doi: 10.1080/00016489.2022.2058082] [Medline: 35382691]
55. Ma X, Shen J, Wang W, Wang L, Sun J, He K, et al. Test-retest reliability of subjective visual vertical from different head-tilt angles in young healthy adults. Lin Chuang Er Bi Yan Hou Tou Jing Wai Ke Za Zhi 2024;38(8):692-696;702. [doi: 10.13201/j.issn.2096-7993.2024.08.004] [Medline: 39118506]
56. Hayashi S, Kamo T, Ogihara H, Tani Y, Hoshino K, Kobayashi K, et al. Test-retest reliability of dynamic subjective visual vertical and visual dependency in older adults using virtual reality methods. Percept Mot Skills 2024;131(6):2069-2084. [doi: 10.1177/00315125241292094] [Medline: 39410748]
57. Lee CH, Sun TL. Evaluation of postural stability based on a force plate and inertial sensor during static balance measurements. J Physiol Anthropol 2018;37(1):27 [FREE Full text] [doi: 10.1186/s40101-018-0187-5] [Medline: 30545421]
58. Slunecko M, Csapo R. Assessment of postural sway with accelerometers: a comparison with force plate measures. Curr Issues Sport Sci 2024;9(4):058. [doi: 10.36950/2024.4ciss058]
59. Rosiak O, Puzio A, Kaminska D, Zwolinski G, Jozefowicz-Korczynska M. Virtual reality-a supplement to posturography or a novel balance assessment tool? Sensors (Basel) 2022;22(20):7904 [FREE Full text] [doi: 10.3390/s22207904] [Medline: 36298254]
60. Cetin H, Kose N, Gurses E, Turkmen C, Aksoy S, Oge HK. Effects of virtual reality for postural control in chronic neck pain: a single-blind, randomized controlled study. J Manipulative Physiol Ther 2023;46(4):191-200. [doi: 10.1016/j.jmpt.2024.02.006] [Medline: 38944805]
61. Freeman HR, Chander H, Kodithuwakku Arachchige SNK, Turner AJ, Jones JA, Pan Z, et al. Postural control behavior in a virtual moving room paradigm. Biomechanics 2023;3(4):539-551. [doi: 10.3390/biomechanics3040043]
62. Pau M, Arippa F, Porta M, Casu G, Leban B, Cocco E. Associations between perceived discomfort, postural control alterations and disability in people with multiple sclerosis exposed to immersive virtual reality. : IEEE; 2024 Presented at: IEEE Gaming, Entertainment, and Media Conference (GEM); 2024 June 5-7; Turin, Italy. [doi: 10.1109/gem61861.2024.10585386]

Abbreviations

CNP: chronic neck pain
FOE: figure of eight
HTR: head-tilt response
ICC: intraclass correlation coefficient
IMU: inertial measurement unit
JPE: joint position error
MDC: minimal detectable change
PS: postural sway
ROM: range of motion
SA: sway area
SEM: standard error of measurement
SMC: sensorimotor control
TPL: total path length
VR: virtual reality

Edited by S Munce; submitted 28.Nov.2025; peer-reviewed by A Otero, A Claus; comments to author 23.Mar.2026; accepted 05.May.2026; published 21.May.2026.

Please cite as:
Chiu YL, Lee PY, Lim PL, Cheng KC, Yin ZX, Tsai YJ
Validity and Reliability of an Immersive Virtual Reality System for Multidimensional Assessment of Cervical Sensorimotor Control: Cross-Sectional Study
JMIR Rehabil Assist Technol 2026;13:e88498
URL: https://rehab.jmir.org/2026/1/e88498
doi:10.2196/88498
PMID:

XSL•FO
RenderX

©Ya-Lan Chiu, Pei-Yun Lee, Pooi-Ling Lim, Kai-Chia Cheng, Zong-Xian Yin, Yi-Ju Tsai. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 21.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Measurement Properties of the European Portuguese Michigan Hand Outcomes Questionnaire (MHQ) in Paper and Electronic Formats: Instrument Validation Study

Eugénio Moita Gonçalves[1,2,3,4], MD; Maria Catarina Silva[2], MS; Henrique Ribeiro[1,4], OT; Mariana Almeida[1,4], MD; Bernardo Filipe[1,4], MD; Sandra Branco[1,4], OT; Filipa Henriques[1,4], OT; Lilibeth Campos[1,4], MD; Ana Campolargo[1,4], MD; Pedro Cantista[5], MD, PhD; Miguel Bernardes[6,7], MD, PhD; Cristina Jácome[2], PT, MSc, PhD

[1]Physical Medicine and Rehabilitation Department, Unidade Local de Saúde de Gaia e Espinho, Porto, Portugal, Porto, Portugal

[2]RISE-Health, Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine of University of Porto, Porto, Portugal, Porto, Porto, Portugal

[3]Physical Medicine and Rehabilitation Department, Hospital da Prelada, Porto, Portugal

[4]Clinical Academic Center Egas Moniz Health Alliance, Aveiro, Portugal

[5]Physical Medicine and Rehabilitation Department, Unidade Local de Saúde de Santo António, Porto, Portugal

[6]Rheumatology Department, Unidade Local de Saúde de São João, Porto, Portugal

[7]Department of Medicine, Faculty of Medicine of University of Porto, Porto, Portugal

Corresponding Author:
Eugénio Moita Gonçalves, MD
Physical Medicine and Rehabilitation Department, Hospital da Prelada
Rua de Sarmento de Beires, 153
Porto, 4250-449
Portugal
Phone: 351 22 833 0770
Email: eugenio.goncalves@hospitaldaprelada.pt

Abstract

Background: The Michigan Hand Outcomes Questionnaire (MHQ) is a key instrument for assessing hand and wrist conditions, but a validated version for European Portuguese speakers is unavailable. Furthermore, the potential of an electronic format to improve data collection and clinical workflows has not been formally evaluated.

Objective: This study aimed to translate and culturally adapt the MHQ for European Portuguese and to evaluate the measurement properties and comparability of its paper and electronic formats.

Methods: An observational study was conducted in three phases: (1) translation and expert review, (2) content validation through cognitive interviews, and (3) psychometric evaluation in a cohort of 81 adult patients with hand or wrist conditions. Exploratory known-groups analyses were performed according to clinical, sociodemographic, and digital-access variables. Analysis also included internal consistency (Cronbach α and McDonald ω), measurement error (standard error of measurement [SEM]), relative and absolute reliability (intraclass correlation coefficient [ICC], Spearman correlation coefficient [ρ], and Bland-Altman analysis). The electronic format's usability was assessed with the System Usability Scale (SUS).

Results: The translation and cognitive interviews confirmed the content validity and conceptual equivalence of the European Portuguese MHQ. The instrument demonstrated excellent internal consistency (Cronbach α=0.85-0.98; McDonald ω=0.88-0.99). Exploratory known-groups analyses showed significant gender-related differences and suggested side-specific clinical discrimination in fracture cases. While group-level relative reliability between formats was strong (ρ=0.68-0.91; ICC=0.76-0.95), individual-level absolute reliability was poor, with wide 95% limits of agreement spanning a 32.83-point interval. Format-related score discrepancies were most pronounced among older participants, those with lower educational attainment, and those with low digital literacy. The electronic MHQ was the preferred format (43 out of 74, 58%), showed acceptable usability (median SUS score 70, IQR 66-88), and had lower SEM in pain and function domains.

Conclusions: The European Portuguese MHQ is a relevant and reliable instrument with adequate measurement properties for both clinical practice and research. The electronic version demonstrated superior precision and user preference. However, maintaining the paper format as a viable alternative is essential to ensure equitable access for patients with digital barriers. Due

XSL•FO
RenderX

to the overall low absolute reliability between formats, once a patient-preferred format is adopted, it should be used consistently for longitudinal patient monitoring.

(JMIR Rehabil Assist Technol 2026;13:e83563) doi:10.2196/83563

KEYWORDS

Michigan Hand Outcomes Questionnaire; patient-reported outcome measures; digital technology; rehabilitation; psychometrics

Introduction

Hand conditions, whether acute or chronic, are impactful due to the essential role of the hand in activities of daily living (ADL) and skilled tasks [1]. Acute hand injuries account for approximately 30% of emergency department visits, with common injuries including tendon damage, fractures, degloving injuries, and amputations [2]. The global incidence rate of hand and wrist fractures is 179 per 100,000 people, and, in Portugal, it is significantly higher at 271.2 per 100,000 inhabitants [3]. In addition, a burn center study revealed that 41.5% of burn victims had hand injuries, often resulting in long-term functional limitations [4].

Beyond acute trauma, the hand is also a primary site for manifestations of systemic and chronic conditions. For instance, rheumatoid arthritis may affect the hand and wrist, with a reported impairment prevalence of 94% [5,6]. Other prevalent hand conditions include osteoarthritis, DeQuervain's tenosynovitis, carpal tunnel syndrome, trigger finger, and Dupuytren's contracture, each of which can also significantly impair hand function [7]. Given the hand's essential role in ADL and its high susceptibility to injury, clinical assessment must integrate physical examination findings with patient-reported outcomes. This dual approach prioritizes patient-centered functional goals and guides treatment strategies [4,8].

There is a growing emphasis on value-based, patient-centered care, where patient-reported outcome measures (PROMs) complement physical assessments. By capturing patients' functional abilities, symptoms, and quality of life, PROMs enhance shared clinical decision-making, treatment monitoring, patient adherence, and overall care effectiveness and experience [9-13]. The Michigan Hand Outcomes Questionnaire (MHQ) is a widely recognized PROM for its comprehensive evaluation of hand and wrist conditions, assessing function, pain, ADL, work performance, hand appearance, and patient satisfaction. Notably, it is the only hand assessment PROM that effectively integrates components of the International Classification of Functioning, Disability, and Health framework, making it a valuable resource for complete patient evaluation [14,15]. The MHQ has also robust measurement properties, including minimal clinically important difference (MCID) values, which range from 8 to 13, with subscale-specific values varying depending on the diagnosis and treatment strategy [16,17]. It is often used in clinical research and included in standardized outcome sets, having received provisional endorsement from Outcome Measures in Rheumatology (OMERACT) and adoption by the International Consortium for Health Outcomes Measurement (ICHOM) for hand and wrist conditions [18,19]. Despite being validated in multiple languages, a European Portuguese version is still needed. Such an adaptation would ensure cultural and linguistic accuracy for the Portuguese population, enable comprehensive assessment of hand outcomes following conservative or surgical treatment in health settings, and support participation in international multicenter research [14,20,21].

With the increasing integration of digital technology into daily life and clinical practice, there is an opportunity for the implementation of electronic PROMs [22,23]. To the best of our knowledge, an electronic format of MHQ has not yet been published in peer-reviewed literature. This format could overcome the limitations of a paper version [20,24], support MHQ integration into clinical workflows, and improve patient engagement.

Thus, our study aimed to translate, culturally adapt, and validate the MHQ for European Portuguese, comparing its paper and electronic formats while evaluating its measurement properties and comparability.

Methods

Study Design

An observational study was conducted between February and December 2024 at the Physical Medicine and Rehabilitation (PMR) Department of the Unidade Local de Saúde de Gaia e Espinho. The study was divided into three phases: phase 1 – translation and expert panel review, phase 2 – content validity through cognitive debriefing, and phase 3 – measurement properties and comparability between formats. The study design followed the research principles outlined in the Declaration of Helsinki, with detailed procedures described in the Ethical Considerations section [25]. The study was reported according to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [26].

Participants

Patients were enrolled only in phases 2 and 3 and were recruited by convenience sampling at our department. Eligible patients were aged 18 years or older, proficient in European Portuguese, and diagnosed with any hand and wrist conditions with functional impairment. We excluded participants who presented with hand conditions secondary to central nervous system disorders, as well as those with severe cognitive or visual impairments.

Phase 1: MHQ Translation and Expert Panel Review

The MHQ is a comprehensive PROM assessing each hand using 37 core items across six subscales, I to VI, respectively: I - overall hand function (5 items per hand; eg, "Overall, how well

XSL•FO
RenderX

did your right hand work?"), II - ADL (5 items per hand and 7 for two-handed activities; eg, "Carry grocery bags"), III - work performance (5 items; eg, "How often were you unable to do your work because of problems with your hand(s)/wrist(s)?"), IV - pain (5 items per hand; eg, "How often did the pain in your right hand/wrist interfere with your sleep?"), V - appearance (4 items per hand; eg, "I am satisfied with the appearance (look) of my right hand"), and VI - patient satisfaction with hand function (6 items per hand; eg, "Motion of the fingers in your right hand") [27]. Except for work performance, all subscales are evaluated separately for each hand [14,15]. Items are scored on a 1-5 Likert scale, summed, and converted to a 0-100 scale. Higher subscale scores indicate better outcomes, except for pain.

For unilateral conditions, the score of the affected hand is used. For bilateral conditions (eg, rheumatoid arthritis), the right- and left-hand scores are averaged. The total score is the average of all subscales, with the pain score reversed. At the end of the MHQ, a set of questions proposed by the original authors collects demographic information, including dominant hand, ethnicity, job-related details, gender, education level, and annual family income. The questionnaire takes approximately 15 minutes to complete and can be self-administered by patients or aided by health care professionals [20].

A license for academic research was obtained from the MHQ copyright holders to translate the questionnaire into European Portuguese and to develop an electronic format. Following international guidelines, the translation and cross-cultural adaptation process began with two independent translations conducted by native Portuguese-speaking PMR residents fluent in English (EMG and MA) [28,29]. The socioeconomic questions of the sociodemographic section were also adapted to ensure cultural relevance. The 2 translations were merged and disagreements identified. Both translations were reviewed item by item during an online meeting with a panel of experts. The expert panel included the 2 translators, a PMR physician (PC), a rheumatologist (MB), and a PROM expert (CJ). A synthesis of the discussion produced a preliminary Portuguese version. In accordance with the Beaton et al [28] methodology, the preliminary version was back-translated into English by two translators: a native Canadian English speaker and a native Portuguese speaker with professional fluency in English from over 25 years of living and working in English-speaking countries. Both back-translators were non–health care professionals and were unaware of the original MHQ. The expert panel reviewed and resolved discrepancies between the two back-translated versions. The final European Portuguese MHQ and the back-translated version were approved by the original authors.

Phase 2: Content Validity Through Cognitive Debriefing

Cognitive debriefing interviews were conducted with patients with hand conditions, following established qualitative research guidelines and the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) recommendations [30,31]. A structured interview guide was used, incorporating think-aloud techniques and probing questions to assess the MHQ relevance, comprehensiveness, and comprehensibility.

The interviews were conducted by two trained PMR residents (EMG or BF); both received specific training in cognitive debriefing techniques prior to the study to ensure procedural consistency. Participants were encouraged to clarify and confirm their interpretations, ensuring accurate data collection. The interviews were audio-recorded, and field notes were taken to capture nonverbal cues. Verbatim transcriptions were thematically analyzed by two researchers (EMG, male, and CS, female) to identify ambiguities and key issues [31]. An initial sample size of 8 to 10 patients was planned, but data collection continued until thematic saturation was reached. This was defined as the point at which no new themes or insights emerged in 2 subsequent interviews, and responses became redundant, occurring at the 10th interview [32].

Participants also reported their clinical diagnosis and provided sociodemographic information, including age, ethnicity, gender, educational level, annual family income, and dominant hand.

Following cognitive debriefing, patient feedback was reviewed by the expert panel to determine whether revisions were required. Suggested changes were considered only when they improved clarity without compromising the original construct. Final decisions were made by consensus after reviewing the interview findings and the corresponding original item content.

Phase 3: Measurement Properties and Comparability Between Formats

In line with COSMIN recommendations, at least 50 participants were recruited and randomly assigned to complete either the paper or electronic format first [29,33]. To mitigate recall bias, participants completed the alternate format within one week of the initial MHQ administration to ensure consistency in their reported clinical and functional status [34].

The electronic format was developed in accordance with the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) best practices [33,35]. The migration from the paper format involved specific design choices to optimize user experience while adhering to specific migration principles. The visual layout was modified from the paper grid layout to a vertical list in the electronic format, where each item was followed by its corresponding radio-button response options. All items belonging to a single subscale were presented together on one screen, a design intended to maintain context for the participant. The instrument's skip logic for the pain subscale was preserved using conditional logic; if a participant selected "Never" (*Nunca*) in response to the initial pain frequency question, the subsequent branched questions about pain characteristics were automatically hidden and could not be answered. The format is hosted on a secure, web-based platform (Jotform) compliant with the European General Data Protection Regulation (GDPR). It was designed to ensure accessibility, provide real-time scoring for each MHQ domain and the total score, and enable efficient data analysis. Indeed, upon completing the electronic format on an 11-inch tablet device, patients were redirected to an interface that uses OpenAI's ChatGPT-4o, an AI tool, to deliver personalized MHQ results

XSL•FO
RenderX

with a report. This report includes comparisons to normative data from US patients with and without hand conditions [36-38]. A demonstration video is available in Multimedia Appendix 1.

After completing the electronic format, participants assessed its usability using the European Portuguese version of the System Usability Scale (SUS) [39]. The SUS is a standard 10-item questionnaire that uses alternating positive and negative statements to avoid response bias. Examples of positive items include "I think that I would like to use this system frequently" and "I thought the system was easy to use," while negative items include statements such as "I found the system unnecessarily complex." SUS scores range from 0 to 100, with scores of 68 or above indicating acceptable [39]. Participants also indicated their format preference: paper, electronic, or no preference.

Participants also provided clinical and sociodemographic data, as described in phase 2, and reported their digital literacy. Digital literacy was evaluated using the text messaging and smartphone usage subscales from the validated European Portuguese version of the Media and Technology Usage and Attitudes Scale (MTUAS) [40,41]. This assessment consisted of 11 questions assessing activities such as sending and receiving text messages, browsing the web, and using mobile apps. Each item is rated on a 10-point Likert scale (1=never to 10=all the time), reflecting the frequency of use. The digital literacy score was determined by calculating the average of all item scores, yielding a final score ranging from 1 to 10.

Data Analysis

Participants' characteristics in phases 2 and 3, as well as MHQ scores, were described using descriptive statistics. Normality was assessed through the Kolmogorov-Smirnov test and histogram inspection. Normally distributed data were presented as mean (SD), while nonnormally distributed data were summarized using median and quartiles (Q1 and Q3).

Missing data were handled according to their type. Structural missing data, which arose from the MHQ's intended skip-pattern logic in the pain subscale (IV-A and IV-B), were preserved as valid nonresponses. Little's test indicated that sporadic missing data were not missing completely at random. Therefore, multiple imputation using the predictive mean matching method was performed [42].

Exploratory known-groups analyses were conducted to examine whether affected-hand MHQ scores differed across clinical (dominant-hand involvement), sociodemographic (age group, income level, educational attainment, and gender), and digital-access (digital literacy and usability) subgroups. Between-group comparisons were performed using Mann-Whitney *U* tests. To further assess comparability between administration formats, paired comparisons between paper and electronic scores were performed within each subgroup using Wilcoxon signed-rank tests.

Internal consistency was assessed using Cronbach (α>0.70 considered acceptable), and McDonald ω as a more robust measure [34,43].

Comparability between formats was evaluated using several methods. At the group level, relative reliability was assessed using the intraclass correlation coefficient (ICC) based on a 2-way random-effects, single-rater, absolute agreement model, with values interpreted as poor (<0.50), moderate (0.50-0.74), good (0.75-0.89), or excellent (≥0.90); and Spearman rank correlation coefficient (ρ) for association between format scores [44]. For individual-level agreement, the analysis of absolute reliability was conducted to visualize systematic bias and quantify random error, using the Bland-Altman method by calculating the mean difference and 95% limits of agreement (LoA) and through standard error of measurement (SEM), calculated separately for each format to estimate the margin of error for an individual's score. Participants who completed only one format were excluded from these analyses.

Format preferences were explored according to gender, age group, educational attainment, and digital literacy using Mann-Whitney *U* tests.

Statistical analyses were performed using IBM SPSS Statistics (version 29.0) and R (version 4.2.0), with a *P* value less than .05 considered statistically significant.

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki [25]. Ethical approval was granted by the Health Ethics Committee of the Unidade Local de Saúde de Gaia e Espinho (reference number 62/2023, addendum numbers 64/2024-1 and 187/2024). All participants received clear explanations of the study objectives and provided written informed consent prior to any data collection. To protect participant privacy and confidentiality, all collected data were strictly deidentified and stored securely. Participants did not receive any financial compensation for their involvement in this study.

Results

Phase 1: Translation and Expert Panel Review

The proposed adaptation maintains the original MHQ context while incorporating linguistic and cultural refinements to ensure clarity and accessibility. These modifications, informed by expert panel discussions, enhanced the questionnaire's relevance for Portuguese-speaking users, regardless of their educational background. For example, "gestures" (*gestos*) was chosen over "activities" (*atividades*) to better capture the nature of hand movements required for the tasks described in the questionnaire items related to the left and right hand, respectively. Similarly, a few response options were modified to reflect expressions commonly used in Portuguese, for instance, "very poor" (*muito pobre*) to "very bad" (*muito mal*). Additionally, the term "normal work" (*trabalho normal*) was translated as "usual work" (*trabalho habitual*), as it more accurately conveys the concept of routine professional and daily tasks in Portuguese. Some English terms also lacked a direct Portuguese equivalent that would be appropriate in the given context. For instance, "sensation (feeling)" was translated as "sensitivity to touch" (*sensibilidade [ao toque]*), since the literal Portuguese equivalent of "feeling" (*sentimento*) is primarily associated with emotions rather than tactile perception. Similarly, "look" was translated as "aspect" (*aspecto*), as the word "look" (*olhar*) in

XSL•FO
RenderX

Portuguese refers to the act of looking rather than physical appearance. Additionally, "pain level" (*nível de dor*) was translated as "pain intensity" (*intensidade da dor*), which is a more commonly known term among Portuguese-speaking patients. Adjustments were also made in the demographic section. Demographic categories, including ethnicity, gender, job-related questions, education, and income, were adapted to align with Portuguese cultural and statistical standards, using terms adopted from the Portuguese National Institute of Statistics.

Phase 2: Content Validity Through Cognitive Debriefing

A total of 12 patients participated in the cognitive debriefing phase (Table 1 and Figure 1), and interviews lasted between 20 and 40 minutes. Participants had a median age of 54, IQR 49-57 years, were predominantly female (8 out of 12, 67%), and mostly non-Hispanic White (11 out of 12, 92%). Most patients demonstrated a clear understanding of the questionnaire's content. However, adjustments were needed based on participants' suggestions. As highlighted in Table 2, some expressions raised doubts regarding interpretation. Regarding finger movement (subscales I and VI), 2 patients asked for help before answering, as not all fingers were equally affected by stiffness and immobility. The same issue applied to the term "sensibility," where patients were unsure which specific finger they should base their answer on, given that sensory perception varies across different digits. In subscale V, several patients noted that the response system was not as intuitive as in previous subscales, considering that the items in these tables were in the affirmative form and not in the interrogative form.

All findings were discussed with the expert panel. Because these are common challenges when completing standardized questionnaires, and a primary goal was to preserve the structural integrity of the original instrument as much as possible to ensure comparability, modifications were kept subtle and limited to minor wording adjustments, with no changes made to the original MHQ structure.

XSL•FO
RenderX

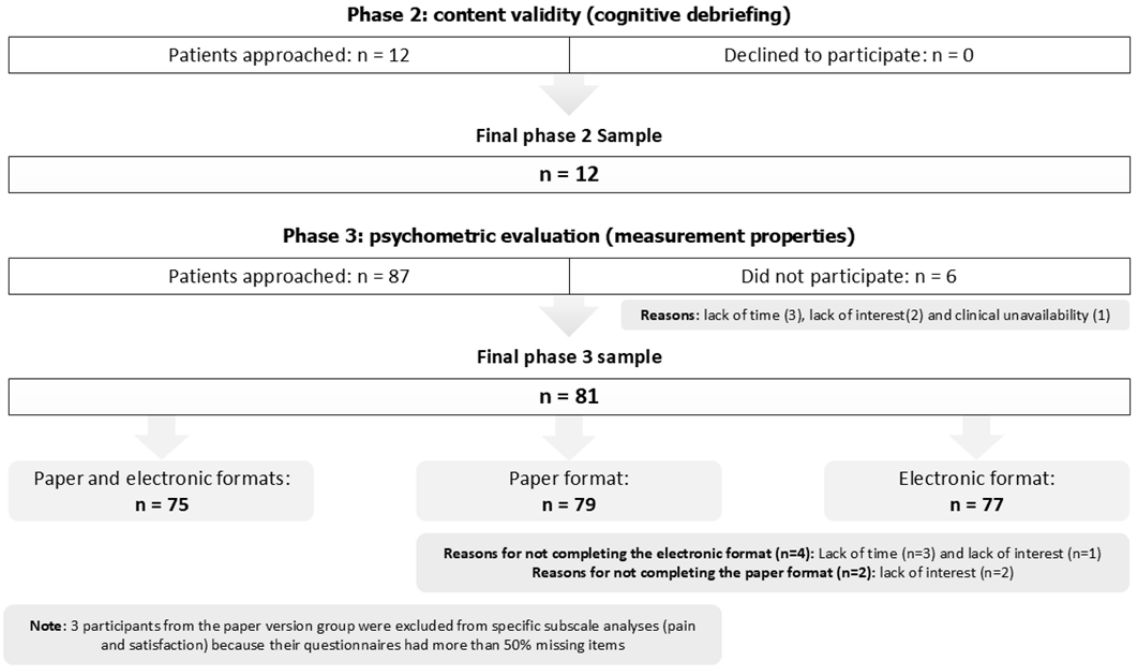
Table 1. Patients' sociodemographic and clinical data (N=12).

Variables	Values
Age (years), median (IQR)	54 (49-57)
Ethnicity, n (%)	
Non-Hispanic White	11 (91.7)
Mixed origin or heritage	1 (8.3)
Sex, n (%)	
Female	8 (66.7)
Male	4 (33.3)
Education level, n (%)	
Primary education (1st cycle)	2 (16.7)
Secondary education	5 (41.7)
Bachelor's degree	1 (8.3)
Licentiate degree	4 (33.3)
Annual family income (in Euros)[a], n (%)	
0 to 5000	4 (33.3)
5001 to 10,000	2 (16.7)
19,001 or above	5 (41.7)
Prefer not to answer	1 (8.3)
Dominant hand, n (%)	
Right-handed	12 (100)
Disability laterality, n (%)	
Right	6 (50)
Left	2 (16.7)
Bilateral	4 (33.3)
Diagnostic group, n (%)	
Fractures and dislocations	5 (41.7)
Peripheral neuropathies	2 (16.7)
Systemic disease	3 (25)
Degenerative joint disease	1 (8.3)
Muscle and/or tendon injury	1 (8.3)
Diagnosis, n (%)	
Distal radius fracture	3 (25)
Carpal tunnel syndrome	2 (16.7)
Metacarpal fracture	1 (8.3)
Rheumatoid arthritis	2 (16.7)
Carpal bone fracture	1 (8.3)
Tenosynovitis	1 (8.3)
Systemic sclerosis	1 (8.3)
Hand osteoarthritis	1 (8.3)

[a]The average exchange rate at the time of data collection was 1 EUR=US $1.08.

XSL•FO
RenderX

Figure 1. Flowchart outlines patient enrollment and retention across phase 2 (content validity) and phase 3 (psychometric evaluation). Phase 2 included all 12 approached patients. In phase 3, 81 of 87 approached patients participated, with reasons provided for the 6 nonparticipants. The final phase 3 sample (n=81) details Michigan Hand Outcomes Questionnaire format completion rates: 75 completed both formats, 79 completed the paper format, and 77 completed the electronic format. Additionally, the flowchart notes the exclusion of 3 participants from specific subscale analyses due to missing data.



XSL•FO
RenderX

Table 2. Examples of controversial expressions identified by patients during cognitive interviews.

MHQ[a] subscale	Original expression	Rationale	Revised expression
Activities of Daily Living Item II[b]-C4	"Carry a grocery bag"	Although there is a direct translation in Portuguese for "carry a grocery bag" (*carregar um saco de mercearia*), this expression does not reflect the need to use both hands as it was intended by the original item. Since most patients usually carry a grocery bag in one hand only, we considered more appropriate to replace it with "carry grocery bags" (*carregar sacos de mercearia*) to translate a bilateral activity.	"Carry grocery bags" (*Carregar sacos de mercearia*)
Activities of Daily Living Item II-C5	"Wash dishes"	It was not considered a common chore. The alternative "loading the dishwasher" (*colocar a loiça na máquina*) was suggested to be added. The expert panel considered that would involve a different hand activity and could alter the construct assessed by the original item.	No revision was made
Work Performance Subscale III	—[c]	Items 3, 4 and 5 were perceived as identical. However, no patient suggested a more suitable alternative.	No revision was made
Pain Items IV-A2 and IV-B2	"Please describe the pain you had in your hand/wrist" (*Por favor, descreva a dor que teve na sua mão/punho*)	There were doubts about whether the statement referred to the maximum pain the patient felt or the average. Furthermore, it was not understandable if these items were related to the pain while at rest or during activity. As the identified ambiguity is also present in the original instrument, the item was intentionally left unchanged.	No revision was made
Pain Items IV-A5 and IV-B5	"Unhappy"	As this word carries a substantial meaning, some patients suggested the alternative "sad" (*triste*).	"Sad" (*triste*)
Satisfaction Items VI-A1 and VI-B1	"Overall function"	The translated expression "overall function" (*função global*) was unclear to some participants, particularly older patients, who did not understand its meaning. The alternative "general function" (*função geral*) was suggested by one participant, but the participant also understood the wording "*função global*."	No revision was made

[a]Michigan Hand Outcomes Questionnaire.

[b]Roman numerals (I-VI) indicate Michigan Hand Outcomes Questionnaire (MHQ) subscales (eg, II=activities of daily living, IV=pain), letters denote the hand assessed (A=right, B=left, C=both), and numbers represent the question order within the subscale (eg, IV-A2=2nd Pain question for the right hand).

[c]Not applicable.

Phase 3: Measurement Properties and Comparability Between Formats

Participants' Characteristics

A total of 81 participants were included in this study phase. Their demographic and clinical characteristics are summarized in Table 3. The median age was 55 (45-67) years, with most participants being women (51 out of 81, 63%). Median digital literacy score was 5.45 (3-6.6), and almost half (39 out of 81, 48%) completed primary education. Right-hand dominance was high (76 out of 81, 94%), with disability affecting mostly the right (36 out of 81, 44%). Fractures and dislocations (42 out of 81, 52%), especially distal radius fractures (29 out of 81, 36%), were the most frequent conditions.

XSL•FO
RenderX

Table 3. Patients' demographic and clinical data (N=81).

Variables	Values[a]
Age (years), median (Q1-Q3)	55 (45-67)
Digital literacy, median (Q1-Q3)[b]	5.4 (3-6.6)
Ethnicity	
Non-Hispanic White	75 (92.6)
Non-Hispanic Black	2 (2.5)
Mixed origin or heritage	1 (1.2)
Prefer not to answer	1 (1.2)
Missing	2 (2.5)
Sex, n (%)	
Female	51 (63)
Male	28 (34.6)
Prefer not to answer	1 (1.2)
Missing	1 (1.2)
Education level, n (%)	
Primary education (first cycle)	20 (24.7)
Primary education (second cycle)	10 (12.3)
Primary education (third cycle)	9 (11.1)
Secondary education	19 (23.5)
Bachelor's degree	2 (2.5)
Licentiate degree	12 (14.8)
Master's degree	6 (7.4)
Missing	3 (3.7)
Annual family income (in Euros)[c], n (%)	
0 to 5000	11 (13.6)
5001 to 10,000	8 (9.9)
13,501 to 19,000	12 (14.8)
19,001 to 27,500	9 (11.1)
Prefer not to answer	21 (25.9)
Others	16 (19.8)
Missing	4 (4.9)
Dominant hand, n (%)	
Right-handed	76 (93.8)
Ambidextrous	1 (1.2)
Missing	4 (4.9)
Disability laterality, n (%)	
Right	36 (44.4)
Left	33 (40.7)
Bilateral	12 (14.8)
Diagnostic group[d], n (%)	
Fractures and dislocations	42 (51.9)
Peripheral neuropathies	12 (14.8)

XSL•FO
RenderX

Variables	Values[a]
Muscle and/or tendon injury	6 (7.4)
Systemic disease	8 (9.9)
Others	15 (18.5)
Diagnosis[d], n (%)	
Distal radius fracture	29 (35.8)
Carpal tunnel syndrome	10 (12.3)
Metacarpal fracture	7 (8.6)
Rheumatoid arthritis	5 (6.2)
Carpal bone fracture	4 (4.9)
Tenosynovitis	4 (4.9)
Dupuytren's disease	3 (3.7)
De-gloving injury	3 (3.7)
Others	24 (29.6)

[a]Values shown as n (%) unless otherwise indicated; n: absolute frequency; %: relative frequency; Q1-Q3: IQR.

[b]Digital literacy was calculated using the Media and Technology Usage and Attitudes Scale (MTUAS), with scores ranging from 1 to 10 points.

[c]The average exchange rate at the time of data collection was 1 EUR=US $1.08.

[d]Total percentage exceeds 100% as participants could have more than one diagnosis.

Three participants' paper questionnaires exhibited more than 50% of missing items in the pain (2) and satisfaction (1) subscales, which were excluded from specific subscale analysis. Isolated missing values were present in two paper questionnaires related to items II-B2 ("pick up a coin") and VI-A5 ("pain level of your right hand"). No missing values were observed in the electronic MHQ, as fields were mandatory by design.

MHQ Scores and Known-Group Analyses

The distribution of scores for the paper and electronic formats was highly similar across nearly all subscales, with comparable median scores and IQRs, although subgroup analyses showed selected format-related differences (Table 4).

Table 4. Descriptive statistics of European Portuguese Michigan Hand Outcomes Questionnaire (MHQ) scores in both formats.

MHQ subscales	Paper format, median (IQR)	Electronic format, median (IQR)
Overall hand function of the right hand	70 (50-90)	70 (50-90)
Overall hand function of the left hand	75 (45-100)	70 (50-95)
ADL[a] with the right hand	75 (50-100)	75 (55-100)
ADL with the left hand	75 (40-100)	75 (50-100)
ADL with both hands	60.7 (35.7-82.1)	57.1 (35.7-75)
Work performance	50 (35-70)	50 (25-67.5)
Pain in the right hand	22.5 (0-51.3)	25 (0-50)
Pain in the left hand	17.5 (0-45)	10 (0-45)
Appearance of the right hand	68.8 (37.5-100)	81.3 (50-96.9)
Appearance of the left hand	75 (31.3-100)	75 (43.8-100)
Satisfaction with the right hand	75 (45.8-92.7)	75 (39.6-100)
Satisfaction with the left hand	75 (36.5-100)	75 (39.6-100)
Total score of the right hand	68.9 (47.3-83.1)	69.8 (47.1-83.3)
Total score of the left hand	68.5 (48.6-80.9)	69.7 (49.5-84.1)
Total score of both hands	63.90 (53.7-77.4)	64 (54.4-76.1)

[a]ADL: activities of daily living.

Exploratory known-groups analyses showed a clinically interpretable pattern in the distal radius fracture subgroup. Dominant-hand involvement alone did not consistently differentiate MHQ scores. However, hand-specific analyses by

XSL•FO
RenderX

fracture side indicated side-specific clinical discrimination, particularly in the paper format, where total scores were significantly lower for the fractured hand than for the contralateral hand in both right- and left-sided fractures. This pattern was only partially reproduced in the electronic format, where the left-hand total score remained significantly lower among participants with left-sided fractures. Contralateral, uninjured-hand scores remained below the maximum possible value at the group level in both formats.

Among sociodemographic subgroups, gender showed the most consistent differences across formats, with men reporting higher scores in several MHQ domains, including function, ADL, work performance, satisfaction, and total score. Education-related differences were observed mainly in the paper format, where participants with lower educational attainment reported lower work performance, appearance, and total scores. Format-related differences were more frequent among older participants, those with lower educational attainment, lower digital literacy, and lower perceived usability, particularly for appearance and total scores.

Internal Consistency

As shown in Table 5, internal consistency was good to excellent across all subscales as demonstrated by Cronbach α values ranging from 0.85 to 0.98, while McDonald ω values ranged from .88 to .99. The highest internal consistency was observed in the "Satisfaction with the left hand" subscale for both coefficients.

Table 5. Internal consistency of the European Portuguese Michigan Hand Outcomes Questionnaire (MHQ) and reliability between formats.

MHQ subscales	Cronbach α	McDonald ω	ρ [a]	ICC[b] (95% CI)	SEM[c]	
					Paper format	Electronic format
Overall hand function of the right hand	0.95	0.97	0.68	0.80 (0.68-0.87)	2.29	1.28
Overall hand function of the left hand	0.96	0.97	0.89	0.93 (0.90-0.96)	2.31	1.12
ADL with the right hand	0.95	0.96	0.76	0.85 (0.76-0.90)	1.72	1.75
ADL with the left hand	0.96	0.98	0.91	0.95 (0.93-0.97)	2.71	2.97
ADL with both hands	0.95	0.97	0.87	0.92 (0.87-0.95)	4.05	2.95
Work performance	0.95	0.96	0.82	0.89 (0.83-0.93)	2.41	2.13
Pain in the right hand	0.94	0.99	0.72	0.85 (0.76-0.90)	3.88	2.21
Pain in the left hand	0.94	0.99	0.88	0.80 (0.69-0.88)	3.44	1.97
Appearance of the right hand	0.86	0.89	0.76	0.80 (0.66-0.88)	1.35	1.39
Appearance of the left hand	0.85	0.88	0.74	0.86 (0.78-0.91)	2.04	2.16
Satisfaction with the right hand	0.95	0.98	0.71	0.76 (0.62-0.85)	3.45	2.76
Satisfaction with the left hand	0.98	0.99	0.75	0.86 (0.78-0.91)	3.95	2.91
Total score of the right hand	__[d]	—	0.85	0.81 (0.71-0.88)	—	—
Total score of the left hand	—	—	0.89	0.91 (0.86-0.94)	—	—
Total score of both hands	—	—	0.83	0.88 (0.82-0.93)	—	—

[a]All Spearman rank correlation coefficients (ρ) were statistically significant (*P*<.01).

[b]ICC: intraclass correlation coefficient.

[c]SEM: standard error of measurement.

[d]Not applicable.

Reliability and Comparability Between Formats

The analysis of relative reliability between the paper and electronic formats revealed a good to excellent reliability, with ICC ranging from 0.76 to 0.95 (Table 5). These findings were supported by strong and statistically significant Spearman coefficients (ρ=0.68 to 0.91; *P*<.01).

In contrast, the absolute reliability analysis demonstrated slight measurement error at the individual level. While Bland-Altman analysis (Figures 2-4) confirmed negligible systematic bias, it revealed wide 95% LoA. For the total score of both hands, this interval spanned a substantial 32.83 points (from –18.26 to 14.57). Reinforcing the differences in measurement error, the analysis of the SEM showed the electronic version to be metrically superior. For instance, the SEM for pain in the right hand was 3.88 for the paper format compared to a much lower 2.21 for the electronic format (Table 5).

XSL•FO
RenderX

Figure 2. Bland-Altman plot comparing paper and electronic Michigan Hand Outcomes Questionnaire total score for the left hand. The mean difference is shown with 95% limits of agreement calculated as mean (SD 1.96).

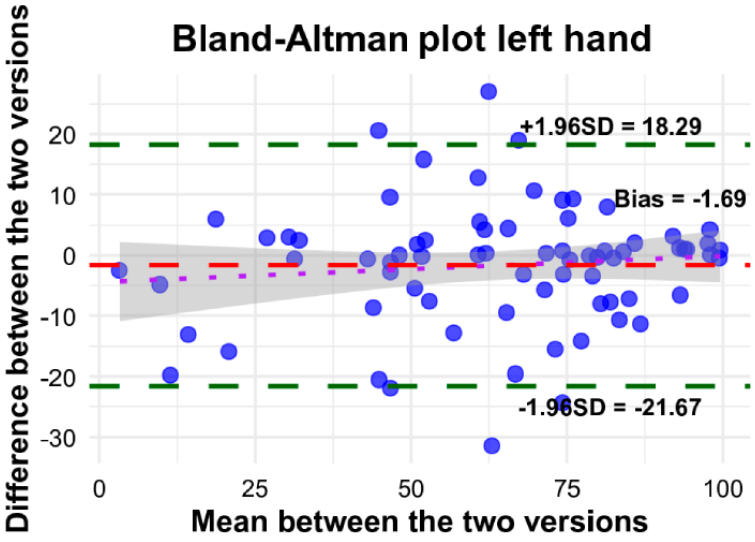
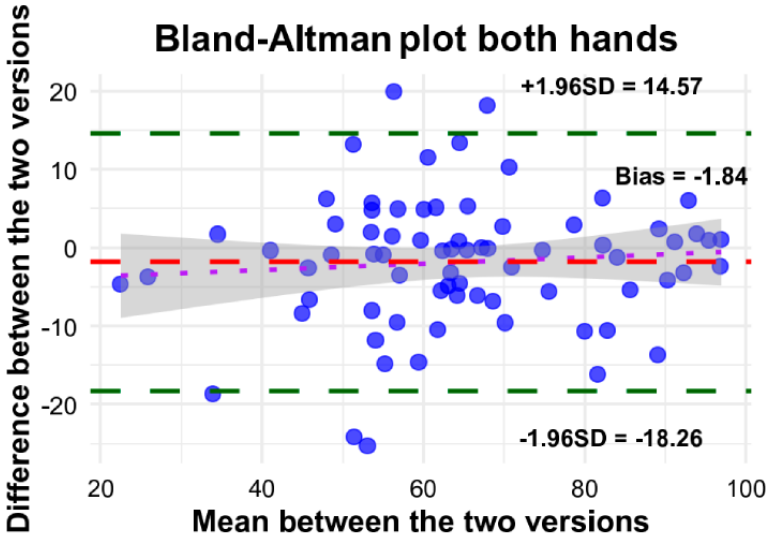


Figure 3. Bland-Altman plot comparing paper and electronic Michigan Hand Outcomes Questionnaire total score for the right hand. The mean difference is shown with 95% limits of agreement calculated as mean (SD 1.96).



XSL•FO
RenderX

Figure 4. Bland-Altman plot comparing paper and electronic Michigan Hand Outcomes Questionnaire total score for both hands. The mean difference is shown with 95% limits of agreement calculated as mean (SD 1.96).



Administration Preference and Usability

Most participants (43 out of 74, 58%) preferred the electronic format, whereas 15 (20%) preferred the paper version, and 16 (22%) were indifferent. This preference was more pronounced among individuals with higher digital literacy (24 out of 36, 67%) and higher education background (13 out of 18, 72%). It was also favored by less frequent technology users (19 out of 38, 50%) and those with less formal education (30 out of 56, 54%). Notably, more than half (12 out of 22, 55%) of participants aged 65 and older also favored the electronic format, closely tracking the preference seen in younger participants (31 out of 52, 60%). Similarly, gender did not significantly influence preference, with the electronic format being the top choice for both male (18 out of 27, 67%) and female (25 out of 47, 53%) participants. Regarding usability, this format scored a median of 70 (66-88) on the SUS.

Discussion

Principal Findings

This study translated and culturally adapted the MHQ for European Portuguese and conducted an evaluation of its measurement properties in both paper and electronic formats. The adaptation process, informed by expert review and patient cognitive interviews, yielded an instrument that was well-understood and relevant. Exploratory known-groups analyses provided partial evidence of clinical relevance. The subsequent psychometric evaluation revealed high internal consistency across all subscales in both formats and an overall adequate relative and absolute reliability between the paper and electronic formats supporting their use in clinical practice and research.

Content Validity and Clinical Relevance

The translation and cultural adaptation process confirmed the relevance and comprehensibility of the European Portuguese MHQ. Cognitive interviews showed that most participants understood the questionnaire, although some items generated uncertainty, particularly those involving abstract concepts such as “overall function,” hand appearance, and satisfaction. These findings are clinically important because they show that even when a PROM is linguistically accurate, some concepts may still be interpreted differently by patients depending on their condition, education, and lived experiences. Importantly, no structural changes were made to the original MHQ, preserving comparability with other validated versions [31].

Exploratory known-groups analyses provided evidence of clinical relevance. Gender showed consistent sociodemographic discrimination across both administration formats, with men generally reporting better function and lower pain burden than women. However, these findings should not be interpreted as reflecting a direct effect of gender alone, as they may also reflect differences in diagnosis, disease severity, occupational exposure, pain perception, treatment stage, or other unmeasured clinical factors [45].

In the distal-radius-fracture subgroup, affected-hand MHQ scores did not differ consistently according to dominance status, suggesting that dominance alone was not a major determinant of perceived disability. More clinically relevant was the observation that contralateral, uninjured-hand scores remained below the maximum possible score. This supports previous MHQ normative findings indicating that unilateral hand or wrist problems may also reduce perceived function of the unaffected hand, likely because many daily tasks require bilateral hand use or because impairment in one hand can increase demand on the other [46].

Therefore, contralateral scores after unilateral distal radius fracture should not necessarily be interpreted as fully “normal” and bilateral summary scores may dilute side-specific impairment. The absence of clear work-performance differences between right- and left-sided fracture groups further suggests that occupational impact may depend on factors beyond injury side or dominance alone, including task demands, compensatory strategies, pain, rehabilitation stage, and postinjury adaptation.

XSL•FO
RenderX

Reliability and Comparability Between Formats

This study provides evidence for the comparability between paper and electronic formats. The high Spearman correlations (ρ up to 0.91) and excellent ICC values (up to 0.95) indicate that, on average, the two formats produce highly comparable scores. This finding supports data collection via either format within a single study. While the formats are equivalent for groups, the Bland-Altman analysis showed they might not be interchangeable for individuals. The 95% LoA for the total score of both hands spanned a 32.83-point interval. This is a substantial degree of random error when compared to the MHQ's established MCID, which is 8 to 13 points [16]. For a patient, a score difference generated simply by switching from paper to electronic format could be 2 to 3 times larger than a genuine, clinically meaningful change. This finding strongly cautions against using the two formats interchangeably to monitor a single patient's progress. Format-related differences were more likely to appear among participants with lower educational attainment, lower digital literacy, lower perceived usability, older age, and female gender. The inherently subjective appearance domain was the most consistently affected by administration format, and total score differed between formats in selected subgroups [47].

Furthermore, the analysis of measurement precision via SEM suggests that the electronic version is not merely an alternative but, in some respects, a metrically superior instrument. The electronic format demonstrated a significantly lower SEM for the crucial domains of hand function and pain, meaning it has less inherent measurement "noise" and can therefore detect smaller true changes more reliably than the paper format.

Given its precision, lack of missing data, and patient preference, the electronic format appears suitable as the preferred administration format in many settings. Nevertheless, maintaining a paper alternative remains advisable to ensure access for patients with lower digital familiarity, lower perceived usability, or specific technological barriers. For clinical care involving longitudinal assessments, the two formats should never be used interchangeably for the same patient.

Strengths, Limitations, and Contribution to Practice

This study's strengths include its adherence to COSMIN and ISPOR guidelines, the direct comparison between paper and electronic formats, and a mixed methods approach that allowed qualitative findings to explain quantitative results. The development of a preferred electronic version with acceptable usability is a step toward modernizing data collection in Portuguese rehabilitation settings [35,48]. Patient preference for electronic MHQ was strongly influenced by age, education, and digital literacy. While younger, highly educated patients showed the most preference, acceptance among older individuals points to a shrinking digital divide. These findings highlight the need for user-friendly electronic tools alongside traditional paper forms to guarantee accessibility [37].

Nonetheless, limitations must be acknowledged. The use of a convenience sample from a single center may limit generalizability. Structural validity was not examined because of the modest sample size and multidimensional structure of the MHQ. Convergent and divergent validity were also not formally assessed because no external comparator instrument was administered. Both paper and electronic formats were completed in the presence of a health care professional, allowing clarification when needed. However, assistance during completion was not systematically recorded. Based on clinical observation, assistance appeared more frequent with the paper format, possibly because writing, turning pages, and handling the questionnaire may be challenging for patients with hand or wrist functional impairment [49]. Therefore, the frequency and potential influence of assistance on responses could not be quantified. The electronic format was tested only on an 11-inch tablet in a controlled clinical setting, and usability may differ in unsupervised or device-diverse contexts. Although mandatory response fields prevented missing data, they may have introduced response bias by forcing uncertain participants to select an answer. Digital literacy was assessed using selected MTUAS subscales as a proxy. While it is practical to reduce respondent burden, this approach primarily measures technology use frequency rather than directly evaluating digital skills, confidence, or the independent ability to complete an electronic platform [40]. High internal consistency coefficients in several subscales may indicate item redundancy, supporting future validation of a shorter European Portuguese MHQ [43,50,51].

Finally, in the Portuguese context, this work provides a valuable tool for patient-reported outcomes in hand and wrist conditions. While other instruments are available, the MHQ offers a more comprehensive evaluation, uniquely assessing each hand separately and including domains like appearance and satisfaction [52,53]. The European Portuguese MHQ is a relevant and reliable instrument with adequate measurement properties for both clinical practice and research. The electronic version demonstrated superior precision and user preference. However, maintaining the paper format as a viable alternative is essential to ensure equitable access for patients with digital barriers. Due to the overall low absolute reliability between formats, once a patient-preferred format is adopted, it should be used consistently for longitudinal patient monitoring. Future research should assess the instrument in larger and more diverse samples, particularly regarding structural validity, external validity, and the performance of the electronic format in routine clinical and remote-care settings.

Acknowledgments

The authors would like to thank Dr. Lilibeth Campos for her valuable support and contributions to this work.

XSL•FO
RenderX

Data Availability

The data that support the findings of this study are not openly available due to patient privacy and confidentiality.

Funding

This research was supported by grant awards from the Grünenthal Foundation and Egas Moniz Health Alliance.

Authors' Contributions

Conceptualization: EMG, MCS, CJ
Methodology: EMG, MCS, CJ, MB
Investigation: EMG, HR, SB, FH, LC
Formal analysis: EMG, MCS
Writing – original draft: EMG, MCS, SB, MB, CJ
Writing – review and editing: EMG, MCS, HR, MA, BF, SB, FH, LC, AC, PC, MB, CJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
This video file is a demonstration of the electronic version of the European Portuguese Michigan Hand Outcomes Questionnaire (MHQ). It shows the entire user process, from entering an ID and providing consent to answering the questions on the web-based platform. After completion, the video demonstrates the system's ability to provide instant, real-time scoring and an automatically generated, personalized report. This report, created with the assistance of an AI tool (ChatGPT-4o), summarizes the patient's results and compares them to normative data.
[MP4 File (MP4 Video), 142497 KB - rehab_v13i1e83563_app1.mp4]

References

1. O'Connell C, Guo M, Soucy B, Calder M, Sparks J, Plamondon S. All hands on deck: the multidisciplinary rehabilitation assessment and management of hand function in persons with neuromuscular disorders. Muscle Nerve 2025;71(5):869-888. [doi: 10.1002/mus.28167] [Medline: 38845187]
2. Moellhoff N, Throner V, Frank K, Benne A, Coenen M, Giunta RE, et al. Epidemiology of hand injuries that presented to a tertiary care facility in Germany: a study including 435 patients. Arch Orthop Trauma Surg 2023;143(3):1715-1724 [FREE Full text] [doi: 10.1007/s00402-022-04617-9] [Medline: 36138241]
3. Crowe CS, Massenburg BB, Morrison SD, Chang J, Friedrich JB, Abady GG, et al. Global trends of hand and wrist trauma: a systematic analysis of fracture and digit amputation using the global burden of disease 2017 study. Inj Prev 2020;26(Supp 1):i115-i124 [FREE Full text] [doi: 10.1136/injuryprev-2019-043495] [Medline: 32169973]
4. Mata-Ribeiro L, Vieira L, Vilela M. Epidemiology and outcome assessment of hand burns: a 3-year retrospective analysis in a burn unit. Ann Burns Fire Disasters 2022;35(1):18-25 [FREE Full text] [Medline: 35582085]
5. Sudoł-Szopińska I, Teh J, Cotten A. Rheumatoid hand and other hand-deforming rheumatic conditions. Semin Musculoskelet Radiol 2021;25(2):232-245. [doi: 10.1055/s-0041-1729150] [Medline: 34082449]
6. Horsten NCA, Ursum J, Roorda LD, van Schaardenburg D, Dekker J, Hoeksma AF. Prevalence of hand symptoms, impairments and activity limitations in rheumatoid arthritis in relation to disease duration. J Rehabil Med 2010;42(10):916-921 [FREE Full text] [doi: 10.2340/16501977-0619] [Medline: 21031287]
7. Currie KB, Tadisina KK, Mackinnon SE. Common hand conditions: a review. JAMA 2022;327(24):2434-2445. [doi: 10.1001/jama.2022.8481] [Medline: 35762992]
8. Mohammadi AA, Mohammadi S. Hand aesthetic, an annoying problem for the burn patients, but commonly overlooked issue by the burn surgeons. Burns 2017;43(5):1130-1131. [doi: 10.1016/j.burns.2017.03.003] [Medline: 28408144]
9. Bertakis KD, Azari R. Determinants and outcomes of patient-centered care. Patient Educ Couns 2011;85(1):46-52. [doi: 10.1016/j.pec.2010.08.001] [Medline: 20801601]
10. Porter ME. What is value in health care? N Engl J Med 2010;363(26):2477-2481. [doi: 10.1056/NEJMp1011024] [Medline: 21142528]
11. Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013;346:f167 [FREE Full text] [doi: 10.1136/bmj.f167] [Medline: 23358487]
12. Lee SJ, Kavanaugh A, Lenert L. Electronic and computer-generated patient questionnaires in standard care. Best Pract Res Clin Rheumatol 2007;21(4):637-647. [doi: 10.1016/j.berh.2007.02.001] [Medline: 17678825]
13. Meirte J, Hellemans N, Anthonissen M, Denteneer L, Maertens K, Moortgat P, et al. Benefits and disadvantages of electronic patient-reported outcome measures: systematic review. JMIR Perioper Med 2020;3(1):e15588 [FREE Full text] [doi: 10.2196/15588] [Medline: 33393920]

XSL•FO
RenderX

14. Martínez-Fernández MV, Sarabia-Cobo CM, Sánchez-Labraca N. Cross-cultural adaptation, reliability, validity and responsiveness of the Michigan hand outcomes questionnaire (MHQ-Sp) in Spain. J Orthop Surg Res 2024;19(1):256 [FREE Full text] [doi: 10.1186/s13018-024-04723-x] [Medline: 38649996]
15. Shauver MJ, Chung KC. The Michigan hand outcomes questionnaire after 15 years of field trial. Plast Reconstr Surg 2013;131(5):779e-787e [FREE Full text] [doi: 10.1097/PRS.0b013e3182865d83] [Medline: 23629117]
16. London DA, Stepan JG, Calfee RP. Determining the Michigan hand outcomes questionnaire minimal clinically important difference by means of three methods. Plast Reconstr Surg 2014;133(3):616-625. [doi: 10.1097/PRS.0000000000000034] [Medline: 24572853]
17. Shauver MJ, Chung KC. The minimal clinically important difference of the Michigan hand outcomes questionnaire. J Hand Surg Am 2009;34(3):509-514 [FREE Full text] [doi: 10.1016/j.jhsa.2008.11.001] [Medline: 19258150]
18. Wouters RM, Jobi-Odeneye AO, de la Torre A, Joseph A, ICHOM HandWrist Working Group, Hovius SER. A standard set for outcome measurement in patients with hand and wrist conditions: consensus by the international consortium for health outcomes measurement hand and wrist working group. J Hand Surg Am 2021;46(10):841-855.e7 [FREE Full text] [doi: 10.1016/j.jhsa.2021.06.004] [Medline: 34325941]
19. Kroon FPB, van der Heijde D, Maxwell LJ, Beaton DE, Abishek A, Berenbaum F, et al. Core outcome measurement instrument selection for physical function in hand osteoarthritis using the OMERACT Filter 2.1 process. Semin Arthritis Rheum 2021;51(6):1311-1319 [FREE Full text] [doi: 10.1016/j.semarthrit.2021.08.014] [Medline: 34538514]
20. Blomstrand J, Karlsson J, Fagevik Olsén M, Kjellby Wendt G. The Michigan hand outcomes questionnaire (MHQ-Swe) in patients with distal radius fractures-cross-cultural adaptation to Swedish, validation and reliability. J Orthop Surg Res 2021;16(1):442 [FREE Full text] [doi: 10.1186/s13018-021-02571-7] [Medline: 34233700]
21. Meireles SM, Natour J, Batista DA, Lopes M, Skare TL. Cross-cultural adaptation and validation of the Michigan hand outcomes questionnaire (MHQ) for Brazil: validation study. Sao Paulo Med J 2014;132(6):339-347 [FREE Full text] [doi: 10.1590/1516-3180-2014-1326701] [Medline: 25351754]
22. Number of smartphone users in Portugal from 2010 to 2050. Statista. URL: https://www.statista.com/statistics/567279/predicted-number-of-mobile-internet-users-in-portugal/ [accessed 2026-07-29]
23. Total population of Portugal from 1980 to 2031 (in millions). Statista. URL: https://www.statista.com/statistics/372162/total-population-of-portugal/ [accessed 2026-07-29]
24. Campbell N, Ali F, Finlay AY, Salek SS. Equivalence of electronic and paper-based patient-reported outcome measures. Qual Life Res 2015;24(8):1949-1961. [doi: 10.1007/s11136-015-0937-3] [Medline: 25702266]
25. World Medical Association. World medical association declaration of Helsinki. Ethical principles for medical research involving human subjects. Bull World Health Organ 2001;79(4):373-374 [FREE Full text] [Medline: 11357217]
26. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
27. Arwert H, Schut S, Boiten J, Vliet Vlieland T, Meesters J. Patient reported outcomes of hand function three years after stroke. Top Stroke Rehabil 2018;25(1):13-19. [doi: 10.1080/10749357.2017.1385232] [Medline: 29025365]
28. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
29. Mokkink BL, Prinsen AC, Patrick DL, Alonso J, Bouter LM, de VC. COSMIN study design checklist for patient-reported outcome measurement instruments. Amsterdam: Amsterdam Public Health Research Institute; 2019. URL: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf [accessed 2023-12-01]
30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
31. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res 2018;27(5):1159-1170 [FREE Full text] [doi: 10.1007/s11136-018-1829-0] [Medline: 29550964]
32. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
33. O'Donohoe P, Reasner DS, Kovacs SM, Byrom B, Eremenco S, Barsdorf AI, et al. Updated recommendations on evidence needed to support measurement comparability among modes of data collection for patient-reported outcome measures: a good practices report of an ISPOR task force. Value Health 2023;26(5):623-633 [FREE Full text] [doi: 10.1016/j.jval.2023.01.001] [Medline: 37121630]
34. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60(1):34-42. [doi: 10.1016/j.jclinepi.2006.03.012] [Medline: 17161752]
35. Mowlem FD, Elash CA, Dumais KM, Haenel E, O'Donohoe P, Olt J, Electronic Clinical Outcome Assessment Consortium. Best practices for the electronic implementation and migration of patient-reported outcome measures. Value Health 2024;27(1):79-94 [FREE Full text] [doi: 10.1016/j.jval.2023.10.007] [Medline: 37879401]

XSL•FO
RenderX

36. Nolte MT, Shauver MJ, Chung KC. Normative values of the Michigan hand outcomes questionnaire for patients with and without hand conditions. Plast Reconstr Surg 2017;140(3):425e-433e [FREE Full text] [doi: 10.1097/PRS.0000000000003581] [Medline: 28841615]
37. Aiyegbusi OL. Key methodological considerations for usability testing of electronic patient-reported outcome (ePRO) systems. Qual Life Res 2020;29(2):325-333 [FREE Full text] [doi: 10.1007/s11136-019-02329-z] [Medline: 31691202]
38. Web content accessibility guidelines (WCAG) 2.1. World Wide Web Consortium. URL: https://www.w3.org/TR/WCAG21/ [accessed 2026-07-29]
39. Martins AI, Rosa AF, Queirós A, Silva A, Rocha NP. European Portuguese validation of the system usability scale (SUS). Procedia Computer Science 2015;67:293-300.
40. Neves AL, Jácome C, Taveira-Gomes T, Pereira AM, Almeida R, Amaral R, Neves, et al. Determinants of the use of health and fitness mobile apps by patients with asthma: secondary analysis of observational studies. J Med Internet Res 2021;23(9):e25472 [FREE Full text] [doi: 10.2196/25472] [Medline: 34550077]
41. Costa J, Matos A, Salvador MC, Vale-Dias ML, Castilho P, Pinheiro MR, et al. MTUAS-PA: Adaptação da Escala de Uso e Atitudes face às TIC e aos Media Sociais para adultos portugueses. In: Diagnóstico e Avaliação Psicológica : atas do 10º Congresso da AIDAP-AIDEP. Lisbon, Portugal: Associação Iberoamericana de Diagnóstico e Avaliação Psicológica (AIDAP-AIDEP); 2018:118-139.
42. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Br J Psychol 2014;105(3):399-412. [doi: 10.1111/bjop.12046] [Medline: 24844115]
43. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011;2:53-55 [FREE Full text] [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
44. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
45. Mehta S, Karagiannopoulos C, Pepin M, Ballantyne B, Michlovitz S, MacDermid J. Distal radius fracture rehabilitation: clinical practice guidelines linked to the international classification of functioning, disability, and healt. J Orthop Sports Phys Ther 2024;54(9):CPG1-CPG78 [FREE Full text]
46. Iitsuka T, Kurumadani H, Inagaki Y, Ota H. Recovery in the symmetry of hand use after distal radius fracture. Hand Ther 2025;30(2):72-81. [doi: 10.1177/17589983251319030] [Medline: 39959705]
47. Jayaram M, Wang C, Wang L, Chung KC. Validating the Michigan hand outcomes questionnaire in patients with rheumatoid arthritis using rasch analysis. PLoS One 2021;16(7):e0254984 [FREE Full text] [doi: 10.1371/journal.pone.0254984] [Medline: 34293037]
48. Muehlhausen W, Doll H, Quadri N, Fordham B, O'Donohoe P, Dogar N, et al. Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013. Health Qual Life Outcomes 2015;13:167 [FREE Full text] [doi: 10.1186/s12955-015-0362-x] [Medline: 26446159]
49. Yaffe M, Goyal N, Kokmeyer D, Merrell GA. The use of an iPad to collect patient-reported functional outcome measures in hand surgery. Hand (N Y) 2015;10(3):522-528 [FREE Full text] [doi: 10.1007/s11552-014-9731-x] [Medline: 26330789]
50. Chung BT, Morris SF. Confirmatory factor analysis of the Michigan hand questionnaire. Ann Plast Surg 2015;74(2):176-181. [doi: 10.1097/SAP.0b013e3182956659] [Medline: 23797023]
51. Waljee JF, Kim HM, Burns PB, Chung KC. Development of a brief, 12-item version of the Michigan hand questionnaire. Plast Reconstr Surg 2011;128(1):208-220 [FREE Full text] [doi: 10.1097/PRS.0b013e318218fc51] [Medline: 21701336]
52. Wormald JCR, Geoghegan L, Sierakowski K, Price A, Peters M, Jain A, et al. Site-specific patient-reported outcome measures for hand conditions: systematic review of development and psychometric properties. Plast Reconstr Surg Glob Open 2019;7(5):e2256 [FREE Full text] [doi: 10.1097/GOX.0000000000002256] [Medline: 31333975]
53. Dacombe PJ, Amirfeyz R, Davis T. Patient-reported outcome measures for hand and wrist trauma: is there sufficient evidence of reliability, validity, and responsiveness? Hand (N Y) 2016;11(1):11-21 [FREE Full text] [doi: 10.1177/1558944715614855] [Medline: 27418884]

Abbreviations

ADL: activities of daily living
COSMIN: Consensus-based Standards for the selection of health Measurement Instruments
GDPR: General Data Protection Regulation
ICC: intraclass correlation coefficient
ICHOM: International Consortium for Health Outcomes Measurement
ISPOR: International Society for Pharmacoeconomics and Outcomes Research
LoA: limits of agreement
MCID: minimal clinically important difference
MHQ: Michigan Hand Outcomes Questionnaire
MTUAS: Media and Technology Usage and Attitudes Scale

XSL•FO
RenderX

OMERACT: Outcome Measures in Rheumatology
PMR: physical medicine and rehabilitation
PROM: patient-reported outcome measure
SEM: standard error of measurement
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
SUS: System Usability Scale

Edited by S Munce; submitted 04.Sep.2025; peer-reviewed by P Larionow, A Syahid; comments to author 05.May.2026; revised version received 15.May.2026; accepted 19.Jun.2026; published 30.Sep.2026.

Please cite as:
Moita Gonçalves E, Silva MC, Ribeiro H, Almeida M, Filipe B, Branco S, Henriques F, Campos L, Campolargo A, Cantista P, Bernardes M, Jácome C
Measurement Properties of the European Portuguese Michigan Hand Outcomes Questionnaire (MHQ) in Paper and Electronic Formats: Instrument Validation Study
JMIR Rehabil Assist Technol 2026;13:e83563
URL: https://rehab.jmir.org/2026/1/e83563
doi:10.2196/83563
PMID:42814995

©Eugénio Moita Gonçalves, Maria Catarina Silva, Henrique Ribeiro, Mariana Almeida, Bernardo Filipe, Sandra Branco, Filipa Henriques, Lilibeth Campos, Ana Campolargo, Pedro Cantista, Miguel Bernardes, Cristina Jácome. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Technologies in Postsurgical Orthopedic Rehabilitation: Scoping Review Mapping Interventions to the CAMS Framework (2020 2025)

Ngan Kuen Lai[1], PhD; Hai Huang[2], PhD; Tet Khuan Chen[1], PhD; Muhammad Nur Affendy Nor'a[1], MPhil; Jeffery Sijore[1], MSc; Wei Yang Chen[1], MSc; YongSheng Luo[2], MSc

[1]School of Computing, Asia Pacific University of Technology and Innovation (APU), Jalan Teknologi 5, Taman Teknologi Malaysia, Wilayah Persekutuan Kuala Lumpur, Malaysia

[2]School of Data Science, Zhuhai College of Science and Technology, Zhuhai, China

Corresponding Author:
Ngan Kuen Lai, PhD
School of Computing, Asia Pacific University of Technology and Innovation (APU), Jalan Teknologi 5, Taman Teknologi Malaysia, Wilayah Persekutuan Kuala Lumpur, Malaysia

Abstract

Background: Digital technologies have become common in postoperative orthopedic rehabilitation. The tendency toward this adoption appeared in the past decade. However, no standard method exists to classify their functional capabilities. This absence hinders practical implementation. Clinicians currently lack a structured framework to distinguish between the functional capabilities of different digital rehabilitation technologies effectively.

Objective: This scoping review aims to map digital interventions used in postsurgical orthopedic rehabilitation to a proposed Coordination, Alignment, Motivation, and Strength (CAMS) framework and to identify research trends, functional coverage, and gaps in current evidence.

Methods: A systematic search was conducted across 5 databases: MEDLINE (via PubMed), Scopus, IEEE Xplore, Cochrane Central Register of Controlled Trials (CENTRAL), and the Physiotherapy Evidence Database (PEDro), following PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. Studies were mapped to the proposed CAMS framework pillars. Multiple validated appraisal tools were applied descriptively according to study design: version 2 of the Cochrane risk-of-bias tool for randomized trials (RoB 2; n=32), the Newcastle-Ottawa Scale (NOS) for observational studies (n=2), and the Methodological Index for Non-Randomized Studies (MINORS) for nonrandomized designs (n=1).

Results: A total of 35 studies met the inclusion criteria. Coordination was the most frequently addressed pillar (29/35, 83%), followed by Motivation (26/35, 74%). Alignment was addressed in 37% (13/35) of studies, while Strength was the least represented pillar (12/35, 34%). Overall, 49% (17/35) of studies were rated “low risk” overall on the primary appraisal tool, 46% (16/35) demonstrated “some concerns,” and 6% (2/35) were rated “high risk,” with performance bias identified as a consistent methodological characteristic across the corpus. Delivery models were predominantly hybrid (14/35, 40%) or home-based (13/35, 37%), with 8 studies (8/35, 23%) using an exclusively clinical model.

Conclusions: Digital health has shown a notable increase in the kinematic monitoring feature, offering effective solutions to movement quality (Coordination and Alignment) and psychological engagement (Motivation). Nevertheless, there is a consistent gap in strength-related studies, as current decentralized tools lack the capability to quantify physiological load. The next phase of innovations should be a shift to integrated 4-pillar solutions that would coordinate resistance-based training with high-precision biomechanical monitoring. The CAMS framework is proposed as an exploratory classification tool to guide future research design and technology evaluation in postsurgical digital rehabilitation.

(JMIR Rehabil Assist Technol 2026;13:e100548) doi:10.2196/100548

KEYWORDS

telerehabilitation; orthopedic surgery; CAMS framework; virtual reality; AI; digital rehabilitation; Coordination, Alignment, Motivation, and Strength framework

XSL•FO
RenderX

Introduction

The Paradigm Shift Toward Decentralized Care

Traditionally, musculoskeletal rehabilitation has been localized in traditional ambulatory clinics, where it is hand-supervised and facilitated by a clinician. Nevertheless, the increasing clinical expenses, geographic access limitations, and the need for customized care have dictated a shift toward decentralized telerehabilitation. Digital health interventions have ceased to be alternative solutions, but they have already become a part of the continuum of care, helping to transform hospital-based acute recovery into independent, home-based maintenance [1]. Although this type of therapy has been rapidly adopted, there is still no standard approach to assessing the functional breadth of these tools.

Immersive Technologies: The Coordination-Motivation Axis

Gamification has transformed the engagement of patients by integrating virtual reality (VR) and augmented reality (AR). Immersive VR environments have been associated with reduced pain perception and perceived exertion during physical activity in the included literature [2], though this review's descriptive design does not establish a causal or clinical analgesic effect. AR systems, on the contrary, improve the rehabilitation experience through superimposing digital instructions directly onto the physical surroundings of the patient [3]. These technologies guide users through accurate movement paths by giving spatial-sensitive and real-time instructions. The tools are particularly efficient in the Coordination and Motivation pillars of the Coordination, Alignment, Motivation, and Strength (CAMS) structure, providing instant visual feedback required when learning a neuromuscular motor skill [4].

Markerless AI and Computer Vision: Ensuring Alignment

The invention of the markerless pose estimation using cameras has transformed alignment tracking. Contrary to wearable sensors, whose use can be affected by sensor drift or even being moved around during intense activity, AI-based computer vision measures the angles of joints and postural balance through normal smartphone cameras [5]. The technology plays a vital role in detecting compensatory movements, that is, patients use secondary muscular groups to avoid pain or weakness, thus providing the clinical quality of the rehabilitation setting in home-based unsupervised environments [6].

Robotics and the "Strength Gap"

The Strength pillar uses robotics and haptics to monitor movement quality via computer vision. Assist-as-needed strategies are used by the robotic devices to facilitate adaptive plasticity and functional recovery [7]. The high price and mechanical nature of these tools, however, tend to limit their application to the clinical environment and introduce a major technological gap in the realm of home-based strength training [8].

The CAMS Framework: A Proposed Taxonomy

Although there are numerous digital tools that provide generic activity tracking, they tend to overlook the subtle issues of clinical recovery. The CAMS framework is proposed in this review as an exploratory conceptual taxonomy to classify digital rehabilitation technologies according to the primary functional domains they address. The framework is not presented as a validated or standardized clinical tool; rather, it serves as a structured lens to support evidence mapping and gap identification. The classification of interventions based on these pillars helps us to understand whether the device just measures the quantity (eg, the number of repetitions) or the quality of recovery by means of biomechanical alignment and neuromuscular coordination. The framework positions motivation as the primary driver for adherence, ensuring that digital movement is viewed as a holistic clinical event rather than an isolated mechanical task. The rationale for selecting these 4 pillars reflects the authors' analysis of the digital rehabilitation literature: CAMS represent the functional domains most commonly and directly targetable by current digital health technologies. Other clinically relevant domains, including pain, range of motion, balance, gait function, and return to sport, are treated as downstream composite outcomes that emerge from the interplay of these 4 primary domains, rather than being directly addressed by a single technology class. This framing is the authors' conceptual contribution and is subject to empirical validation.

Methods

Study Design and Protocol Registration

This scoping review aimed to address the following question:

1. What types of digital technologies are used in postsurgical orthopedic rehabilitation?
2. How can they be classified across functional domains using the CAMS framework?
3. What research gaps exist in the functional coverage of current digital rehabilitation tools?

This study was conducted as a scoping review following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines [9]. A scoping review methodology was selected due to the heterogeneity of intervention types, outcome measures, and study designs, and the aim to map the breadth and characteristics of existing digital rehabilitation technologies rather than to evaluate their clinical effectiveness or synthesize outcome data. The protocol in the review was prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO) with a registration number of CRD420251062131.

Search Strategy and Information Sources

A comprehensive systematic search was conducted across 5 electronic databases, namely MEDLINE (via PubMed), Scopus, IEEE Xplore, Cochrane Central Register of Controlled Trials (CENTRAL) and Physiotherapy Evidence Database (PEDro). Web of Science was not accessible through the authors'

XSL•FO
RenderX

institutional subscriptions and was therefore not included in the search; this is acknowledged as a limitation.

The initial systematic search, conducted on December 15, 2025, identified 1078 records across the 5 databases, supplemented by 55 additional records identified through bidirectional citation tracking (Google Scholar snowballing), yielding 1133 records in total. Duplicate records were identified and removed using EndNote (Clarivate), supplemented by manual verification, reducing the total to an estimated 818 records for title and abstract screening. Searches covered literature published between January 2020 and December 2025, restricted to English-language publications. The search strategy used MeSH and free-text terms related to digital rehabilitation, orthopedic surgery, and postoperative care. The complete search strings for all 5 databases were as follows.

1. MEDLINE (via PubMed): ("telerehabilitation"[MeSH Terms] OR "Telemedicine"[MeSH Terms] OR "Mobile Applications"[MeSH Terms] OR "virtual reality"[MeSH Terms] OR "virtual rehabilitation"[tiab] OR "wearable electronic devices"[MeSH Terms] OR "augmented reality"[tiab] OR "mHealth"[tiab] OR "mobile health"[tiab] OR "digital rehabilitation"[tiab] OR "digital health"[tiab] OR "smartphone"[tiab]) AND ("orthopedic procedures"[MeSH Terms] OR "Orthopedics"[MeSH Terms] OR "arthroplasty, replacement, knee"[MeSH Terms] OR "arthroplasty, replacement, hip"[MeSH Terms] OR "anterior cruciate ligament reconstruction"[MeSH Terms] OR "rotator cuff"[MeSH Terms] OR "orthopedic rehabilitation"[tiab] OR "post-surgical rehabilitation"[tiab] OR "orthopedic surgery"[tiab]) AND ("postoperative care"[MeSH Terms] OR "rehabilitation"[MeSH Terms] OR "recovery of function"[MeSH Terms]). Filters: 2020 - 2025, English language, Human subjects.
2. Scopus: TITLE-ABS-KEY(("telerehabilitation" OR "mobile health" OR "virtual reality" OR "wearable electronic device*" OR "augmented reality" OR "mHealth" OR "digital rehabilitation") AND ("orthopedic procedure*" OR "knee arthroplasty" OR "hip arthroplasty" OR "anterior cruciate ligament reconstruction" OR "rotator cuff" OR "orthopedic rehabilitation" OR "post-surgical rehabilitation") AND ("postoperative care" OR "rehabilitation" OR "recovery of function")) AND PUBYEAR>2019 AND PUBYEAR<2026 AND LANGUAGE(english).
3. IEEE Xplore: ("telerehabilitation" OR "mobile health" OR "virtual reality" OR "wearable device" OR "augmented reality" OR "mHealth" OR "digital rehabilitation") AND ("orthopedic" OR "knee arthroplasty" OR "hip arthroplasty" OR "ACL reconstruction" OR "rotator cuff" OR "orthopedic rehabilitation") AND ("postoperative" OR "rehabilitation" OR "recovery of function"), limited to 2020 - 2025.
4. Cochrane CENTRAL: #1 telerehabilitation OR "mobile health" OR "virtual reality" OR "wearable electronic devices" OR "augmented reality" OR mHealth OR "digital rehabilitation"; #2 "orthopedic procedures" OR "knee arthroplasty" OR "hip arthroplasty" OR "anterior cruciate ligament reconstruction" OR "rotator cuff" OR "orthopedic rehabilitation"; #3 "postoperative care" OR rehabilitation OR "recovery of function"; #1 AND #2 AND #3, limited to 2020 - 2025.
5. PEDro: Advanced search conducted using the following filters: Body part=knee, hip, shoulder; Subdiscipline=orthopedics; Method=clinical trial; combined with free-text terms "telerehabilitation," "digital rehabilitation," "wearable," and "virtual reality" in the title or abstract field, limited to 2020 - 2025.

The search string provided above represents a reconstruction of the search logic applied, developed to document the conceptual scope of the original search (technology type AND orthopedic procedure AND rehabilitation or recovery outcome) for reproducibility purposes. Minor variation in exact result counts between the original search and this reconstruction may reflect index growth in MEDLINE (via PubMed) since the original search date, as well as evolving MeSH term indexing.

The CAMS framework was not used as a search term, as it is a novel taxonomy proposed in this review; rather, CAMS domain concepts informed post hoc mapping of retrieved studies. Google Scholar was used to perform bidirectional citation tracking (forward and backward snowballing), yielding 55 additional records (40 backward and 15 forward), which are documented as a separate identification source in the PRISMA-ScR flow diagram ().

Eligibility Criteria

Studies were included based on the population-concept-context (PCC) framework recommended for scoping reviews:

- Population: Adults undergoing postsurgical orthopedic rehabilitation following procedures such as anterior cruciate ligament (ACL) reconstruction, total knee or hip arthroplasty, or rotator cuff repair.
- Concept: Digital health technologies used as primary or adjunct rehabilitation tools, including VR, AR, AI-based computer vision, wearable sensors, or mobile health (mHealth) applications, and their functional classification across the CAMS domains.
- Context: Primary peer-reviewed experimental studies, including randomized controlled trials (RCTs), feasibility RCTs, prospective cohort studies, and nonrandomized controlled trials with an active comparator (which may include usual or standard care), published in English between January 2020 and December 2025, conducted in clinical, hybrid, or home-based settings. The full inclusion and exclusion criteria are presented in Textbox 1.

Studies were excluded if they involved nonsurgical populations, healthy participants, neurorehabilitation (eg, stroke), purely observational designs without an active digital intervention, or secondary research (reviews and meta-analyses). Only peer-reviewed articles published in English were included. A total of 3 studies identified during full-text review were subsequently excluded: 1 published in a nonindexed journal, 1 available only as an unpublished preprint, and 1 identified as a duplicate publication of an already-included trial. These exclusions are documented in the PRISMA-ScR flow diagram.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for a scoping review of digital technologies in postsurgical orthopedic rehabilitation, covering studies published between January 2020 and December 2025. (Population-concept-context framework applied per Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews guidance.)

Inclusion Criteria

- Adults (≥18 years) undergoing postsurgical orthopedic rehabilitation following anterior cruciate ligament (ACL) reconstruction, total knee or hip arthroplasty, or rotator cuff repair
- Active digital health technology as primary or adjunct intervention (virtual reality, augmented reality, AI computer vision, wearable sensors, or mHealth application)
- Peer-reviewed experimental studies: randomized controlled trials (RCTs), feasibility RCTs, prospective cohort studies, or nonrandomized controlled trials with an active comparator (which may include usual or standard care)
- Published in English, January 2020 to December 2025
- Conducted in clinical, hybrid, or home-based settings
- Reported ≥1 functional rehabilitation outcome measurable by the digital intervention

Exclusion Criteria

- Nonsurgical or healthy participant populations
- Neurorehabilitation populations (eg, stroke and traumatic brain injury)
- Passive or diagnostic-only technologies without an active rehabilitation component
- Secondary research (systematic reviews, meta-analyses, and narrative reviews)
- Conference abstracts, editorials, or nonpeer-reviewed grey literature
- Unpublished preprints
- Non-English language publications

Study Selection and Data Charting

Titles and abstracts were independently screened by 2 reviewers (WYC and YSL), followed by full-text assessment of potentially eligible studies. Discrepancies were resolved through discussion and consensus.

Data charting was conducted using a standardized form developed and pilot-tested by the research team prior to full extraction. Charting was performed independently by 2 reviewers (WYC and YSL), with discrepancies resolved through consensus.

The following variables were extracted: study design, patient population, type of digital intervention, rehabilitation phase, rehabilitation setting (clinical, home-based, or hybrid), anatomical focus (upper limb, lower limb, or full body), and reported outcome measures. Interventions were subsequently categorized according to the proposed CAMS framework for structured analysis.

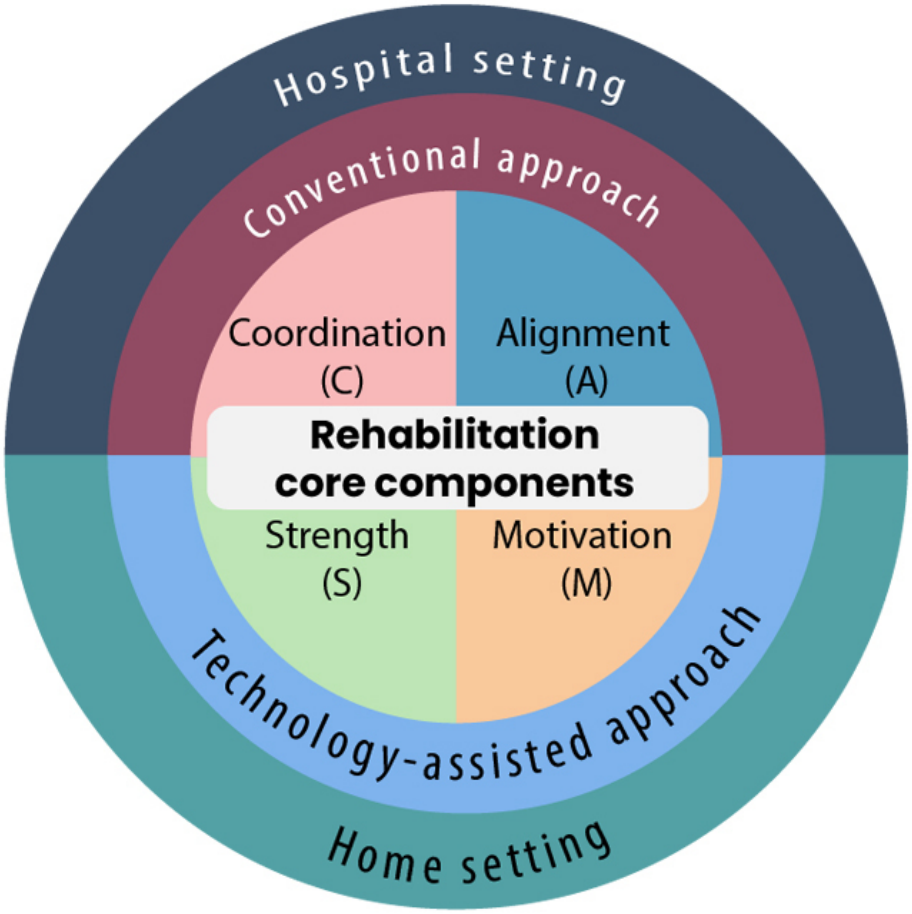
CAMS Framework Classification

To characterize the clinical focal point of each technology, interventions were mapped to the 4 pillars of the CAMS framework (Figure 1):

- Coordination: focus on neuromuscular control, movement sequencing, and timing.
- Alignment: focus on postural integrity and the mitigation of compensatory movement patterns.
- Motivation: focus on patient engagement, adherence, and psychological involvement.
- Strength: focus on mechanical loading, resistance training, and force production.

XSL•FO
RenderX

Figure 1. The Coordination-Alignment-Motivation-Strength (CAMS) framework was applied to classify digital rehabilitation technologies in postsurgical orthopedic rehabilitation, based on 35 studies [1,3,10-42] published between 2020 and 2025.



Interventions were permitted to map to multiple CAMS domains where the digital tool addressed more than 1 functional construct (eg, an AR system providing joint angle biofeedback [Alignment] while tracking repetition speed [Coordination]). CAMS domain classification followed the same independent dual-reviewer process described above, using predefined operational coding criteria. Each reviewer (WYC and YSL) independently assigned binary codes (present or absent) to each of the 4 CAMS pillars for every included study. Overall percentage agreement across all pillar-study combinations (n=140 binary decisions: 35 studies × 4 pillars) was 94.3%, with discrepancies resolved through discussion and adjudicated by a third senior reviewer (NKL). Given the high prevalence of the Coordination pillar (29/35, 83%), Cohen kappa was not reported due to the well-documented prevalence-bias paradox [43]. The CAMS framework was applied post hoc after study selection was complete and was not used as an eligibility criterion at any stage of screening or selection. Coding in Table 1 is based on the primary intervention design and reported functional outcomes of each study; an X indicates the digital tool explicitly addressed that CAMS domain as a primary functional target.

The CAMS framework is proposed as an initial conceptual classification model to organize and interpret the functional domains targeted by digital rehabilitation technologies. As a newly proposed framework, CAMS has not yet undergone formal validation and is presented here as an exploratory tool for structured analysis. The CAMS pillars were selected because they represent key modifiable domains commonly targeted by digital rehabilitation technologies. Other clinically relevant outcomes such as pain, balance, and quality of life were intentionally treated as downstream or composite outcomes influenced by these 4 primary domains rather than independent intervention targets. It is not intended as a validated evaluative tool but rather as a structured lens to support synthesis and identify gaps in this research. Future work should focus on empirical validation of the framework across clinical and technological contexts.

Methodological Quality Appraisal

Methodological quality appraisal is optional in scoping reviews and was conducted here to characterize the methodological profile of the included studies, not as a basis for exclusion. Given the heterogeneity of study designs in the final corpus, multiple appraisal tools were applied according to study design: the Cochrane Risk of Bias 2 (RoB 2) tool was applied to all randomized controlled trials (n=32) [1,3,10-22,24-30,32-38,40-42] including feasibility RCTs across 5 domains (randomization process, deviations from intended interventions, missing outcome data, outcome measurement, and selection of reported results); the Newcastle-Ottawa Scale (NOS) was applied to prospective cohort and retrospective observational studies (n=2) [23,31]; and the Methodological Index for Non-Randomized Studies (MINORS) was applied to quasi-experimental and nonrandomized controlled trials (n=1)

XSL•FO
RenderX

[39]. PEDro scale scores were additionally reported to enable comparability with prior physiotherapy rehabilitation literature. Two reviewers (WYC and YSL) independently appraised each study, with discrepancies resolved by the same third senior (NKL) reviewer noted above. Appraisal results are summarized descriptively in Table 2 and were not used to weight or exclude studies.

Data Synthesis and Analysis

A narrative synthesis was conducted because of a high level of heterogeneity in digital modalities, intervention protocols, and outcome measures. The synthesis of findings was based on a thematic mapping approach that is congruent with the CAMS framework. Descriptive statistics were used to summarize the use of digital tools in terms of setting and anatomical targets. Temporal trends and heatmaps were used to depict how the digital orthopedic environment would evolve and identify some technological gaps in the time frame from 2020 to 2025.

Results

Overview

The initial systematic search identified 1078 records across the 5 databases, supplemented by 55 additional records identified through citation tracking, yielding 1133 records in total. The total number of records after the elimination of duplicates was an estimated 818 to be subjected to screening in relation to title and abstract. After this step, an estimated 692 records were eliminated because they failed to meet the main inclusion criteria. Among the estimated 126 reports that had been requested to be retrieved, 5 were unavailable in full text. As a result, an estimated 121 reports were mapped as being eligible.

After a rigorous full-text review, an estimated 86 studies were excluded: 32 involved nonsurgical cohorts, 26 used only diagnostic or passive technologies, 25 did not report any active functional rehabilitation outcome measurable by the digital intervention (eg, range of motion, gait parameters, adherence, strength, or patient-reported functional outcomes), and 3 were excluded following post hoc eligibility verification (1 published in a nonindexed journal; 1 an unpublished preprint; 1 duplicate publication). In total, 35 studies [1,3,10-42] met all inclusion criteria and were included in the final synthesis, representing an inclusion rate of 3.1% (35/1133). This rate is consistent with and higher than comparable reviews in the digital rehabilitation field: a systematic review of digital home rehabilitation for stroke survivors reported 0.13% (13/9873) [44], and a scoping review of AI-driven virtual rehabilitation reported 0.38% (13/3458) [45]. The CAMS framework was applied post hoc after study selection was complete and was not used as an eligibility criterion at any stage of the screening or selection process. The PRISMA-ScR flow diagram (Figure 2) describes the full selection process.

XSL•FO
RenderX

Figure 2. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flow diagram of study identification, screening, and inclusion for a scoping review of digital technologies in postsurgical orthopedic rehabilitation, 2020 - 2025. CAMS: Coordination, Alignment, Motivation, and Strength.

XSL•FO
RenderX

Study Selection and Characteristics

The reviewed cohort is a pool of evidence that is specialized on the orthopedic postsurgical populations applying active digital rehabilitation interventions. Table 1 represents an overall outline of the studies included and their mapping to the pillars of CAMS.

XSL•FO
RenderX

Table . Anatomical target, delivery setting, technology type, and CAMS pillar mapping among 35 studies of digital postsurgical orthopedic rehabilitation technologies published between 2020 and 2025. Note: "X" indicates the digital tool explicitly addressed the corresponding CAMS domain as a primary functional target, based on independent dual-reviewer coding (94.3% overall percentage agreement; see Methods). Row numbers (No.) are sequential table identifiers and do not correspond to reference list numbers; full citations for each study are provided in the reference list, numbered in order of first citation in the text.

No.	Year	Ref. no.	Anatomical target	Setting	Technology type	C[a]	A[b]	M[c]	S[d]
1	2025	Chen et al [10]	Lower limb (knee)	Home	Cloud prescription	✓	—[e]	✓	X
2	2025	Dajpratham et al [11]	Lower limb (hip fracture)	Hybrid	Teleresistance	✓	—	✓	✓
3	2025	O'Reilly et al [12]	Upper limb (shoulder)	Hybrid	Telerehabilitation	✓	—	✓	—
4	2025	Sadiq et al [13]	Lower limb (knee)	Hybrid	Web-based+sensor	✓	✓	✓	✓
5	2025	Yuan et al [14]	Lower limb (ACL[f])	Hybrid	Telerehabilitation	✓	—	✓	✓
6	2025	Zhou et al [15]	Lower limb (hip)	Hybrid	Mobile app	✓	—	✓	—
7	2024	Christiansen et al [16]	Lower limb (knee)	Hybrid	Telehealth+sensor	✓	—	✓	—
8	2024	Lim et al [3]	Lower limb (ACL)	Home	AR[g]	✓	✓	✓	✓
9	2024	Xie et al [17]	Lower limb (knee)	Clinical	Intelligent monitor	✓	✓	—	—
10	2024	Zhao et al [18]	Lower limb (knee)	Home	Smartphone app	—	—	✓	✓
11	2024	Shin et al [19]	Lower limb (knee)	Clinical	Immersive VR[h]	✓	✓	—	—
12	2023	Guo et al [20]	Lower limb (ACL)	Home	mHealth[i] app	✓	—	✓	—
13	2023	Duong et al [21]	Lower limb (knee)	Hybrid	Digital interventions	✓	—	✓	—
14	2023	Höher et al [22]	Lower limb (knee)	Home	Sensor-based system	—	—	✓	—
15	2023	Lebleu et al [23]	Lower limb (knee)	Hybrid	Digital rehabilitation	✓	—	✓	—
16	2023	Osterloh et al [24]	Lower limb (hip or knee)	Hybrid	Digital group rehab	✓	—	✓	✓
17	2023	Shim et al [25]	Upper limb (shoulder)	Hybrid	Digital health system	✓	✓	✓	—
18	2023	Shim et al [26]	Lower limb (knee)	Hybrid	AR Digital health care	✓	✓	—	✓
19	2023	Gsangya et al [27]	Lower limb (ACL)	Clinical	Immersive VR	✓	✓	—	✓
20	2023	Gür and Başar [28]	Lower limb (knee)	Clinical	VR	—	—	✓	—
21	2023	Nuevo et al [29]	Lower limb (knee)	Hybrid	Telerehab platform (IMU[j])	✓	—	✓	✓

XSL•FO
RenderX

No.	Year	Ref. no.	Anatomical target	Setting	Technology type	C[a]	A[b]	M[c]	S[d]
22	2023	Mojica et al [30]	Lower limb (meniscectomy)	Home	Video telerehab	✓	—	—	—
23	2022	Hong et al [31]	Lower limb (knee)	Home	Digital platform	✓	—	✓	—
24	2022	Marley et al [32]	Upper limb (shoulder)	Home	Gamification or sensors	✓	✓	✓	—
25	2022	Pournajaf et al [33]	Lower limb (knee)	Clinical	VR Serious games	✓	✓	—	—
26	2022	Fuchs et al [34]	Lower limb (knee)	Clinical	VR	—	—	✓	—
27	2022	Zavala-González [35]	Lower limb (hip)	Clinical	VR (Wii balance)	✓	✓	—	—
28	2022	Fascio et al [36]	Lower limb (hip)	Home	Tablet+wearable	✓	—	✓	—
29	2022	Şahi˙n et al [37]	Lower limb (knee)	Home	SMS or call reminders	—	—	✓	—
30	2021	Dunphy et al [38]	Lower limb (ACL)	Hybrid	TRAK-ACL[k] app	✓	—	✓	—
31	2020	Prvu Bettger et al [1]	Lower limb (knee)	Hybrid	Virtual therapy	✓	✓	✓	—
32	2020	Winjen et al [39]	Lower limb (hip)	Home	Tablet app or coaching	✓	—	✓	—
33	2020	Bell et al [40]	Lower limb (knee)	Home	Wearable motion tracking	✓	—	—	—
34	2020	Gianola et al [41]	Lower limb (knee)	Clinical	VR (inpatient)	—	✓	—	✓
35	2020	Nelson et al [42]	Lower limb (hip)	Home	Telerehab app (iPad; Apple Inc.)	✓	✓	✓	✓

[a]C: coordination.

[b]A: alignment.

[c]M: motivation.

[d]S: strength.

[e]Not applicable.

[f]ACL: anterior cruciate ligament.

[g]AR: augmented reality

[h]VR: virtual reality.

[i]mHealth: mobile health.

[j]IMU: inertial measurement unit.

[k]TRAK: Taxonomy for the Rehabilitation of Knee Conditions.

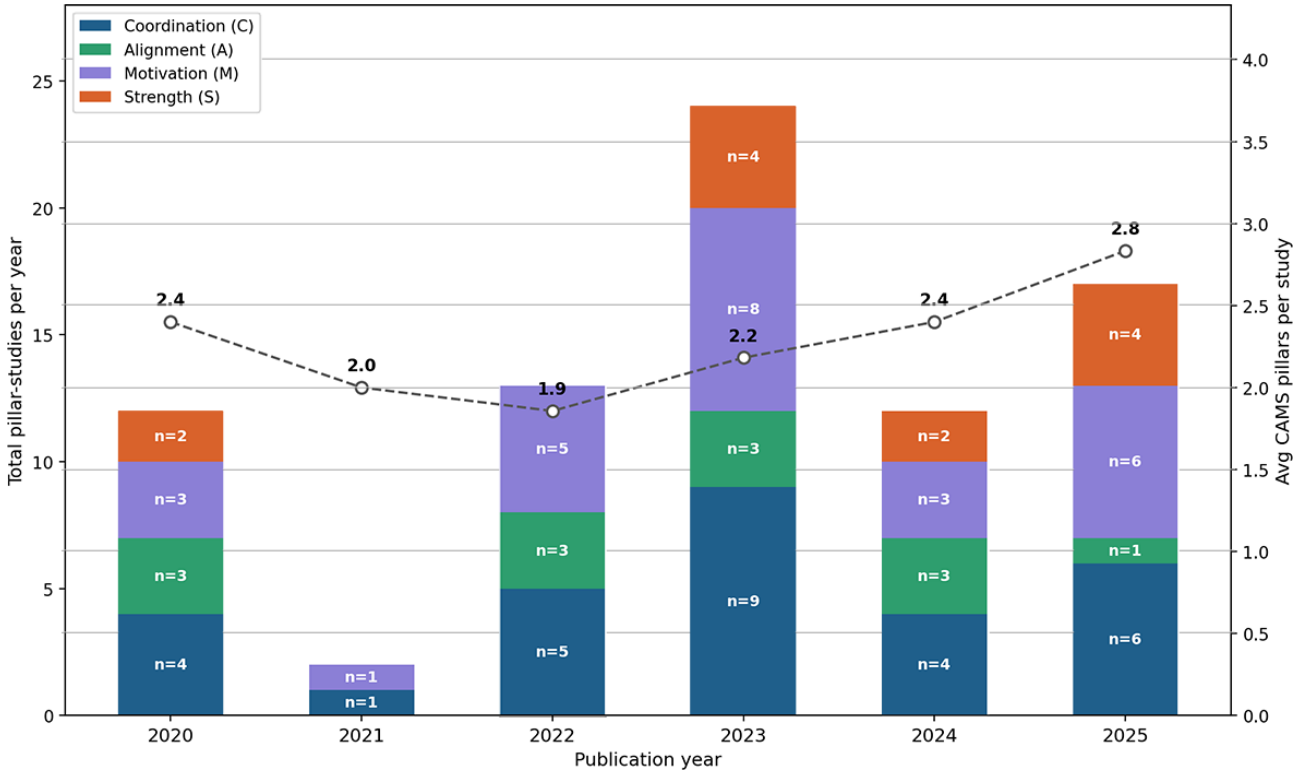
The included studies represent a diverse geographical distribution, with significant contributions from Asia (eg, [24,25,31]), Europe (eg, [1,32,38]), and North America (eg, [10,39]). The participant cohorts were predominantly postarthroplasty (total knee arthroplasty [TKA] and total hip arthroplasty [THA]) patients (eg, [15,20,21]), with sample sizes ranging from small-scale pilot feasibility trials to large-scale multicenter randomized controlled trials (n=300, [1]). In terms of technological deployment, wearable inertial measurement units (IMUs) and mobile health (mHealth) applications were the most frequently used modalities, reflecting a clinical priority for portable and scalable solutions.

The temporal distribution of included studies shows a gradual increase in publication volume across the review period (Figures 3 and 4): 5 studies [1,39-42] in 2020, 1 in 2021, 7 in 2022, 11 in 2023, 5 in 2024, and 6 in 2025, with 2023 - 2025 collectively accounting for 63% (22/35) of the corpus. This pattern suggests growing research activity in the postpandemic period, though

XSL•FO
RenderX

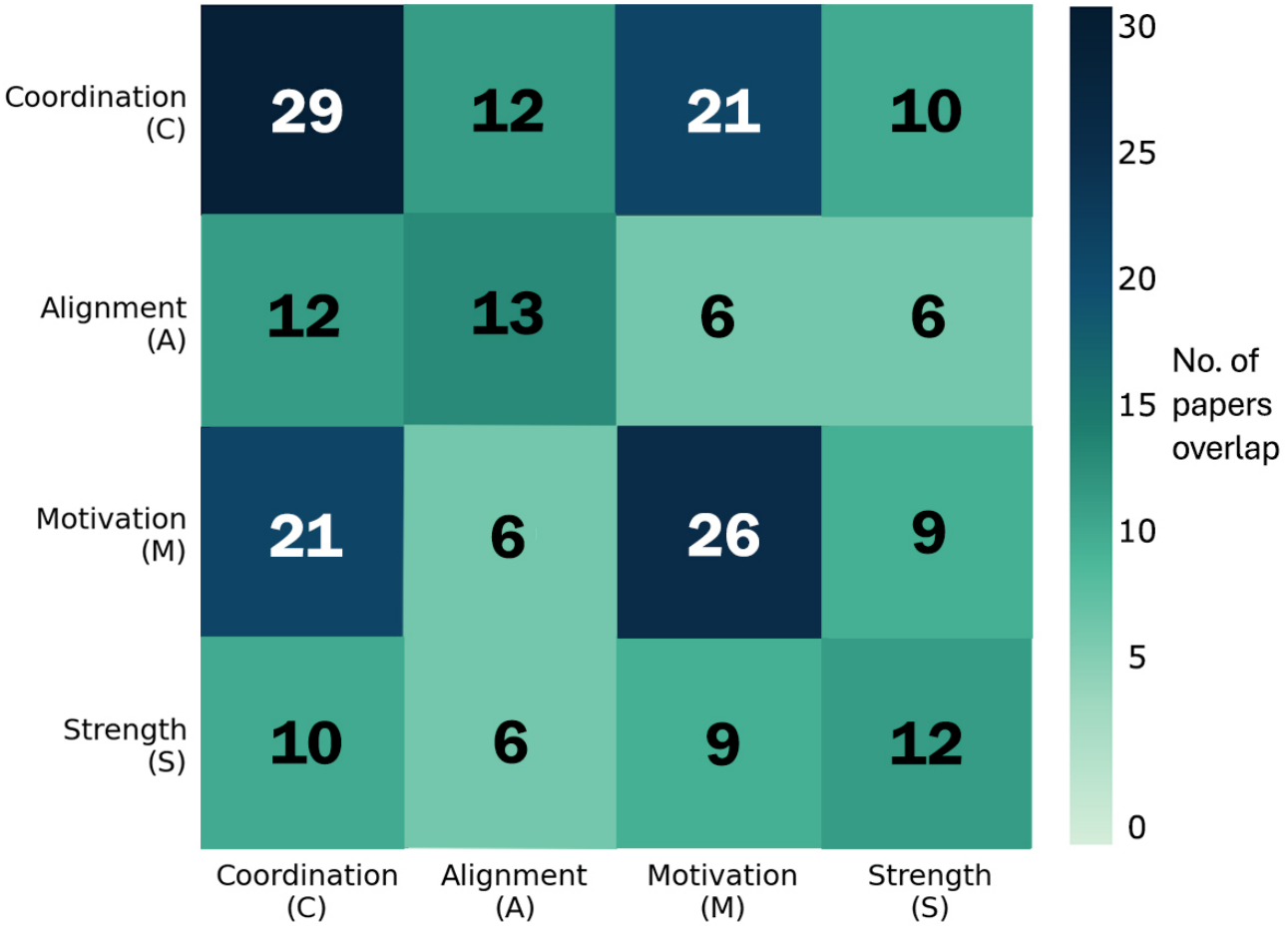
the limited number of studies per year precludes formal trend analysis.

Figure 3. Longitudinal evolution and thematic density of Coordination-Alignment-Motivation-Strength (CAMS) pillars among 35 studies [1,3,10-42] of digital postsurgical orthopedic rehabilitation technologies published between 2020 and 2025. Avg: average.



XSL•FO
RenderX

Figure 4. Heatmap of synergy and overlap between Coordination-Alignment-Motivation-Strength (CAMS) pillars among 35 studies [1,3,10-42] of digital postsurgical orthopedic rehabilitation technologies published between 2020 and 2025.



Methodological Quality and RoB

The methodological profile of the included studies was characterized using design-appropriate appraisal tools (RoB 2 for RCTs, NOS for observational studies, MINORS for nonrandomized designs). As detailed in Table 2, 49% (17/35) of the studies were rated low risk overall on their respective primary appraisal tool, 46% (16/35) demonstrated some concerns, and 6% (2/35) were rated high risk. The temporal distribution of RoB 2 ratings (Table 3) shows that studies published between 2024 and 2025 (eg, [12,16]) more commonly achieved low risk overall ratings compared to earlier pilot studies, though the small number of studies per year means this observation is descriptive rather than inferential.

XSL•FO
RenderX

Table . Methodological quality appraisal of 35 studies [1,3,10-42] of digital postsurgical orthopedic rehabilitation technologies published between 2020 and 2025, assessed using the Cochrane risk of bias 2 tool (n=32 randomized trials) [1,3,10-22,24-30,32-38,40-42], the Newcastle-Ottawa Scale (n=2 observational studies) [23,31], and the Methodological Index for Non-Randomized Studies (n=1 nonrandomized designs) [39].

No.	Reference no.	Design	Tool	D1[a]	D2[b]	D3[c]	D4[d]	D5[e]	Overall	PEDro[f] score	Key methodological characteristic
1	Chen et al [10]	RCT[g]	RoB2[h]	L[i]	L	L	L	L	L	8	—[j]
2	Dajpratham et al [11]	RCT	RoB2	L	L	L	L	L	L	8	—
3	O'Reilly et al [12]	RCT	RoB2	L	L	L	L	L	L	8	—
4	Sadiq et al [13]	RCT	RoB2	L	S[k]	L	L	L	L	9	Lack of participant or therapist blinding
5	Yuan et al [14]	RCT	RoB2	L	L	L	L	L	L	8	—
6	Zhou et al [15]	RCT	RoB2	L	S	L	S	L	S	7	Reliance on subjective PROMs[l] (HHS[m])
7	Christiansen et al [16]	RCT	RoB2	L	L	L	L	L	L	8	—
8	Lim et al [3]	RCT	RoB2	L	L	L	L	L	L	9	—
9	Xie et al [17]	RCT	RoB2	L	S	L	L	L	L	7	Inherent lack of participant or therapist blinding
10	Zhao et al [18]	RCT	RoB2	S	S	S	S	L	S	6	Unclear allocation concealment and pilot sample size
11	Shin and Hong [19]	RCT	RoB2	S	L	L	L	L	S	—	Randomization sequence detail not specified
12	Guo et al [20]	RCT	RoB2	L	L	L	L	L	L	8	High attrition rate in intervention group
13	Duong et al [21]	RCT	RoB2	L	L	L	L	L	L	8	—

XSL•FO
RenderX

No.	Reference no.	Design	Tool	D1[a]	D2[b]	D3[c]	D4[d]	D5[e]	Overall	PEDro[f] score	Key methodological characteristic
14	Höher et al [22]	RCT	RoB2	L	S	L	S	L	S	6	Monitoring focus with subjective outcomes
15	Lebleu et al [23]	Cohort	NOS	—	—	—	—	—	L	8	Multicenter cohort; no randomization
16	Osterloh et al [24]	RCT	RoB2	L	L	L	S	L	L	7	Use of self-reported functional outcome
17	Shim et al [25]	RCT	RoB2	L	L	L	L	L	L	8	—
18	Shim et al [26]	RCT	RoB2	S	S	S	L	S	S	—	Abstract-only evidence; domains largely unclear
19	Gsangaya et al [27]	RCT	RoB2	S	L	L	S	L	S	—	Unblinded investigators and subjects for subjective outcomes
20	Gür et al [28]	RCT	RoB2	L	H[n]	S	S	H	H	—	Confirmed selective reporting concern (independent SR[o] verification)
21	Nuevo et al [29]	RCT	RoB2	L	S	L	L	L	S	—	Unblinded treating clinicians with dashboard access
22	Mojica et al [30]	RCT	RoB2	S	S	S	S	L	S	—	Randomization or blinding detail limited
23	Hong et al [31]	Retrospective	NOS[p]	—	—	—	—	—	S	7	Reliance on patient-reported WOMAC[q] scores

XSL•FO
RenderX

No.	Reference no.	Design	Tool	D1[a]	D2[b]	D3[c]	D4[d]	D5[e]	Overall	PEDro[f] score	Key methodological characteristic
24	Marley et al [32]	RCT	RoB2	L	S	L	S	L	S	7	Detection bias in gamified feedback delivery
25	Pournajaf et al [33]	RCT	RoB2	S	L	S	L	L	S	—	Randomization sequence or attrition detail limited
26	Fuchs et al [34]	RCT	RoB2	H	S	S	S	L	H	—	Quasi-random (time-based) allocation; unblinded subjective outcomes
27	Zavala-González et al [35]	RCT	RoB2	L	L	L	L	L	L	—	—
28	Fascio et al [36]	RCT	RoB2	S	S	S	S	S	S	—	Abstract-only evidence; domains largely unclear
29	Şahi˙n et al [37]	RCT	RoB2	S	S	S	S	L	S	—	Abstract-only evidence; domains largely unclear
30	Dunphy et al [38]	Feasibility RCT	RoB2	L	L	L	L	L	L	8	—
31	Prvu Bettger et al [1]	RCT	RoB2	L	L	L	L	L	L	8	—
32	Wijnen et al [39]	Non-RCT	MINORS[r]	—	—	—	—	—	S	7	Nonrandomized; historical controls
33	Bell et al [40]	RCT	RoB2	L	S	S	S	L	S	—	Unblinded outcome assessor; substantial pre-baseline attrition
34	Gianola et al [41]	RCT	RoB2	L	L	L	L	L	L	—	—

XSL•FO
RenderX

No.	Reference no.	Design	Tool	D1[a]	D2[b]	D3[c]	D4[d]	D5[e]	Overall	PEDro[f] score	Key methodological characteristic
35	Nelson et al [42]	RCT	RoB2	S	L	S	L	L	S	—	Allocation concealment method not specified

[a]D1: randomization process.

[b]D2: deviations from intended interventions.

[c]D3: missing outcome data.

[d]D4: measurement of the outcome.

[e]D5: selection of the reported result.

[f]PEDro: Physiotherapy Evidence Database.

[g]RCT: randomized controlled trial.

[h]RoB2: risk of bias 2 tool.

[i]L: low risk.

[j]Nonapplicable domain (Newcastle-Ottawa Scale or Methodological Index for Non-Randomized Studies tools do not use Risk of Bias 2 domain structure) or Physiotherapy Evidence Database score not independently verified.

[k]S: some concerns.

[l]PROM: patient-reported outcome measure.

[m]HHS: Harris Hip Score.

[n]H: high risk.

[o]SR: systematic review.

[p]NOS: Newcastle-Ottawa Scale.

[q]WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

[r]MINORS: Methodological Index for Non-Randomized Studies.

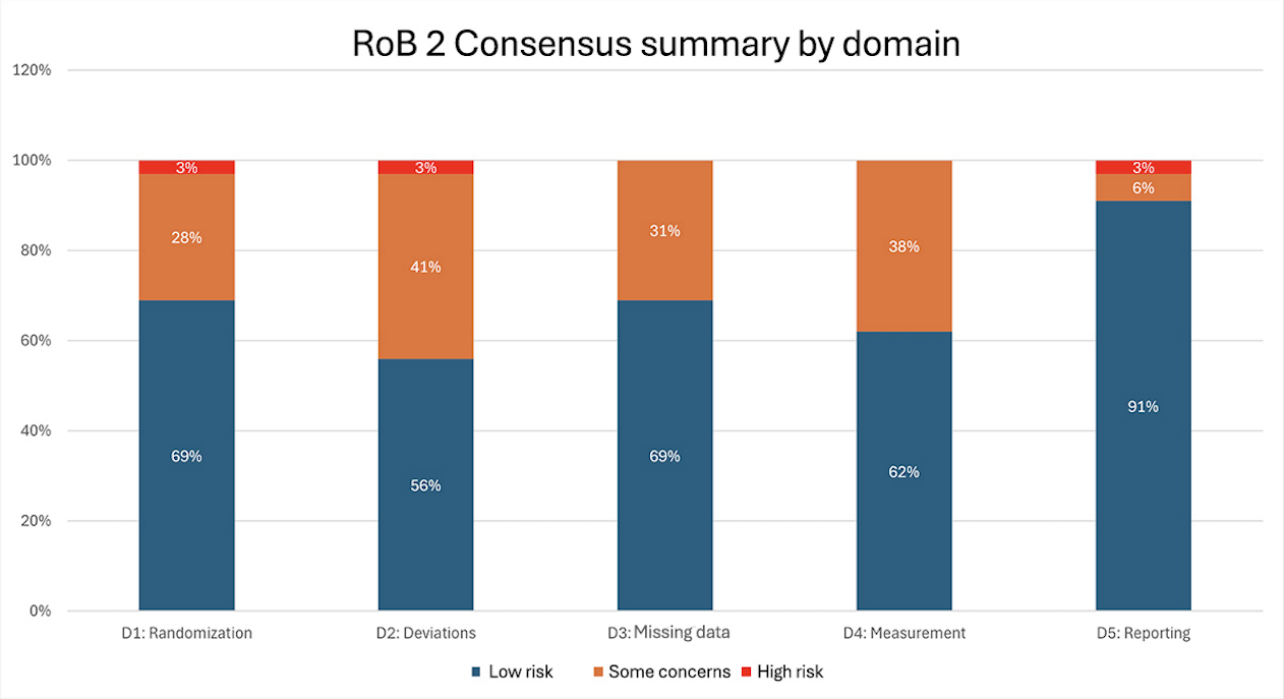
Table . Longitudinal evolution of methodological rigor among 35 [1,3,10-42] studies of digital postsurgical orthopedic rehabilitation technologies published between 2020 and 2025, based on overall risk of bias ratings by publication year.

Year	Low risk, n (%)	Some concerns, n (%)	High risk, n (%)	Total studies, n
2020	2 (40)	3 (60)	0 (0)	5
2021	1 (100)	0 (0)	0 (0)	1
2022	1 (14)	5 (71)	1 (14)	7
2023	5 (45)	5 (45)	1 (9)	11
2024	3 (60)	2 (40)	0 (0)	5
2025	5 (83)	1 (17)	0 (0)	6
Total	17(49)	16 (46)	2 (6)	35

Domain-based analysis (Figure 5) shows that selection and reporting biases were minimal; albeit performance bias (Domain 2) is an intrinsic issue with digital health trials because blinding participants and the therapists is technically impossible (eg, [32]). However, the large mean of PEDro (7.7) indicates that there is a serious focus on objective outcome measurement and high-statistic reporting throughout the corpus. It is important to note that most of the “some concern” ratings were signaled by the use of patient-reported outcome measures (PROMs) in unblinded designs, as opposed to inherent weaknesses in randomization or data integrity.

XSL•FO
RenderX

Figure 5. Risk of bias 2 domain-level ratings among the 32 randomized controlled trials [1,3,10-22,24-30,32-38,40-42] of digital postsurgical orthopedic rehabilitation technologies; the remaining 3 nonrandomized studies were appraised using different tools with noncomparable domain structures (Table 2). RoB2: Risk of Bias 2



Each included study contributed an independent patient cohort—no studies shared participants. Patient cohorts were predominantly postarthroplasty (total knee or hip arthroplasty; n=26 studies; Table 1) or post-ACL reconstruction (n=5 studies), with the remaining studies covering rotator cuff repair, meniscectomy, hip fracture, or mixed lower limb populations (n=4 studies). The range of digital interventions was broad: wearable sensor or IMU systems and VR or AR platforms were most common (n=10 each), followed by mHealth applications (n=9) and other telerehabilitation platforms (n=6). Several studies used multimodal combinations of these modalities. Rehabilitation duration across studies ranged from single inpatient sessions (2 - 4 days postoperatively) to 24 weeks, with home- and hybrid-based interventions spanning a wide range of postoperative timeframes, from brief early interventions to extended multiweek programs. Intervention dosing varied considerably—from predominantly unsupervised home sessions (eg, [3,15]) to more clinician-supported telehealth models (eg, [11,20])—reflecting the absence of standardized digital rehabilitation protocols in this literature. These characteristics are summarized in Table 1.

The Lower Limb Dominance: Biomechanical and Economic Drivers

The anatomy of the covered studies focused on the lower limb (32/35, 91%) and targeted the recovery interventions for total knee and hip arthroplasty. This distribution may reflect 2 convergent factors, though causal inference is beyond the scope of this review. First, from a biomechanical perspective, lower limb rehabilitation often prioritizes repetitive, sagittal-plane movements (eg, knee flexion or extension) and gait metrics, which are more easily quantified by current IMUs and markerless vision systems than the complex, multiplanar kinematics of the upper limb (eg, shoulder circumduction). Second, health economic pressures play a significant role; the high volume of lower limb arthroplasty procedures globally has created an urgent clinical need for scalable, automated solutions to reduce hospital length of stay and postoperative physiotherapy costs.

The Shift Toward Decentralized Care: From Supervision to Autonomy

Regarding the clinical delivery environment, a decisive shift from hospital-centric to decentralized care is evident. The hybrid model was the most frequently used approach (14/35, 40%). This model reflects a “safety-first” innovation strategy, where initial in-clinic supervision ensures patient safety and device calibration, followed by a transitioned digital home program to promote long-term autonomy.

The emergence of protocols that are purely home-based and remote (13/35, 37%) is the second wave of digital orthopedics, which can be enabled through the enhancements in user interface design and the reliability of remote monitoring. This pattern coincides with the accelerated adoption of telerehabilitation during the COVID-19 pandemic, during which remote-only pathways demonstrated clinical feasibility across multiple orthopedic contexts [31,38]. On the other hand, 8 studies used an exclusively clinical delivery model (8/35, 23%), suggesting that digital tools are increasingly deployed to extend rather than replace in-person care but as force multipliers that can be deployed in the domestic setting of the patient to close the critical continuity-of-care gap after discharge.

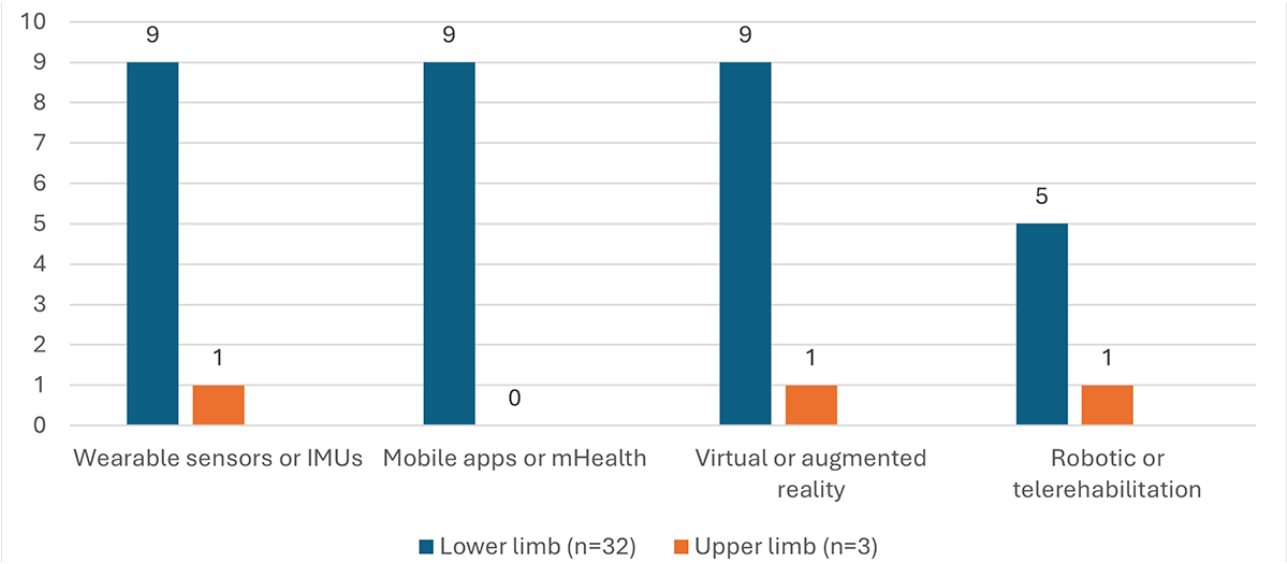
Technological Modalities and System Architecture

The technological modalities analysis (Figure 6) indicates that there is a hierarchical deployment of hardware. Wearable IMUs are widely used for objective monitoring and are used in the

XSL•FO
RenderX

majority of studies of the lower limbs to measure the angle of joints and gait parameters. Nonetheless, one of the major trends is the use of these sensors with mHealth ecosystems.

Figure 6. Distribution of technological modalities across anatomical targets among 35 studies of digital postsurgical orthopedic rehabilitation technologies published between 2020 and 2025. IMU: inertial measurement unit; mHealth: mobile health.



The most innovative change that was observed in the collected studies, for the time frame of 2024 - 2025, is the emergence of markerless computer vision and AR. These systems (eg, [3]) can reduce the patient's physical load, that is, by substituting physical sensors with camera-based pose estimation. The more recent studies (2023 - 2025) showed increasing architectural diversity, ranging from cloud-based asynchronous monitoring platforms to immersive VR or AR systems and synchronous telehealth models, reflecting the field's ongoing exploration of different digital delivery paradigms rather than convergence on a single dominant architecture.

Longitudinal Evolution of the CAMS Pillars

Figure 3 illustrates the distribution of CAMS pillars addressed by publication year. Between 2020 and 2022, when annual study counts ranged from 1 to 7 per year, interventions most frequently addressed Coordination (10/13, 77%) and Motivation (9/13, 69%), while Alignment (6/13, 46%) and, in particular, Strength (2/13, 15%) were less commonly addressed. Between 2023 and 2025, Coordination (19/22, 86%), Motivation (17/22, 77%), and Strength (10/22, 45%) coding were more frequent than in the earlier period, whereas Alignment coding was less frequent (7/22, 32%); however, given the small number of studies per year, these patterns are reported descriptively and should not be interpreted as a defined developmental sequence. The average number of CAMS pillars addressed per study ranged from 1.9 (SD 0.7) in 2022 to 2.8 (SD 0.8) in 2025, with no clear monotonic trend across the review period (Figure 3). This variability reflects heterogeneity in study scope and intervention type, though the data are consistent with a modest shift toward more multidimensional digital interventions in recent years.

Thematic Synthesis of CAMS Pillars

This section synthesizes the functional outcomes and technological architecture across the n=35 corpus. The results reveal that the current digital landscape is defined by multimodal integration. As illustrated in the synergy heatmap (Figure 4), Coordination and Motivation were the most frequently coaddressed pillars, appearing together in 60% (21/35) of included studies, suggesting this pairing functions as the functional baseline for current digital rehabilitation tools.

Motivation: The Digital Anchor

The most widespread pillar (26/35, 74%) was motivation, which became the main factor in influencing longitudinal adherence in the unsupervised domestic environment. The synthesis reveals an evolution from extrinsic motivators (eg, automated SMS reminders and adherence logs) to intrinsic engagement through immersive gamification.

VR and high-engagement digital interventions (eg, [2]) have been associated with reduced pain perception and perceived exertion during physical activity in the included literature, though this review's descriptive design does not establish a causal or clinical analgesic effect. Leveraging the "distraction effect," digital interventions may redirect emphasis on the postsurgical mechanical discomfort of movement toward goal-oriented play; these digital health technologies (DHTs) may thereby help address psychological obstacles to rehabilitation. This engagement set the basis of the high-frequency repetition volume needed to stimulate effective tissue remodeling and functional recovery.

Coordination-Motivation Axis: The Engagement-Precision Loop

Coordination and Motivation are identified as the dominant synergy. The clinical importance of this relationship is high, as high-frequency repetition is an element of the precision-based motor learning (Coordination) that could be retained only with the help of psychological engagement (Motivation). By integrating real-time biofeedback, such as AR-based movement guidance (eg, [3]), with gamified reward structures, DHTs support precise motor execution within an engaging framework. This process may reduce the cognitive load on the patient,

XSL•FO
RenderX

transitioning exercise from a monitored chore to a goal-oriented task, which is a prerequisite for achieving the high dose of training required for neural plasticity.

Alignment as a Gatekeeper for Coordination

While Coordination tracks the path of movement, Alignment is intended to assess the integrity of that path. The 17% (6/35) overlap between these pillars (eg, [3,13,25]) indicates an emerging trend toward "Gait-Aware" systems. Without Alignment, Coordination tracking is at risk of validating "compensatory movements"—where a patient achieves a coordinate goal by using maladaptive postures. AI-driven markerless vision may function as a monitoring tool for postural alignment, flagging compensatory movement patterns before Coordination metrics are recorded; whether this translates to reduced risk of abnormal loading at the surgical site was not assessed by the included studies.

The "Strength Isolation": A Technological Bottleneck

According to the heatmap, the Strength pillar had the lowest co-occurrence with Alignment (6/35, 17%), indicating that studies addressing strength rarely also addressed postural alignment. This isolation may point to a technological bottleneck: current digital platforms appear to demonstrate greater capability for monitoring kinematics (tracking the quality of motion) than for kinetic measurement (force quantification). Strength interventions currently function as standalone silos (eg, [24]), lacking the integrated biofeedback seen in the coordination-motivation axis. This gap may also reflect clinical and safety factors, including conservative protocols for early postoperative resistance training and publication trends favoring adherence studies over strength training trials.

Discussion

Summary of Principal Findings

Examined through the CAMS framework, the current corpus of digital orthopedic rehabilitation studies reveals a descriptive imbalance between behavioral engagement and physical load assessment. Coordination (29/35, 83%) and Motivation (26/35, 74%) were the most commonly addressed domains, reflecting the maturation of engagement-focused and movement-tracking technologies in home-based settings. By contrast, Strength-focused interventions appeared in only 34% (12/35) of studies, most of which required clinically supervised resistance hardware, limiting their applicability in decentralized care models. This gap, between kinematic monitoring (tracking the quality of movement) and kinetic measurement (quantifying mechanical force), is the central finding of this review.

The shift toward hybrid (14/35, 40%) and home-based (13/35, 37%) delivery models across the corpus is consistent with the broader adoption of telerehabilitation during and after the COVID-19 pandemic. This distributional pattern suggests that decentralized care has moved from being a contingency measure to an established component of orthopedic rehabilitation pathways, though the heterogeneity of included interventions and outcome measures means these observations are descriptive rather than conclusive.

From Kinetic Monitoring to Kinetic Integration

The co-occurrence of 60% (21/35) between Coordination and Motivation pillars (Figure 4) suggests that this pairing forms the functional core of current digital rehabilitation tools. Conventional physiotherapy relies on direct clinician observation for corrective feedback; current AR-based and IMU-based systems (eg, [3]) replicate elements of this closed-loop feedback in nonsupervised home settings, supporting the neuromuscular repetition volumes required for motor relearning.

Clinical Implications

The prevalence of hybrid delivery (14/35, 40%) suggests that a phased model, initial in-clinic assessment followed by digital home-based maintenance, may represent an emerging standard of care for postsurgical orthopedic rehabilitation. This model balances safety and therapist oversight in the acute phase with the scalability and patient autonomy of digital tools during extended recovery. The growing availability of markerless AI-based alignment monitoring [3,13] further supports this model, removing the need for sensor placement and enabling higher-quality alignment tracking for patients with limited technological literacy.

Based on the evidence mapped in this review, the following principles are suggested for clinical consideration. Platforms addressing the Coordination-Motivation axis alongside Alignment offer broader functional coverage and may be preferable to single-pillar adherence tools. Markerless vision systems should be considered for early detection of compensatory movement patterns in home settings. Clinicians should recognize that current DHTs are predominantly kinematic — direct strength measurement requires supplementary in-clinic assessment until affordable kinetic wearables become available. Hybrid delivery models offer a pragmatic balance between clinical oversight and home-based scalability. These suggestions are based on descriptive mapping findings and require validation through comparative clinical trials.

Limitation and Future Directions

This review has several limitations. First, the CAMS framework is a novel, unvalidated taxonomy proposed here as an exploratory classification tool; its categories reflect the authors' conceptual mapping rather than empirically derived dimensions. Second, the high degree of heterogeneity across intervention types, outcome measures, dosing protocols, and follow-up periods precluded formal meta-analysis, and the descriptive findings cannot be interpreted causally. Third, performance bias is an inherent challenge in digital health trials where blinding participants and therapists is not feasible, and most included studies relied at least partially on patient-reported outcomes. Fourth, Web of Science was not searched due to lack of institutional access, which may have resulted in missed records not indexed in the 5 databases used, though substantial overlap with Scopus and IEEE Xplore is expected for this multidisciplinary technology-focused literature. Finally, intervention durations varied considerably across studies, ranging from single inpatient sessions to programs extending several months postsurgery, leaving long-term functional durability, including joint longevity and secondary injury

XSL•FO
RenderX

prevention, largely unaddressed in the literature, regardless of intervention length.

Future research should prioritize 3 directions. First, empirical validation of the CAMS framework across diverse clinical and technological settings is needed to establish its utility as a classification and prescriptive tool. Second, trials specifically targeting the Strength pillar, particularly studies integrating smart resistance hardware with kinematic monitoring systems, are required to close the gap identified in this review. Third, longitudinal studies with follow-up extending well beyond the early postoperative period are needed to evaluate whether digital rehabilitation pathways support durable functional recovery and safe return to sport.

Conclusions

This scoping review mapped 35 [1,3,10-42] studies of digital technologies used in postsurgical orthopedic rehabilitation between 2020 and 2025, using the proposed CAMS framework to characterize their functional coverage. Within this literature, Coordination and Motivation were the most frequently addressed domains, while Alignment and, in particular, Strength were less consistently represented. This pattern may reflect current technological capabilities for quantifying postural correction and mechanical load, though the descriptive nature of this review precludes definitive conclusions about the underlying causes. Hybrid delivery models, combining initial clinical supervision with subsequent home-based digital support, were the most common approach across the included studies. These findings suggest several directions for future research, including empirical validation of the CAMS framework, development of accessible tools for quantifying physiological load, and longitudinal studies evaluating functional outcomes beyond the early postoperative period. As an exploratory taxonomy, CAMS may offer a useful starting point for classifying and comparing digital rehabilitation technologies in future research.

Acknowledgments

The authors would like to thank Dr Chan Kin Yuen, orthopedic surgeon, for his invaluable guidance and clinical insights throughout this study. This manuscript was prepared with the assistance of generative AI tools (ChatGPT, OpenAI; Claude, Anthropic) for the purposes of language editing, proofreading, text refinement, structural organization, and reference formatting. All scientific content, interpretation of findings, conceptual framework development, data extraction, quality appraisal, and final manuscript decisions were performed and verified by the authors. The authors take full responsibility for the integrity and accuracy of the work.

Funding

This study was supported by the Asia Pacific University of Technology and Innovation Research Development Internal Grant (RDIG/07/2024) and the Guangdong Key Disciplines Project (2024ZDJS137). The funders had no role in study design, data collection, analysis, interpretation, or the decision to publish.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: NKL
Methodology: NKL
Formal analysis: MNA, JJS
Data curation: WYC, YSL
Validation: NKL
Writing – original draft: NKL
Writing – review & editing: HH, TKC
Project administration: HH
Supervision: NKL

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR checklist.
[PDF File, 132 KB - rehab_v13i1e100548_app1.pdf]

References

XSL•FO
RenderX

1. Prvu Bettger J, Green CL, Holmes DN, et al. Effects of virtual exercise rehabilitation in-home therapy compared with traditional care after total knee arthroplasty: VERITAS, a randomized controlled trial. J Bone Joint Surg Am 2020 Jan 15;102(2):101-109. [doi: 10.2106/JBJS.19.00695] [Medline: 31743238]
2. Zitti M, Regazzetti M, Federico S, et al. Effectiveness of virtual reality for pain management in musculoskeletal disorders across anatomical regions: a systematic review and meta-analysis. Musculoskeletal Care 2025 Mar;23(1):e70041. [doi: 10.1002/msc.70041] [Medline: 39754331]
3. Lim JY, Yu HJ, Kim SH, et al. Effectiveness of in-home, augmented reality-based telerehabilitation after anterior cruciate ligament reconstruction: a randomized controlled trial. Orthop J Sports Med 2024 Oct;12(10):23259671241271729. [doi: 10.1177/23259671241271729] [Medline: 39415796]
4. Kim N. Emerging role of visual biofeedback in rehabilitation following ACL reconstruction. Asian J Kinesiol 2025 Oct;27(4):44-48. [doi: 10.15758/ajk.2025.27.4.44]
5. Hu R, Diao Y, Wang Y, et al. Effective evaluation of HGcnMLP method for markerless 3D pose estimation of musculoskeletal diseases patients based on smartphone monocular video. Front Bioeng Biotechnol 2023;11:1335251. [doi: 10.3389/fbioe.2023.1335251] [Medline: 38264579]
6. Seo NJ, Coupland K, Finetto C, Scronce G. Wearable sensor to monitor quality of upper limb task practice for stroke survivors at home. Sensors (Basel) 2024 Jan 16;24(2):554. [doi: 10.3390/s24020554] [Medline: 38257646]
7. Khalid S, Alnajjar F, Gochoo M, Renawi A, Shimoda S. Robotic assistive and rehabilitation devices leading to motor recovery in upper limb: a systematic review. Disabil Rehabil Assist Technol 2023 Jul;18(5):658-672. [doi: 10.1080/17483107.2021.1906960] [Medline: 33861684]
8. Aprile IG, Quaglini S, Turchetti G, et al. Rehabilitation robotics and allied digital technologies: opportunities, barriers and solutions for improving their clinical implementation. A position paper from the Fit for Medical Robotics Initiative. Front Robot AI 2025;12:1531067. [doi: 10.3389/frobt.2025.1531067] [Medline: 40678490]
9. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
10. Chen W, Qin X, Yang H, Tan Z, Liu P. Impact of cloud-based follow-up exercise prescription on knee function and quality of life in homebound patients post-total knee arthroplasty: a randomized controlled trial. J Orthop Surg Res 2025 Aug 30;20(1):809. [doi: 10.1186/s13018-025-06187-z] [Medline: 40885961]
11. Dajpratham P, Komas J, Yamthed R, et al. Feasibility and efficacy of real-time teleresistance exercise programs for physical function in elderly patients after hip fracture surgery: a randomized controlled trial. BMC Geriatr 2025 Aug 20;25(1):647. [doi: 10.1186/s12877-025-06230-y] [Medline: 40836222]
12. O'Reilly OC, Bozoghlian M, Glass N, et al. Physical therapy following shoulder arthroplasty: an assessment of telerehabilitation vs. in-person physical therapy. Semin Arthroplasty JSES 2025 Sep;35(3):505-516. [doi: 10.1053/j.sart.2025.04.009]
13. Sadiq S, Noor R, Akram R. Effect of lifestyle modification through web-based telerehabilitation monitoring combined with supervised sensorimotor training after total knee arthroplasty: randomized controlled trial. JMIR Mhealth Uhealth 2025 Oct 2;13:e64643. [doi: 10.2196/64643] [Medline: 41037782]
14. Yuan C, Xu C, Huang L, He Y, Lu S. The efficacy and cost-effectiveness analysis of telerehabilitation for patients after arthroscopic ACL-reconstruction: a non-inferiority randomized controlled trial. J Orthop Surg Res 2025 Dec 24;20(1):1085. [doi: 10.1186/s13018-025-06403-w] [Medline: 41444889]
15. Zhou Y, Lyu Y, Wang Q, Ma Y, Huang L, Zhang X. Mobile-based in-home telerehabilitation compared with in-hospital face-to-face rehabilitation for elderly patients after total hip arthroplasty in China's level 1 trauma center: a noninferiority randomized controlled trial. Front Surg 2024;11:1536579. [doi: 10.3389/fsurg.2024.1536579] [Medline: 39877311]
16. Christiansen CL, Kline PW, Anderson CB, et al. Optimizing total knee arthroplasty rehabilitation with telehealth physical activity behavior change intervention: a randomized clinical trial. Phys Ther 2024 Oct 2;104(10):pzae088. [doi: 10.1093/ptj/pzae088] [Medline: 39001713]
17. Xie W, He M, Zheng S, et al. Clinical application research of intelligent monitoring system for knee rehabilitation: a randomized controlled trial. J Orthop Surg Res 2024 Aug 13;19(1):477. [doi: 10.1186/s13018-024-04982-8] [Medline: 39138567]
18. Zhao R, Cheng L, Zheng Q, et al. A smartphone application-based remote rehabilitation system for post-total knee arthroplasty rehabilitation: a randomized controlled trial. J Arthroplasty 2024 Mar;39(3):575-581. [doi: 10.1016/j.arth.2023.08.019] [Medline: 37572720]
19. Shin D, Hong S. A study on the effectiveness of VR rehabilitation training content for older individuals with total knee replacement: pilot study. Health Care (Don Mills) 2024 Jul 29;12(15):1500. [doi: 10.3390/healthcare12151500]
20. Guo Y, Li D, Wu YB, Sun X, Sun XY, Yang YP. Mobile health-based home rehabilitation education improving early outcomes after anterior cruciate ligament reconstruction: a randomized controlled clinical trial. Front Public Health 2023 Jan 12;10:1042167. [doi: 10.3389/fpubh.2022.1042167]
21. Duong V, Robbins SR, Dennis S, Venkatesha V, Ferreira ML, Hunter DJ. Combined digital interventions for pain reduction in patients undergoing knee replacement: a randomized clinical trial. JAMA Netw Open 2023 Sep 5;6(9):e2333172. [doi: 10.1001/jamanetworkopen.2023.33172] [Medline: 37713201]

XSL•FO
RenderX

22. Höher J, Lischke B, Petersen W, et al. Sensor-based telerehabilitation system increases patient adherence after knee surgery. PLOS Digit Health 2023 Feb;2(2):e0000175. [doi: 10.1371/journal.pdig.0000175] [Medline: 36812639]
23. Lebleu J, Pauwels A, Anract P, Parratte S, Van Overschelde P, Van Onsem S. Digital rehabilitation after knee arthroplasty: a multi-center prospective longitudinal cohort study. J Pers Med 2023 May 13;13(5):824. [doi: 10.3390/jpm13050824] [Medline: 37240994]
24. Osterloh J, Knaack F, Bader R, et al. The effect of a digital-assisted group rehabilitation on clinical and functional outcomes after total hip and knee arthroplasty-a prospective randomized controlled pilot study. BMC Musculoskelet Disord 2023 Mar 14;24(1):190. [doi: 10.1186/s12891-023-06270-8] [Medline: 36915066]
25. Shim GY, Kim EH, Baek YJ, et al. A randomized controlled trial of postoperative rehabilitation using digital healthcare system after rotator cuff repair. NPJ Digit Med 2023 May 23;6(1):95. [doi: 10.1038/s41746-023-00842-7] [Medline: 37221303]
26. Shim GY, Kim EH, Lee SJ, et al. Postoperative rehabilitation using a digital healthcare system in patients with total knee arthroplasty: a randomized controlled trial. Arch Orthop Trauma Surg 2023 Oct;143(10):6361-6370. [doi: 10.1007/s00402-023-04894-y] [Medline: 37129691]
27. Gsangaya MR, Htwe O, Selvi Naicker A, et al. Comparison between the effect of immersive virtual reality training versus conventional rehabilitation on limb loading and functional outcomes in patients after anterior cruciate ligament reconstruction: a prospective randomized controlled trial. Asia Pac J Sports Med Arthrosc Rehabil Technol 2023 Oct;34:28-37. [doi: 10.1016/j.asmart.2023.09.002] [Medline: 37841643]
28. Gür O, Başar S. The effect of virtual reality on pain, kinesiophobia and function in total knee arthroplasty patients: a randomized controlled trial. Knee 2023 Dec;45:187-197. [doi: 10.1016/j.knee.2023.09.012] [Medline: 37931366]
29. Nuevo M, Rodríguez-Rodríguez D, Jauregui R, et al. Telerehabilitation following fast-track total knee arthroplasty is effective and safe: a randomized controlled trial with the ReHub® platform. Disabil Rehabil 2024 Jun;46(12):2629-2639. [doi: 10.1080/09638288.2023.2228689] [Medline: 37403684]
30. Mojica ES, Vasavada K, Hurley ET, et al. Despite equivalent clinical outcomes, patients report Less satisfaction with telerehabilitation versus standard in-office rehabilitation after arthroscopic meniscectomy: a randomized controlled trial. Arthrosc Sports Med Rehabil 2023 Apr;5(2):e395-e401. [doi: 10.1016/j.asmr.2023.01.006] [Medline: 37101883]
31. Hong M, Loeb J, Yang M, Bailey JF. Postoperative outcomes of a digital rehabilitation program after Total knee arthroplasty: retrospective, observational feasibility study. JMIR Form Res 2022 Sep 19;6(9):e40703. [doi: 10.2196/40703] [Medline: 36121690]
32. Marley WD, Barratt A, Pigott T, Granat M, Wilson JD, Roy B. A multicenter randomized controlled trial comparing gamification with remote monitoring against standard rehabilitation for patients after arthroscopic shoulder surgery. J Shoulder Elbow Surg 2022 Jan;31(1):8-16. [doi: 10.1016/j.jse.2021.08.019] [Medline: 34587534]
33. Pournajaf S, Goffredo M, Pellicciari L, et al. Effect of balance training using virtual reality-based serious games in individuals with total knee replacement: a randomized controlled trial. Ann Phys Rehabil Med 2022 Nov;65(6):101609. [doi: 10.1016/j.rehab.2021.101609] [Medline: 34839056]
34. Fuchs L, Kluska A, Novak D, Kosashvili Y. The influence of early virtual reality intervention on pain, anxiety, and function following primary total knee arthroplasty. Complement Ther Clin Pract 2022 Nov;49:101687. [doi: 10.1016/j.ctcp.2022.101687] [Medline: 36347152]
35. Zavala-González J, Martínez D, Gutiérrez-Espinoza H. Effectiveness of adding virtual reality to physiotherapeutic treatment in patients with total hip arthroplasty. A randomized controlled trial. Clin Rehabil 2022 May;36(5):660-668. [doi: 10.1177/02692155221080546] [Medline: 35166608]
36. Fascio E, Vitale JA, Sirtori P, Peretti G, Banfi G, Mangiavini L. Early virtual-reality-based home rehabilitation after total hip arthroplasty: a randomized controlled trial. J Clin Med 2022 Mar 22;11(7):1766. [doi: 10.3390/jcm11071766] [Medline: 35407373]
37. Şahi˙n A, Agar A, Ertürk C. The effect of telerehabilitation on early outcomes in patients undergoing primary total knee replacement: a prospective randomized study. J Surg Med 2022 Feb;6(2):139-143. [doi: 10.28982/josam.1035076]
38. Dunphy E, Button K, Hamilton F, Williams J, Spasic I, Murray E. Feasibility randomised controlled trial comparing TRAK-ACL digital rehabilitation intervention plus treatment as usual versus treatment as usual for patients following anterior cruciate ligament reconstruction. BMJ Open Sport Exerc Med 2021;7(2):e001002. [doi: 10.1136/bmjsem-2020-001002] [Medline: 34035951]
39. Wijnen A, Hoogland J, Munsterman T, et al. Effectiveness of a home-based rehabilitation program after total hip arthroplasty driven by a tablet app and remote coaching: nonrandomized controlled trial combining a single-arm intervention cohort with historical controls. JMIR Rehabil Assist Technol 2020 Apr 27;7(1):e14139. [doi: 10.2196/14139] [Medline: 32338621]
40. Bell KM, Onyeukwu C, Smith CN, et al. A portable system for remote rehabilitation following a total knee replacement: a pilot randomized controlled clinical study. Sensors (Basel) 2020 Oct 27;20(21):6118. [doi: 10.3390/s20216118] [Medline: 33121204]
41. Gianola S, Stucovitz E, Castellini G, et al. Effects of early virtual reality-based rehabilitation in patients with total knee arthroplasty: a randomized controlled trial. Medicine (Baltimore) 2020 Feb;99(7):e19136. [doi: 10.1097/MD.0000000000019136] [Medline: 32049833]

XSL•FO
RenderX

42. Nelson M, Bourke M, Crossley K, Russell T. Telerehabilitation is non-inferior to usual care following total hip replacement - a randomized controlled non-inferiority trial. Physiotherapy 2020 Jun;107:19-27. [doi: 10.1016/j.physio.2019.06.006] [Medline: 32026820]
43. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 1990;43(6):543-549. [doi: 10.1016/0895-4356(90)90158-l] [Medline: 2348207]
44. Hestetun-Mandrup AM, Toh ZA, Oh HX, He HG, Martinsen ACT, Pikkarainen M. Effectiveness of digital home rehabilitation and supervision for stroke survivors: a systematic review and meta-analysis. Digit Health 2024;10:20552076241256861. [doi: 10.1177/20552076241256861] [Medline: 38832099]
45. Abedi A, Colella TJF, Pakosh M, Khan SS. Artificial intelligence-driven virtual rehabilitation for people living in the community: a scoping review. NPJ Digit Med 2024 Feb 3;7(1):25. [doi: 10.1038/s41746-024-00998-w] [Medline: 38310158]

Abbreviations

ACL: anterior cruciate ligament
AR: augmented reality
CAMS : Coordination, Alignment, Motivation, and Strength
CENTRAL: Cochrane Central Register of Controlled Trials
DHT: digital health technology
IMU: inertial measurement unit
mHealth: mobile health
MINORS: Methodological Index for Non-Randomized Studies
NOS: Newcastle-Ottawa Scale
PCC: population-concept-context
PEDro: Physiotherapy Evidence Database
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PROM: patient-reported outcome measure
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
RoB2: Cochrane risk-of-bias tool, version 2, for randomized trials
THA: total hip arthroplasty
TKA: total knee arthroplasty
TRAK: Taxonomy for the Rehabilitation of Knee Conditions
VR: virtual reality

Edited by J Sarvestan; submitted 07.May.2026; peer-reviewed by N Meng, Y Han; revised version received 09.Aug.2026; accepted 11.Aug.2026; published 20.Oct.2026.

Please cite as:
Lai NK, Huang H, Chen TK, Nor'a MNA, Sijore J, Chen WY, Luo Y
Digital Technologies in Postsurgical Orthopedic Rehabilitation: Scoping Review Mapping Interventions to the CAMS Framework (2020 - 2025)
JMIR Rehabil Assist Technol 2026;13:e100548
URL: https://rehab.jmir.org/2026/1/e100548
doi:10.2196/100548

© Ngan Kuen Lai, Hai Huang, Tet Khuan Chen, Muhammad Nur Affendy Nor'a, Jeffery Sijore, Wei Yang Chen, YongSheng Luo. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Advances in Machine Learning Paradigms for Reliable Swimming Pool Drowning Detection: Systematic Review of Methods and Implementations

Mercy Chepkorir[1*], MSc; Idah Orowe[1*], PhD; Elisha Abade[2], PhD; Timothy Kamanu[1], PhD

[1]Department of Mathematics, University of Nairobi, P.O Box 30197, Nairobi, Kenya

[2]Department of Computer Science, University of Nairobi, Nairobi, Kenya

[*]these authors contributed equally

Corresponding Author:
Mercy Chepkorir, MSc
Department of Mathematics, University of Nairobi, P.O Box 30197, Nairobi, Kenya

Abstract

Background: Drowning is a leading cause of unintentional injury-related mortality worldwide, particularly in swimming pool environments, where rapid detection is critical to prevent fatalities. Traditional monitoring approaches rely on lifeguards and rule-based systems, which are often limited by human error, delayed response, and environmental challenges such as occlusion, glare, and multiple swimmers. Machine learning (ML) and computer vision methods have emerged as promising alternatives for automated drowning detection; however, their methodological rigor, real-world applicability, and reliability remain unclear.

Objective: This study aimed to systematically review advancements in ML-based drowning detection systems over the past 15 years. The review focused on detection methods, sensing modalities, evaluation metrics, and validation practices. It also examined how studies addressed real-world challenges such as occlusion, lighting variability, and multiswimmer environments, and identified key gaps related to generalizability, reliability, and privacy.

Methods: A systematic literature review was conducted using major electronic databases, including IEEE Xplore, Scopus, PubMed, Web of Science, ACM Digital Library, and Google Scholar. A total of 4927 records were initially identified and screened by 2 independent reviewers using predefined inclusion and exclusion criteria. A final set of 30 studies was included for data extraction and synthesis. The methodological quality and risk of bias were assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST), while reporting transparency was evaluated using the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines. Data analysis was conducted in R.

Results: The findings indicate that drowning detection research is dominated by supervised, vision-based approaches (23/30, 76.7%), with convolutional neural networks (5/30, 16.7%) and You Only Look Once–based architectures (7/30, 23.3%) being the most commonly used methods. While reported performance is generally high, with mean accuracy exceeding 90.9%, evaluation practices are heavily centered on accuracy, with limited reporting of precision, recall, false alarm rate, and latency. One of 30 (3.3%) studies reported validation under real-world operational conditions, with an additional 5 (16.7%) studies using real-world data without formal real-world validation, and no studies conducted cross-dataset evaluation. In addition, critical real-world challenges such as occlusion, glare, and crowding were rarely addressed (<10% of studies), and all studies relied on private datasets. The PROBAST assessment further indicated a high risk of bias in 96.7% (29/30) of studies, primarily due to limited validation and nonrepresentative data sources.

Conclusions: ML-based drowning detection systems have demonstrated substantial progress in algorithm development, particularly in vision-based models. However, this progress is not matched by equivalent advances in validation, robustness, and real-world applicability. The field is characterized by a significant gap between experimental performance and deployment readiness. Future research should prioritize standardized evaluation frameworks, external validation, comprehensive performance reporting, and the integration of multimodal and privacy-preserving approaches to enable reliable real-world implementation.

(JMIR Rehabil Assist Technol 2026;13:e100939) doi:10.2196/100939

KEYWORDS

drowning detection; machine learning; computer vision; deep learning; real-time detection; multimodal sensing; evaluation metrics; privacy; water safety

XSL•FO
RenderX

Introduction

Drowning remains a major global public health concern and a leading cause of unintentional injury-related mortality worldwide. Drowning is a major but preventable public health problem worldwide [1]. Each year, about 300,000 people die from unintentional drowning [2]. It is a leading cause of injury-related deaths. Young people are most affected. Over half of victims are under 25 years. Children aged 1 to 4 years are at the highest risk [3]. Most deaths, about 90%, occur in low- and middle-income countries due to limited safety measures and supervision.

Drowning is difficult to prevent because it is rapid and often silent, and victims show the instinctive drowning response. This includes vertical positioning and little movement. They are often unable to call for help. The time to intervene is very short, often seconds to minutes. Lifeguards play an important role but face limitations. Their performance is affected by fatigue, attention limits, and crowded environments [4]. Occlusion and high swimmer-to-lifeguard ratios make detection harder. These challenges highlight the need for automated, real-time monitoring systems to support human supervision.

Early technological approaches relied on rule-based systems and handcrafted features derived from visual observations. These methods attempted to detect distress through predefined patterns such as vertical body positioning and irregular motion. However, they lacked robustness and performed poorly in dynamic and real-world environments.

The emergence of machine learning (ML) and computer vision significantly advanced drowning detection research [5]. Convolutional neural networks (CNNs) were widely used for feature extraction from visual data [6]. Object detection models such as YOLO (You Only Look Once) and region-based approaches like Faster region-based CNN (R-CNN) improved real-time detection capabilities [7-10]. Recurrent neural networks and long short-term memory networks were applied to capture temporal patterns in swimmer behavior [11]. More recent studies explored spatiotemporal architectures and transformer-based models to model complex motion dynamics. Hybrid approaches combining multiple models further improved detection performance.

Despite substantial methodological progress, the field lacks a comprehensive and rigorous synthesis of existing evidence. Current studies are highly heterogeneous, with inconsistent reporting of evaluation metrics, limited use of operational performance indicators such as false alarm rates and latency, and minimal external validation. Most systems are evaluated in controlled environments using simulated scenarios, with little evidence of real-world deployment or cross-dataset generalization. These limitations make it difficult to assess the true effectiveness and readiness of existing systems for practical use.

This study therefore aimed to systematically review advancements in ML-based drowning detection methods published over the past 15 years. This time period was selected to capture the period during which modern ML, particularly deep learning and computer vision, emerged and rapidly evolved. The review examined detection models, data modalities, evaluation metrics, and deployment strategies. It also assessed how studies addressed real-world challenges and evaluated their methodological rigor.

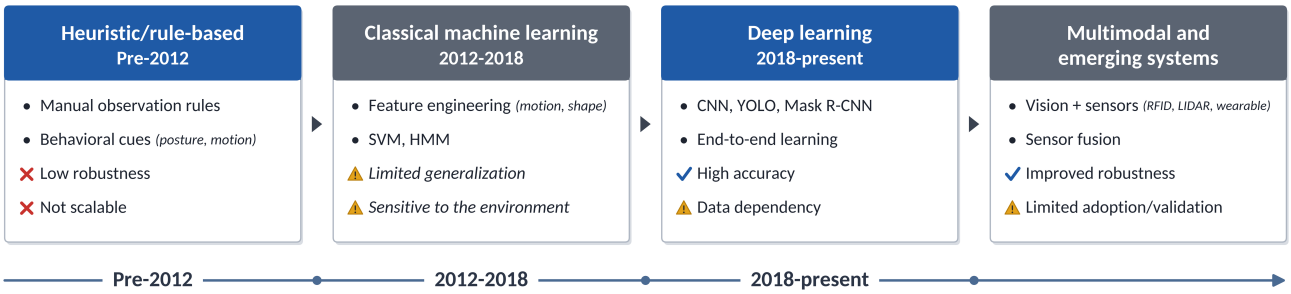
Building on this trajectory, the remainder of this section traces the field's technical evolution in greater depth. Automated drowning detection has evolved from heuristic observation models to data-driven ML systems. Early approaches identified drowning based on observable behavioral cues such as vertical body posture, limited forward motion, and irregular limb activity [12]. While these rule-based systems provided initial insights into drowning behavior, they were constrained by predefined thresholds and lacked robustness to environmental variability, including reflections, occlusion, and crowding. The introduction of classical ML methods improved automation by leveraging engineered features derived from motion and appearance. However, these approaches remained dependent on feature design and exhibited limited generalizability in complex aquatic environments [5].

The adoption of deep learning marked a significant shift in the field, enabling end-to-end learning of spatial and temporal representations directly from visual data. CNNs became widely used for feature extraction, while object detection frameworks such as YOLO have been extensively applied for real-time swimmer localization and monitoring [13,14]. More advanced architectures, including feature pyramid networks, Mask R-CNN, and pose estimation models, have further improved detection performance under varying conditions [11]. Despite these advances, most systems rely on supervised learning, which requires labeled datasets that are scarce for rare events such as drowning [11,15]. Research in this domain remains dominated by frame-based detection approaches, with comparatively limited exploration of temporal modeling or anomaly detection strategies.

Recent work has also explored alternative sensing modalities, including underwater cameras, wearable devices, radio-frequency identification (RFID) systems, and depth-based approaches. While these modalities aim to improve robustness under challenging conditions, vision-based systems remain dominant due to their practicality and scalability [16]. Across the literature, key challenges persist, including sensitivity to environmental factors such as glare and occlusion, limited validation in real-world settings, and the absence of standardized datasets. These limitations highlight the need for a systematic synthesis of existing approaches, particularly in relation to methodological trends, evaluation practices, and deployment readiness. The evolution practices have been illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Evolution of drowning detection. CNN: convolutional neural network; HMM: hidden Markov model; LiDAR: light detection and ranging; Mask R-CNN: mask region-based convolutional neural network; RFID: radio-frequency identification; SVM: support vector machine; YOLO: You Only Look Once.



Methods

Goal and Review Questions

Goal of the Study

The main goal of this study was to systematically map, review, and synthesize studies that focused on ML-based drowning detection. This included an in-depth examination of detection models, data modalities, deployment strategies, and practical applications of these systems. The study also examined how different approaches addressed real-world challenges such as occlusion, glare, and the presence of multiple swimmers.

This study aimed to identify recent trends, coverage, and directions in drowning detection research. It also aimed to highlight key limitations related to generalizability, reliability, and privacy. These insights were intended to inform future research and support the development of robust and deployable detection systems. The protocol was defined before data collection, thereby minimizing the risk of bias in study selection and reporting. The protocol has been registered with the Open Science Framework before data collection. Details of the registered protocol may be found in [17]. This ensured absolute transparency, reproducibility, and minimized risk of bias in study selection and reporting.

Based on established guidelines for conducting systematic literature reviews and reporting findings, the selection and review process was carefully designed and documented. Based on this goal, the following review questions (RQs) were formulated and adopted.

Category 1: General RQs for the Systematic Review

The following research questions were general to all systematic literature review studies:

1. RQ 1.1. Contributions presented in the studies: What were the main contributions of the reviewed studies? This question focused on identifying whether the studies contributed through the development of models, algorithms, systems, datasets, or practical applications for drowning detection.
2. RQ 1.2. Research methods used in the studies: How were the studies categorized based on research methods? This question aimed to classify the studies based on whether they employed empirical, experimental, or hybrid approaches, given the multidisciplinary nature of drowning detection systems.

Category 2: RQs Specific to the Topic of Systematic Review

The following RQs were specific to the topic of advances in ML-based drowning detection:

1. RQ 2.1. Methods and models employed for drowning detection: What methods and models were employed for detecting drowning incidents? This question aimed to identify whether studies used classical approaches, ML models, deep learning architectures, or hybrid methods. It also examined the use of specific models such as CNNs, object detection frameworks, pose estimation, and spatiotemporal models.
2. RQ 2.2. Data modalities and feature representation: What types of data and feature representations were used in the studies? This question focused on whether studies used vision-based data, sensor-based data, or multimodal approaches. It also examined how spatial, temporal, and motion-related features were represented and used for detection.
3. RQ 2.3. Types and sources of datasets: What datasets were used in the studies? Were they primary datasets collected by researchers or secondary datasets obtained from existing sources?
4. RQ 2.4. Evaluation metrics and reported performance: What evaluation metrics and performance measures were reported? This question examined metrics such as accuracy, precision, recall, F_1-score, false alarm rate, and latency. It aimed to compare the effectiveness of different models and assess consistency in reporting.
5. RQ 2.5. Real-time capability and deployment: Were the proposed models capable of real-time detection, and were they deployed in practical settings? This question evaluated whether systems were implemented in real-world environments such as swimming pools, surveillance systems, or embedded devices, and how they handled operational constraints.
6. RQ 2.6. Handling of real-world challenges: How did the studies address real-world challenges such as occlusion, glare, lighting variability, and multiple swimmers? This question examined whether the models were robust under complex and dynamic conditions.
7. RQ 2.7. Generalizability and external validation: To what extent did the studies evaluate generalizability? This question assessed whether models were tested across different datasets, environments, or populations, and whether external validation was performed.

XSL•FO
RenderX

8. RQ 2.8. Privacy and data security considerations: How did the studies address privacy and data security? This question examined whether privacy-preserving approaches, such as federated learning, were used or discussed.
9. RQ 2.9. Limitations and risk of bias: What limitations and risks of bias were identified in the studies? This question focused on issues related to data quality, study design, lack of validation, and incomplete reporting, as observed in the risk of bias assessment.

Search Strategy

To ensure transparency and completeness in our systematic review, we followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for conducting systematic literature reviews as outlined by relevant studies. Our search spanned 6 databases, covering literature from 2012 to 2025, with the aim of identifying relevant articles on the methodologies, deployment strategies, and practical applications in drowning detection. Records published between 2012 and 2018 were identified during the search but did not meet the study's inclusion criteria at the full-text screening stage, primarily because they described rule-based or handcrafted-feature detection approaches rather than ML-based systems, and were therefore excluded under the exclusion criterion. All 30 included studies fall within the 2019 to 2025 period, consistent with the rapid emergence of deep learning–based approaches in this field during that time.

The databases used for this review include IEEE Xplore, Web of Science Core Collection, Scopus, PubMed, ACM Digital Library, and Google Scholar. The databases and full search terms are provided in Multimedia Appendix 1. The search strings were developed and structured using the PICO framework, which considers the Problem, Intervention, Comparison, and Outcome components (Multimedia Appendix 2) [18]. It was designed to address the overarching question: "Can advanced ML methodologies enhance drowning detection and improve real-time response in water environments?" We tailored the PICO structure as follows:

1. Comparison: Due to the specificity of our research objectives, direct comparison studies were not applicable in this context.
2. Outcome: We sought outcomes related to accuracy, deployment, and practical applications of drowning detection methodologies, incorporating terms such as "accuracy" OR "precision" OR "recall" OR "F_1-score" OR "false alarm rate" OR "real-time" OR "deployment" OR "latency."

By employing "AND" and "OR" operators, we formulated a general search phrase:

- ("drowning detection" OR "drowning recognition" OR "swimmer distress" OR "near-drowning" OR "water safety monitoring") AND ("machine learning" OR "deep learning" OR "computer vision" OR "neural networks" OR "YOLO" OR "CNN" OR "LSTM" OR "pose estimation") AND ("accuracy" OR "precision" OR "recall" OR "F_1-score" OR "false alarm rate" OR "real-time" OR "deployment" OR "latency")

Eligibility Criteria

The eligibility criteria for inclusion were established to comprehensively evaluate studies that focused on ML-based drowning detection. The criteria considered a wide range of methodologies, data modalities, and application contexts. They were carefully designed to ensure the inclusion of studies aligned with the objectives of this review.

The studies were evaluated to determine whether they addressed key aspects of drowning detection, including detection methods, deployment strategies, real-time capabilities, and practical applications. Emphasis was placed on studies that examined challenges such as occlusion, lighting variability, and multiswimmer environments. The criteria also ensured the inclusion of studies that contributed to understanding trends, limitations, and future research directions in drowning detection.

Initial Screening

During this preliminary phase, articles identified through database searches and manual exploration underwent early screening based on titles and abstracts. The aim was to ensure that the screened studies aligned with the goal of systematically reviewing ML-based drowning detection.

Emphasis was placed on filtering out irrelevant studies and identifying those contributing to the analysis of detection methods, data modalities, deployment strategies, performance evaluation, and practical applications. Studies not related to drowning detection or not involving computational or ML approaches were excluded at this stage.

Adherence to established systematic review practices ensured comprehensive documentation of the selection process. This supported the formulation of the research questions and the synthesis of findings.

Inclusion and Exclusion

Out of 4927 search results, a total of 30 studies were included for review. Studies that were included had to fulfill the following criteria: (1) studies that focused on ML-based drowning or swimmer distress detection, (2) studies that detailed the methodologies or models used for detection, and (3) studies that clearly defined evaluation metrics, deployment strategies, or practical applications of drowning detection techniques.

To ensure precise and robust findings, this study used exclusion criteria guided by specific thematic considerations. (1) Irrelevant topics: literature investigating other aspects of aquatic monitoring not related to drowning or distress detection. (2) Incomplete focus: studies focused solely on other processing tasks, such as preprocessing or noise reduction, without subsequent detection or classification. (3) Lack of technological interventions: literature that did not apply ML, computer vision, or related computational approaches, such as studies focusing purely on theoretical or descriptive aspects without implemented detection systems. (4) Insufficient application: literature that only described swimmer behavior, features, or observations without using them in detection models. (5) Limited scope: studies that focused solely on general surveillance, swimmer tracking, or water safety without addressing drowning or distress detection. (6) Non-English publications: studies that were not

XSL•FO
RenderX

available in English due to language proficiency constraints of the review team.

A 2-stage screening process was used to ensure thoroughness and reduce the bias arising. Two reviewers independently screened all titles and the corresponding abstracts in the first step to assess relevance based on the inclusion criteria. Afterward, the papers were advanced to the next stage if and when at least 1 reviewer deemed them relevant. In cases where consensus at the title and abstract level could not be achieved between the 2 reviewers, an adjudication by a third person was sought on a consultation basis and on the purpose. During the second stage, 2 independent reviewers performed a full-text blind review. Consensus was obtained through deliberation between reviewers when required. Consensus meetings were held to resolve any disagreements and uncertainties. The objectivity of the criteria was assessed through reviewer agreement during the screening stages. The study selection process is presented in Figure 1.

Data Synthesis and Comprehensive Analytical Approach

In alignment with the RQs, the following information points were recorded and analyzed: (1) year of publication, (2) contribution of each study, (3) categorization of each study based on research approach, (4) ML methods and model architectures, (5) data modalities employed (eg, red, green, and blue [RGB] video, underwater video, sensor-based, or multimodal data), (6) deployment characteristics including real-time capability and hardware requirements, (7) practical application settings (eg, swimming pools, open water environments, surveillance systems), (8) reported performance metrics such as accuracy, precision, recall, F_1-score, false alarm rate, and latency, and (9) challenges addressed, including occlusion, lighting variability, multiswimmer environments, and viewpoint variation.

Tables and graphical summaries were developed to synthesize these information points, providing a comprehensive overview of advancements in drowning detection methods, system design, and real-world applicability. The temporal distribution of publications indicated a growing research interest in recent years, reflecting the increasing importance of automated drowning detection systems and the advancement of ML techniques in safety-critical environments. Meta-analysis was not undertaken because the included studies were heterogeneous in model, evaluation metrics reported, dataset type, and study design. This degree of methodological and outcome heterogeneity precluded meaningful statistical pooling, consistent with PRISMA guidance favoring narrative and tabular synthesis under such conditions [19].

Quality Assessment

The objective of this quality assessment was to evaluate the significance and validity of each selected document in the context of ML-based drowning detection. Although the quality rating did not significantly influence the selection of the primary studies, the assessment aimed to reflect the overall validity and reliability of the included studies.

The TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) [20] standard was employed as a measure of reporting quality. Each article was scored based on adherence to the TRIPOD checklist, with binary scores assigned as either 1 (compliant) or 0 (noncompliant). Moreover, all reviewed studies were expected to meet a minimum threshold of reporting quality to ensure clarity and reproducibility. A detailed checklist of the TRIPOD assessment is provided in .

The TRIPOD checklist was used to evaluate the reporting clarity and comprehensiveness across all selected studies. Each study was assessed based on key TRIPOD components, including data description, model development, validation procedures, and reporting of results. The assessment highlighted variability in reporting practices across studies, particularly in relation to external validation, dataset transparency, and performance metric reporting. Overall, the application of TRIPOD ensured a structured evaluation of reporting quality and supported the reproducibility of the findings.

The study additionally employed the PROBAST (Prediction Model Risk of Bias Assessment Tool) to assess the risk of bias in the included studies [20]. Each study was scored across 5 core TRIPOD reporting domains: Title/Abstract, Introduction, Methods, Results, and Discussion, rather than scoring each of the checklist's individual items separately. Each domain was scored as 1 (adequately reported) or 0 (not adequately reported), yielding a total domain-level score out of 5 per study; this approach was adopted to provide a tractable, reproducible summary measure of reporting quality across a heterogeneous set of study designs, while a detailed item-level checklist is provided in Checklist 1 [21]. The PROBAST framework was considered appropriate given the predictive modeling nature of drowning detection systems. It is acknowledged that PROBAST and TRIPOD were originally developed for clinical multivariable prediction models used for individual prognosis or diagnosis, and that applying them to computer vision object detectors, including YOLO variants, Mask R-CNN, and wearable sensor systems, represents an extension beyond their original scope. This extension was considered justified because the included studies, despite differing in architecture, share the core structure of a prediction model. They take input data such as images, video frames, or sensor signals and output a binary or probabilistic prediction (drowning vs nondrowning), which is the fundamental unit of analysis PROBAST and TRIPOD were designed to appraise. Domain-specific tools for AI/ML diagnostic studies, such as Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2), Checklist for Artificial Intelligence in Medical Imaging (CLAIM), and CLEAR, were considered as alternatives; however, PROBAST and TRIPOD were retained because they offer the most established, validated domain structure for systematically assessing risk of bias and reporting transparency across heterogeneous prediction-model designs, and because comparable systematic reviews in adjacent safety-critical ML domains have applied the same tools. The risk of bias assessment covered 4 key domains: (1) bias related to participants and datasets, (2) bias in predictor selection and feature representation, (3) bias in outcome definition and measurement, and (4) bias in model development and analysis.

XSL•FO
RenderX

The results of the quality and risk of bias assessments were summarized in Table 1, with detailed individual evaluations provided in Multimedia Appendix 3.

Table . Risk of bias assessment (N=30).

Domain	Low risk, n (%)	Some concerns, n (%)	High risk, n (%)
Data source bias	4 (13.3)	10 (33.3)	16 (53.3)
Prediction bias	27 (90)	3 (10)	0 (0.0)
Outcome assessment bias	1 (3.3)	28 (93.3)	1 (3.3)
Analysis bias	11 (36.7)	10 (33.3)	9 (30)
Overall risk	0 (0.0)	1 (3.3)	29 (96.7)

The assessment of data source bias indicated that a substantial proportion of studies, 53.3% (16/30)**,** were considered to have a high risk of bias, largely due to reliance on limited, nonrepresentative, or synthetic datasets. A further 33.3% (10/30) of studies raised some concerns, while only 13.3% (4/30) were assessed as low risk. These findings suggest that data-related limitations remain a major challenge, potentially affecting the generalizability and robustness of the reported results.

The assessment of prediction bias showed that 90% (27/30) of studies had a low risk of bias, indicating that most studies employed appropriate modeling techniques and prediction strategies. Only 10% (3/30) of studies presented some concerns, suggesting that predictive modeling approaches were generally well implemented across the reviewed literature.

The assessment of outcome assessment bias revealed a significant area of concern, with 93.3% (28/30) of studies classified as having some concerns, primarily due to inconsistencies in evaluation protocols and lack of standardized performance metrics. Only 3.3% (1/30) of studies were considered low risk, while 3.3% (1/30) were categorized as high risk. These findings indicate that outcome evaluation remains insufficiently standardized, limiting comparability across studies.

The assessment of analysis bias indicated a relatively balanced distribution, with 36.7% (11/30) of studies categorized as low risk, 33.3% (10/30) showing some concerns, and 30% (9/30) identified as high risk. This variability reflects differences in validation strategies, with several studies lacking external validation or relying on limited experimental setups, which may introduce overfitting and reduce reliability.

The overall assessment of risk of bias indicated that 96.7% (29/30) of the studies posed a high risk, while only 3.3% (1/30) showed some concerns. No studies were classified as low overall risk. These findings highlight substantial methodological limitations across the literature, particularly related to dataset quality, evaluation consistency, and real-world validation.

The TRIPOD assessment indicates generally strong reporting practices, with an average score of 4.07 (81%) out of a maximum of 5 (one point per reporting domain: Title/Abstract, Introduction, Methods, Results, and Discussion). Domain-level compliance varied substantially: Title/Abstract and Introduction were adequately reported in 100% (30/30) of studies, Discussion in 86.7% (26/30), Results in 70.0% (21/30), and Methods in only 50% (15/30), the weakest-performing domain, consistent with this review's broader finding of limited methodological transparency. Overall**,** 24 studies were classified as having high reporting quality (domain score of 4/5 or 5/5), 5 studies as having moderate reporting quality (3/5), and 1 study as having low reporting quality (2/5 or below). This suggests that most studies provided clear descriptions of model development and reported performance metrics adequately. However, important gaps remained, particularly in relation to external validation, dataset transparency, and reproducibility details. These limitations affect the interpretability and generalizability of the findings despite strong reporting compliance.

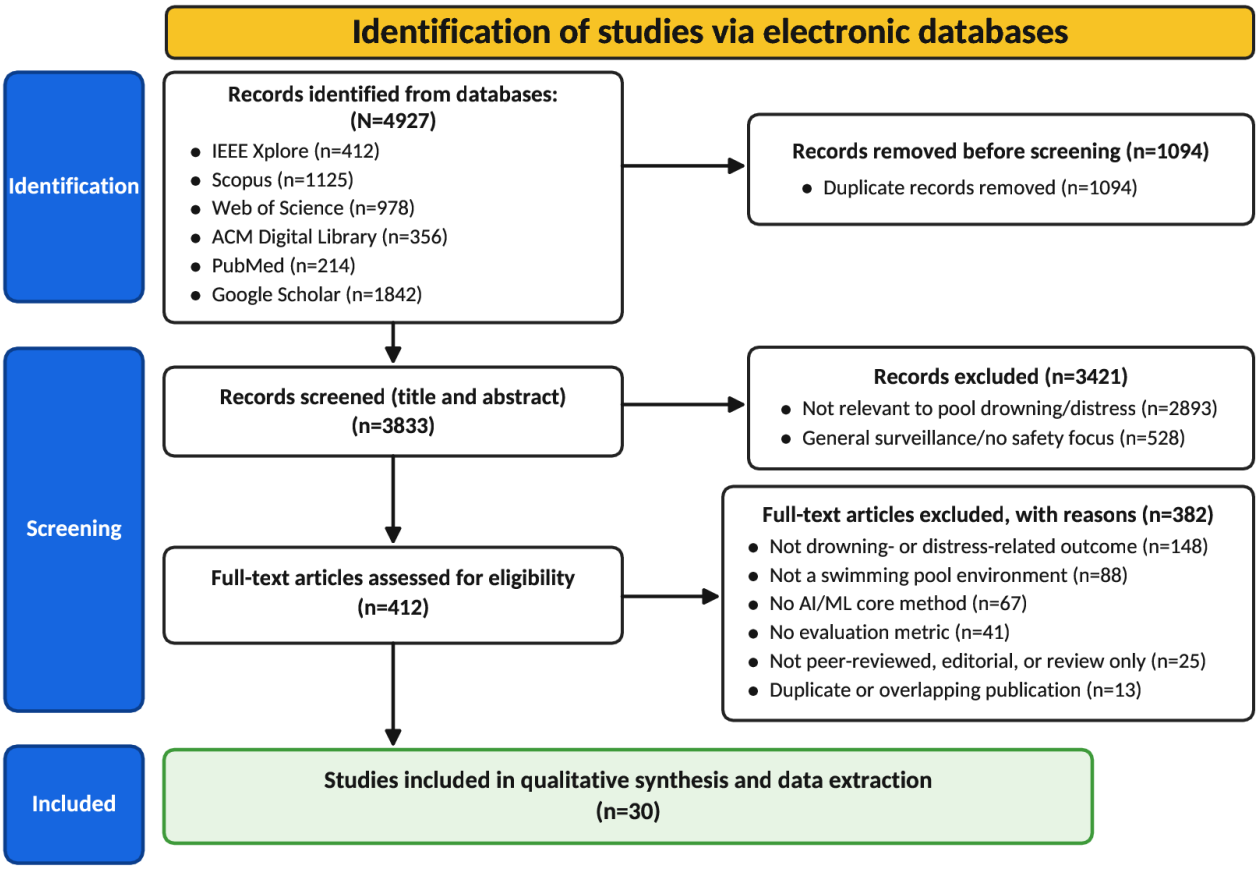
The combined risk of bias and reporting quality assessments indicate that, while the field demonstrates strong methodological progress in model development and reporting clarity, there are significant gaps in data quality, evaluation standardization, and real-world validation.

Results

The findings in this section are structured according to the classification of the reviewed studies and the key technical considerations relevant to contextualizing drowning detection systems (Figure 2).

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. ML: machine learning.



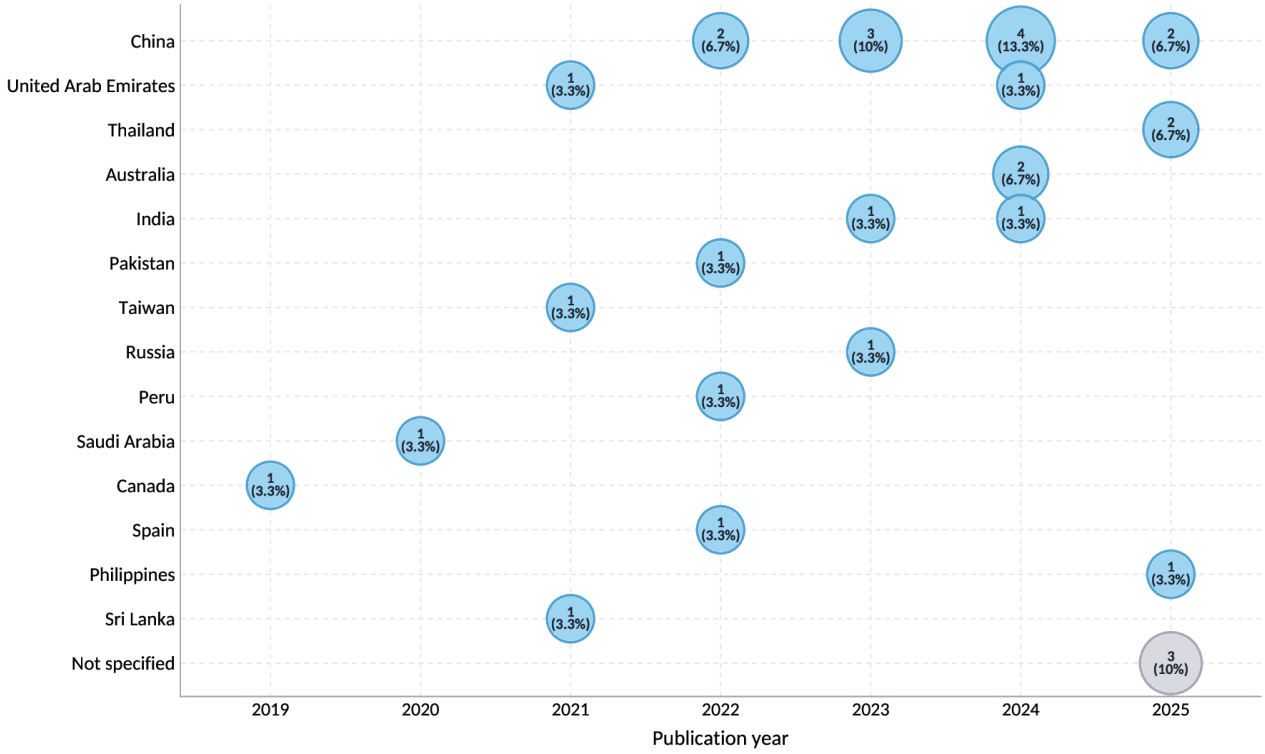
Overview of Included Studies

A total of 30 studies published between 2019 and 2025 were included in this review. The temporal distribution demonstrates a clear and rapid increase in research activity over time. Publication output remained minimal in the early years, with 3.3% (n=1) of studies published in both 2019 and 2020, followed by a sharp increase in subsequent years and peaking in 2024. This trend reflects the accelerating interest in ML-based drowning detection systems, particularly in recent years.

Geographically, research output was highly concentrated. China accounted for 36.7% (11/30) of the included studies, emerging as the dominant contributor. This was followed by the United Arab Emirates, Thailand, Australia, and India, each contributing 6.7% (2/30). Nine additional countries, Saudi Arabia, Canada, Spain, Sri Lanka, Pakistan, Taiwan, Russia, Peru, and the Philippines, each contributed a single study (3.3%). The dataset's country of origin could not be determined for 3 (10%) studies, and one additional study was a China and Pakistan collaboration counted toward both countries, which is why the percentages sum to slightly more than 100%. As illustrated in Figure 3, the recent surge in publications is largely driven by contributions from China across 2023 to 2025.

XSL•FO
RenderX

Figure 3. Distribution of included studies by country and publication year. Bubble size represents number of studies; labels show study count and percentage of total studies (N=30).



These findings indicate that while the field is rapidly expanding, it remains geographically concentrated. This concentration suggests that current advancements are shaped by regions with greater access to technological infrastructure and data resources. The diversity of environmental contexts represented in the literature remains limited, raising concerns about the generalizability and real-world applicability of existing drowning detection systems.

Methodological Approaches

The methodological landscape of the included studies was heavily dominated by detection-oriented approaches, which accounted for 70% (21/30) of all studies. Systems combining detection with early warning mechanisms represented 10% (3/30), while classification or recognition-based approaches constituted only 6.7% (2/30). Other methodological paradigms, including pose-based detection, anomaly-assisted detection, multimodal systems, and dataset-focused studies, were each represented by 3.3% (1/30) of the included studies. This distribution highlights a strong emphasis within the literature on direct event detection and real-time alert generation, reflecting the practical importance of timely intervention in drowning scenarios [10,22,23]. However, it also reveals a relative underrepresentation of approaches focused on behavioral modeling, temporal dynamics, and probabilistic inference, which are critical for capturing the progressive and often subtle nature of drowning events.

Despite the predominance of detection-based methods, most studies addressed only a limited subset of real-world operational challenges. In particular, key issues such as crowd density and multiswimmer interactions, as well as class imbalance arising from the rarity of true drowning events, were rarely incorporated into model design or evaluation frameworks [24]. Additionally, environmental distortions, including underwater visibility challenges, were predominantly handled in studies specifically tailored to underwater settings [15,25]. This suggests that robustness in current systems is often context-specific rather than generalizable across diverse aquatic environments.

Learning Paradigms and Model Architecture

Supervised learning overwhelmingly dominated the methodological landscape, accounting for 96.7% (29/30) of the included studies. Only a single study (1/30, 3.3%) employed a semisupervised anomaly detection framework, highlighting the limited exploration of alternative learning paradigms such as unsupervised or self-supervised approaches [26].

In terms of model architecture, the YOLO family of object detection models emerged as the most frequently utilized, appearing in 23.3% (7/30) of studies. This included variants such as YOLOv5, YOLOv7, YOLOv8, and YOLOv5n, reflecting the field's strong preference for real-time, detection-oriented pipelines. Generic CNN architectures were also widely used, accounting for 16.7% (5/30) of studies. Similarly, 16.7% (5/30) of studies reported the use of deep learning approaches without clearly specifying the underlying architecture, indicating a lack of transparency in model reporting.

Other architectures were less frequently represented and included Mask R-CNN, feature pyramid networks, transfer learning–based systems, neural network–based classifiers, YOLOv8-Pose, and YPMNet, each contributing 3.3% (1/30) of the total studies [7,27]. Notably, 20% (6/30) of studies did not provide sufficient detail regarding their model architecture.

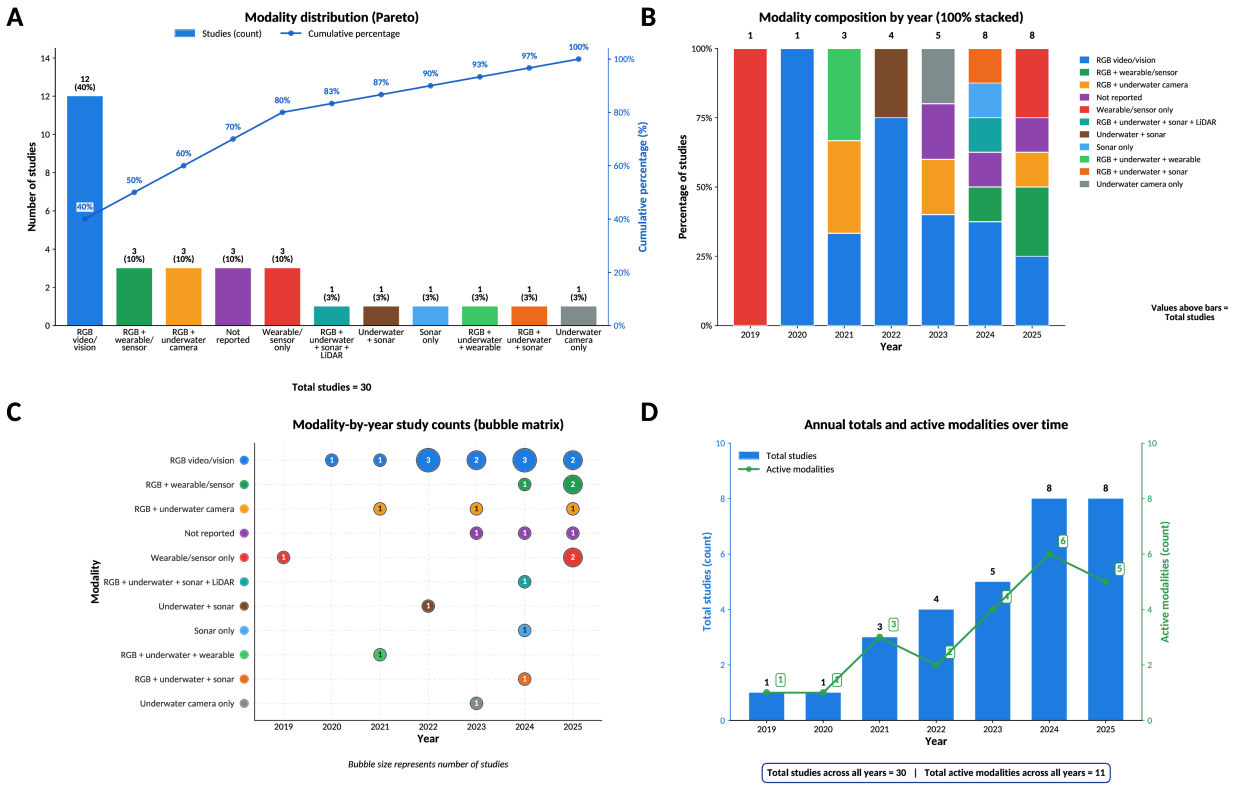
XSL•FO
RenderX

Sensing Modalities

RGB video was the most prevalent sensing modality, used either alone in 40% (12/30) of the included studies or in combination with one or more additional modalities in a further 30% (9/30) of studies (6 studies paired RGB video with either wearable/sensor data or underwater camera footage; 3 studies combined RGB video with two further modalities, underwater camera plus wearable sensing, sonar, or sonar and light detection and ranging).

Non-RGB sensing modalities were used in a further 20% (6/30) of studies: wearable or smartphone-based sensing alone (n=3), underwater camera alone (n=1), sonar alone (n=1), and underwater camera combined with sonar (n=1). The sensing modality could not be determined for the remaining 10% (n=3), as shown in Figure 4.

Figure 4. Distribution of sensing modality over time. LiDAR: light detection and ranging; RGB: red, green, and blue.



These findings indicate that the field is heavily dominated by vision-based modalities, particularly RGB and underwater imaging [13,28]. In contrast, multimodal and sensor-fusion approaches remain underexplored [29]. Notably, robustness to underwater distortion appears largely confined to systems explicitly designed for underwater environments, reinforcing the modality-specific nature of current solutions and highlighting a gap in generalizable, cross-modal detection frameworks [16].

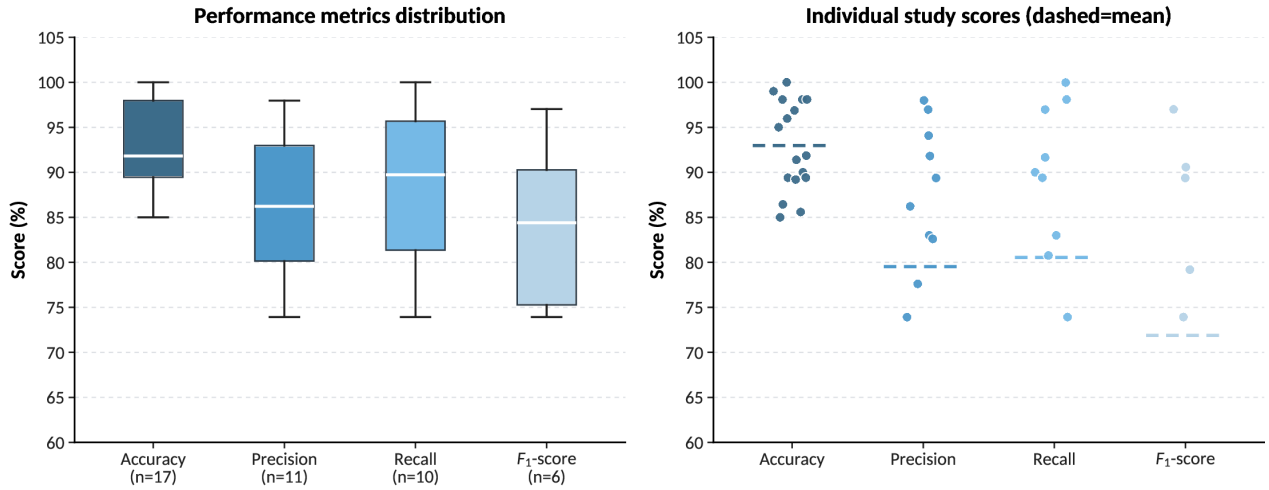
Performance Metrics

Accuracy was the most frequently reported evaluation metric, appearing in 56.7% (17/30) of the included studies. Among these studies, the mean reported accuracy was 90.9%, with a median of 90%, and values ranging from 72.17% to 100%. The highest reported accuracies included 100%, 99.4%, and 99%, indicating generally strong classification performance across studies.

Precision was reported in 36.7% (11/30) of studies, with values ranging from 73.91% to 98%**,** while recall was reported in 33.3% (10/30). In contrast, F_1-score and related classification metric negative trade-offs were reported inconsistently, limiting direct comparability across studies as shown in Figure 5.

XSL•FO
RenderX

Figure 5. Reported performance metrics distribution.



Overall, the metric profile of the field is heavily centered on accuracy, despite drowning being a rare event relative to normal swimming activity. This imbalance raises concerns about the adequacy of accuracy as a standalone metric. The limited and inconsistent reporting of precision, recall, and F_1-score constrains the ability to evaluate false positive and false negative trade-offs, which are critical factors for real-world deployment.

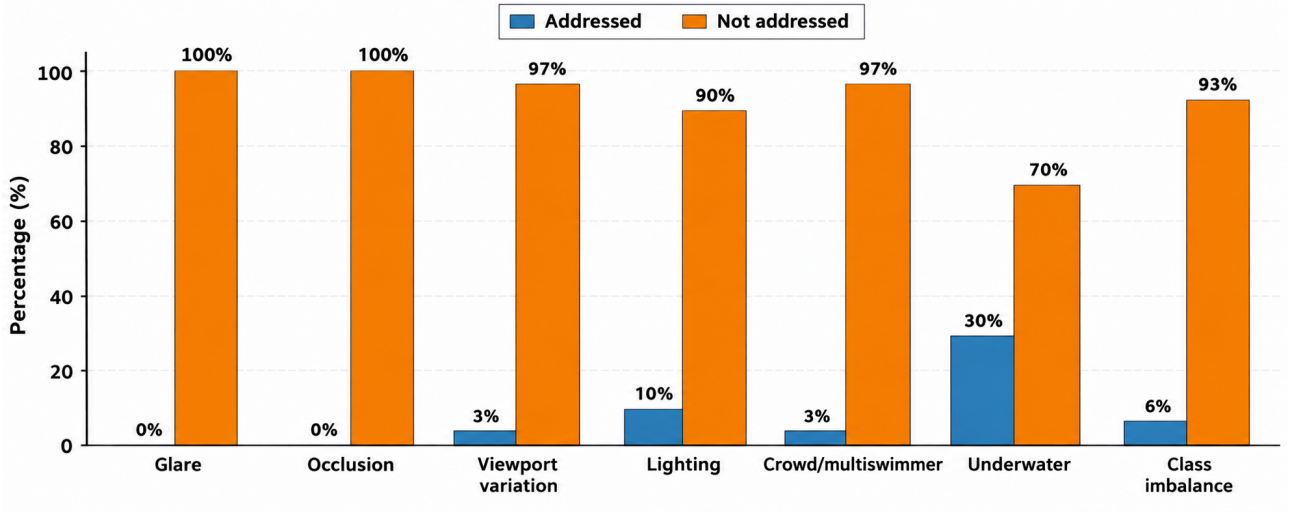
Operational performance metrics were notably scarce. Latency was reported in only 3.3% (1/30) of studies, while frame rate was reported in 20% (6/30), with values including 118.6 FPS, 31 FPS, 30 FPS, 22.9 FPS, and 5 FPS. This suggests that the literature remains largely focused on classification performance, with limited emphasis on real-time feasibility and deployment constraints.

Real-Time Capability and Validation

A total of 46.7% (14/30) of the included studies reported real-time capability. However, only 3.3% (1/30) provided empirical validation demonstrating both real-time performance and evaluation under real-world operational conditions.

The remaining 43.3% (13/30) of studies claimed or implied real-time operation without providing sufficient validation evidence, while 53.3% (16/30) made no reference to real-time capability, as shown in Figure 6.

Figure 6. Real-world challenges addressed by studies.



This gap between reported capability and validated performance is further reflected in the limited reporting of operational metrics. Latency and frame rate (FPS) were inconsistently documented across studies, restricting the ability to assess whether proposed systems meet real-time deployment requirements. These findings indicate that while real-time detection is frequently presented as a core objective, it remains largely unverified in practice.

Dataset Description and Reporting Practices

The reporting of dataset characteristics across the included studies was limited and inconsistent, raising concerns about transparency and reproducibility. A substantial majority of studies, 80% (24/30), did not clearly specify whether their datasets were real, simulated, or synthetic. Only 16.7% (5/30) explicitly reported the use of real-world data, while 3.3% (1/30) used simulated data.

XSL•FO
RenderX

This lack of clarity limits the ability to assess the ecological validity of reported performance outcomes and undermines confidence in how well these systems would generalize to real-world settings [25].

In addition, all studies reviewed relied exclusively on private datasets, with no use of publicly available benchmark datasets. The absence of shared datasets restricts comparability across studies and significantly limits reproducibility, both of which are critical for advancing the field.

Sample size reporting was also inconsistent. A total of 43.3% (13/30) of studies did not report sample size. Among those that did, there was substantial variability, ranging from as few as 17 samples to as many as 8572 images, indicating a lack of standardization in dataset scale and evaluation practice (Table 2).

XSL•FO
RenderX

Table . Summary table.

Method category	Representative approaches	Key achievements	Key limitations (as reported in studies)
Object detection (YOLO[a]-based)	YOLOv5, YOLOv7, YOLOv8, YOLOv8-Pose [23,27,29-32]	Achieved high detection performance, for example, accuracy up to 98%, high mean average precision values, demonstrated suitability for real-time detection; strong object localization capabilities	Several studies explicitly reported reliance on self-collected datasets and lack of external validation; key limitations include absence of false alarm rate reporting, limited real-world evaluation, and generalizability concerns
CNN[b]-based	CNN, Mask R-CNN[c], ResNet-based models [22,33-40]	Demonstrated strong classification performance (85%-96% accuracy); effective spatial feature learning; widely adopted for drowning recognition tasks	Studies reported lack of real-world deployment, small or simulated datasets, and absence of comprehensive evaluation metrics (eg, AUC[d], latency); limitations also include restricted generalizability across environments
Pose-based detection	YOLOv8-Pose, pose estimation methods [31,41,42]	Improved representation of human posture and motion; enhanced interpretability of drowning-related behaviors	Reported limitations include sensitivity to occlusion, difficulty handling multiple swimmers, and lack of real-world testing scenarios; studies also note limited evaluation datasets
Underwater computer vision	Underwater CNNs, YOLO-based underwater models [16,42-44]	Achieved high accuracy in underwater datasets (up to ~96%); effective handling of underwater distortion and visibility conditions	Studies reported limited applicability outside underwater settings, lack of cross-environment validation, and restricted dataset diversity; also noted absence of real-time validation in practical settings
Temporal/spatiotemporal models	Temporal anomaly detection, motion-based models [25,33]	Capture temporal dynamics of drowning behavior; suitable for detecting abnormal motion patterns over time	Reported limitations include computational complexity, lack of standardized evaluation metrics, and limited validation on real-world continuous video streams; also insufficient reporting of latency and deployment feasibility
Multimodal systems	Vision+LiDAR[e]/multimodal fusion [29]	Improved robustness through sensor fusion; enhanced contextual understanding; potential to reduce false positives	Study explicitly reported system complexity, lack of large-scale evaluation, and absence of real-world deployment validation; also noted limited benchmarking against standard vision-based approaches
Sensor-based/wearable systems	RFID[f], wearable devices, embedded sensing [33,41,45-48]	Robust to visual challenges such as glare and occlusion; enables continuous monitoring independent of camera systems	Studies reported very small sample sizes, limited scalability, and dependency on user compliance; also lack of comprehensive performance evaluation metrics and absence of real-world deployment validation
Anomaly detection/semisupervised learning	Semisupervised and anomaly-based models [13,49]	Suitable for rare-event detection; reduces reliance on labeled drowning data; captures deviations from normal swimming patterns	Reported limitations include lack of interpretability, absence of standardized evaluation benchmarks, and limited comparative analysis with supervised methods; also minimal real-world validation

[a]YOLO: You Only Look Once.

[b]CNN: convolutional neural network.

[c]Mask R-CNN: mask region-based convolutional neural network.

[d]AUC: area under the curve.

[e]LiDAR: light detection and ranging.

[f]RFID: radio-frequency identification.

XSL•FO
RenderX

Table 3 summarizes the reported accuracy across the ML architecture categories included in this review. Mean reported accuracy ranged from 85.7% for YOLO-based approaches to 100% for the RFID-based approach. However, the latter was based on a single study with a small sample size. Similarly, the R-CNN–based and underwater computer vision categories each comprised only one study. Overall, the findings indicate that a range of ML approaches have achieved high reported accuracies, although direct comparisons should be interpreted with caution because the studies were conducted using different datasets and evaluation settings. The studies differed in ML architectures, sensing modalities (vision, sonar, RFID, and multimodal systems), datasets, evaluation protocols, and reported performance metrics. In addition, accuracy was the only performance metric reported consistently across a sufficient number of studies, whereas precision, recall, F_1-score, and inference latency were reported inconsistently. Consequently, a descriptive synthesis was considered more appropriate than a statistical meta-analysis.

Table . Summary of reported accuracy by machine learning architecture.

Method category	Studies, n	Studies reporting accuracy, n	Accuracy range, %	Mean reported accuracy, %
YOLO[a]-based	7	4	72.2 - 98.1	85.7
CNN[b]-based	5	3	85.6 - 96	93.5
Pose-based	2	2	89.4 - 95	92.2
Sonar-based	2	2	82 - 99.4	90.7
Multimodal (vision+sensor)	2	2	85 - 90	87.5
R-CNN–based	1	1	94.1	94.1
RFID[c]-based	1	1	100	100
Underwater computer vision (device-based)	1	0	Not reported	—[d]
Architecture not clearly specified	9	2	89.4 - 99	94.2

[a]YOLO: You Only Look Once.

[b]CNN: convolutional neural network.

[c]RFID: radio-frequency identification.

[d]Not applicable.

Discussion

Overview of Key Findings

This review synthesizes current evidence on ML-based drowning detection and highlights both methodological progress and critical limitations. The findings show that the field is dominated by detection-based systems using supervised learning, with CNN and YOLO-based visual architectures being the most commonly adopted approaches [8,11,22]. Reported performance is generally high, particularly in terms of accuracy; however, evaluation practices are heavily skewed toward this single metric, with limited reporting of precision, recall, latency, and false alarm rates.

In addition, most systems are developed and evaluated using private datasets under controlled conditions, with minimal real-world validation and no evidence of external or cross-dataset testing [24,28]. Real-world challenges such as glare, occlusion, and crowding are rarely incorporated into evaluation frameworks, and multimodal approaches remain limited [25,33]. Together, these findings indicate that while methodological development is advancing, evidence of real-world reliability remains insufficient.

Methodological Patterns and Evaluation Limitations

The dominance of detection-oriented, supervised learning systems suggests that the field has converged around a relatively narrow set of approaches. While these methods have demonstrated strong performance, this convergence may limit innovation and reduce the ability of systems to capture the complexity of drowning behavior, which is inherently subtle and temporally dynamic [13,49,50].

Only one of the 30 included studies evaluated vision-language models (VLMs), specifically contrastive language-image pretraining and bootstrapping language-image pretraining for swimmer-behavior question answering with light detection and ranging depth data, while the remaining studies used CNN- and YOLO-based architectures. This suggests that VLMs are still underrepresented in the current literature and warrant further investigation for real-time drowning detection.

At the same time, evaluation practices reveal a critical limitation. Accuracy remains the most frequently reported metric, despite drowning detection being a rare-event problem where false positives and false negatives have serious consequences [43]. Other key metrics, including precision, recall, F_1-score, latency, and frame rate, are inconsistently reported. As a result, many systems demonstrate strong classification performance without

XSL•FO
RenderX

providing sufficient evidence of operational effectiveness, particularly in terms of timely detection and false alarm control.

Research-to-Deployment Gap and Real-World Robustness

A central finding of this review is the persistent gap between laboratory performance and real-world readiness. Although many studies claim real-time capability, only a very small number provide validation under real-world conditions, and none report external or cross-dataset evaluation. This discrepancy arises because real-time performance is typically assessed using inference speed (eg, frames per second or latency) on offline video datasets rather than end-to-end system performance, which includes alert generation, communication, and rescue response under operational conditions [13,33,45,49,51]. As a result, high inference speed alone does not guarantee effective real-world deployment. Similar challenges have been observed in other safety-critical domains, where staged validation frameworks have been adopted to bridge the gap between laboratory performance and operational use. For example, autonomous vehicle perception systems undergo progressive evaluation through simulation, controlled environments, and supervised real-world testing before deployment, while AI-based medical devices require prospective real-world validation before regulatory approval. In contrast, drowning detection research lacks an equivalent validation framework. Establishing standardized deployment-oriented evaluation protocols, such as supervised pilot studies with lifeguard oversight, would improve the translation of laboratory findings into reliable real-world systems.

This gap is further reinforced by the limited consideration of real-world challenges in model development and testing. Key disturbances such as glare, occlusion, crowding, and viewpoint variation are rarely addressed, despite being common in operational environments. Even when specific challenges are considered, they are often handled in isolation rather than as part of a comprehensive robustness framework.

In addition, the dominance of RGB and underwater vision systems reflects practical advantages but also highlights limitations in robustness, particularly when modalities are evaluated independently [24,28,33,34]. The limited adoption of multimodal approaches suggests that the field has not yet fully leveraged strategies that could improve detection across diverse and challenging conditions. One potential barrier is the lack of a standardized, publicly available multimodal benchmark dataset. Collecting synchronized RGB video, underwater imagery, sonar, and sensor data requires considerably more infrastructure and resources than single-camera studies, forcing individual research groups to develop their own datasets. Similar challenges were observed in autonomous driving before the release of synchronized multimodal benchmarks, such as the Waymo Open Dataset [52], which substantially accelerated multimodal perception research by providing shared, standardized datasets. Developing a comparable benchmark for drowning detection, even in a controlled swimming pool environment with synchronized RGB, underwater, and sensor data, could lower barriers to research, improve comparability across studies, and accelerate the development of robust multimodal detection systems.

Reproducibility, Ethics, and Future Research Directions

The review also identifies significant challenges related to reproducibility and data transparency. All included studies relied on private datasets, and many did not clearly report dataset characteristics such as sample size or data type [28,49,51]. This lack of standardization limits comparability across studies and hinders independent validation, indicating that advances in model development have not been matched by progress in evaluation infrastructure.

Privacy and ethical considerations also remain underaddressed, despite their importance for real-world deployment. Only a minority of studies explicitly discussed privacy concerns, even though many systems rely on continuous visual monitoring in sensitive environments [25]. Nonvisual sensing approaches may offer more privacy-preserving alternatives and should be explored further.

Future research should prioritize the development of standardized evaluation frameworks, public benchmark datasets, and external validation protocols. In addition, greater emphasis is needed on comprehensive performance reporting, incorporation of real-world challenges, multimodal system design, and privacy-aware methodologies. Addressing these gaps will be essential for translating current advances into reliable and deployable drowning detection systems. Beyond the single VLM-based study identified, general-purpose VLMs such as Gemini, MiniCPM-V, and the Qwen-VL series remain largely unexplored for drowning detection. Given their strong zero- and few-shot capabilities, they represent a promising direction for future research. However, their suitability for real-time, safety-critical deployment remains unverified due to their high computational demands and the lack of benchmarking on inference latency and false alarm rates.

Conclusion

This review demonstrates that while ML has achieved remarkable progress in drowning detection, current systems remain largely experimental and fall short of real-world reliability. The widespread reliance on curated datasets, limited external validation, and inconsistent evaluation practices exposes a critical gap between reported performance and practical deployment. Bridging this gap demands a shift toward rigorous validation, standardized evaluation frameworks, and robust, privacy-aware, real-time systems. Advancing in this direction is not merely a technical necessity but a public health imperative, essential for translating promising innovations into scalable, life-saving drowning prevention solutions.

XSL•FO
RenderX

Acknowledgments

The authors would like to acknowledge the support and guidance received during the course of this research. JMIR Publications provided article processing fee support for the publication of this article.

Funding

This research was funded in part by the Advancing Public Health Research in Eastern Africa through Data Science Training (APHREA-DST) project (Grant No. 1U2RTW012123), with support from the U.S. National Institutes of Health (NIH).

Authors' Contributions

MC was responsible for the conceptualization of the systematic review and contributed significantly to the drafting of the introduction, methodology, results, discussion, and conclusion sections. MC also prepared the original manuscript and took part in its validation. IO, EA, and TK contributed through manuscript review and editing, provided oversight for the study, and assisted in validating the research and advising.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Databases and search terms used in the literature search.
[DOCX File, 16 KB - rehab_v13i1e100939_app1.docx]

Multimedia Appendix 2
PICO (Problem, Intervention, Comparison, Outcome) framework of the research questions.
[DOCX File, 14 KB - rehab_v13i1e100939_app2.docx]

Multimedia Appendix 3
PROBAST: assessment of the risk of bias for each study.
[DOCX File, 19 KB - rehab_v13i1e100939_app3.docx]

Checklist 1
PRISMA checklist.
[PDF File, 191 KB - rehab_v13i1e100939_app4.pdf]

Checklist 2
TRIPOD Checklist
[DOCX File, 20 KB - rehab_v13i1e100939_app5.docx]

References

1. Global status report on drowning prevention 2024. : World Health Organization; 2024 URL: https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/global-report-on-drowning-prevention [accessed 2026-07-12]
2. Franklin RC, Peden AE, Hamilton EB, et al. The burden of unintentional drowning: global, regional and national estimates of mortality from the Global Burden of Disease 2017 Study. Inj Prev 2020 Oct;26(Supp 1):i83-i95. [doi: 10.1136/injuryprev-2019-043484] [Medline: 32079663]
3. Janson S. Drowning is a complex but preventable cause of child mortality. Acta Paediatr 2021 Jul;110(7):1981-1982. [doi: 10.1111/apa.15674] [Medline: 33301198]
4. Tyler MD, Richards DB, Reske-Nielsen C, et al. The epidemiology of drowning in low- and middle-income countries: a systematic review. BMC Public Health 2017 May 8;17(1):413. [doi: 10.1186/s12889-017-4239-2] [Medline: 28482868]
5. Jegham I, Ben Khalifa A, Alouani I, Mahjoub MA. Vision-based human action recognition: an overview and real world challenge. Forensic Sci Int Digit Investig 2020 Mar;32:200901. [doi: 10.1016/j.fsidi.2019.200901]
6. Dulhare UN, Hussam Ali M. Underwater human detection using Faster R-CNN with data augmentation. Mater Today 2023;80:1940-1945. [doi: 10.1016/j.matpr.2021.05.653]
7. Gao Q, Pan J, Zhang X, Chen X. Design of drowning detection method based on YOLOv8. IEEE 6th Int Conf Civil Aviat Saf Inf Technol (ICCASIT) :1635-1638. [doi: 10.1109/ICCASIT62299.2024.10828012]
8. Wei E, Ren S. Drowning detection based on YOLOv8 improved by GP-GAN augmentation. Presented at: The First Tiny Papers Track at ICLR 2023; May 5, 2023 URL: https://dblp.org/rec/conf/iclr/WeiR23.html?utm_source [accessed 2026-07-12]

XSL•FO
RenderX

9. Bai B, Chen L, Li X. Improved YOLOv7 fusion detection line for swimming pool drowning detection. IEEE 16th Int Conf Electron Meas Instrum (ICEMI) 2023. [doi: 10.1109/ICEMI59194.2023.10270676]
10. Zhou Y, Wang J, Li L, Wang W. YOLOv5-FRNet: a real-time drowning detection method with multi-model cascade. J Big Data Comput ;3(2):58-64. [doi: 10.62517/jbdc.202501210]
11. Hayat MA, Yang G, Iqbal A. Mask R-CNN based real time near drowning person detection system in swimming pools. 2022 Mohammad Ali Jinnah Univ Int Conf Comput (MAJICC) 2022. [doi: 10.1109/MAJICC56935.2022.9994135]
12. Laxton V, Howard CJ, Guest D, Crundall D. Intense classification training to increase the detection of drowning swimmers. Appl Cogn Psychol 2023;37(1):189-199. [doi: 10.1002/acp.4038]
13. Shatnawi M, Albreiki F, Alkhoori A, Alhebshi M. Deep learning and vision-based early drowning detection. Information 2023;14(1):52. [doi: 10.3390/info14010052]
14. He Q, Mei Z, Zhang H, Xu X. Automatic real-time detection of infant drowning using YOLOv5 and Faster R-CNN models based on video surveillance. J Soc Comput 2023;4(1):62-73. [doi: 10.23919/JSC.2023.0006]
15. Zhang W, Chen L, Shi J. A pool drowning detection model based on improved YOLO. Sensors 2025;25(17):5552. [doi: 10.3390/s25175552] [Medline: 40942980]
16. Alnajjar K, Ramadan OH. AI-enhanced drowned-detection system using information theory for rapid rescue operations. AIP Conf Proc 2025;3320:070001. [doi: 10.1063/5.0286144]
17. Advances in machine learning paradigms for reliable swimming pool drowning detection: A systematic review of methods and implementations. OSF. URL: https://osf.io/hp6au/ [accessed 2026-09-23]
18. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Serv Res 2014 Nov 21;14:579. [doi: 10.1186/s12913-014-0579-0] [Medline: 25413154]
19. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
20. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019 Jan 1;170(1):51-58. [doi: 10.7326/M18-1376] [Medline: 30596875]
21. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015 Jan 6;162(1):W1-73. [doi: 10.7326/M14-0698] [Medline: 25560730]
22. Kumari N, Yadav R, Sisodia A, Mishra A. Integrated real-time pool operation and maintenance platform. Int J Comput Tech ;12 [FREE Full text]
23. Amer DA, Ibrahim NY, Ibrahim IK, Mohamed AM, Soliman SA. Intelligent eyes on water: YOLOv11-based real-time drowning detection system. J Supercomput 2025;81(12):1242. [doi: 10.1007/s11227-025-07732-7]
24. Song Q, Yao B, Xue Y, Ji S. MS-YOLO: a lightweight and high-precision YOLO model for drowning detection. Sensors (Basel) 2024 Oct 30;24(21):6955. [doi: 10.3390/s24216955] [Medline: 39517851]
25. Yang R, Wang K, Yang L. An improved YOLOv5 algorithm for drowning detection in the indoor swimming pool. Appl Sci 2024;14(1):200. [doi: 10.3390/app14010200]
26. Shatnawi M, Albreiki F, Alkhoori A, Alhebshi M, Shatnawi A. Advances and challenges in automated drowning detection and prevention systems. Information 2024;15(11):721. [doi: 10.3390/info15110721]
27. Chang L, Zhang Z. Drowning detection based on YOLOv8-pose and human pose estimation. 2024 10th Int Conf Syst Inform (ICSAI) 2024. [doi: 10.1109/ICSAI65059.2024.10893825]
28. He T, Ye X, Wang M. An improved swimming pool drowning detection method based on YOLOv8. 2023 IEEE 7th Inf Technol Mechatron Eng Conf (ITOEC) 2023:835-839. [doi: 10.1109/ITOEC57671.2023.10291322]
29. Khan MZ, Gaddam A, Thiruvady D, Story A. Swim VQA: a multimodal approach to drowning detection by fusing VQA and LiDAR. Int Conf Sens Technol (ICST) 2024:1-6. [doi: 10.1109/ICST62759.2024.10992006]
30. Wang W, Wang Z, Wang Q, Wang Y. Intelligent monitoring drowning early warning system. Proc 6th Int Conf Adv Image Process (ICAIP) 2022:91-96. [doi: 10.1145/3577117.3577120]
31. Cepeda-Pacheco JC, Domingo MC. Deep learning and 5G and beyond for child drowning prevention in swimming pools. Sensors (Basel) 2022 Oct 10;22(19):7684. [doi: 10.3390/s22197684] [Medline: 36236782]
32. Sowmya P, Shettigar R, Nishanth N, Kumar D. AquaRescue: intelligent drowning detection and swift rescue system using integrated approach. 4th Int Conf Technol Adv Comput Sci (ICTACS) 2024:1455-1459. [doi: 10.1109/ICTACS62700.2024.10840686]
33. Liu T, He X, He L, Yuan F. A video drowning detection device based on underwater computer vision. IET Image Process 2023 May;17(6):1905-1918. [doi: 10.1049/ipr2.12765]
34. Karimi Nejad M, Ghahreman Tabrizi K, Sharifian E, Rahimi A. Identification and prioritization of Internet of Things applications in indoor swimming pools of Kerman province [Article in Persian]. Inf Technol Sport 2025 Mar;2(1):103-128. [doi: 10.22091/its.2025.13385.1019]
35. Hasan S, Joy J, Ahsan F, Khambaty H, Agarwal M, Mounsef J. A water behavior dataset for an image-based drowning solution. 2021 IEEE Green Energy Smart Syst Conf (IGESSC) 2021. [doi: 10.1109/IGESSC53124.2021.9618700]

XSL•FO
RenderX

36. Iampanit T, Somprakij P, Siriwanna N, Iampanit N, Jirakittayakorn N. A novel drowning detection system and its performance analysis. 2025 22nd Int Conf Electr Eng Electron Comput Telecommun Inf Technol (ECTI-CON) 2025. [doi: 10.1109/ECTI-CON64996.2025.11100770]
37. Vestnikov R, Stepanov D, Bakhshiev A. Development of neural network algorithms for early detection of drowning in swimming pools. 2023 Int Conf Ind Eng Appl Manuf (ICIEAM) 2023:820-824. [doi: 10.1109/ICIEAM57311.2023.10139153]
38. Urruchi C, Cervantes-Chauca D, Huamanchahua D. Proposal of a swimming pool drowning detection system using cameras and raspberry pi based on machine learning. 2022 2nd Int Conf Robot Autom Artif Intell (RAAI) 2022:178-181. [doi: 10.1109/RAAI56146.2022.10092956]
39. Arriesgado GNC, Monserate GS, Uytengsu BIY. Enhancing water safety in the Philippines: early detection of drowning patterns through wearable devices and machine learning algorithms. 2025 IEEE Conf Comput Appl (ICCA) 2025. [doi: 10.1109/ICCA65395.2025.11011077]
40. Belshia Jebamalar G, Adlin LJA, Saranya S, Mabel Rose R, Nickel S, et al. To detect active drowning using deep learning algorithms. 2023 9th Int Conf Smart Struct Syst (ICSSS) 2023:1-7. [doi: 10.1109/ICSSS58085.2023.10407065]
41. Yang G, Yi Q, Wang Z, Liu T. Water motion safety recognition and monitoring system based on the wireless signal transmission. IOP Conf Ser: Earth Environ Sci 2020;440(5):052040. [doi: 10.1088/1755-1315/440/5/052040]
42. Jian JX, Wang CM. Deep learning used to recognize swimmers drowning. 2021 IEEE/ACIS 22nd Int Conf Softw Eng Artif Intell Netw Parallel/Distrib Comput (SNPD) 2021. [doi: 10.1109/SNPD51163.2021.9704884]
43. Azeem MM, Al-Kaabi A, Al-Jaber S. A distributed edge-intelligent IoT network framework for hybrid vision and physiological drowning detection. SSRN. Preprint posted online on Dec 24, 2025. [doi: 10.2139/ssrn.5941677]
44. Sukkuea A, Ubonkan S, Phetsri T, Limprasert I, Akkajit P, Xing X. Comparative analysis of underwater drowning detection using convolutional neural networks and transfer learning. Proc 4th Int Conf Front Artif Intell Mach Learn (FAIML) 2025. [doi: 10.1145/3748382.3748385]
45. Fan L, Wang Y, Li K, Wei X, Zhao N, Huang Y. A drowning detection model by integrating multi-scale enhanced receptive fields with the YOLOv8 framework. 2025 IEEE 8th Inf Technol Mechatron Eng Conf (ITOEC) 2025. [doi: 10.1109/ITOEC63606.2025.10969024]
46. Dehbashi F, Ahmed N, Mehra M, et al. SwimTrack:drowning detection using RFID. SIGCOMM Posters and Demos '19: Proc ACM SIGCOMM Conf Posters Demos 2019:161-162. [doi: 10.1145/3342280.3342345]
47. He L, Hou H, Yan Z, Xing G. Demo abstract: an underwater sonar-based drowning detection system. 2022 21st ACM/IEEE Int Conf Inf Process Sensor Netw (IPSN) 2022. [doi: 10.1109/IPSN54338.2022.00047]
48. Hou H, Zheng B, Wu P, Xing G, Yan Z. Poster: AquaGuard: a sonar-based pool monitoring system. Proc 22nd ACM Conf Embed Netw Sens Syst (SENSYS '24) 2024 Nov 4. [doi: 10.1145/3666025.3699426]
49. Handalage U, Nikapotha N, Subasinghe C, Prasanga T, Thilakarthna T, Kasthurirathna D. Computer vision-enabled drowning detection system. 2021 3rd Int Conf Adv Comput (ICAC) 2021. [doi: 10.1109/ICAC54203.2021.9671126]
50. Alotaibi A. Automated and intelligent system for monitoring swimming pool safety based on the IoT and transfer learning. Electronics 2020;9(12):2082. [doi: 10.3390/electronics9122082]
51. Ambre D, Sureshkumar H, Trivedi P. Cygnus: a vision-based drone system for drowning detection using IoT. 2024 2nd Int Conf Recent Trends Microelectron Autom Comput Commun Syst (ICMACC) 2024:184-189. [doi: 10.1109/ICMACC62921.2024.10894649]
52. Sun P, Kretzschmar H, Dotiwalla X, et al. Scalability in perception for autonomous driving: Waymo Open Dataset. IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR) 2020:2446-2454. [doi: 10.1109/CVPR42600.2020.00252]

Abbreviations

CLAIM: Checklist for Artificial Intelligence in Medical Imaging
CNN: convolutional neural network
ML: machine learning
PICO: Problem, Intervention, Comparison, and Outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROBAST: Prediction Model Risk of Bias Assessment Tool
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies 2
R-CNN: region-based convolutional neural network
RFID: radio-frequency identification
RGB: red, green, and blue
RQ: review question
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
VLM: vision-language model
YOLO: You Only Look Once

XSL•FO
RenderX

Edited by S Munce; submitted 10.May.2026; peer-reviewed by T Zhang, X Shen; revised version received 23.Aug.2026; accepted 28.Aug.2026; published 03.Nov.2026.

Please cite as:
Chepkorir M, Orowe I, Abade E, Kamanu T
Advances in Machine Learning Paradigms for Reliable Swimming Pool Drowning Detection: Systematic Review of Methods and Implementations
JMIR Rehabil Assist Technol 2026;13:e100939
URL: https://rehab.jmir.org/2026/1/e100939
doi:10.2196/100939

© Mercy Chepkorir, Idah Orowe, Elisha Abade, Timothy Kamanu. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 3.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Methods and Effectiveness of Auditory-Cognitive Training for Older Adults: Scoping Review and Meta-Analysis

Mengyuan Li[1*], PhD; Chifen Ma[2*], PhD; Piers Dawes[3], PhD; Yanni Yang[1], PhD; Weiwei Liu[1], PhD

[1]School of Nursing, Army Medical University, No.30 Gaotanyanzhengjie Street, Shapingba District, Chongqing, China

[2]College of Health Services and Management, Xuzhou, China

[3]University of Queensland Centre for Hearing Research (CHEAR), School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Queensland, Australia

[*]these authors contributed equally

Corresponding Author:
Mengyuan Li, PhD
School of Nursing, Army Medical University, No.30 Gaotanyanzhengjie Street, Shapingba District, Chongqing, China

Abstract

Background: Auditory-cognitive training (ACT) combines auditory and cognitive exercises to improve listening and cognition in individuals with hearing difficulties; yet, it lacks standardized implementation and consistent evidence.

Objective: This mixed-evidence synthesis aimed to map ACT's implementation characteristics and summarize its reported short-term and maintenance effects on trained and nontrained cognitive and auditory outcomes, including hearing function in daily life.

Methods: The review followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines and the Arksey and O'Malley scoping review framework. Databases searched included Embase, the Cochrane Library, PubMed, and APA PsycNet from inception to July 30, 2026. Articles were included or excluded based on specific criteria, and included articles were evaluated for risk of bias.

Results: ACT showed significant pooled effects on speech-in-noise performance (mean difference [MD]=−1.23 dB, 95% CI −2.35 to −0.10; *P*=.03), attention (standardized mean difference [SMD]=0.50, 95% CI 0.23-0.77; *P*<.001), memory (SMD=1.08, 95% CI 0.25-1.92; *P*=.01), and global cognition (MD=3.27, 95% CI 0.27-6.26; *P*=.03). Effects on pure-tone audiometry, subjective hearing function, working memory, and executive function were limited and nonsignificant. Current evidence is limited by unstandardized training dosage/duration, limited linguistic generalizability across non-English languages, and insufficient assessment of patient-reported real-life listening outcomes.

Conclusions: ACT may improve selected auditory and cognitive outcomes in older adults, particularly speech-in-noise perception, and attention. The evidence remains limited by substantial heterogeneity, a small number of studies, and low certainty for most outcomes. Standardized protocols, clinically meaningful outcome measures, and longer follow-up are needed to determine its clinical effectiveness and long-term benefits.

(JMIR Rehabil Assist Technol 2026;13:e101632) doi:10.2196/101632

KEYWORDS

hearing loss; auditory-cognitive training; older adults; cognitive decline; auditory training

Introduction

Hearing function is critical for daily social connections and often declines with age. Globally, 1.57 billion people have hearing loss (HL), and this number is projected to reach 2.17 billion by 2040 [1]. In China, over 70% of older adults are affected [2]. Beyond auditory impairment, age-related HL is a modifiable risk factor for adverse outcomes in older adults, including accelerated cognitive decline, dementia, social isolation, depression, and reduced quality of life [3-6]. Severe HL increases dementia risk 5-fold compared to normal hearing [7].

Hearing aids (HAs) are the primary treatment; however, only 7.7% of patients with HL use them [8], mainly due to persistent speech-in-noise (SIN) difficulties [9]. SIN perception relies on auditory and cognitive functions (eg, attention and working memory), which synergistically affect listening in older adults [9,10]. This has shifted interventions from "bottom-up" auditory training to integrated auditory-cognitive training (ACT) [11].

The theoretical underpinnings of ACT are rooted in neuroplasticity and the Ease of Language Understanding (ELU) model. Neuroplasticity theory suggests that ACT induces neural remodeling to enhance auditory-cognitive coordination [12], while the ELU model highlights working memory's role in

XSL•FO
RenderX

resolving ambiguous auditory input—challenging for older adults [13]. Unlike HAs, which only amplify sound, ACT addresses cognitive deficits underlying SIN perception.

Preliminary ACT studies in older adults with HL show dual benefits for SIN perception [14-17] and cognitive functions (eg, memory, attention, and processing speed) [14,15,18]. However, findings are heterogeneous due to variable study designs, interventions, and outcomes, and the magnitude and consistency of the effects are unclear.

No consensus has yet been established regarding the optimal implementation of ACT. Existing studies vary considerably in terms of training domains, intervention dosage, and delivery modalities. For example, training duration ranges widely, from as short as 2 weeks [17] to as long as 8 weeks [19]. The targeted training domains are heterogeneous: some studies focus on auditory or speech perception and recognition [14,18], whereas others adopt integrated approaches combining auditory training with cognitive components such as attention, working memory, or executive function [20]. Such variability hinders the quantification of effect sizes, limits cross-study comparisons, and impedes the development of standardized, replicable ACT protocols.

Thus, this study systematically maps ACT implementation characteristics (modalities, dosage, and participants' hearing status) and quantitatively synthesizes its effects on auditory and cognitive outcomes via meta-analysis. It aims to provide a comprehensive overview of current ACT interventions and evidence-based guidance for future design and clinical applications.

This review aimed to map existing literature and characterize ACT interventions for older adults with or without HL. It synthesized ACT implementation characteristics and summarized evidence for effects on hearing and cognitive outcomes. Four research questions guided the review:

1. What are the hearing and cognitive profiles of older adults for whom ACT is applicable in existing studies?
2. How should ACT be designed and delivered (eg, content, modality, and difficulty adjustment) for older adults?
3. What is the evidence for the effects of ACT on hearing and cognitive outcomes, and what are the pooled effect sizes from meta-analyses?
4. What key limitations and research gaps exist in the current ACT literature?

Methods

Overview

This review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines, PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines, and the Arksey and O'Malley scoping review framework [21-23]. The protocol, including synthesis methods, was reregistered on Open Science Framework (OSF) before data collection [24]. The PRISMA 2020 and PRISMA-ScR checklists are provided in Checklists 1 and 2 [25].

Search Strategy

A systematic literature search was conducted in Embase, the Cochrane Library, PubMed, and APA PsycNet, supplemented by manual screening of the reference lists of included studies and relevant reviews. The core search concepts included older adults, HL, auditory training, and cognitive training, with database-specific subject headings, keywords, Boolean operators, and search syntax adapted for each database. No restrictions on study design were applied during the initial search. Studies published in English and available through the final search date of July 30, 2026, were considered eligible. We also manually screened the reference lists of the included studies and relevant reviews to identify potentially eligible studies that may not have been retrieved through the database searches. The date of each database search and the complete database-specific search strategies, including all search terms, field specifications, and database-specific operators, are provided in Multimedia Appendix 1.

Eligibility Criteria

Studies were eligible for inclusion in the scoping review if they met the following criteria: participants were older adults, with or without HL. The review focused on ACT interventions, defined as those that explicitly integrate auditory processing training with cognitive components (eg, attention, working memory, executive function, or cognitive control) and have the primary aim of improving speech perception, auditory processing, or communication abilities in complex acoustic environments. No restrictions were applied regarding study design. Studies published in English before August 2026 were considered eligible.

For the subsequent meta-analysis, eligibility criteria were further refined. Only randomized controlled trials (RCTs) and quasi-experimental studies were included. Eligible studies were required to compare ACT interventions with a control condition, including education-based controls, waitlist controls, usual care, or traditional auditory rehabilitation. Studies were included if they reported quantitative outcomes related to auditory function (eg, speech perception, SIN performance, and auditory processing) and/or cognitive function (eg, attention, executive function, working memory, or global cognition), and provided sufficient data for effect size calculation (eg, means, SDs, and sample sizes).

Studies were excluded if they did not involve an explicit integrative auditory-cognitive intervention, focused solely on auditory training or on cognitive training without integration, lacked full-text availability (eg, protocols and conference abstracts), or did not provide sufficient data for meta-analysis.

Extracting and Charting the Data

Data were extracted independently by 2 reviewers (ML and CM) using a standardized form, with discrepancies resolved by consensus. Extracted data included study characteristics, participant demographics, hearing and cognitive status, ACT design, delivery mode, dosage, outcome measures, and key results. These data were synthesized descriptively to provide a comprehensive overview of the characteristics, intervention designs, and outcomes of ACT studies in older adults.

XSL•FO
RenderX

Quality of Evidence

Two reviewers independently assessed the methodological quality of the included studies. For RCTs, the revised Cochrane risk-of-bias tool (RoB 2) [26] was used to evaluate bias across 5 domains: randomization process, deviations from intended interventions, missing outcome data, measurement of outcomes, and selection of the reported result. For quasi-experimental studies, the Joanna Briggs Institute (JBI) critical appraisal checklist for quasi-experimental studies was applied [27]. Disagreements between reviewers were resolved through discussion or consultation with a third reviewer. Among the 24 included publications, 3 articles reported the same underlying study populations as other publications and were considered multiple reports of the same study. Therefore, risk-of-bias assessment was conducted at the study level, including 16 independent RCTs assessed using RoB 2 and 5 quasi-experimental studies assessed using the JBI checklist.

Synthesis Methods

The synthesis methods were preregistered before data collection on the OSF [24]. Findings were integrated through qualitative and quantitative synthesis. Content analysis was used to summarize nonquantitative findings, and a Sankey diagram was used to visualize outcome patterns. Meta-analyses were conducted in R (version 4.4.2; R Foundation for Statistical Computing) using the meta package. For continuous outcomes, intervention effects were calculated using change scores from baseline to postintervention. Mean differences (MDs) were used when outcomes were measured on the same scale, whereas standardized mean differences (SMDs; Hedges *g*) were used when the same construct was assessed using different scales. When change-score SDs were not directly reported, they were calculated from baseline and postintervention SDs assuming a correlation of 0.50. For studies reporting within-group change scores and Cohen *d* but not change-score SDs, the SDs were derived from the reported mean changes and Cohen *d* and were incorporated into the corresponding meta-analyses. Fixed-effects models were applied when heterogeneity was low (I^2<50%), and random-effects models were applied when heterogeneity was moderate to high (I^2≥50%) [28]. Heterogeneity was assessed using Cochran *Q* test and I^2. Forest plots were used to present pooled effect estimates and 95% CIs; the Egger test and funnel plots were used to assess potential publication bias when sufficient studies (≥10) were available.

Results

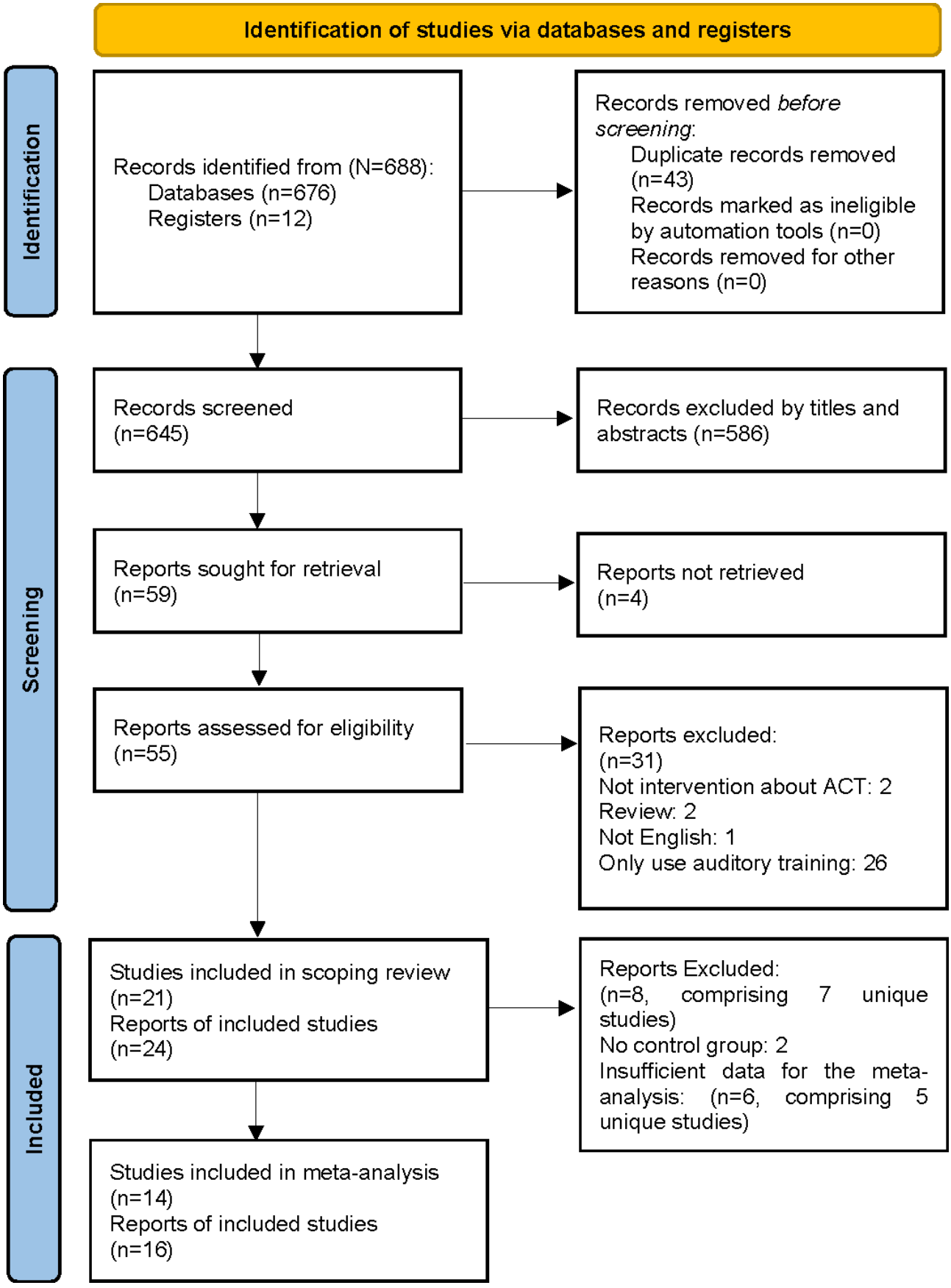
Overview

A total of 688 records were identified. After removing 43 duplicates, 645 records were screened, of which 586 were excluded based on title and abstract. A full-text evaluation of the remaining 59 articles resulted in a further 35 exclusions (non-ACT interventions, reviews, unavailable texts, non-English papers, and single-modal auditory training), yielding 24 eligible articles. Among these, 3 articles were identified as multiple reports of the same intervention study, involving the same participants, group allocation, and intervention protocol. Therefore, the 24 included articles represented 21 independent studies. Eight studies were excluded from the meta-analysis for lack of data. Therefore, 16 reports corresponding to 14 studies were included in the meta-analysis. The screening process is shown in Figure 1.

The 21 included studies were published between 2007 and 2026, with 9 published in the past 5 years. Eleven were conducted in the United States, 2 in the United Kingdom, and 1 each in Switzerland, Japan, Singapore, India, Germany, Turkey, China, and Malaysia. Sixteen studies used RCT designs, and 5 used quasi-experimental designs. Sample sizes of intervention groups ranged from 11 to 63, and participant ages varied from 40 to 89 years. Participant demographic characteristics are summarized in Table 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. ACT: auditory-cognitive training.



XSL•FO
RenderX

Table . Demographic characteristics of included studies.

Study ID	Country	Study design	Sample size (n)		Age (years)	Gender	Hearing status	Cognition status
			T group[a]	Ctrl group[b]				
Abrams et al [29], 2015	United States	RCT[c]	15	14	T group: mean 65.6 (SD 5.53); Ctrl group: mean 61.8 (SD 8.4)	T group: 11 men, 4 women; Ctrl group: 8 men, 6 women	Bilateral symmetrical sensorineural HL[d]; include new and experienced HA[e] users	Not reported
Anderson et al [14], 2013	United States	RCT	29	29	Mean 64.11 (SD 5.78)	Women: 35, men: 23	Normal hearing: PTA[f] (6 Hz) <25 dB (n=28); HL: PTA (8 Hz) <80 dB (n=30); no HA users	Normal IQ, without neurologic conditions
Anderson et al [15], 2013	United States	RCT	35	32	Mean 63.0 (SD 3.7)	Women: 38, men: 29	PTA (4 kHz) ≤45 dB; no HA users	Normal IQ, MoCA[g]≥23
Faust et al [18], 2020	United States	RCT	37	37	Range: 68.0-79.8	Not reported	No reported HL; no HA users	MMSE[h]≥24
Fisher et al [16,30], 2025	United States	RCT with pseudo-randomization	Low cognitive group: 25; high cognitive group: 12	None	Range: 60-83	Women: 19, men: 4	Normal/near-normal: PTA (250 - 4000 Hz) ≤25 dB HL (both ears); no HA users	MoCA≥26
Frei and Giroud [17], 2025	Switzerland	RCT, crossover design	39	None	Mean 70 (SD 4.92)	Women: 19, men: 20	PTA (0.5, 1, 2, 4 kHz) ≤40 dB; no HA users	No preexisting neurological diseases, cognition not reported
Frost et al [31], 2025	United Kingdom	RCT	22	21	Range: 40-75	Women: 28, men: 15	Mild subjective hearing concern; objective: normal/mild-to-moderate HL; no HA users	Mild subjective memory concern, without dementia diagnosis, not suspected to have moderate-to-severe hearing or cognitive impairment
Ferguson et al [32], 2014	United Kingdom	RCT, crossover design	23	DT[i] group: 21	Range: 50-74	Women: 15, men: 29	Symmetrical mild sensorineural HL: PTA (0.5 - 4 kHz, better ear) 21 - 40 dB HL; no HA users	Not reported

XSL•FO
RenderX

Study ID	Country	Study design	Sample size (n)		Age (years)	Gender	Hearing status	Cognition status
			T group[a]	Ctrl group[b]				
Gundmi et al [33], 2026	India	RCT	40	40	Range: 45 - 65	Women: 34, men: 46	Bilateral mild-to-moderate sensorineural HL (PTA 26 - 70 dB HL); all HA users	MMSE≥24
Henderson Sabes [34], 2007; Sweetow and Henderson Sabes [35], 2006	United States	RCT, crossover design	38	27	Mean 64.7 (SD 14.9)	Not reported	56 with HL; 9 with speech understanding difficulties (5 had normal PTA); HA users: n=56	Not reported
Kawata et al [20], 2022	Japan	RCT	AC[j]: 13; auditory: 13; cognitive: 14	12	Mean 68.07 (SD 4.14)	Women: 43, men: 14	Normal-to-mild HL; no HA users	MoCA≥26
Kucuk Ceyhan et al [36], 2022	Turkey	Parallel-group quasi trial	17	16	T group: mean 72.17 (SD 6.94); Ctrl group: mean 71.75 (SD 6.81)	T group: 8 men, 9 women; Ctrl group: 10 men, 6 women	Age-related bilateral symmetrical sensorineural HL, thresholds ≤40 dB HL at 500 Hz, ≤80 dB HL at 4000 Hz; all HA users	MMSE≥24
Lai et al [37], 2023	Singapore	Three-group quota-sampled parallel trial	20	Active: 21; no-training: 23	Experimental: mean 66.7 (SD 4.55); active Ctrl: mean 66.2 (SD 5.79); no-training Ctrl: mean 67.26 (SD 6.11); all aged 60 - 87	Not reported	Mild-to-moderate age-related HL, 4 frequency PTA average 30 - 34 dB HL; no HA users	MMSE≥25
Ockelmann et al [38], 2025	Germany	RCT	27	27	Range: 65-82	Women: 30, men: 30	Mild-to-moderate HL: PTA 20 - 60 dB; HA users: n=21	MoCA≥23
Olson et al [39], 2013	United States	Prospective repeated-measures quasi-experiment	14	7	Overall: 66 - 68; new training, mean 66 (SD 9.6); experienced, mean 68 (SD 6.9); new Ctrl, mean 66 (SD 9.3)	New training: 1 man, 7 women; experienced: 10 men, 4 women; new control: 7 men	Mild-to-severe sensorineural HL; include new and experienced HA users	Not reported
Rao et al [40], 2017	United States	Parallel-group quasi-experiment	11	11	T group: range: 60 - 85, mean 69 (SD 7.3); Ctrl group: range: 49 - 85, mean 65 (SD 8.9)	T group: 6 men, 5 women; Ctrl group: 6 men, 5 women	Mild-to-moderate bilateral sensorineural HL; thresholds ≤40 dB≤1500 Hz, ≤70 dB at high frequencies; all new HA users	MMSE=27 - 30

XSL•FO
RenderX

Study ID	Country	Study design	Sample size (n)		Age (years)	Gender	Hearing status	Cognition status
			T group[a]	Ctrl group[b]				
Rishiq et al [41], 2016	United States	RCT	12	12	T group: range; 62 - 81; Ctrl group: range: 51 - 84	T group: 6 men, 6 women; Ctrl group: 8 men, 4 women	Mild-to-moderate high-frequency sensorineural HL; all HA users	MMSE=27 - 30
Saunders et al [42], 2016; Smith et al [43], 2016	United States	RCT	LACE-C[k]: 61; LACE-DVD: 63	Placebo audiobook training: 69; educational counseling: 70	Mean 68.6 (SD 7.7)	Not reported	Symmetrical mild-to-moderate sensorineural HL; all HA users	Age and educationally appropriate scores on MMSE
Whitton et al [44], 2017	United States	RCT	13	11	Range: 50 - 89, mean 70	Women: 32, men: 16	Mild-to-severe bilateral sensorineural HL; all long-term users	MoCA>25, WASI[l] full IQ>80, GDS[m]<6
Yusof et al [19], 2019	Malaysia	Quasi-experiment	40	36	Range: ≥60	Women: 57, men: 19	Symmetrical: PTA (0.5 - 4.0 kHz, better ear) ≤40 dB HL; no HA users	NC[n] and NCI[o] group
Zhao et al [45], 2026	China	Pilot RCT	30	30	Mean 67.65 (SD 4.78)	Women: 15, men: 45	Mild-to-moderate age-related HL, PTA 20 - 50 dB; no HA users	MoCA≥26

[a]T group: training group.

[b]Ctrl group: control group.

[c]RCT: randomized controlled trial.

[d]HL: hearing loss.

[e]HA: hearing aid.

[f]PTA: pure-tone average.

[g]MoCA: Montreal Cognitive Assessment.

[h]MMSE: Mini-Mental State Examination.

[i]DT: delay training.

[j]AC: auditory-cognitive.

[k]LACE:listening and communication enhancement.

[l]WASI: Wechsler's Abbreviated Scale of Intelligence.

[m]GDS: Geriatric Depression Scale.

[n]NC: normal cognition.

[o]NCI: neurocognitive impairment.

Quality of Evidence

Risk-of-bias assessments are presented in Multimedia Appendix 2. Among the 16 RCTs assessed using RoB 2, overall risk of bias was rated as low risk in 18.8% (n=3) of studies, some concerns in 75% (n=12), and high risk in 6.3% (n=1). The majority of concerns were related to the randomization process and selection of the reported results. Specifically, 31.3% (n=5) of studies were judged as having low risk, 62.5% (n=10) as having some concerns, and 6.3% (n=1) as having high risk regarding the randomization process. For deviations from intended interventions, most studies were rated as low risk (n=14, 87.5%). All studies showed low risk for the measurement of the outcome domain, whereas most studies had some concerns regarding selection of the reported results (n=13, 81.3%).

For the 5 quasi-experimental studies assessed using the JBI checklist, most studies demonstrated adequate methodological quality. However, several studies had potential concerns regarding selection and allocation procedures, temporal precedence, and control of confounding factors. Detailed assessments for each domain are presented in Multimedia Appendix 2 (JBI critical appraisal qualitative evidence syntheses).

XSL•FO
RenderX

Participants' Hearing and Cognitive Characteristics

Hearing characteristics were heterogeneous across studies. Four studies included participants with normal or near-normal hearing, 7 studies focused on mild HL, and 9 studies investigated mild-to-moderate HL. Five studies included mixed hearing profiles, combining participants with different degrees of hearing ability. Because several studies recruited heterogeneous samples, some studies were classified into multiple categories. HA use was uncommon; only 2 studies included HA users, whereas most studies enrolled nonusers.

Twelve studies included cognitively normal older adults, defined via cutoffs on Mini-Mental State Examination (MMSE; MMSE scores: ≥24 - 25 or between 27 and 30)/Montreal Cognitive Assessment (MoCA; MoCA scores: ≥26 or ≥23) or IQ testing and a clinical interview (n=2; normal IQ). One study enrolled participants with educationally appropriate scores on the MMSE, one enrolled participants with subjective cognitive decline, and one included both cognitively normal and mild cognitive impairment groups. The remaining 4 studies did not report sufficient cognitive information.

ACT Summary

Table 2 provides an overview of ACT delivery methods, modalities, content, and dosage across the included studies; detailed intervention and control characteristics, follow-up, and measurements are provided in Table S1, Multimedia Appendix 1.

Table . Summary of auditory-cognitive training intervention characteristics[a].

Domain	Summary of specific characteristics
Training system	• Commercial integrated cognitive training systems, including Brain Fitness, Insight, CogniFit, ActivEar, and ReadMyQuips • Commercial auditory rehabilitation system: LACE[b] AI Pro and LACE DVD/online • Researcher-developed training systems, including APEX[c] 4, HELIX[d], Closed-Loop, and other customized platforms
Training content	• Pure auditory training: tone discrimination, phoneme discrimination, syllable sequence recall, speech comprehension, sound source localization, and speech recognition under noise-masking conditions • Auditory-cognitive integrated training: auditory processing speed training, instruction-following tasks, adaptive auditory memory training, word or sentence recall (short-term memory), backward digit recall (working memory), audio-based vocabulary span training, and Go/No-Go tasks (attention) • Cognitive training: tracking moving targets, spatial memory tasks, rapid motion direction identification, and word memory training
Training delivery modality	• Computer-based training (14 studies), mobile or tablet-based apps (4 studies), immersive virtual reality–based training (1 study), and one-to-one supervised training (2 studies)

[a]Training frequency and dosage characteristics were as follows: training period, median 4 (IQR 1-12) weeks; sessions per week, median 5 (IQR 1-7) sessions; session duration, median 30 (IQR 10-60) minutes; total number of sessions, median 20 (IQR 2-60) sessions; and total training duration, median 10 (IQR 1.5-60) hours.

[b]LACE:listening and communication enhancement.

[c]APEX: Audiological Platform for Experimental Research.

[d]HELIX:Hearing and Listening Experience.

Training Methods and Delivery Modalities

Fourteen studies used computer-based training; 4 used mobile app–based training; 1 used immersive virtual reality (VR); and 2 used face-to-face supervised training. Training tools included 5 commercial cognitive platforms (Brain Fitness [Posit Science Corp], Insight [Posit Science Corp], CogniFit [CogniFit Inc], ActivEar [CogniFit Inc], and ReadMyQuips [Sense Synergy, Inc]), a commercial auditory rehabilitation program (LACE AI Pro [Neurotone]), and researcher-developed systems (eg, APEX [Audiological Platform for Experimental Research] 4, HELIX [Hearing and Listening Experience], and Closed-Loop). Notably, all the training systems featured adaptive difficulty for individualized training.

Training Content

With respect to training content, ACT interventions were classified into auditory training, auditory-cognitive combined training, and cognitive training. Auditory training targeted auditory perception and speech processing, including tone and phoneme discrimination, syllable sequence recall, speech comprehension, sound source localization, and speech recognition under noise-masking conditions. Auditory-cognitive training integrated auditory tasks with cognitive demands, such as auditory processing speed training, instruction-following tasks, adaptive auditory memory exercises, word or sentence recall targeting short-term memory, backward digit recall tasks engaging working memory, audio-based vocabulary span training, and attention-related paradigms such as Go/No-Go tasks. Cognitive training components included tracking moving targets, spatial memory tasks, rapid motion direction identification, and word memory training.

Training Dosage and Intensity

The median training period was 4 (IQR 1-12) weeks. The median weekly frequency was 5 (IQR 1-7) sessions, and the

XSL•FO
RenderX

median session duration was 30 (IQR 10-60) minutes. The median number of total sessions was 20 (IQR 2-60) sessions, and the median cumulative training time was 10 (IQR 1.5-60) hours.

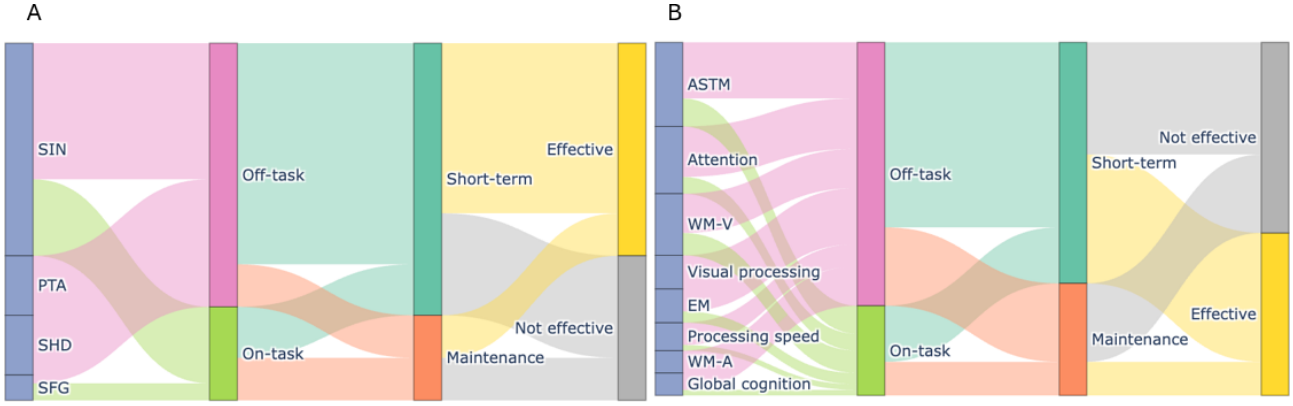
ACT Training Outcomes Summary

Overview

Figure 2 summarizes the main outcomes, task effect categories, short-term and maintenance effects, and corresponding intervention effectiveness. A narrative summary of key findings, task effects, time effects, and auditory/cognitive/neural outcomes is presented in Table S2, Multimedia Appendix 1. The detailed narrative summary of each outcome, including the studies assessing short-term and follow-up effects, task-effect classification (on-task: the outcome was trained during the intervention; off-task: the outcome was not trained during the intervention), and the effectiveness of ACT in the intervention group, is provided in Multimedia Appendix 3.

Overall, off-task, nontrained outcomes (15 studies) were assessed more frequently than on-task, trained outcomes (6 studies). Among the auditory-related outcomes, more than half of the short-term effects improved significantly, and half of the maintenance effects were significant. For the cognitive-related outcomes, almost half of the short-term effects were significant, but only one-third of the maintenance effects were significant.

Figure 2. (A) Auditory-related outcomes and (B) cognitive-related outcomes: auditory and cognitive effects, task attributes, and effectiveness. ASTM: auditory short-term memory; EM: episodic memory; HE: hearing effort; PTA: pure-tone average; SFG: auditory stream segregation; SHD: self-report hearing difficulty; SIN: speech-in-noise; WM-A: working memory (auditory); WM-V: working memory (visual).



Auditory-Related Outcomes

Among auditory outcomes, SIN recognition was the most commonly assessed measure (19 studies), including 4 on-task and 15 off-task evaluations, with 6 studies reporting maintenance effects. Short-term SIN improvements were significant in 15 studies. Out of 6 follow-up studies, 4 [19,31,34,38] showed significant maintained benefits, while 2 [32,42,43] did not.

Pure-tone average (PTA) was assessed as an off-task outcome in 6 studies. Only 1 study [20] reported a significant short-term improvement ($F_{1,42}$=3.12; P=.02; adjusted P=.06); the others showed no significant changes, and only 1 study reported a nonsignificant maintenance effect [38]. Insufficient data precluded meta-analysis.

Five studies evaluated self-reported hearing function outcomes. Three studies used the Speech, Spatial, and Qualities of Hearing Scale (SSQ); 2 studies found [32,38] no significant effects, and one found a significant effect [39]. Another 2 studies reported significant short-term improvements in subjective hearing difficulty (Hearing Handicap Inventory and Communication Scale for Older Adults for the Elderly/Adults) [39,45].

Auditory stream segregation task was evaluated in 2 studies [30,34], both reporting significant short-term improvements; only 1 study that included follow-up data confirmed the maintenance of this effect [34]. There were sufficient data to conduct a meta-analysis for this indicator.

Cognitive-Related Outcomes

Attention was evaluated in 9 studies. Six studies reported significant short-term improvements, whereas 3 found no significant effects. Three studies assessed the maintenance of treatment effects, with 2 [32,38] reporting sustained benefits and 1 [18] reporting no significant maintenance effect.

Short-term auditory memory was examined in 10 studies, with 6 reporting significant improvements; maintenance effects were mixed, with 2 studies showing retained benefits [19,38], and 3 studies showing no significant maintenance effect [18,32,42,43]. Four studies assessed episodic/verbal memory [18,20,31,38], mostly with nonsignificant short-term and maintenance effects.

Working memory outcomes showed domain-specific patterns. For auditory working memory, 3 of the 5 studies reported significant short-term improvements, whereas 2 [18,38] found no significant effects. Three studies assessed maintenance effects, with 2 [18,38] demonstrating sustained benefits and 1 [19] finding no significant maintenance effect. Similarly, for visual working memory, 3 of the 5 studies reported significant short-term improvements, whereas 2 [17,18] found no significant effects. Two studies [18,32] assessed maintenance, and 1 [32] reported a significant sustained benefit.

Processing speed was assessed in 4 studies: 3 reported significant short-term improvements [15,33,42,43], 1 did not [31]; 1 study found no sustained maintenance effects [31].

Global cognition was assessed in 3 studies, all of which used researcher-developed ACT training systems. One study reported

XSL•FO
RenderX

a significant short-term improvement in MoCA scores [19], whereas the other 2 studies found no significant effects, with one assessing the RBANS (Repeatable Battery for Assessment of Neuropsychological Status) total score [18] and the other assessing MoCA scores [45]. One study assessed maintenance effects at follow-up but found no significant maintenance effects [18].

Meta-Analysis Results

Hearing-Related Outcomes

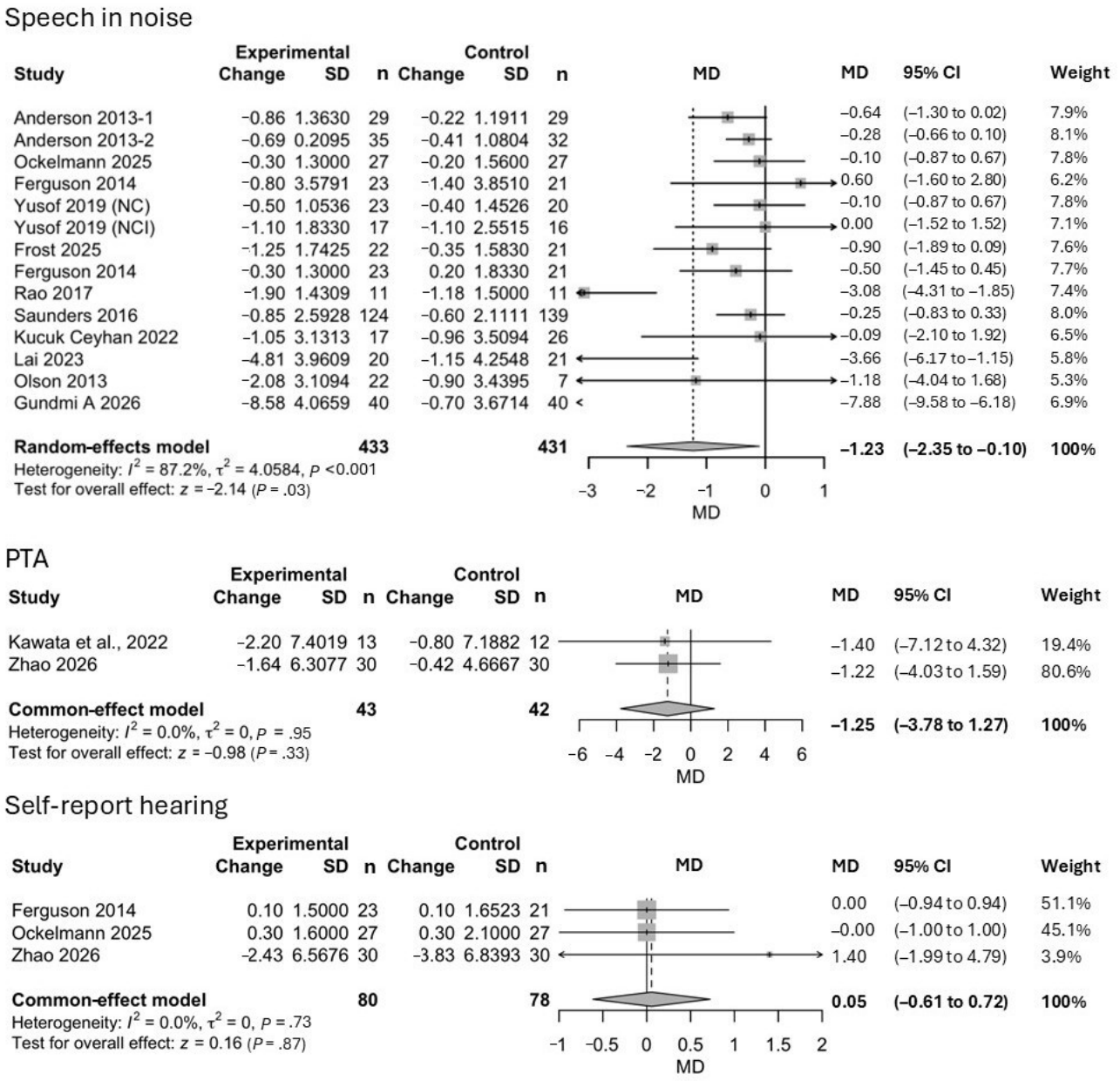
The forest plots for the meta-analyses of auditory-related outcomes are presented in Figure 3. Fourteen studies assessing SIN perception were included in the meta-analysis. A random-effects model demonstrated a significant pooled effect (MD=−1.23, 95% CI −2.35 to −0.10; *P*=.03), with substantial heterogeneity (I^2=87.2%). Egger regression test showed no significant evidence of funnel plot asymmetry (*t*=−1.33; *P*=.21; Figure S1 in Multimedia Appendix 1).

We also conducted a subgroup analysis to explore whether HA use contributed to the heterogeneity in the effects of ACT on SIN performance. Studies were stratified according to HA use into 2 subgroups: HA use (Yes) and no HA use (No). Among participants who did not use HAs, ACT was associated with a significant improvement in SIN performance (SMD=−0.32, 95% CI −0.52 to −0.11; *P*=.006), with low heterogeneity across studies (I^2=9.3%). In contrast, among HA users, the pooled effect was not statistically significant (SMD=−0.87, 95% CI −1.76 to 0.02; *P*=.06), with substantial heterogeneity (I^2=92.0%). The test for subgroup differences was not statistically significant (χ^2_1=1.38; *P*=.24), indicating no evidence of a statistically significant difference in the effects of ACT between participants who used HAs and those who did not (Figure S2 in Multimedia Appendix 1).

Two studies assessed PTA, with no significant pooled effect under a fixed-effect model (MD=−1.25, 95% CI −3.78 to 1.27; *P*=.33; I^2=0%). Three studies assessed self-reported hearing outcomes, which also showed no significant pooled effect (MD=0.05, 95% CI −0.61 to 0.72; *P*=.87; I^2=0%).

XSL•FO
RenderX

Figure 3. Forest plots for the meta-analyses of auditory-cognitive training on hearing-related outcomes [14,15,19,20,31-33,37-40,42,45]. MD: mean difference; PTA: pure-tone average.

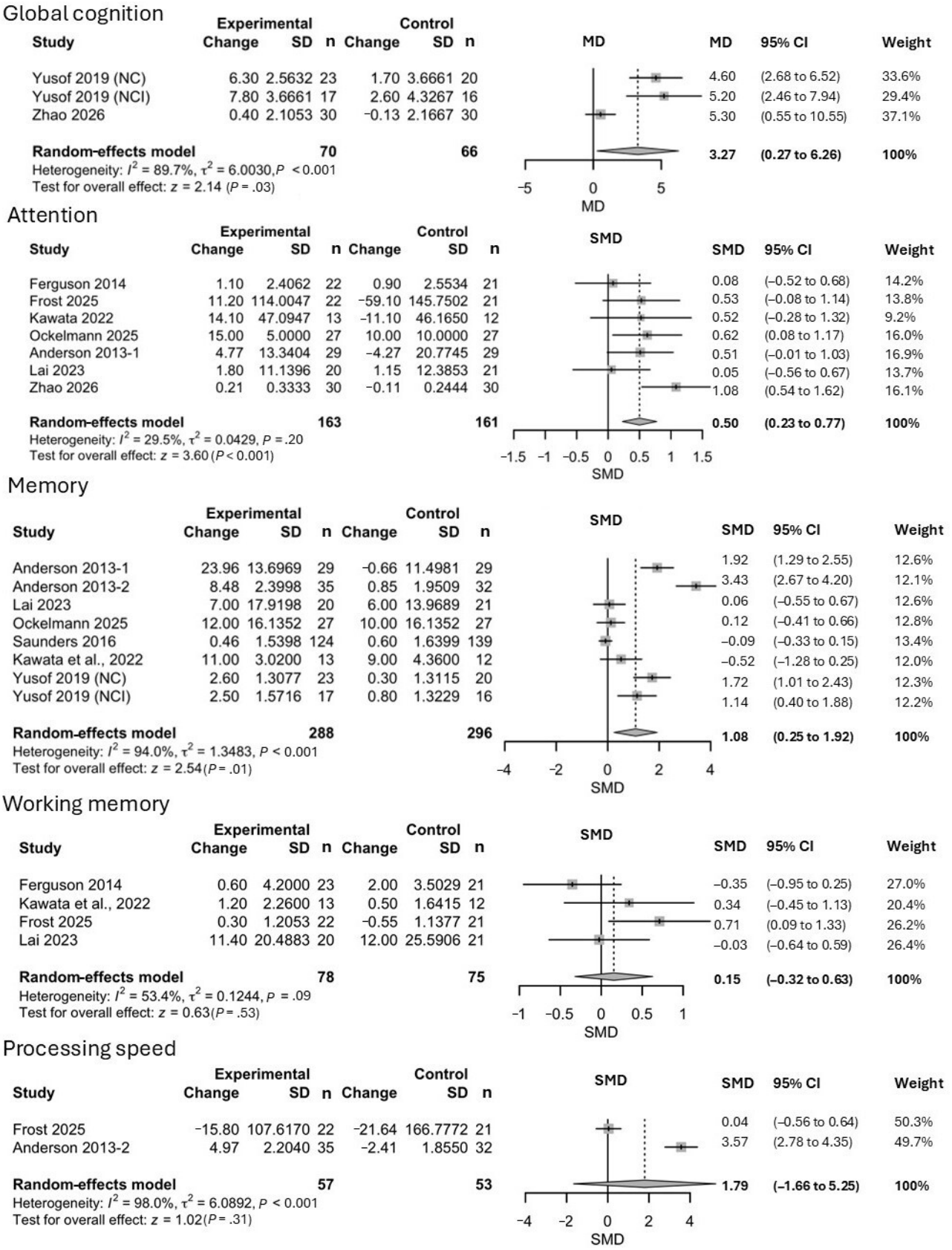


Cognitive-Related Outcomes

The forest plots for the meta-analyses of cognitive-related outcomes are presented in Figure 4. Three studies assessing global cognition using the MoCA were included, with a significant pooled effect favoring the intervention group (MD=3.27, 95% CI 0.27-6.26; P=.03; I^2=89.7%). Seven studies assessing attention showed a significant overall effect (SMD=0.50, 95% CI 0.23-0.77; P<.001; I^2=29.5%). Eight studies assessing memory also demonstrated a significant pooled effect (SMD=1.08, 95% CI 0.25-1.92; P=.01), with substantial heterogeneity (I^2=94.0%). Subgroup analysis showed positive pooled effects for both short-term memory (SMD=1.07, 95% CI −0.27 to 2.41; I^2=96.0%) and delayed memory (SMD=0.96, 95% CI −0.01 to 1.93; I^2=84.4%), but neither subgroup reached statistical significance. There was no significant difference between the 2 subgroups (P=.90). Leave-one-out sensitivity analysis did not identify any single study as a major contributor to the observed heterogeneity. In contrast, four studies assessing working memory showed no significant overall effect (SMD=0.15, 95% CI −0.32 to 0.63; P=.53; I^2=53.4%). Two studies assessing processing speed also showed no significant pooled effect (SMD=1.79, 95% CI −1.66 to 5.25; P=.31; I^2=98.0%).

XSL•FO
RenderX

Figure 4. Forest plots for meta-analyses of auditory-cognitive training on cognitive-related outcomes [14,15,19,20,31,32,37,38,42,45]. MD: mean difference; SMD: standardized mean difference.



Summary of Findings

Table 3 summarizes the certainty of evidence for the main outcomes assessed using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) approach [46]. The certainty of evidence ranged from very low to moderate. Attention was rated as moderate certainty, whereas SIN performance, memory, and self-reported hearing were rated

XSL•FO
RenderX

as low certainty. Global cognition, PTA, and processing speed were rated as very low certainty. The certainty of evidence was mainly downgraded because of risk of bias, substantial heterogeneity, and imprecision due to limited sample sizes and wide CIs.

Table . Summary of meta-analysis findings.

Outcomes	Studies (n)	Participants (n)		Effect estimate (95% CI)	Certainty of evidence[b]
		ACT[a] group	Control group		
Speech in noise	14	433	431	MD[c]: –1.23 (–2.35 to –0.10)	⊕⊕◯◯
Pure-tone average (dB)	2	43	42	MD: −1.25 (−3.78 to 1.27)	⊕◯◯◯
Self-reported hearing	3	80	78	MD: 0.05 (−0.61 to 0.72)	⊕⊕◯◯
Global cognition	3	70	66	MD: 3.27 (0.27 to 6.26)	⊕◯◯◯
Attention	7	163	161	SMD[d]: 0.50 (0.23 to 0.77)	⊕⊕⊕◯
Memory	8	288	296	SMD: 1.08 (0.25 to 1.92)	⊕⊕◯◯
Working memory	4	78	75	SMD: 0.15 (−0.32 to 0.63)	⊕◯◯◯
Processing speed	2	57	53	SMD: 1.79 (−1.66 to 5.25)	⊕◯◯◯

[a]ACT: auditory-cognitive training.

[b]GRADE certainty of evidence: ⊕⊕⊕◯, moderate certainty; ⊕⊕◯◯, low certainty; and ⊕◯◯◯, very low certainty.

[c]MD: mean difference.

[d]SMD: standardized mean difference.

Discussion

Principal Findings

This mixed synthesis is the first to summarize ACT in older adults. A total of 21 studies (24 articles) involving 1300 older adults from 10 countries (the United States, United Kingdom, Switzerland, Japan, Malaysia, Germany, Turkey, Singapore, India, and China) were included. This review comprehensively summarized key features of existing ACT interventions, including participants' hearing characteristics, delivery modes, training content, frequency and dosage, as well as outcome-specific effects. Considerable heterogeneity was observed in target populations' hearing status, and ACT was delivered via diverse modalities including computer-based, mobile app, VR, and face-to-face training. Training content fell into 3 categories: auditory training, auditory-cognitive integrated training, and cognitive training, with a median total training dosage of 10 (IQR 1.5-60) hours. Narrative evidence indicated consistent short-term benefits of ACT for SIN recognition, attention, and auditory memory, with some benefits maintained at follow-up. Meta-analyses further supported significant beneficial effects of ACT on SIN recognition, attention, overall memory, and global cognition, although substantial heterogeneity was observed for SIN, memory, and global cognition. In contrast, pooled effects were not significant for working memory, processing speed, self-reported hearing outcomes, or PTA, while evidence for auditory stream segregation was limited but consistently positive.

Current evidence shows that ACT has been applied to older adults across a range of hearing levels for both preventive and rehabilitative purposes. It has been used not only in those with mild-to-moderate HL to improve speech perception but also in older adults with normal audiometric thresholds who report subjective hearing difficulties to reduce listening effort and cognitive load. ACT has also been used as an adjunct to audiological rehabilitation before or after HA fitting, suggesting potential value in boosting auditory and cognitive outcomes. Most existing ACT studies have been conducted in older adults with normal cognition or mild cognitive impairment. Notably, Ferguson et al [32] reported that cognitively normal participants achieved greater ACT-related improvements in SIN perception, gap detection, and pitch pattern sequencing than did those with mild cognitive impairment. However, this between-group difference should be interpreted cautiously, as it may partly reflect regression to the mean or baseline differences rather than differential responsiveness to ACT. Thus, the extent to which baseline cognitive status moderates ACT effectiveness remains uncertain. For those with existing cognitive impairment, higher training dosage and broader multidomain interventions (eg, physical activity and psychosocial approaches) may warrant consideration [47], highlighting the need for tailored protocols based on baseline cognitive function.

XSL•FO
RenderX

To date, no standardized ACT protocols exist, with substantial variability in delivery modalities, content, and dosage. Computer-based and mobile app–based interventions are more accessible and flexible than face-to-face or VR training, supporting higher adherence through shorter, more frequent sessions, with a median total dosage of 10 hours. Integrated ACT appears superior to unimodal auditory or cognitive training alone. For example, Faust et al [18] found that separate auditory or visual cognitive training improved modality-specific functions, but simple coadministration did not produce synergistic benefits, emphasizing the importance of genuine integration (eg, combining speech perception with memory or attention demands) rather than parallel task delivery. Common integrated designs pair auditory stimuli with demands on short-term/working memory, attention, processing speed, or executive function. All included ACT interventions used adaptive difficulty. Although only one study [20] directly compared adaptive and fixed difficulty, adaptive protocols may better accommodate individual learning rates and heterogeneous baseline abilities in older adults, thereby improving engagement and training outcomes.

Regarding intervention effects, the most consistent improvement was observed in SIN recognition, with benefits maintained for 1 to 2 months postintervention [34,38], and short-term effects supported by a meta-analysis. A potential mechanism is that ACT may reduce excessive subcortical encoding of speech envelopes, strengthen top-down attentional control, and enhance functional connectivity between the prefrontal and auditory cortices, thereby improving temporal processing and SIN perception [14,15]. Nevertheless, the pooled estimate showed substantial heterogeneity (I^2=87.2%), which may partly reflect differences in SIN assessment paradigms, including HINT (Hearing-in-Noise Test), WIN (Words-in-Noise test), QuickSIN (Quick Speech-in-Noise), speech-in-babble, and competing-speaker tasks, as well as differences in participants' baseline hearing status and HA use. The pooled effect indicated a significant reduction in signal-to-noise ratio/speech reception threshold of approximately 1.23 dB following ACT (MD=−1.23, 95% CI −2.35 to −0.10 dB), suggesting a clinically meaningful improvement in speech recognition under noisy conditions [48]. However, the high between-study heterogeneity indicates that the magnitude of benefit varied considerably across studies. Subgroup analysis by HA use provided some additional insight into this heterogeneity. Among participants who did not use HAs, the pooled effect was significant and showed low heterogeneity (I^2=9.3%), whereas the effect among HA users was not statistically significant and was accompanied by substantial heterogeneity (I^2=92.0%). Although the test for subgroup differences was not statistically significant (P=.24), the marked difference in within-subgroup heterogeneity suggests that HA use may be one potential contributor to the variability in treatment effects across studies. This modest improvement may also have limited implications for real-world hearing function, consistent with the nonsignificant pooled effect for self-reported hearing outcomes (SSQ). By contrast, ACT showed minimal effects on pure-tone thresholds, consistent with the limited reversibility of age-related cochlear degeneration [49]. This may be because ACT primarily targets central auditory and cognitive processing rather than peripheral hearing sensitivity.

For cognitive outcomes, the meta-analysis showed significant pooled effects for global cognition, attention, and memory, whereas the effects for working memory and processing speed were not statistically significant. Among these outcomes, the evidence for attention was the most consistent, with a significant effect (SMD=0.50, 95% CI 0.23-0.77) and low heterogeneity (I^2=29.5%); moreover, the certainty of evidence was rated as moderate. This finding suggests that ACT may have a relatively robust beneficial effect on attention, potentially because the integration of auditory processing with cognitive control, selective attention, and task monitoring repeatedly engages overlapping auditory-cognitive processes. Such integrated training may strengthen the ability to allocate and maintain attention to relevant auditory information, particularly under cognitively demanding listening conditions. The pooled effect for memory was statistically significant but accompanied by substantial heterogeneity, and neither subgroup analysis by memory domain nor leave-one-out sensitivity analysis identified a major contributor to this heterogeneity. Therefore, the memory findings should be interpreted cautiously, given the substantial between-study variability and uncertainty regarding the magnitude and generalizability of the effect. For global cognition, the significant pooled effect should also be interpreted cautiously because it was based on a small number of studies and was accompanied by substantial heterogeneity. Similarly, the nonsignificant findings for working memory and processing speed provide insufficient evidence to determine whether ACT produces consistent benefits in these domains, particularly given the limited number of studies and variability in outcome measures and intervention characteristics. Overall, the current evidence suggests that the cognitive effects of ACT may vary across domains, with relatively stronger and more consistent evidence for attention than for broader or more specific cognitive functions. Evidence regarding the longer-term maintenance of cognitive benefits remains limited, particularly because of the scarcity of follow-up assessments, heterogeneous measures, and small sample sizes. Further well-designed trials using standardized cognitive outcomes, larger samples, and longer follow-up periods are needed to clarify the magnitude, generalizability, and durability of these cognitive effects.

We note several limitations when interpreting both the included evidence and our synthesis. At the evidence level, most studies used passive or waitlist control groups, making it difficult to distinguish the specific effects of ACT from those of single-domain auditory or cognitive training. Commercial training platforms (eg, CogniFit) may offer limited customization for SIN tasks and limited crosslinguistic adaptability, which may reduce ecological validity and generalizability. Few studies considered stimulus delivery modes (eg, headphones vs loudspeakers), which may affect listening comfort and effort in older adults with HL. In addition, most studies enrolled cognitively normal participants; only one study directly compared responses between cognitively normal and cognitively impaired groups. Thus, the efficacy of ACT for older adults with cognitive impairment remains uncertain. Few studies used patient-reported real-life listening outcomes, such

XSL•FO
RenderX

as subjective hearing difficulty and daily listening function, as primary end points. Given the importance of subjective hearing experience for communication confidence, social participation, mental health, and quality of life [50], future studies should incorporate more patient-reported measures and longer follow-up periods to assess the persistence and real-life relevance of training benefits.

Our review also has its own limitations. First, we restricted inclusion to English-language publications, which may have introduced language bias and led to the omission of relevant evidence published in other languages. Although the included evidence covered several non–English-speaking countries, including Japan, Germany, and Malaysia, relevant studies published in local languages may have been overlooked. This limitation may reduce the completeness and crosslinguistic representativeness of the evidence base and should be considered when interpreting the generalizability of the findings. Second, although meta-analyses were performed for key auditory and cognitive outcomes, pooled effects should be interpreted cautiously because of substantial heterogeneity and the small number of studies contributing to some outcomes. Variability in construct definitions, training dosage, delivery modes, and assessment time points may have reduced comparability across studies. Third, limited follow-up data prevented quantitative synthesis of maintenance effects across domains. Despite these limitations, this synthesis integrated narrative and meta-analytic evidence to characterize outcome patterns, quantify short-term effects, and identify key research gaps, thereby providing a basis for the standardization, refinement, and mechanistic investigation of ACT interventions for older adults.

Conclusions

This review systematically synthesizes the current evidence on ACT for older adults through a combined scoping review and meta-analytic approach. This review indicates that ACT may improve selected auditory and cognitive outcomes in older adults, particularly SIN perception, attention, memory, and global cognition. However, the certainty of the evidence was low for most outcomes, with substantial heterogeneity observed for several pooled effects. Evidence for changes in hearing thresholds, self-reported hearing function, working memory, and processing speed remains limited or inconclusive, and evidence for long-term maintenance is insufficient. These findings support the potential of ACT as a complementary approach to promote auditory-cognitive function in older adults, but do not yet establish its clinical effectiveness. Future trials should use standardized and adaptable ACT protocols, include patient-reported real-life listening outcomes, recruit more diverse samples with varied cognitive and hearing profiles, and incorporate longer follow-up periods. Establishing evidence-based protocols and clinically meaningful outcome benchmarks will be important for determining how ACT can be integrated into individualized hearing and cognitive health interventions.

Generative AI Use

During the study, Doubao (an AI language model) was used to assist with the preliminary screening of article titles and abstracts, the extraction of bibliographic and study-related information, and the assessment of methodological quality. All AI-assisted outputs were independently reviewed and verified by the authors, and final decisions regarding study eligibility, data extraction, and quality assessment were made by the authors. During manuscript preparation, Doubao was also used for language polishing and editing. The research concept, study design, literature search, study selection, data extraction and verification, data organization, statistical analyses, interpretation of results, and discussion were conducted by the authors. The authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Funding

This work was supported by the Qinglan Project of Colleges and Universities in Jiangsu Province (Project 2025) and the Key Sponsored Project of the Jiangsu Provincial Social Science Application Research Excellence Project (Special Program for Social Education and Social Science Popularization: 25SJA-19).

Data Availability

All data generated or analyzed during this study are included in this published article and its multimedia appendices. Multimedia Appendices were uploaded to the Open Science Framework database [25].

Authors' Contributions

ML contributed to the conception and design of the work; conducted the literature search; extracted, analyzed, and interpreted the data; and drafted and revised the manuscript. CM conducted the literature search; extracted, analyzed, and interpreted the data; and drafted and revised the manuscript. PD contributed to the conception of the work, interpreted the data, and revised the manuscript. YY contributed to the conception of the work, interpreted the data, and revised the manuscript. WL conducted the literature search, interpreted the data, and revised the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Database searching strategies and results; intervention and control characteristics of included studies (Table S1); detailed outcomes of included studies (Table S2); funnel plot for SIN (Figure S1); and subgroup analysis of ACT effects on SIN performance by hearing aid use (Figure S2).
[DOCX File, 742 KB - rehab_v13i1e101632_app1.docx]

Multimedia Appendix 2
Quality assessment results (RoB-2 for RCT and JBI critical appraisal qualitative Evidence Syntheses).
[PDF File, 308 KB - rehab_v13i1e101632_app2.pdf]

Multimedia Appendix 3
The detailed narrative summary of each outcome.
[PDF File, 308 KB - rehab_v13i1e101632_app3.pdf]

Checklist 1
PRISMA checklist.
[PDF File, 172 KB - rehab_v13i1e101632_app4.pdf]

Checklist 2
PRISMA-ScR checklist.
[PDF File, 163 KB - rehab_v13i1e101632_app5.pdf]

References

1. Jiang CY, Han K, Yang F, et al. Global, regional, and national prevalence of hearing loss from 1990 to 2019: a trend and health inequality analyses based on the Global Burden of Disease Study 2019. Ageing Res Rev 2023 Dec;92:102124. [doi: 10.1016/j.arr.2023.102124] [Medline: 37972859]
2. Zhao F, Wang Z, Wu Z, et al. Joint association of combined healthy lifestyle factors and hearing loss with cognitive impairment in China. J Gerontol A Biol Sci Med Sci 2024 Nov 1;79(11):glae226. [doi: 10.1093/gerona/glae226] [Medline: 39276141]
3. Dixon PR, Feeny D, Tomlinson G, Cushing S, Chen JM, Krahn MD. Health-related quality of life changes associated with hearing loss. JAMA Otolaryngol Head Neck Surg 2020 Jul 1;146(7):630-638. [doi: 10.1001/jamaoto.2020.0674] [Medline: 32407468]
4. Rutherford BR, Brewster K, Golub JS, Kim AH, Roose SP. Sensation and psychiatry: linking age-related hearing loss to late-life depression and cognitive decline. Am J Psychiatry 2018 Mar 1;175(3):215-224. [doi: 10.1176/appi.ajp.2017.17040423] [Medline: 29202654]
5. Cantuaria ML, Pedersen ER, Waldorff FB, et al. Hearing loss, hearing aid use, and risk of dementia in older adults. JAMA Otolaryngol Head Neck Surg 2024 Feb 1;150(2):157-164. [doi: 10.1001/jamaoto.2023.3509] [Medline: 38175662]
6. Wen Z, Peng S, Yang L, et al. Factors associated with social isolation in older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2023 Mar;24(3):322-330. [doi: 10.1016/j.jamda.2022.11.008] [Medline: 36549651]
7. Nieman CL, Oh ES. Hearing loss. Ann Intern Med 2020 Dec 1;173(11):ITC81-ITC96. [doi: 10.7326/AITC202012010] [Medline: 33253610]
8. Döge J, Hackenberg B, O Brien K, et al. The prevalence of hearing loss and provision with hearing aids in the Gutenberg Health Study. Dtsch Arztebl Int 2023 Feb 17;120(Forthcoming):99-106. [doi: 10.3238/arztebl.m2022.0385] [Medline: 36519221]
9. Ferguson M, Maidment D, Henshaw H, Heffernan E. Evidence-based interventions for adult aural rehabilitation: that was then, this is now. Semin Hear 2019 Feb;40(1):68-84. [doi: 10.1055/s-0038-1676784] [Medline: 30728650]
10. Ferguson M, Henshaw H. How does auditory training work? Joined-up thinking and listening. Semin Hear 2015 Nov;36(4):237-249. [doi: 10.1055/s-0035-1564456] [Medline: 27587911]
11. Ferguson MA, Henshaw H. Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss. Front Psychol 2015;6:556. [doi: 10.3389/fpsyg.2015.00556] [Medline: 26074826]
12. Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. Hear Res 2013 Jun;300:18-32. [doi: 10.1016/j.heares.2013.03.006] [Medline: 23541911]
13. Rönnberg J, Lunner T, Zekveld A, et al. The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. Front Syst Neurosci 2013;7:31. [doi: 10.3389/fnsys.2013.00031] [Medline: 23874273]
14. Anderson S, White-Schwoch T, Choi HJ, Kraus N. Training changes processing of speech cues in older adults with hearing loss. Front Syst Neurosci 2013;7:97. [doi: 10.3389/fnsys.2013.00097] [Medline: 24348347]
15. Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. Reversal of age-related neural timing delays with training. Proc Natl Acad Sci U S A 2013 Mar 12;110(11):4357-4362. [doi: 10.1073/pnas.1213555110] [Medline: 23401541]

XSL•FO
RenderX

16. Fisher C, Johns MA, Decruy LP, et al. The impacts of two adaptive auditory-cognitive training paradigms on listening to competing talkers. Semin Hear 2025;46(2):141-160. [doi: 10.1055/s-0045-1811532] [Medline: 41036264]
17. Frei V, Giroud N. Immersive auditory-cognitive training improves speech-in-noise perception in older adults with varying hearing and working memory. NPJ Sci Learn 2025 Mar 7;10(1):12. [doi: 10.1038/s41539-025-00306-5] [Medline: 40055345]
18. Faust ME, Multhaup KS, Ong MS, Demakis GJ, Balz KG. Exploring the specificity, synergy, and durability of auditory and visual computer gameplay transfer effects in healthy older adults. J Gerontol B Psychol Sci Soc Sci 2020 Jun 2;75(6):1170-1180. [doi: 10.1093/geronb/gbz096] [Medline: 31410475]
19. Yusof Y, Mukari S, Dzulkifli MA, et al. Efficacy of a newly developed auditory-cognitive training system on speech recognition, central auditory processing and cognitive ability among older adults with normal cognition and with neurocognitive impairment. Geriatr Gerontol Int 2019 Aug;19(8):768-773. [doi: 10.1111/ggi.13710] [Medline: 31237107]
20. Kawata NYS, Nouchi R, Oba K, Matsuzaki Y, Kawashima R. Auditory cognitive training improves brain plasticity in healthy older adults: evidence from a randomized controlled trial. Front Aging Neurosci 2022;14:826672. [doi: 10.3389/fnagi.2022.826672] [Medline: 35431898]
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
22. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
23. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
24. Auditory-cognitive training for older adults: a scoping review. Open Science Framework (OSF). URL: https://osf.io/e8w9y/overview [accessed 2026-10-17]
25. The methods and effectiveness of auditory-cognitive training for older adults: a scoping review and meta-analysis. Open Science Framework (OSF). URL: https://osf.io/6u7ns/overview?view_only=e930df61e07248058a888ac46292070c [accessed 2026-10-17]
26. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
27. Barker TH, Habibi N, Aromataris E, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for quasi-experimental studies. JBI Evid Synth 2024 Mar 1;22(3):378-388. [doi: 10.11124/JBIES-23-00268] [Medline: 38287725]
28. Li M, Ma C, Wu C. Facial emotion recognition in sleep deprivation: a systematic review and meta-analysis. Int Rev Soc Psychol 2023;36:9. [doi: 10.5334/irsp.679] [Medline: 40951795]
29. Abrams HB, Bock K, Irey RL. Can a remotely delivered auditory training program improve speech-in-noise understanding? Am J Audiol 2015 Sep;24(3):333-337. [doi: 10.1044/2015_AJA-15-0002] [Medline: 26649542]
30. Fisher C, Karunathilake IMD, Johns MA, et al. Neural and behavioral changes in older adults from auditory-cognitive training. Annu Int Conf IEEE Eng Med Biol Soc 2025 Jul;2025:1-6. [doi: 10.1109/EMBC58623.2025.11254829] [Medline: 41335864]
31. Frost E, Malhotra P, Porat T, Poole K, Menon A, Picinali L. HEaring and LIstening eXperience (HELIX): evaluation of a co-designed serious game for auditory-cognitive training. Q J Exp Psychol (Hove) 2025 Feb 19:17470218251324168. [doi: 10.1177/17470218251324168] [Medline: 39972554]
32. Ferguson MA, Henshaw H, Clark DPA, Moore DR. Benefits of phoneme discrimination training in a randomized controlled trial of 50- to 74-year-olds with mild hearing loss. Ear Hear 2014;35(4):e110-e121. [doi: 10.1097/AUD.0000000000000020] [Medline: 24752284]
33. Gundmi A, Yerraguntla K, Rajashekhar B, Guddattu V. Efficacy of a structured auditory training module in older individuals with hearing loss: a randomized controlled trial. Eur Arch Otorhinolaryngol 2026 Jun;283(6):3585-3594. [doi: 10.1007/s00405-026-10058-6] [Medline: 41838150]
34. Henderson Sabes J, Sweetow RW. Variables predicting outcomes on listening and communication enhancement (LACE) training. Int J Audiol 2007 Jul;46(7):374-383. [doi: 10.1080/14992020701297565] [Medline: 17680469]
35. Sweetow RW, Sabes JH. The need for and development of an adaptive listening and communication enhancement (LACE) program. J Am Acad Audiol 2006 Sep;17(8):538-558. [doi: 10.3766/jaaa.17.8.2] [Medline: 16999250]
36. Kucuk Ceyhan A, Dere HH, Mujdeci B. Evaluating the effectiveness of a new auditory training program on the speech recognition skills and auditory event-related potentials in elderly hearing aid users. Audiol Neurootol 2022;27(5):368-376. [doi: 10.1159/000523807] [Medline: 35398843]
37. Lai CYY, Ng PS, Chan AHD, Wong FCK. Effects of auditory training in older adults. J Speech Lang Hear Res 2023 Oct 4;66(10):4137-4149. [doi: 10.1044/2023_JSLHR-22-00621] [Medline: 37656601]
38. Ockelmann J, Scherpiet S, Stropahl M, Giroud N. Personalized and gamified auditory-cognitive training improves naturalistic speech-in-noise comprehension in older adults with hearing loss. NPJ Sci Learn 2025 Nov 19;10(1):80. [doi: 10.1038/s41539-025-00369-4] [Medline: 41258026]
39. Olson AD, Preminger JE, Shinn JB. The effect of LACE DVD training in new and experienced hearing aid users. J Am Acad Audiol 2013 Mar;24(3):214-230. [doi: 10.3766/jaaa.24.3.7] [Medline: 23506666]

XSL•FO
RenderX

40. Rao A, Rishiq D, Yu L, Zhang Y, Abrams H. Neural correlates of selective attention with hearing aid use followed by ReadMyQuips auditory training program. Ear Hear 2017;38(1):28-41. [doi: 10.1097/AUD.0000000000000348] [Medline: 27556531]
41. Rishiq D, Rao A, Koerner T, Abrams H. Can a commercially available auditory training program improve audiovisual speech performance? Am J Audiol 2016 Oct 1;25(3S):308-312. [doi: 10.1044/2016_AJA-16-0017] [Medline: 27768194]
42. Saunders GH, Smith SL, Chisolm TH, Frederick MT, McArdle RA, Wilson RH. A randomized control trial: supplementing hearing aid use with listening and communication enhancement (LACE) auditory training. Ear Hear 2016;37(4):381-396. [doi: 10.1097/AUD.0000000000000283] [Medline: 26901263]
43. Smith SL, Saunders GH, Chisolm TH, Frederick M, Bailey BA. Examination of individual differences in outcomes from a randomized controlled clinical trial comparing formal and informal individual auditory training programs. J Speech Lang Hear Res 2016 Aug 1;59(4):876-886. [doi: 10.1044/2016_JSLHR-H-15-0162] [Medline: 27567017]
44. Whitton JP, Hancock KE, Shannon JM, Polley DB. Audiomotor perceptual training enhances speech intelligibility in background noise. Curr Biol 2017 Nov 6;27(21):3237-3247. [doi: 10.1016/j.cub.2017.09.014] [Medline: 29056453]
45. Zhao IY, Leung AYM, Li C, et al. Web-based gamified auditory-cognitive dual-task training for older adults with age-related hearing loss: pilot randomized controlled trial. JMIR Aging 2026 Jun 16;9:e84083. [doi: 10.2196/84083] [Medline: 42302249]
46. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol 2011 Dec;64(12):1283-1293. [doi: 10.1016/j.jclinepi.2011.01.012] [Medline: 21839614]
47. Ma C, Wu C. Effects of a combined intervention of music-based movement, neuromodulation-synchronized cognitive training, and emotion regulation on cognitive function in community-dwelling older adults: a randomized controlled trial. Arch Gerontol Geriatr 2025 Dec;139:106000. [doi: 10.1016/j.archger.2025.106000] [Medline: 40865406]
48. Jigo M, Carmel JB, Wang Q, Rodenkirch C. Transcutaneous cervical vagus nerve stimulation improves speech comprehension in noise: a crossover, placebo-controlled study. Neuromodulation 2025 Dec;28(8):1384-1395. [doi: 10.1016/j.neurom.2025.04.007] [Medline: 40396942]
49. Pichora-Fuller MK, Mick P, Reed M. Hearing, cognition, and healthy aging: social and public health implications of the links between age-related declines in hearing and cognition. Semin Hear 2015 Aug;36(3):122-139. [doi: 10.1055/s-0035-1555116] [Medline: 27516713]
50. Henderson N, Hodgson S, Mulhern B, Page K, Sampson C. A qualitative systematic review of the impact of hearing on quality of life. Qual Life Res 2025 Apr;34(4):879-892. [doi: 10.1007/s11136-024-03851-5] [Medline: 39579270]

Abbreviations

ACT: auditory-cognitive training
APEX: Audiological Platform for Experimental Research
ELU: Ease of Language Understanding
GRADE: Grading of Recommendations, Assessment, Development, and Evaluation
HA: hearing aid
HELIX: Hearing and Listening Experience
HINT: Hearing-in-Noise Test
HL: hearing loss
JBI: Joanna Briggs Institute
MD: mean difference
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
OSF: Open Science Framework
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PTA: pure-tone average
QuickSIN: Quick Speech-in-Noise
RBANS: Repeatable Battery for Assessment of Neuropsychological Status
RCT: randomized controlled trial
RoB 2: revised Cochrane risk-of-bias tool
SIN: speech-in-noise
SMD: standardized mean difference
SSQ: Speech, Spatial, and Qualities of Hearing Scale
VR: virtual reality
WIN: Words-in-Noise test

XSL•FO
RenderX

Edited by J Sarvestan; submitted 17.May.2026; peer-reviewed by A Gundmi, V Manchaiah; revised version received 22.Sep.2026; accepted 05.Oct.2026; published 09.Nov.2026.

Please cite as:
Li M, Ma C, Dawes P, Yang Y, Liu W
Methods and Effectiveness of Auditory-Cognitive Training for Older Adults: Scoping Review and Meta-Analysis
JMIR Rehabil Assist Technol 2026;13:e101632
URL: https://rehab.jmir.org/2026/1/e101632
doi:10.2196/101632

© Mengyuan Li, Chifen Ma, Piers Dawes, Yanni Yang, Weiwei Liu. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 9.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Emerging Technologies on Outpatient Care Quality: Mixed Methods Study

Inas Al Khatib, DPhil; Mahmoud Awad, DPhil; Abdulrahim Shamayleh, DPhil

Department of Industrial Engineering, College of Engineering, American University of Sharjah, University City, Sharjah, United Arab Emirates

Corresponding Author:
Inas Al Khatib, DPhil
Department of Industrial Engineering, College of Engineering, American University of Sharjah, University City, Sharjah, United Arab Emirates

Abstract

Background: Digital transformation in health care, including telemedicine, wearable devices, robotic process automation (RPA), and electronic medical records (EMRs), is reshaping outpatient care delivery. However, consolidated evidence on how these technologies influence patient care quality (PCQ) remains limited.

Objective: This study evaluated the impact of emerging digital health technologies on PCQ dimensions in outpatient clinics through a systematic literature review combined with expert perspectives.

Methods: A systematic review following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines was conducted across ABI/INFORM, ScienceDirect, MEDLINE, and Google Scholar to identify studies published between 2017 and 2022. Studies examining digital technologies and their effects on outpatient PCQ were included. After screening and eligibility assessment, 53 studies were retained for synthesis. In parallel, semistructured interviews were conducted with 19 subject matter experts from public health, health informatics, and outpatient clinical practice. Findings were thematically analyzed using SERVQUAL (Service Quality Model) dimensions: responsiveness, reliability, empathy, communication, and tangibles, and triangulated across both datasets.

Results: RPA led to substantial reductions in waiting times, while telemedicine improved accessibility and continuity of care, as highlighted by 14 of 19 experts. Wearable devices enhanced remote monitoring and patient engagement but required adequate digital literacy. Several studies (n=7) reported that EMR usability challenges negatively affected patient-physician communication. Overall, improvements were most evident in responsiveness and communication, while empathy and patient-centeredness remained comparatively weaker without targeted training and workflow adjustments.

Conclusions: Emerging digital health technologies can enhance outpatient PCQ, particularly in operational efficiency, access to care, and communication. This review is innovative in that it integrates evidence across telemedicine, wearable monitoring, RPA, and EMRs within a unified SERVQUAL framework, offering practical guidance for implementing patient-centered digital health strategies in outpatient settings.

(JMIR Rehabil Assist Technol 2026;13:e81847) doi:10.2196/81847

KEYWORDS

digital health; telemedicine; robotic process automation; patient care quality; outpatient care; SERVQUAL

Introduction

Background

The surge in health care demand directly correlates with expenses and resource allocation, underscoring the imperative to prioritize enhancing service quality on the management agendas of health care operators. This prioritization is crucial due to its profound impact on patients' perceived value, satisfaction levels, and loyalty [1]. A comprehensive comprehension of service quality in health care commences with grasping its conceptualization. Despite its importance, there is no unified common understanding of patient care quality (PCQ). Instead, there are several interpretations, starting with a complex and abstract definition that defies standardization and quantification, acknowledged as fundamental to customer satisfaction and contentment [2]. Second, it involves clients' perceptions and assessments of the service, elucidating the disparity between expectations and experiences [3]. This perspective resonates with those defining service quality as the differential between clients' initial expectations and those developed postservice interaction [4]. Third, service quality encompasses a service provider's efficacy in satisfying consumers effectively, thereby fostering heightened profits, market share, sustainable customer satisfaction, and corporate marketing prowess [5,6]. Moreover, service quality is intricately intertwined with service delivery, evaluated through diverse determinants governing performance and consumer benefits

XSL•FO
RenderX

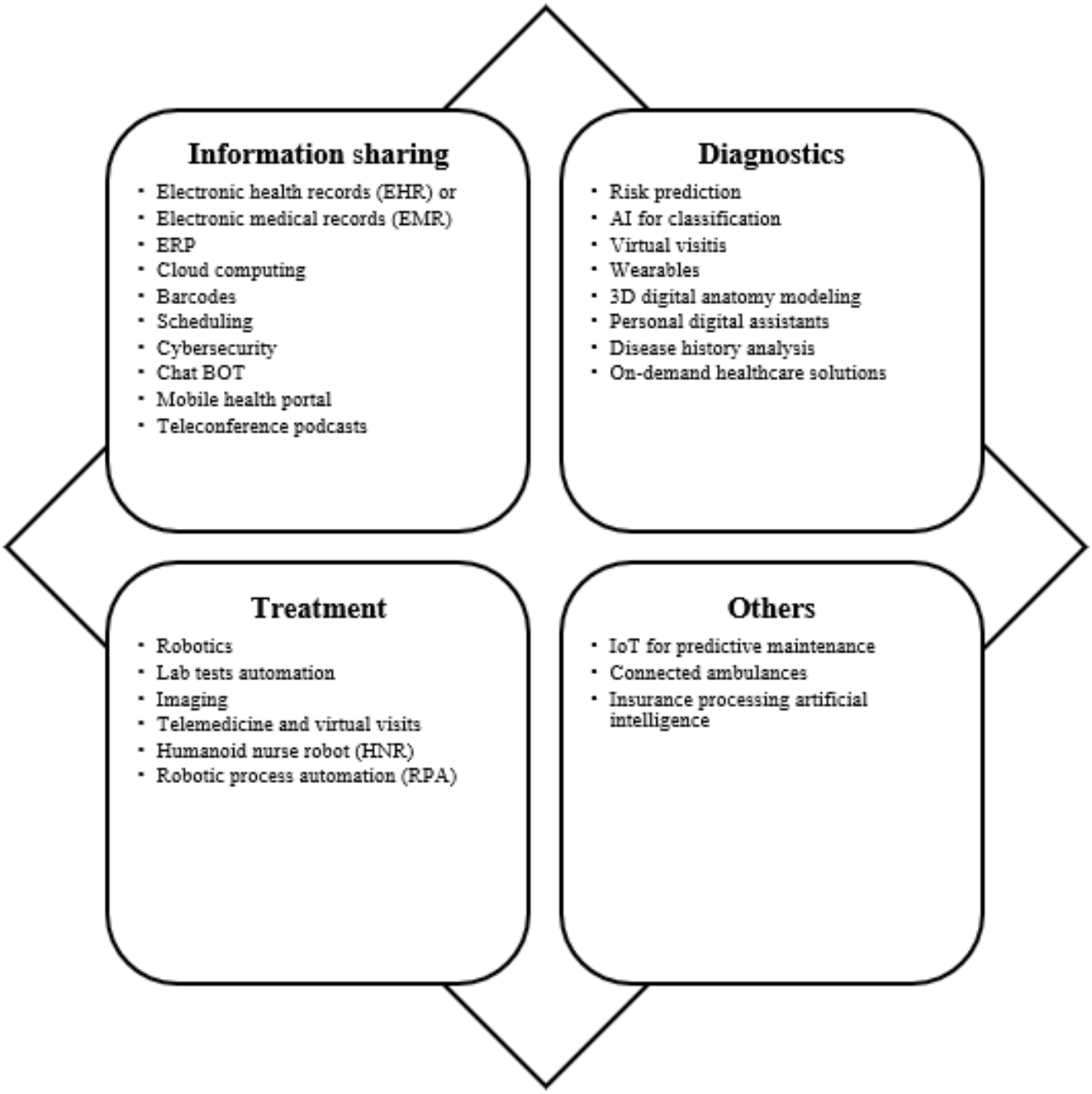
[7]. Finally, customers significantly influence service delivery, assessing quality via a 5-gap model encompassing knowledge, standards, communication, delivery, and the differential between expected and actual service quality [2].

There is an increasing demand for outpatient services, leading to an increase in the patient volumes seen, increased waiting times through the patient journeys, overworked caregivers, and increased overtime hours worked. Evaluating the patient's satisfaction with the quality of outpatient clinic service delivery [8]. Health care operators realized the need to change the way they work [9]. Digital transformation improves 4 key operational decision-making areas: patient flow, staffing, scheduling, and supply chain [10].

To address health care challenges and intense competition, technology adaptation has become a priority for decision-makers. Figure 1 illustrates some of the technologies used by the health care sector to tackle these various challenges. Digital transformation involves a process aimed at enhancing an entity by instigating significant changes to its attributes through the integration of information, computing, communication, and connectivity technologies [11]. As highlighted by Allen [12], it serves as a pivotal step in preparing for a consumer-centric "Future of Health." Cloud adoption, including Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Software as a Service (SaaS) [12], along with advancements in AI and the Internet of Things (IoT), are actively driving global digital transformation efforts [13]. Esteemed health systems perceive digital transformation to become more customer-oriented while simultaneously evolving their operations, work culture, and technological landscape. This includes outpatient clinics leveraging digital capabilities to fundamentally reshape their client relationships, focusing on incremental milestones to deliver value. However, successful transformation faces several challenges, including talent acquisition, data management, establishing key performance indicators, and budget allocation. Notably, having executive leadership dedicated to championing digital transformation is essential for success [14]. Barriers in health care digital transformation encompass the cybersecurity workforce shortage and reliance on legacy systems [13]. Key technologies driving digital transformation include automation, connected ambulances, on-demand health care solutions, telemedicine, virtual visits, patient portals, health care wearables, and disease history analysis. These innovations aim to streamline workflows for clinicians and hospitals, ensuring accurate patient data and health performance indicators to facilitate more effective treatment plans in shorter timeframes. Hospitals stand to benefit from optimized workflows, improved client interaction, cost reduction, secure electronic medical records (EMR) databases, and enhanced communication with medical staff [15]. Patients, on the other hand, gain access to personalized health care services, simplified access to personal health records, convenient appointment scheduling, improved online communication with doctors, and real-time tracking of crucial health metrics [16].

XSL•FO
RenderX

Figure 1. Available technologies for health care. ERP: enterprise resource planning; IoT: Internet of Things.



Rationale

Consequently, the primary focus of this review article revolves around investigating the impact of adopting emerging technologies on PCQ, encompassing both clinical and nonclinical dimensions, for assessment and monitoring purposes from the perspective of health practitioners operating within an outpatient setting to determine their impact on patient experience. These dimensions, as highlighted by Abbasi-Moghaddam et al [17] and Berwick [18], include safety, effectiveness, patient-centeredness, timeliness, efficiency, and equity. Assessing health care quality proves challenging due to its intangible nature, with evaluations varying based on factors such as the specialization, size, and complexity of health care organizations [5]. It involves strategic decision-making across various disciplines within health care, with perceptions differing among different professional fields [19].

Substandard health care quality can lead to patient dissatisfaction, stress, and despair [7]. It is estimated that poor-quality health care contributes significantly to health care costs, accounting for approximately 30%, owing to reduced productivity, wasteful clinical practices, high rates of medical errors, and costly health issues stemming from individual behaviors [20]. The World Health Organization [21] identifies factors such as inaccurate diagnosis, unnecessary treatments, unsafe clinical practices, and inadequately trained health care professionals as drivers of escalating illness burden and health care costs, commonly referred to as the cost of poor quality (COPQ). The broader economic and social costs of substandard care amount to trillions of dollars annually on a global scale due to long-term disability and lost productivity.

Despite extensive research exploring the determinants of health care quality and their impact on various operational aspects, there remains a gap in understanding the evolution of these

XSL•FO
RenderX

determinants over time and the role of emerging technology adoption in enhancing PCQ. This research seeks to address this gap by examining whether investments in emerging technologies aimed at improving outpatient clinic quality determinants result in positive patient experience from the perspective of health care professionals operating in outpatient clinic settings. By examining a wide range of published articles and existing literature, this review aims to provide a comprehensive understanding of patients' perspectives. Specifically, the integration of technologies such as robotic process automation (RPA), telemedicine and virtual visits, mobile health (mHealth) portal, videoconferencing, other educational technology, 3D digital anatomy modeling, insurance processing AI, health care wearables, personal digital assistants (PDAs), disease history analysis, teleconference podcasts, EMRs, on-demand health care solutions, and connected ambulance is being evaluated. These technologies have the potential to streamline processes, improve efficiency, and enhance the overall patient experience throughout their treatment journey.

Thus, the main research question to be explored through further methodologies is does the use of emerging technologies in outpatient clinics lead to better quality determinants and enhanced patient experiences?

This study was guided by the following hypothesis: emerging technology adoption and integration enhances all dimensions of PCQ in outpatient health clinics.

Recent review studies have examined the role of individual digital health technologies, such as telemedicine, wearable monitoring devices, and EMRs, in improving health care delivery and patient outcomes. However, these reviews typically focus on single technologies or specific clinical contexts, providing limited insight into how multiple digital health innovations collectively influence PCQ in outpatient settings. As digital transformation increasingly involves the simultaneous implementation of interconnected technologies, a more integrated understanding of their combined effects on health care service quality is needed. To address this gap, this systematic review synthesizes evidence across multiple digital health domains and evaluates their impact on PCQ using SERVQUAL (Service Quality Model) dimensions. By consolidating fragmented evidence and incorporating expert perspectives, this review offers a more comprehensive assessment of how emerging digital technologies shape our PCQ and provides practical insights for health care decision-makers [22-25].

Objectives

To answer the research question and test the research hypothesis, the following objectives are identified:

1. Identify the dimensions that influence the perception of PCQ in outpatient clinics
2. Assess the present emerging technology endeavors integrated into determinants of health service quality in outpatient clinics
3. Authenticate the favorable patient experience stemming from emerging technology initiatives implemented within determinants of health service quality in outpatient clinics

This research will directly benefit health care providers and quality experts, particularly those working in hospitals with outpatient clinics, medical centers, or satellite clinics. Additionally, academic scholars, research scholars, accreditation professionals, continuous improvement specialists, regulatory bodies, and occupational health and safety professionals will also find value in this study.

Methods

Overview

This study followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines for systematic reviews (Checklist 1) and the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist for expert interviews (Checklist 2). The methodology followed in this research consists of 2 main steps: systematic literature review (SLR) and subject matter experts (SMEs) interviews. The 2 steps were intentionally designed to provide complementary perspectives on the central research question, each contributing a distinct layer of insight. The systematic literature review established the theoretical foundation, identifying key quality dimensions across health care technology interventions. The expert interviews offered contextual validation from stakeholders actively involved in implementation. The SLR is used to gather secondary data and analyze its relevance to the study objectives. This method was selected due to its ability to offer a comprehensive, evidence-based synthesis of prior research, aligning with best practices recommended by Nunn and Chang [26] and Siddaway et al [27]. Systematic reviews are particularly well-suited for health care–related inquiries because they apply a transparent, replicable methodology that ensures rigor in identifying, integrating, and evaluating diverse sources.

Eligibility Criteria

Studies were selected according to predefined inclusion and exclusion criteria to ensure relevance, methodological consistency, and alignment with the review objectives, as shown in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Outpatient or ambulatory care setting
- Use of emerging technologies (telemedicine, robotic process automation, electronic medical record or electronic health record, AI, wearables, and educational tech)
- Reported patient care quality or satisfaction outcomes
- Peer-reviewed empirical or systematic review studies

Exclusion criteria

- Inpatient or hospital-only focus
- Non–peer-reviewed (editorials and conference abstracts)
- Technology not linked to patient care or service quality

Information Sources

A systematic search was conducted across 7 major databases: ABI/INFORM Global, ScienceDirect, MEDLINE (via PubMed), Google Scholar, Embase, Cochrane Library, and CINAHL. Studies published between 2017 and 2022 were prioritized. The search strategy was designed to capture digital health innovations in outpatient care and their impact on patient experience, quality, and satisfaction. Database-specific search syntax, MeSH and Emtree terms, synonyms, and field codes were used. To ensure validity and reliability, studies published between 2017 and 2022 were prioritized.

Search Strategy

Overview

The search included three key concept clusters:

1. Digital health technologies: such as telemedicine, EMR and electronic health record (EHR), AI, mHealth apps, and wearables
2. Outpatient care: such as ambulatory care, outpatient clinic, and primary care
3. Patient-centered outcomes: patient experience, patient satisfaction, and quality of care

Boolean operators AND and OR combined these clusters, and truncation and wildcards were applied where supported. For each database, search strings were adapted to reflect its indexing and search syntax. Detailed search strings and database queries are provided in this section.

Boolean Search Strings Across Databases

The following Boolean search strings were used for the database searches:

1. (“digital health” OR “eHealth” OR “telehealth” OR “mobile health” OR “mHealth” OR “health IT” OR “digital innovation”) AND (“outpatient care” OR “ambulatory care” OR “primary care clinic” OR “ambulatory services”) AND (“patient experience” OR “patient satisfaction” OR “quality of care” OR “patient-centered care”)
2. (“telemedicine” OR “virtual care” OR “remote consultation” OR “video consultation”) AND (“patient satisfaction” OR “patient perception” OR “patient experience”)
3. (“robotic process automation” OR “RPA”) AND (“healthcare” OR “medical services”) AND (“outpatient” OR “ambulatory care”)
4. (“electronic health record” OR “EHR” OR “EMR” OR “HER”) AND (“patient-centered care” OR “quality of care” OR “service quality”)
5. (“wearable devices” OR “healthcare wearables” OR “mHealth apps” OR “mobile health portal” OR “smartwatch health app”) AND (“outpatient clinic” OR “ambulatory services”)
6. (“artificial intelligence” OR “AI” OR “machine learning” OR “deep learning”) AND (“insurance processing” OR “clinical decision support” OR “diagnostic support”) AND (“outpatient” OR “ambulatory care”)
7. (“3D anatomy modeling” OR “educational technology” OR “digital simulation” OR “virtual reality training”) AND (“patient education” OR “health education”) AND (“outpatient setting” OR “ambulatory clinic”)
8. (“personal digital assistant” OR “clinical decision support system” OR “mobile clinical assistant”) AND (“physician competence” OR “clinical performance”) AND (“outpatient” OR “ambulatory care”)

Database-Specific Notes

The following provides database-specific notes on searches.

- ABI/INFORM Global: all Boolean strings applied with limits, including English language, peer-reviewed journals, and publication years 2017 - 2022.
- ScienceDirect: used Boolean strings with filters for research articles and review articles, publication years 2017 - 2022.
- Google Scholar: keywords applied sequentially due to limitations on advanced Boolean combinations; the first 200 results for each query were screened.
- Medline (PubMed):
  - Used MeSH terms: “Telemedicine”[MeSH], “Electronic Health Records”[MeSH], “Patient Satisfaction”[MeSH]
  - Combined with title and abstract searches: (digital health[tiab] OR eHealth[tiab] OR mHealth[tiab])
  - Filters: English language, human studies, and publication years 2017 - 2022
- Embase:

XSL•FO
RenderX

  - Used Emtree terms: “telemedicine”/exp, “electronic health record”/exp, “patient satisfaction”/exp
  - Field modifiers: ti,ab for title/abstract
  - Filters: English language and publication years 2017 - 2022
- ABI/INFORM Global, ScienceDirect, CINAHL, Cochrane Library, and Google Scholar:
  - Boolean operators applied with syntax adjustments per database
  - First 200 results screened in Google Scholar due to limits
- Notes on reproducibility:
  - All database-specific queries, including MeSH and Emtree terms, truncation, and field codes, are fully documented
  - Any item not applicable to a database was noted (such as MeSH terms in Google Scholar).

This detailed reporting ensures transparency and reproducibility of our search methodology, in accordance with PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension) guidelines.

Data Synthesis

Overview

To supplement the literature review, semistructured interviews were conducted with 19 SMEs across the fields of public health, health care informatics, and blockchain-based health technology. The experts were selected using purposive sampling, ensuring diversity in terms of geographic location (United Arab Emirates, Europe, and North America), professional role (policymakers, IT developers, and medical practitioners), and years of experience (ranging from 8 to 25 years). The interview protocol covered themes such as data privacy, trust in decentralized systems, donor engagement, and integration challenges within health care infrastructures. Responses were coded thematically and triangulated with findings from the literature review.

This study was reviewed and approved by the Institutional Review Committee (IRB) at the American University of Sharjah in the United Arab Emirates. All expert participants provided written informed consent for interview participation.

Risk of Bias Assessment

To ensure methodological rigor and transparency, the risk of bias of all included studies was assessed using established appraisal tools appropriate to the study design. Randomized studies were evaluated using the Cochrane Risk of Bias Tool 2, while observational studies were assessed using the Newcastle-Ottawa Scale. For nonrandomized intervention studies, the ROBINS-I was applied where relevant. Two reviewers independently conducted the quality and risk-of-bias assessments, and any disagreements were resolved through discussion and consensus to ensure consistency and reliability. The results of the risk-of-bias evaluation were documented in summary tables and considered during the synthesis and interpretation of findings. Studies identified as having a higher risk of bias were carefully examined during analysis to ensure that conclusions regarding the impact of digital health technologies on outpatient PCQ were not disproportionately influenced by lower-quality evidence.

Screening and Selection

Screening followed PRISMA guidelines, with a flow diagram illustrating identification, screening, eligibility, and inclusion steps.

Data Collection

To organize the findings of the literature and interview questions, several models focused on suggesting dimensions to measure PCQ were reviewed. The SERVQUAL model was selected for this study due to its widespread application, simplicity, and robust structure in assessing service quality across diverse sectors, including health care [28]. It operationalizes service quality through 5 distinct dimensions: tangibility, reliability, responsiveness, assurance, and empathy, which align closely with patient-centered care indicators [29]. Several researchers advocated for implementing the SERVQUAL model as a robust method for assessing service quality in outpatient health care facilities [2]. While alternative models such as the HEALTHQUAL [30] and HCAHPS [31] also assess health care quality, SERVQUAL offers a more generalized and flexible structure that allows for comparative analysis across settings and time, making it suitable for exploratory research in dynamic or cross-cultural health care environments [32]. Furthermore, unlike HCAHPS, which is limited to inpatient experiences and largely administrative reporting, SERVQUAL captures perceived service gaps directly from users, making it more relevant for studies focused on user satisfaction and perceived service deficiencies [33]. Thus, its adaptability, empirical grounding, and comparative advantage in measuring perceived service quality support its use in this context.

The SERVQUAL model is a widely used framework for measuring and managing service quality across various industries. It was developed based on the premise that customers’ perceptions of service quality are influenced by the gap between their expectations and their actual experiences with a service. SERVQUAL breaks down service quality into 5 dimensions, often represented by the acronym RATER. First, reliability, which refers to the ability of a service provider to deliver its services consistently and accurately. Customers expect services to be performed dependably and accurately. Second, assurance relates to the knowledge and courtesy of employees and their ability to inspire trust and confidence. It includes factors such as employees’ competence, credibility, courtesy, and ability to convey trust and confidence. Third, tangibles refer to the physical evidence and appearance of facilities, equipment, personnel, and communication materials. This dimension encompasses the appearance of physical facilities, equipment, personnel, and communication materials. Fourth, empathy which involves the degree of caring, individualized attention, and understanding that the service provider gives to its customers. It relates to the provider’s ability to understand and empathize with customers’ needs and concerns. Fifth, responsiveness, which refers to the willingness of the service provider to help customers and provide prompt

XSL•FO
RenderX

service. It involves aspects such as a willingness to assist customers promptly, attentiveness, and responsiveness to customer requests or inquiries.

The SERVQUAL model is typically assessed using a survey questionnaire that measures customers' perceptions and expectations across these 5 dimensions. Customers are asked to rate both their expectations of an ideal service (ie, how they believe a service should be) and their perceptions of the actual service received. The difference between customers' expectations and perceptions provides insights into the gaps in service quality. This survey questionnaire was used in interviewing 23 SMEs, yet only feedback from 19 experts operating in an outpatient clinic setting was considered. Those respondents operate in different capacities as outlined in Table 1 below.

Through this research review paper, ethical measures are followed in terms of honesty in data reporting, objectivity, confidentiality, integrity and sincerity, human error and negligence avoidance, and transparency in results sharing.

Table . Interviewed SME[a] demographics.

Job title	Age bracket (years)	Years of experience
Continuous improvement project manager	40 - 49	11 - 15
Receptionist	30 - 39	6 - 10
Clinic information officer	30 - 39	3 - 5
Customer care associate	30 - 39	3 - 5
Administrator	40 - 49	6 - 10
Receptionist	30 - 39	11 - 15
Charge nurse (Aviation Medicine Clinic)	40 - 49	16 - 20
Officer	40 - 49	≥20
Registered nurse	50 - 59	≥20
Doctor	40 - 49	16 - 20
Manager clinical systems	40 - 49	11 - 15
Registered nurse	30 - 39	6 - 10
Registered nurse	30 - 39	11 - 15
In-flight nurse	30 - 39	6 - 10
Aviation nurse	40 - 49	11 - 15
Executive director, nursing services	40 - 49	≥20
Respiratory therapist	50 - 59	≥20
Practical nurse	30 - 39	6 - 10
Registered nurse	40 - 49	≥20

[a]SME: subject matter expert.

Ethical Considerations

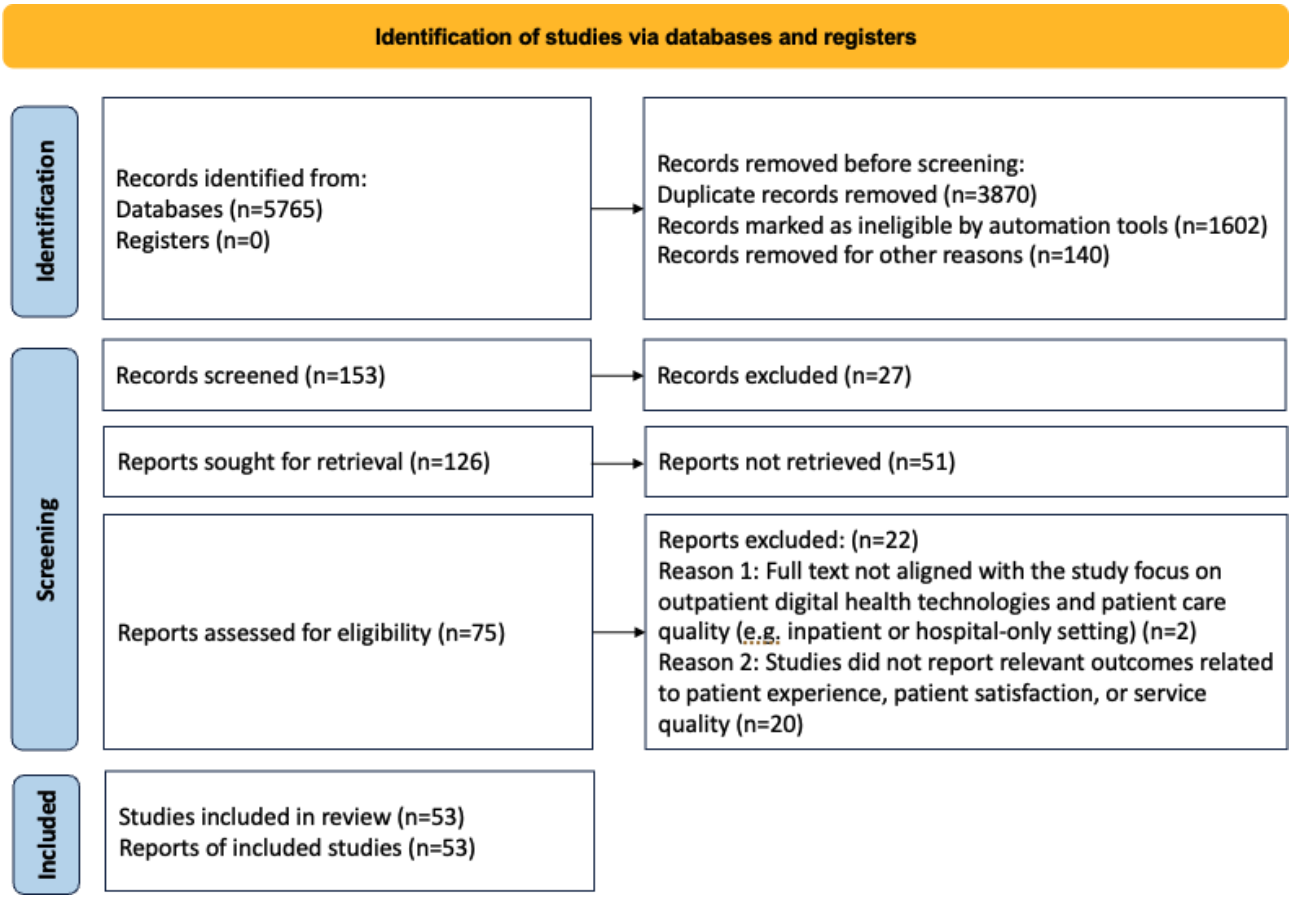
This study was approved by the IRB Committee of the American University of Sharjah in the United Arab Emirates (Protocol Number 25-091). Written informed consent was obtained from all interview participants.

Results

The initial search generated 5765 records, which were screened by title and abstract. After applying inclusion and exclusion criteria and assessing full texts, 53 relevant articles published between 2017 and 2022 were included (Figure 2).

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram.



Literature reviewed part of the SLR consists of research studies and case studies. The review of the case studies was not to offer an exhaustive ethnographic account, but rather to illustrate the practical implications of identified trends. An augmented version of SERVQUAL was applied selectively as a structuring tool during the literature synthesis phase to help organize findings under recognized quality dimensions. Other dimensions such as trust, data governance, and interoperability emerged more prominently in expert and case study findings, which necessitated a broader, more adaptive analytical approach. This divergence reflects the evolving nature of health care service delivery in the digital era and underscores the need for hybrid evaluative frameworks that go beyond traditional service quality models.

Interview respondents confirmed that safety is the most important quality dimension, followed by patient-centeredness from a patient perspective, as demonstrated in Table 2. The study assumes that emerging technologies generally yield positive outcomes across quality dimensions for a favorable patient experience. Tables 3 and 4 provide a summary of the SMEs' feedback on the impact of technology on organizational improvement objectives and PCQ dimensions, respectively. Some of the organizational objectives are focused on PCQ, such as overall satisfaction, patient education, and patient information sharing.

Table . Which quality dimension is considered the most important from a patient perspective?

Dimension	Responses, n
Not disclosed	1
Effectiveness	2
Efficiency	1
Equity	3
Patient-centeredness	4
Safety	5
Timeliness	3
Total	19

XSL•FO
RenderX

Table . Technology impact on organizational improvement objectives.

Technology	Organizational improvement category						
	Patient satisfaction	Caregiver perception	Data collection	Productivity	Reduce cost	Patient education	Patient information sharing
Robotic process automation	1-PI[a]	1-PI	1-PI	2-SPI[b]	2-SPI	1-PI	1-PI
Telemedicine and virtual visits	1-PI	1-PI	0-NI[c]	1-PI	2-SPI[d]	1-PI	1-PI
Robotics	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI
Mobile health portal	2-SPI	1-PI	1-PI	2-SPI	1-PI	1-PI	1-PI
Videoconferencing	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI
Other educational technology	2-SPI	2-SPI[d]	1-PI	1-PI	1-PI	2-SPI[d]	1-PI
3D digital anatomy modeling	2-SPI[d]	2-SPI[d]	2-SPI[d]	0-NI	2-SPI	2-SPI[d]	2-SPI
Insurance processing AI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI
Health care wearables	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI
Personal digital assistants	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI
Disease history analysis	2-SPI	2-SPI[d]	2-SPI	1-PI	1-PI	2-SPI	2-SPI
Teleconference podcasts	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI
Electronic medical records	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI[d]	2-SPI
On-demand health care solutions	2-SPI[d]	2-SPI[d]	2-SPI[d]	2-SPI[d]	1-PI	2-SPI[d]	1-PI
Connected ambulance	2-SPI	2-SPI	2-SPI	2-SPI[d]	1-PI	1-PI[d]	2-SPI

[a]PI: positive impact.

[b]SPI: strong positive impact.

[c]NI: no impact.

[d]When there is a tie in feedback, the highest scale is considered (eg, equal representation of 2-SPI, 1-SPI, and O-NI).

XSL•FO
RenderX

Table . Technology impact on PCQ[a] dimensions legend.

Technology	Quality dimension								
	Professional competence and credibility	Service reliability and assurance	Patient-centeredness and empathy	Responsiveness, waiting, and service time	Nurse care	Access to care/care equality	Service environment and cleanliness	Communication	Technical support
Robotic process automation	2-SPI[b]	2-SPI	0-NI[c]	2-SPI	1-PI[d]	2-SPI	1-PI	1-PI	2-SPI
Telemedicine and virtual visits	1-PI	1-PI	1-PI	2-SPI[e]	1-PI	1-PI	1-PI	1-PI	1-PI
Robotics	2-SPI	1-PI	2-SPI[e]	1-PI	1-PI	2-SPI[e]	1-PI	1-PI	2-SPI[e]
Mobile health portal	2-SPI[e]	1-PI	2-SPI[e]	1-PI	1-PI	1-PI	1-PI	2-SPI	2-SPI
Videoconferencing	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI	2-SPI[e]	1-PI
Other educational technology	1-PI	1-PI	1-PI	1-PI	2-SPI	2-SPI	1-PI	2-SPI[e]	2-SPI
3D digital anatomy modeling	0-NI	1-PI[e]	1-PI[e]	0-NI	1-PI	0-NI	1-PI[e]	1-PI	1-PI
Insurance processing AI	2-SPI[e]	2-SPI	2-SPI[e]	2-SPI	2-SPI[e]	2-SPI	2-SPI*[e]	2-SPI[e]	2-SPI[e]
Health care wearables	1-PI	1-PI	0-NI	0-NI	1-PI	1-PI	1-PI	1-PI	1-PI
Personal digital assistants	1-PI	1-PI	0-NI	1-PI	1-PI	1-PI	1-PI	1-PI	1-PI
Disease history analysis	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI[e]	0-NI	2-SPI	2-SPI
Teleconference podcasts	1-PI	1-PI[e]	1-PI[e]	0-NI	1-PI	1-PI[e]	1-PI	1-PI	1-PI
Electronic medical records	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI	2-SPI
On-demand solutions	2-SPI	2-SPI[e]	2-SPI	2-SPI	2-SPI	2-SPI	1-PI	2-SPI	2-SPI
Connected ambulance	0-NI	2-SPI	2-SPI	2-SPI	1-PI	2-SPI	2-SPI[e]	1-PI	2-SPI[e]

[a]PCQ: patient care quality.

[b]SPI: strong positive impact.

[c]NI: no impact.

[d]PI: positive impact.

[e]When there is a tie in feedback, the highest scale is considered (eg, equal representation of 2-SPI, 1-SPI, and O-NI).

One of the major findings both the SLR and interviews revealed is the patient-centeredness. Patients differ significantly in their attitudes and preferences; some may be submissive, readily accepting prescribed treatments, while others might be technology enthusiasts who prefer to engage with digital health tools. Additionally, there are those who might challenge medical advice and seek more information or alternative options. Recognizing and addressing these variations is crucial for delivering personalized and effective care, especially if technology is used [34].

Discussion

Principal Findings

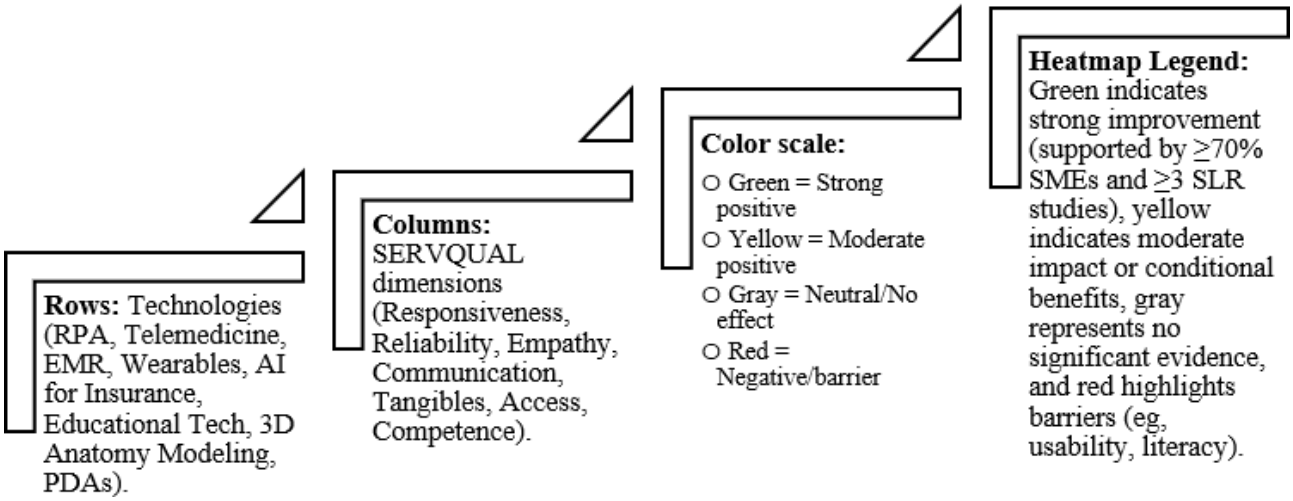
In the next sections, the results are discussed in two stages: (1) SLR findings structured by SERVQUAL dimensions, and (2)

XSL•FO
RenderX

SME insights triangulated with literature trends. Key quantitative outcomes from reviewed studies (eg, reductions in waiting time and improved satisfaction scores) are tabulated, while qualitative themes emphasize usability barriers, empathy gaps in telemedicine, and digital literacy challenges. Findings from the SLR and expert interviews (n=19) were mapped against SERVQUAL dimensions to illustrate the impact of emerging technologies. Figure 3 presents a heatmap showing technology influence across dimensions, while Table 1 details key quantitative effects and expert agreement rates.

This study examined whether the adoption of emerging digital health technologies improves PCQ in outpatient clinics. By synthesizing findings from 53 peer-reviewed studies and triangulating them with insights from 19 SMEs, the results indicate that digital technologies generally enhance several dimensions of outpatient service quality. The most consistent improvements were observed in responsiveness, communication, and operational efficiency through technologies such as robotic process automation, telemedicine, and EMRs. However, the findings also highlight that technological implementation alone does not automatically improve patient-centeredness or empathy, and that factors such as usability, digital literacy, and clinician training influence the extent to which technology translates into improved patient experiences. Overall, the results partially support the study hypothesis that emerging technologies enhance PCQ dimensions, while also emphasizing that human factors and organizational readiness remain critical to achieving patient-centered outcomes.

Figure 3. Heatmap of technology impact on SERVQUAL dimensions. EMR: electronic medical record; PDA: personal digital assistant; RPA: robotic process automation; SERVQUAL: Service Quality Model; SLR: systematic literature review; SME: subject matter expert.



Dimension 1: Responsiveness and Timeliness (Waiting Time and Service Time)

Excessive patient waiting time at outpatient clinics yields undesired levels of patient satisfaction and demands the need to be efficiently managed from a cost and duration perspective [8]. Patients perceive being at the receiving end of service quality when there is an patients strictly follow their scheduled appointments and doctors are available according to the promised time communicated to them [35].

The most promising technologies cited or appraised by SMEs are RPA, EHR, and connected ambulances. RPA relies on a combination of workflow, business rules, and presentation layer integration with information systems to act as a semi-intelligent user of the system, which can be used in preauthorization, updating patient records, and billing (repetitive tasks) [36]. RPA can be used for temporary or permanent outpatient clinics [37]. documented the use of RPA and its positive impact on the waiting time for an urgent care service for COVID-19. RPA was used to develop a web platform digital solution that uses modern Hode-JS, RESTful, and SOAP technologies to process data entries by the nurses, in addition to automating the administrative process at the EHR, replacing human intervention when interpreting data entry. The queuing system number generated by the EHR for the patient and displayed on a visual display, in addition to the patient's federal identification number, are the main field dependencies of this solution. Once the nurse concludes the defaulted health screening, the patient's federal identification number is input whereby the digital solution searches for the data linked to this identifier and completes the medical care registration form with all the mandated information for a complete medical record and links it to the queuing system number. This ensures a definitive linkage to the patient's health record, queuing number after the medical care registration, expediting the patient's journey immediately to the doctor's consultation. This solution is installed and used at the nursing workstations by properly trained nursing caregivers [37]. Despite the success stories of using EHR on responsiveness and accuracy, observational studies highlighted a negative impact on patient-physician communication behaviors when physicians are engaged with computers and ignoring their patients [38]. These moments of silence or screen gaze could be due to complicated systems or lack of training. Also, the frustration of physicians with the EHR system may translate into dissatisfaction, resistance to use, and missing some key information [39].

RPA has proven to be the technology of the future, providing sustainable solutions that reduce costs and delivery time, improve quality, and the speed and operational efficiency of business process complexities [7,36]. Additionally, it integrates different systems using software robots to automatically acquire and integrate data from clinical applications, lab information

XSL•FO
RenderX

systems, third-party portals, insurance portals, radiology information systems, enterprise resource planning (ERP) and human resources (HR) applications, and scheduling applications. RPA frees employees from the repetitive tasks of scheduling so they can apply their skills to scenarios that require human interaction and touch and automate the eligibility requests to access information for better communication to providers and patients [40]. RPA drastically reduced the amount of time staff spent on repetitive and routine activities from appointment scheduling, hence optimizing appointment scheduling, patient account creation and verification, file handling and processing, inventory management, claims processing, and test management [41].

Another form of digital transformation is in the use of connected ambulances, which are vehicles equipped with advanced communication technology that allows them to transmit vital patient information, such as medical history, vital signs, and real-time video, to health care providers before the patient arrives at the hospital. If care is not ensured in the emergency department, follow-up treatment in the outpatient clinics will be negatively impacted. These ambulances are essentially mobile telemedicine platforms, enabling health care professionals to assess and prepare for incoming patients more effectively [42]. The impact of connected ambulances on service reliability in outpatient clinics can be significant. By providing health care providers with detailed information about incoming patients, including their medical condition and needs, connected ambulances facilitate better preparation and coordination of care upon the patient's arrival. This can lead to reduced waiting times, more timely interventions, and improved overall patient outcomes. Overall, the use of connected ambulances can help outpatient clinics improve service reliability by streamlining processes, enhancing communication, and optimizing resource allocation to better meet the needs of patients [43].

Tables 3 and 4 depict respondents' perspectives on the influence of RPA on various organizational aspects, driving its adoption. It also delineates the impact of RPA on quality dimensions (including responsiveness, waiting, and service time), predominantly indicating favorable effects across organizational domains. Regarding connected ambulances, respondents affirmed they would indirectly positively affect responsiveness, waiting times, and service efficiency, as well as other dimensions, except for physicians' professional competence, credibility, and technical performance.

Table 5 demonstrates a strong alignment between expert opinion and the systematic literature regarding the impact of digital technologies on outpatient health care services. RPA received the highest level of SME consensus (84%), followed by telemedicine (74%) and AI-enabled insurance processing (71%), indicating that experts perceive these technologies as having the greatest potential to improve operational efficiency, reduce waiting times, and streamline administrative processes. Telemedicine also had the strongest evidence base in the literature, followed by EMR and EHR systems and RPA, reflecting substantial research support for their effectiveness. Across technologies, reported benefits consistently included enhanced efficiency, improved patient access, strengthened data integration, better remote monitoring, and increased patient education, although challenges such as digital literacy and usability remained important implementation considerations. Overall, the findings suggest broad agreement between SME perspectives and published evidence regarding the value of digital technologies in improving outpatient health care delivery.

Table . SME[a] consensus rates and literature support per technology (19 SMEs and 53 SLR[b] studies).

Technology	SMEs reporting positive impact (%)	Number of SLR studies supporting	Key outcomes reported
RPA[c]	84	7	↓ Waiting times (20% - 40%), ↑ efficiency
Telemedicine and virtual visits	74	11	↑ Access, ↑ satisfaction, digital literacy required
EMR[d] and EHR[e]	63	9	↑ Data integration, usability issues noted
Health care wearables	68	6	↑ Remote monitoring, literacy barriers
AI (insurance processing)	71	5	↑ Claim turnaround, ↓ admin time
Educational technology	60	4	↑ Empathy training outcomes
3D anatomy modeling	58	3	↑ Patient education, trust
PDAs[f] and decision support tools	55	3	↑ Physician decision support

[a]SME: subject matter expert.

[b]SLR: systematic literature review.

[c]RPA: robotic process automation.

[d]EMR: electronic medical record.

[e]EHR: electronic health record.

[f]PDA: personal digital assistant.

XSL•FO
RenderX

Dimension 2: Reliability (Competence)

Applying reliability principles to health care has the potential to help reduce "defects" in care or care processes, increase the consistency with which appropriate care is delivered, and improve patient outcomes. It is a failure-free operation over time and an individual patient's experience over time, particularly effectiveness, timeliness, and patient-centeredness. Providing reliably good care over time requires understanding and addressing the reasons that patients are returning to the clinics [44]. In health care, reliability must be built across, not just within, groups and organizations; it has come to be used as a proxy for safety despite reliability and safety not being equivalent [45].

In outpatient clinics, the application of robotics and automation is broadly designated, such as robot areas, to assist patients and disseminate information about various units of the specialized clinic and hospital as a whole. Those robots can handle several visitors without becoming tired, stressed, or overloaded and direct the patient to the correct physician of choice. They are deemed attractive to patients, especially children, enhancing the overall patient experience. Moreover, humanoid nurse robots (HNRs) are being used to care for older adult patients, offering 24/7 and 7/7 service with minimal cost, inclusive of lifting and therapeutic assistance. Telemedicine robots are helpful in telemedicine applications where a virtual doctor takes all the physiological parameters and diagnoses a disease using audiovisual aids. Those systems are extremely useful in the case of pandemics whereby the patients are located in remote areas and cannot reach physicians for clinical diagnosis and treatment; those robots are used as a low-cost remote interface [46]. Vendors have developed sophisticated remote-controlled cars with advanced cameras, lenses, and connected microphones, which enable the physician to interact with the patient's bedside despite being miles away [47]. Nowadays, different versions of telemedicine such as online health platforms [48] and remote presence virtual and independent telemedicine assistants are widely used by hospitals [49].

Furthermore, an on-demand health care solution refers to a service or platform that provides immediate access to medical care and resources whenever and wherever needed. This could include telemedicine platforms, virtual consultations, remote monitoring devices, or mHealth apps. On-demand health care solutions leverage technology to connect patients with health care providers in real-time, offering convenience, accessibility, and timely assistance [50].

Tables 3 and 4 illustrate the viewpoints of participants regarding how telemedicine and virtual visits influence different organizational aspects, prompting their acceptance of this technology. It outlines the impact of these advancements on 9 quality measures, such as service reliability and assurance, indicating predominantly positive effects across organizational realms. Nonetheless, respondents noted that while telemedicine enhances most quality measures, it falls short in improving data collection capabilities and demands patients to possess technological proficiency for maximal benefits. Moreover, some studies suggest that reducing the visit length using technology is associated with medical errors such as inappropriate antibiotic prescriptions [51]. Conversely, on-demand health care solutions are seen to positively affect all quality measures.

Integrating technical support services with the implementation of EMRs in health care outpatient centers and clinics is essential for ensuring smooth operational performance, maintaining system reliability, and promptly addressing any technical challenges that may arise during system use [52]. Technical support teams play a fundamental role throughout the lifecycle of EMR deployment, beginning with the initial implementation phase and continuing through ongoing system optimization and maintenance.

During the early stages of EMR implementation, technical support teams provide critical assistance with system installation, configuration, and data migration. Their specialized expertise ensures that health care organizations can transition from traditional record systems to digital platforms with minimal disruption to clinical workflows and patient care processes. By carefully managing these technical transitions, support teams help reduce operational interruptions and facilitate a smoother adoption of EMR systems across health care facilities [53].

In addition to implementation support, technical teams are responsible for conducting training sessions and developing educational materials that help health care professionals effectively use EMR systems. These training initiatives typically include guidance on accurate data entry, system navigation, and the use of advanced system functionalities. Well-trained health care personnel are better able to use EMR capabilities efficiently, which contributes to improved accuracy, enhanced documentation quality, and more efficient patient care delivery [54].

Technical support teams also play a key role in the day-to-day operation of EMR systems by providing timely troubleshooting services and resolving technical issues that may occur during routine clinical use. Their responsibilities often include performing regular system maintenance tasks such as applying software updates, installing security patches, and optimizing system performance to ensure reliability and data protection. Furthermore, these teams collaborate closely with health care providers to customize EMR systems according to specific clinical workflows and operational requirements. They also facilitate the integration of EMR systems with other health care information technologies, including laboratory information systems and picture archiving and communication systems (PACS). Such integration enhances interoperability between different health care platforms, enabling more efficient data exchange and improving the overall effectiveness of health care delivery [55,56].

Another critical function of technical support teams involves maintaining the security and regulatory compliance of EMR systems. Health care organizations must adhere to strict data protection regulations, including standards such as the Health Insurance Portability and Accountability Act (HIPAA). Technical teams therefore implement comprehensive security mechanisms, including access control protocols, data encryption, and periodic security audits to safeguard sensitive patient information from unauthorized access or potential breaches. Ensuring regulatory compliance not only protects patient privacy

XSL•FO
RenderX

but also helps health care organizations avoid legal and financial risks associated with data security violations [57].

To ensure continuous system availability, many technical support providers also offer 24-hour helpdesk services and emergency assistance. This round-the-clock technical support is particularly important in health care environments where uninterrupted access to patient data is essential for timely clinical decision-making. Rapid response to system failures or technical issues helps minimize downtime and ensures that health care professionals can access critical patient information whenever needed, thereby supporting effective clinical care and patient safety [57].

Beyond operational support, technical teams continuously monitor the performance and usage patterns of EMR systems to identify opportunities for improvement. They gather feedback from health care staff and collaborate with EMR vendors to implement system updates, enhancements, and new functionalities that align with evolving health care standards and best practices. Through this ongoing monitoring and system refinement process, health care organizations can maintain efficient, reliable, and secure EMR infrastructures. Ultimately, the integration of dedicated technical support services with EMR implementation and maintenance significantly enhances the effectiveness of EHR systems and contributes to improved health care quality and patient outcomes [58].

The adoption of EMRs has significantly transformed health care delivery by improving operational efficiency, documentation accuracy, and accessibility of patient information. As health care systems increasingly transition from paper-based records to digital platforms, EMR systems offer multiple advantages that enhance both clinical and administrative functions within health care organizations.

One of the primary benefits of EMR implementation is the streamlining of administrative tasks. Processes such as patient record management, documentation, and prescription handling become more efficient through digital systems, allowing health care providers to reduce time spent on manual paperwork. As a result, clinicians can allocate more time and attention to direct patient care, ultimately improving the quality of health care services provided [59].

EMRs also provide a comprehensive and accurate overview of a patient's medical history, enabling health care professionals to access complete and organized clinical information in a timely manner. This improved accessibility supports more informed clinical decision-making and reduces the likelihood of medical errors that may arise from illegible handwriting, incomplete documentation, or fragmented paper records. By ensuring that patient data are clearly documented and readily available, EMRs contribute to safer and more reliable health care delivery [60].

In addition, EMR systems facilitate remote access to patient records, which significantly enhances communication and collaboration among health care providers. Physicians, specialists, and other health care professionals can access patient information from different locations, allowing for faster consultations, coordinated care, and improved clinical outcomes [61]. Patient engagement is also strengthened through EMR-integrated patient portals, which enable individuals to view their medical information, communicate electronically with health care providers, and manage appointments through online platforms. These features empower patients to become more actively involved in managing their own health and health care decisions [62].

Furthermore, the digital nature of EMR systems allows health care organizations to aggregate and analyze large volumes of clinical data. This capability supports advancements in population health management, enables more robust clinical research, and facilitates quality improvement initiatives across health care systems. By leveraging EMR-generated data, health care providers and policymakers can identify health trends, evaluate treatment outcomes, and implement strategies that enhance the overall effectiveness and efficiency of health care delivery.

While initial implementation costs can be substantial, long-term cost savings are realized through increased efficiency, reduced paperwork, and better management of chronic conditions [63]. Furthermore, the interoperability of EMRs improves continuity of care by enabling seamless sharing of patient information between different health care providers and systems [60]. Enhanced compliance with regulations such as HIPAA and robust security measures ensures the protection of patient data, further bolstering the benefits of digital transformation in health care using EMRs [64].

Tables 3 and 4 illustrate the perspectives of respondents regarding the influence of EMR on different organizational aspects, motivating their interest in adopting this technology. It also outlines the effects of EMR systems on quality dimensions, such as technical support services, which show improvement with the digitization of medical records and related clinical operations. This primarily indicates a positive impact across various organizational sectors.

Dimension 3: Environment of Service Delivery (Cleanliness)

Cleaning and disinfection of surfaces in clinics are strategies to minimize the occurrence of health care–associated infections. Given that clinics with it increase patient volumes are environments that act as a pool for bacteria and viruses that may spread in the clinic causing infectious disease. For stringent infection control, an intensified cleaning and disinfection process is applied [65]. Moreover, strict policies and procedures to eliminate cross-contamination in outpatient clinics are essential [66], such as disposable glove changes, caregiver's hand washing, clinic surfaces being disinfected often, tools and equipment cleaning and sterilization, and single-use supplies disposal.

Cleaning robots are used in outpatient clinics; they use intelligent navigating dry vacuum pumps for dry or wet mopping activities, being an integral part of disinfecting the physical space from germs and pesticides. Moreover, the UV radiation-based robot is used to disinfect the clinic from microbes and clean washrooms using sensing systems. Spraying and disinfection robots widely spread antiseptic combinations over large indoor and outdoor areas [46].

XSL•FO
RenderX

Tables 3 and 4 depict the respondents' views on how robotics affects various organizational aspects, driving their inclination toward adopting this technology. It also delineates the impact of robotics on quality dimensions, including the environment of delivering service, primarily indicating a favorable influence across organizational sectors. Robotics enhances surgical precision (outcome of surgery), resulting in expedited recovery and rehabilitation, thereby improving surgical outcomes.

Dimension 4: Communication

Interpersonal communication plays a key role in the patient's treatment plan; not only does it impact patient satisfaction levels with the offered clinical service but also it places them in the role of a partner in the journey building an environment of empathy, friendliness, kindness in treatment, safety, and trust. Enhanced communication guarantees that patients have access to their health information and ultimately enhances the quality of patient care within an outpatient clinic setting [67]. Moreover, communication errors can be a profound contributing factor in outpatient clinic sentinel events resulting in poor health outcomes alongside elevated distress levels in patients and their family members. Effective communication skills are deemed a core clinical skill expected from every clinical caregiver when effectively exchanging information between a patient and them and include communication with the family and carer as well. Among the key elements of effective communication in clinics is the ability to foster relationships, the 2-way exchange of information, convey empathy, engage patients in decision-making and care planning, and manage complexity and uncertainty [68].

Patient preferences are leaning toward using technology when communicating about their symptoms with their health providers in clinics especially after discharge from outpatient clinics. Email communication, phone, text messaging, mHealth portal apps, social media, online discussion forums, and videoconferencing [69]. Furthermore, patient education is a crucial element within health care. It is known as forecasters for increased patient engagement levels in shared decision-making, improved medication and treatment adherence, higher levels of satisfaction, and even better overall treatment outcomes. Most patients are simply not capable of processing large amounts of new medical information in a short period of time. Apps for smartphones and tablets have the potential to actively educate patients by providing them with timely information using push notifications [70]. Finally, Owens et al [71] conducted a study to evaluate the use of ambient voice recognition, coupled with natural language processing and AI in patient-physician relationships. The feedback of 288 patients suggests a strong positive correlation between the use of this technology and patient-physician communication, which improved patients' satisfaction.

The impact of on-demand health care solutions on the communication quality dimension is significant. First, these solutions facilitate seamless communication between patients and health care providers, overcoming barriers such as geographical distance or time constraints. Patients can easily connect with health care professionals via video calls, chat messaging, or phone calls, enabling timely access to medical advice and assistance [72]. Moreover, on-demand health care solutions often incorporate features that enhance the quality of communication. These may include secure messaging platforms, multimedia capabilities for sharing images or documents, and real-time data exchange functionalities. By providing clear channels for communication, these solutions enable patients to effectively convey their symptoms, concerns, and medical history to health care providers, leading to more accurate diagnoses and personalized treatment plans [73]. Additionally, on-demand health care solutions can improve communication among health care teams. They enable seamless collaboration and information sharing between primary care providers, specialists, and other health care professionals involved in a patient's care [74]. This interdisciplinary communication fosters coordinated and integrated health care delivery, ultimately improving patient outcomes [75].

Tables 3 and 4 present the viewpoints of the respondents regarding the impact of mHealth portal apps and videoconferencing on various organizational aspects, which motivate their adoption of these technologies. It also delineates the effects of these apps on quality dimensions, notably communication, predominantly indicating a significantly positive influence across organizational sectors. Particularly for mHealth portal apps, an equal number of respondents expressed a perception that, unlike their counterparts, they observe no impact on physicians' professional competence, credibility, technical performance, as well as patient-centeredness and empathy (staff demeanor). However, for the sake of this research, the favorable responses were taken into consideration.

Dimension 5: Empathy (Staff Demeanor)

Terry and Cain [76] stressed the criticality of empathy throughout the patient care journey and optimal patient-caregiver relationship. It increases both the patient's satisfaction and compliance and elevates the caregiver's ability to cater to the clinical needs of their patients. Furthermore, it has a strong correlation and positive impact on the health outcomes of patients and decreases the risk of any potential malpractice liability. Empathy in its core element does not change yet evolves to align with the digital advancements in the channels of provisioning and communicating clinical services.

With the widespread use of telemedicine in outpatient clinics, the construct of digital empathy has become an increased necessity and must be integrated into the academic and training curricula of health professionals [76]. An increasing trend in the use of digital health tools to engage patients and enhance outcomes is seen; with its effectiveness being dependent on the consistency of the patients' use and engagement, with empathy being the foundation of such tools. This has further increased the importance of not only the caregiver's increase in their telemedicine levels but also the importance of architects being more considerate with their home designs, especially with working from home being the future of work, especially within the lifespan; the same homes may have to accommodate chronically ill patients. Hence, making the design of sustainable and healthy homes a new trend, that is, designing for well-being in residential health care. It is more than having a healthy physical body and includes social, spiritual, emotional, and

XSL•FO
RenderX

intellectual well-being. Incorporating nature into the environment through biophilic design creates a sense of serenity, therapeutic, and well-being effect in the homes of patients [77].

Tables 3 and 4 provide insights into how educational technology influences various aspects within outpatient clinics, shedding light on why these clinics are increasingly inclined to adopt such technology. The findings suggest that educational technology has a predominantly positive effect across different areas within outpatient clinics, indicating its potential to enhance patient care, staff interactions, and overall clinic operations. This underscores the importance and potential benefits of integrating educational technology into outpatient clinic settings, not only for improving efficiency and effectiveness but also for fostering positive patient experiences and enhancing the quality of care provided through more empathy levels by the care providers [78]. Tables 3 and 4 depict respondents' perspectives on the influence of RPA on various organizational aspects, driving its adoption. Respondents observed that while RPA improves most quality dimensions, it does not significantly enhance patient-centeredness or empathy (staff demeanor). It also delineates the impact of RPA on quality dimensions (including responsiveness, waiting, and service time), predominantly indicating favorable effects across organizational domains. Such findings are in line with the study by Chen et al [79], who suggested that efficiency improvement that results in shorter visit duration can negatively impact the provision of counseling and the patient-physician relationship and communication.

Dimension 6: Nurse Care

According to the study by Oermann and Templin [80], patients perceive high nursing quality when they are cared for by nurses who are knowledgeable and resourceful. Having direct access to the nurse and can communicate and spend enough time not being rushed during the clinic visit. Moreover, they consider the fact that a nurse educates them and teaches them more about their illness (especially among those who are less educated, have lower income levels, and have been living with chronic illness) and the medications that they require; reiterating the treatment plan and expectation setting alongside coaching them on how to remain healthy was a quality differentiating factor. Additionally, having the ability to reach out at any given time for clarifications, concerns, or guidance is becoming more critical than ever.

Patients are expecting more detailed explanations about their health situation to increase their understanding and minimize their anxiety levels and complaints; making the use of 3D digital anatomy modeling an aid to teaching and its use in patient education and operation planning. Clinicians will be expected to ensure that the information is retained and risks and benefits weighed up in the 3D computer animation. Patients' understanding improved, their trust levels increased in the professional treatment, and their anxiety reduced toward their medical conditions and treatment plans [81].

Tables 3 and 4 illustrate the opinions of participants regarding the influence of 3D digital anatomy modeling on different organizational facets, motivating their interest in embracing this technology. It also outlines how this technology affects 9 quality aspects, including nursing care, mostly suggesting a positive effect across organizational domains and that it permits a more effective 2-way dialogue with the patient. Nonetheless, respondents noted that for 4 dimensions, they observed no impact of this technology on (1) improving productivity, (2) physicians' professional competence, credibility, and technical performance, (3) responsiveness and waiting and service time, and (4) access to care and health care inequality.

Dimension 7: Access to Care and Health Care Equality

Access to medical care is a comparatively multifaceted phenomenon categorized by potential access and realized access. The latter is what matters in outpatient clinics, as it resembles the use of physician time for all specializations, which makes the design of robust policies to optimize this quality determinant. If patients can't reach care, their satisfaction levels will drop negatively, impacting their treatment journey [82].

Digital transformation improves access to health care services and supports health insurance processing large numbers of claims and payments, which upscales the service of outpatient clinics [83]. Davenport and Kalakota [36] pointed out that machine learning, neural networks, deep learning, and precision medicine are where AI predicts what treatment protocols are likely to succeed for a patient based on various patient attributes and his or her treatment context. In addition, Bruce and Hardy [84] added that the use of natural language processing (NLP) is where AI involves the creation, understanding, and classification of clinical documentation and published research, viewing disease history and analyzing unstructured clinical notes on patients, preparing reports, and transcribing patient interactions. An example is IBM's Watson and Google DeepMind Health. AI combined with consumer wearables and other medical devices is applied to oversee early stages of heart diseases, allowing caregivers to monitor and detect potentially life-threatening episodes at an early stage.

Tables 3 and 4 present the perspectives of respondents concerning the impact of AI in insurance processing on various organizational aspects, stimulating their inclination toward adopting this technology. It also delineates the influence of this technology on 9 quality dimensions, with particular emphasis on access to care and health care inequality, generally indicating a positive impact across organizational realms. It is noteworthy that some respondents highlighted its tendency to diminish the soft skills of insurance professionals despite its benefits. Regarding health care wearables, participants affirmed that they observed a beneficial influence across various quality aspects except for patient-centeredness and empathy (staff behavior) and responsiveness, waiting, and service time, as they believe these dimensions remain unaffected by this technology. One main challenge to information-access equality is patients' digital literacy and access. Negative attitudes toward technology, disinterest in learning new skills, physical and mental disability, and financial hardships are among the reasons for technology avoidance [85]. Such challenges may undermine the benefits gained from technological adoption, such as telemedicine, and make it a curse rather than a blessing.

XSL•FO
RenderX

Dimension 8: Physician's Professional Competence, Credibility, and Technical Performance

This dimension is concerned with the physician consultation stage of the patient journey in terms of the way the doctor diagnoses and explains the treatment plan to the patient, reflecting their competency level as a key determinant of quality for their clientele. Moreover, leadership training, systems theory and analysis, cross-disciplinary training and multidisciplinary teams, understanding and respecting the skills of other practitioners, population health management, palliative care and end-of-life and resource management and medical economics, health policy and regulation, less ''captain of the ship'' and more ''member/leader of the team,'' empathy and customer service, time management, conflict management, providing performance feedback, understanding cultural and economic diversity, and emotional intelligence were identified as the core physician skills required for the next generation of health care delivery [86].

Medical information technology throughout the continuum of the physician's lifelong journey has become a widespread phenomenon providing physicians access to teleconference podcasts, consulting resources in PDAs for enhanced patient management decisions at the point of care. Medical simulators are another form of technology that is also increasingly being used for targeted training and assessment that would elevate the technical skill set levels of care providers [87]. A reliable disease history analysis is integral to assessing a physician's professional competence, credibility, and technical performance. Physicians who possess a thorough understanding of their patients' medical histories demonstrate competence in their ability to gather and analyze pertinent information critical to diagnosis and treatment decisions. This competence enhances their credibility among patients and colleagues alike, fostering trust and confidence in their clinical judgment [88]. Furthermore, a meticulous disease history analysis enables physicians to demonstrate technical proficiency in identifying relevant patterns, risk factors, and potential complications associated with a patient's health status. By effectively using this information, physicians can deliver more precise and tailored care, thereby enhancing their overall performance and ensuring optimal patient outcomes [89]. Thus, a reliable disease history analysis is not only essential for providing high-quality health care but also for upholding the standards of professional competence, credibility, and technical proficiency expected of physicians [90].

Tables 3 and 4 showcase respondent viewpoints regarding the effects of PDAs on insurance processing across organizational facets, encouraging proclivity toward its adoption. It outlines the impact of PDAs on 9 quality dimensions, notably focusing on physician's professional competence, credibility, and technical performance, generally indicating a positive influence across organizational domains, except for patient-centeredness and empathy, where respondents perceived no effect from this technology. Notably, some respondents pointed out the low quality of data collected by PDAs and that using PDAs dilutes accountability and responsibility. In examining disease history, participants confirmed a positive impact across all dimensions except for the environment of delivering service (cleanliness). Conversely, in evaluating teleconference podcast technology, all dimensions showed a positive effect except for responsiveness, waiting times, and service efficiency.

In summary, each technology comes with benefits and some negative impact on some of the quality dimensions. Training and good change-management practices can help overcome these negative aspects and better integrate the technology into the patient care system [91]. Table 4 provides a summary of the positive and the negative impacts of technologies adopted in health care based on the literature review and SMEs' feedback.

The study is limited by its qualitative nature and reliance on SME perspectives, which may not generalize across all outpatient contexts. Additionally, limited patient-reported outcomes restrict conclusions on direct patient experience impact. Quantitative evaluations using patient satisfaction surveys, randomized trials of specific digital tools, and cost-effectiveness analyses are needed. Further exploration into digital literacy interventions and human-centered technology design is recommended to optimize patient-centered outcomes.

This systematic review contributes to the literature by providing an integrated synthesis of how emerging digital health technologies, including telemedicine, wearable monitoring, and clinical automation, collectively influence PCQ in outpatient settings. The review is innovative in its holistic approach, moving beyond technology-specific evaluations that dominate existing reviews and instead examining the combined impact of multiple digital health modalities within a unified analytical framework. By organizing and interpreting evidence across fragmented research streams, the study advances theoretical understanding of digital transformation in health care and clarifies the pathways through which digital tools enhance accessibility, efficiency, and patient-centered care delivery. In practical terms, these findings offer actionable insights for health care leaders, policymakers, and technology developers seeking to implement digital solutions that improve service quality and patient outcomes in real-world outpatient care environments.

Taken together, the findings demonstrate that emerging digital technologies influence outpatient service quality through multiple operational and clinical pathways. Automation technologies primarily improve efficiency-related dimensions such as responsiveness and workflow coordination, while digital communication tools strengthen patient-provider interaction and access to care. In contrast, technologies that rely heavily on patient engagement, such as telemedicine platforms and wearable monitoring devices, reveal the importance of digital literacy and user-centered system design. These findings align with recent literature suggesting that digital transformation in health care improves operational performance but requires complementary organizational strategies to fully enhance patient-centered care. The results therefore reinforce the importance of integrating technological innovation with training, workflow redesign, and patient engagement strategies to achieve meaningful improvements in health care quality.

Despite its contributions, this study has several limitations that should be considered when interpreting the findings. First, the systematic review was limited to studies published in English between 2017 and 2022, which may exclude relevant research

XSL•FO
RenderX

published in other languages or earlier foundational studies. Second, the review synthesized evidence across different health care systems and geographical contexts, which may limit the generalizability of specific findings to all outpatient settings. Third, although expert interviews provided valuable contextual insights, the sample size of 19 participants may not capture the full diversity of perspectives across the health care industry. Additionally, some technologies included in the analysis are evolving rapidly, meaning that their impacts on PCQ may continue to change as implementation practices mature. Finally, the study primarily relied on qualitative synthesis rather than meta-analysis due to heterogeneity across study designs and outcome measures, which limits the ability to quantify the magnitude of technology effects on PCQ.

The findings of this review have several implications for both research and health care practice. From a theoretical perspective, the study demonstrates the value of applying service quality frameworks such as SERVQUAL to evaluate the impact of digital transformation on health care delivery. By integrating evidence across multiple technology domains rather than examining individual innovations in isolation, this review provides a more comprehensive understanding of how digital health ecosystems influence PCQ. From a practical standpoint, the results highlight that health care organizations should approach technology adoption as part of a broader transformation strategy that includes workforce training, workflow redesign, and patient engagement initiatives. Future research should further investigate how different technologies interact within health care systems and explore longitudinal evidence on their long-term effects on patient outcomes and health care equity. Ultimately, the effective integration of emerging digital technologies has the potential to strengthen outpatient care delivery and support more responsive, accessible, and patient-centered health care systems.

Conclusion

Using the augmented SERVQUAL model, the analysis presented in this research offers a structured evaluation of various emerging technologies' impacts on organizational aspects within health care and insurance processing domains. Each technology is assessed across 9 quality dimensions, enabling a comprehensive understanding of its effects on service quality, efficiency, and patient care. The findings reveal a predominantly positive influence of technologies such as RPA, telemedicine, robotics, EMR, mHealth portal apps, educational technology, 3D digital anatomy modeling, AI in insurance processing, health care wearables, PDAs, and teleconference podcast technology on organizational performance. These technologies demonstrate strengths across dimensions such as service reliability, communication, and technical support services.

Patients differ significantly in their attitudes and preferences; some may be submissive, readily accepting prescribed treatments, while others might be technology enthusiasts who prefer to engage with digital health tools. Additionally, there are those who might challenge medical advice and seek more information or alternative options. Recognizing and addressing these variations is crucial for delivering personalized and effective care, especially if technology is used [34]. Although automation is an important enabler for productivity boost, it treats every patient the same way, ignoring the fact that patients have diverse expectations and needs.

However, nuanced insights emerge from the analysis. While most technologies exhibit positive effects, challenges such as data quality issues, technological proficiency requirements for users, and limited impact on certain dimensions are identified. For example, while telemedicine enhances service reliability and accessibility, it may require patients to possess adequate technological skills for optimal use. For example, the use of EHR to reduce patient visits, coupled with poor usability of the system, which results in physicians' distraction, may negatively impact patient-centeredness. Proper integration of EHRs in clinical practice may boost the patient-physician communication dynamics, while poor usability and long keyboarding may jeopardize this communication. Moreover, although technology adoption provided significant reductions in visit and waiting time, shorter visit time may be associated with medical errors and lack of patient-centeredness. As a result, optimal visit time is necessary. Table 6 summarizes the components of the literature search strategy, including the databases searched, the timeframe covered, the language restriction, and the study types included in the review.

Table 7 summarizes the potential positive and negative impacts of the technologies identified in the literature, organized according to health care service quality dimensions and supported by evidence from the included studies.

Table . Components of the literature search.

Component	Details
Databases searched	ABI/INFORM Global, ScienceDirect, Medline (via PubMed), and Google Scholar
Timeframe	January 2017-December 2022 (SERVQUAL[a] foundational literature included up to 1992)
Language	English
Study types	Peer-reviewed empirical studies and systematic reviews

[a]SERVQUAL: Service Quality Model.

XSL•FO
RenderX

Table . Summary of potential impact of technologies.

Dimension and technology	Negative impact	Positive impact
Responsiveness		
EHR[a]	• Missing some key information due to EHR literacy and usability [39]	• Appointment scheduling [35] • Data collection accuracy [37]
RPA[b]	—[c]	• Waiting time reductions [36] • Better communication to providers and patients [40] • Improve efficiency and use [41]
Connected ambulances	—	• Improve incoming patients' treatment effectiveness [42] • Enhance communication and use [43]
Reliability		
HNR[d]	—	• Improve diagnoses and care [46]
Telemedicine	• Visit length reduction may increase medical errors such as inappropriate antibiotic prescription [51] • Reduction of data collection (SMEs[e])	• Improve health accessibility and improve use [47-49]
EMR[f]	—	• Maximizing system uptime and addressing issues that may arise [52,54] • Enhancing data security [57,64] • Enabling seamless sharing of patient information between different providers and systems [60]
On-demand health care	—	• Improve convenience, accessibility, and timely assistance [50]
Robotics	—	• Enhances surgical precision resulting in expedited recovery and rehabilitation
Cleanliness		
Cleaning robots	—	• Intensified cleaning and disinfection improvement [46,65]
Communication		
EHR	• Lack of patient-physician communication behaviors due to data entry and keyboarding distraction [38]	• Enhanced health information accessibility and overall PCQ[g] [67] • Patient engagement in decision-making [68]
Videoconferencing	—	• Increased patient engagement and enhanced treatment outcomes [69,72,73]
Apps	• Shorter visit duration can negatively impact provision of counseling and the patient-physician relationship and communication [79]	• Improve patients' education and timely information notifications [70]
Empathy		
Telemedicine	—	• Enhance patients' engagement
Voice recognition	—	• Improve patient-physician communication [71]
RPA	• No significant impact on empathy (SMEs)	—
Nurse care		

XSL•FO
RenderX

Dimension and technology	Negative impact	Positive impact
3D computer animation	—	• Increase trust levels reduce anxiety [81]
Access to care		
AI in health insurance, mobile apps, and wearables	• Digital illiteracy results in a negative attitude toward technology [85] • Diminishing the soft skills of insurance professionals (SMEs)	• Improve serviceability and satisfaction [36,83]
Professional competence		
Medical simulators	—	• Elevate the technical skill set levels of care providers [87]
PDA[h] and EMR	• Low quality of data collected by PDAs dilutes accountability and responsibility (SMEs)	• Fostering trust and confidence in care providers' clinical judgment [88-90]

[a]EHR: electronic health record.

[b]RPA: robotic process automation.

[c]Not available.

[d]HNR: humanoid nurse robot.

[e]SME: subject matter expert.

[f]EMR: electronic medical record.

[g]PCQ: patient care quality.

[h]PDA: personal digital assistant.

The SERVQUAL model facilitates a systematic examination of these technologies' impacts, enabling stakeholders to identify areas of strength and opportunities for improvement. By addressing challenges and leveraging strengths, organizations can enhance their adoption and implementation of emerging technologies, thereby optimizing service delivery, patient care, and operational efficiency within health care and insurance processing sectors.

Despite extensive efforts to identify and evaluate the determinants influencing digital transformation in health care, existing literature provides limited evidence regarding the measurable value, temporal evolution, and relative prioritization of these determinants, particularly from the patient perspective. Current research has largely focused on conceptual assessments and implementation considerations, while the empirical relationship between digital initiatives, return on investment, service quality, and human-centered care remains insufficiently explored. Therefore, future research should adopt quantitative and mixed methods approaches to validate these determinants, assess their impact over time, and establish evidence-based frameworks that balance technological advancement with patient expectations, operational efficiency, and quality of care.

Acknowledgments

No generative AI tool was used.

The work in this study was supported, in part, by the Open Access Program from the American University of Sharjah, United Arab Emirates.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated and analyzed (including anonymized subject matter expert interview transcripts) are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Checklist 1
PRISMA 2020 checklist.

XSL•FO
RenderX

[DOCX File, 24 KB - rehab_v13i1e81847_app1.docx]

Checklist 2
COREQ checklist.
[PDF File, 433 KB - rehab_v13i1e81847_app2.pdf]

References

1. Cronin JJ, Taylor SA. Measuring service quality: a reexamination and extension. J Mark 1992 Jul;56(3):55-68. [doi: 10.1177/002224299205600304]
2. Ko CH, Chou CM. Apply the SERVQUAL instrument to measure service quality for the adaptation of ICT technologies: a case study of nursing homes in Taiwan. Healthcare (Basel) 2020 Apr 24;8(2):1-16. [doi: 10.3390/healthcare8020108] [Medline: 32344589]
3. Taner MT, Antony J. Comparing public and private hospital care service quality in Turkey. Leadersh Health Serv 2006 Apr 1;19(2):1-10. [doi: 10.1108/13660750610664991]
4. Aagja JP, Garg R. Measuring perceived service quality for public hospitals (PubHosQual) in the Indian context. Int J Pharm Healthc Mark 2010 Apr 6;4(1):60-83. [doi: 10.1108/17506121011036033]
5. Naidu A. Factors affecting patient satisfaction and healthcare quality. Int J Health Care Qual Assur 2009;22(4):366-381. [doi: 10.1108/09526860910964834] [Medline: 19725209]
6. Singaraj A. Service quality and its dimensions. EPRA Int J Res Dev 2019;4(2) [FREE Full text]
7. Upadhyai R, Jain AK, Roy H, Pant V. A review of healthcare service quality dimensions and their measurement. J Health Manag 2019 Mar;21(1):102-127. [doi: 10.1177/0972063418822583]
8. Hong TS, Shang PP, Arumugam M, Yusuff RM. Use of simulation to solve outpatient clinic problems: a review of the literature. S Afr J Ind Eng 2013;24(3):27-42. [doi: 10.7166/24-3-551]
9. van Velthoven MH, Cordon C, Challagalla G. Digitization of healthcare organizations: the digital health landscape and information theory. Int J Med Inform 2019 Apr;124:49-57. [doi: 10.1016/j.ijmedinf.2019.01.007] [Medline: 30784426]
10. Kim SH, Song H. How digital transformation can improve hospitals' operational decisions. Harv Bus Rev. 2022 Jan 18. URL: https://hbr.org/2022/01/how-digital-transformation-can-improve-hospitals-operational-decisions [accessed 2026-07-08]
11. Vial G. Understanding digital transformation: a review and a research agenda. J Strateg Inf Syst 2019 Jun;28(2):118-144. [doi: 10.1016/j.jsis.2019.01.003]
12. Allen S. 2022 global health care outlook. : Deloitte; 2022 URL: https://www.deloitte.com/au/en/Industries/life-sciences-health-care/perspectives/global-health-care-sector-outlook.htmlI [accessed 2026-07-23]
13. McKeon J. What is holding healthcare back from digital transformation? Healthtech Security. 2022 Apr 27. URL: https://healthitsecurity.com/features/what-is-holding-healthcare-back-from-digital-transformation [accessed 2026-07-08]
14. Shukla M, Chang C, Hendricks J, Shudes C, Wurz C, Appleby C. Digital transformation: from a buzzword to an imperative for health systems. Deloitte Insights. 2021 Oct 26. URL: https://www2.deloitte.com/us/en/insights/industry/health-care/digital-transformation-in-healthcare.html [accessed 2026-07-07]
15. Sharif AN, Sharif AN, Alzahoof FH, et al. Integrating hospital management systems with electronic medical records: improving efficiency and quality of healthcare. J Int Crisis Risk Commun Research 2024;7(S3):159-164 [FREE Full text]
16. Guide to digital transformation in healthcare overview. Intellectsoft US. 2022. URL: https://www.intellectsoft.net/blog/digital-transformation-in-healthcare/ [accessed 2026-07-23]
17. Abbasi-Moghaddam MA, Zarei E, Bagherzadeh R, Dargahi H, Farrokhi P. Evaluation of service quality from patients' viewpoint. BMC Health Serv Res 2019 Mar 15;19(1):170. [doi: 10.1186/s12913-019-3998-0] [Medline: 30876453]
18. Berwick D. The evolving challenges facing healthcare leaders. Institute for Healthcare Improvement (IHI). 2022 Dec 1. URL: https://www.ihi.org/library/blog/evolving-challenges-facing-health-care-leaders [accessed 2026-07-07]
19. Allen-Duck A, Robinson JC, Stewart MW. Healthcare quality: a concept analysis. Nurs Forum 2017 Oct;52(4):377-386. [doi: 10.1111/nuf.12207] [Medline: 28407249]
20. Piper K. Results-driven healthcare: addressing the staggering cost of poor performance. Am Health Drug Benefits 2008 Jun;1(5):6-7. [Medline: 25126234]
21. Low quality healthcare is increasing the burden of illness and health costs globally. WHO. 2018 Jul 5. URL: https://www.who.int/news/item/05-07-2018-low-quality-healthcare-is-increasing-the-burden-of-illness-and-health-costs-globally [accessed 2026-07-10]
22. Leveraging telemedicine for enhanced healthcare accessibility: a narrative review of patient satisfaction, appointment adherence, and cost-effectiveness. J Health Rehabil Res 2025 Jan 31;5(1):1-6. [doi: 10.61919/jhrr.v5i1.1757]
23. Hatef E, Wilson RF, Zhang A, et al. Effectiveness of telehealth versus in-person care during the COVID-19 pandemic: a systematic review. NPJ Digit Med 2024 Jun 15;7(1):157. [doi: 10.1038/s41746-024-01152-2] [Medline: 38879682]
24. Monshi SS, Maajeeni JE, Aljawi LK, Jaha MJ, Alturkistani MI, Alqurashi YS. Patient experience with health technology in Saudi Arabia: a systematic review. Saudi J Health Syst Res 2025 Sep 9;5(3):155-164. [doi: 10.1159/000545897]

XSL•FO
RenderX

25. Naik N, Talyshinskii A, Rassweiler J, Hameed BMZ, Somani BK. Digital health innovations in urology: telemedicine, wearables, and mobile applications - a systematic review of literature. Curr Opin Urol 2024 Mar 1;34(2):116-127. [doi: 10.1097/MOU.0000000000001153] [Medline: 38038411]
26. Nunn J, Chang S. What are systematic reviews? Wiki J Med 2020;7(1):5. [doi: 10.15347/WJM/2020.005]
27. Siddaway AP, Wood AM, Hedges LV. How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. Annu Rev Psychol 2019 Jan 4;70:747-770. [doi: 10.1146/annurev-psych-010418-102803] [Medline: 30089228]
28. Parasuraman A, Zeithaml VA, Berry LL. SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. J Retail 1988;64(1):12-40 [FREE Full text]
29. Dagger TS, Sweeney JC, Johnson LW. A hierarchical model of health service quality: scale development and investigation of an integrated model. J Serv Res 2007 Nov;10(2):123-142. [doi: 10.1177/1094670507309594]
30. Lee H, Delene LM, Bunda MA, Kim C. Methods of measuring health-care service quality. J Bus Res 2000 Jun;48(3):233-246. [doi: 10.1016/S0148-2963(98)00089-7]
31. HCAHPS: patients' perspectives of care survey. Centers for Medicare & Medicaid Services. 2025. URL: https://www.cms.gov/medicare/quality/initiatives/hospital-quality-initiative/hcahps-patients-perspectives-care-survey [accessed 2026-07-08]
32. Butt MM, de Run EC. Private healthcare quality: applying a SERVQUAL model. Int J Health Care Qual Assur 2010;23(7):658-673. [doi: 10.1108/09526861011071580] [Medline: 21125961]
33. Pakdil F, Harwood TN. Patient satisfaction in a preoperative assessment clinic: an analysis using SERVQUAL dimensions. Tot Qual Manag Bus Excell 2005 Jan;16(1):15-30. [doi: 10.1080/1478336042000255622]
34. Toll K, Spark L, Neo B, et al. Consumer preferences, experiences, and attitudes towards telehealth: qualitative evidence from Australia. PLoS One 2022;17(8):e0273935. [doi: 10.1371/journal.pone.0273935] [Medline: 36044536]
35. Irfan SM, Ijaz A. Comparison of service quality between private and public hospitals: emperical evidance from Pakistan. J Qual Technol Manag 2011 Jun;7(1):1-22 [FREE Full text]
36. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019 Jun;6(2):94-98. [doi: 10.7861/futurehosp.6-2-94] [Medline: 31363513]
37. Bin KJ, Melo AAR, da Rocha J, et al. The impact of artificial intelligence on waiting time for medical care in an urgent care service for COVID-19: single-center prospective study. JMIR Form Res 2022 Feb 1;6(2):e29012. [doi: 10.2196/29012] [Medline: 35103611]
38. Street RL Jr, Liu L, Farber NJ, et al. Keystrokes, mouse clicks, and gazing at the computer: how physician interaction with the EHR affects patient participation. J Gen Intern Med 2018 Apr;33(4):423-428. [doi: 10.1007/s11606-017-4228-2] [Medline: 29188544]
39. Eschenroeder HC, Manzione LC, Adler-Milstein J, et al. Associations of physician burnout with organizational electronic health record support and after-hours charting. J Am Med Inform Assoc 2021 Apr 23;28(5):960-966. [doi: 10.1093/jamia/ocab053] [Medline: 33880534]
40. Bhatnagar R, Jain R. Robotic process automation in healthcare-a review. Int Robot Autom J 2019;5(1):12-14. [doi: 10.15406/iratj.2019.05.00164]
41. Duddridge C. How RPA is closing the gap between healthcare and technology. Healthcare Digital. 2020 Jun 1. URL: https://healthcare-digital.com/technology-and-ai/how-rpa-closing-gap-between-healthcare-and-technology [accessed 2026-07-08]
42. Usman MA, Philip NY, Politis C. 5G enabled mobile healthcare for ambulances. Presented at: 2019 IEEE Globecom Workshops (GC Wkshps); Dec 9-13, 2019; Waikoloa, HI, USA p. 1-6. [doi: 10.1109/GCWkshps45667.2019.9024584]
43. Fong B, Fong AC, Li CK. Internet of things in smart ambulance and emergency medicine. In: Hassan Q, editor. Internet of Things A to Z: Technologies and Applicationa: John Wiley & Sons, Inc; 2018:475-506. [doi: 10.1002/9781119456735]
44. Nolan T, Resar R, Haraden C, Griffin FA. Improving the reliability of health care. : Institute for Healthcare Improvement; 2004 URL: https://www.ihi.org/sites/default/files/ReliabilityWhitePaper2004revJune06.pdf [accessed 2026-07-23]
45. Martelli PF, Rivard PE, Roberts KH. Caveats for high reliability in healthcare. J Health Organ Manag 2018 Aug 20;32(5):674-690. [doi: 10.1108/JHOM-10-2017-0286] [Medline: 30175681]
46. Khan ZH, Siddique A, Lee CW. Robotics utilization for healthcare digitization in global COVID-19 management. Int J Environ Res Public Health 2020 May 28;17(11):1-21. [doi: 10.3390/ijerph17113819] [Medline: 32481547]
47. Sollof J. Telemedicine robots helping to ease pressure on hospitals during Covid-19. Digit Health. 2021. URL: https://www.digitalhealth.net/2021/12/telemedicine-robots-game-changing-healthcare-impact/ [accessed 2026-07-10]
48. Hornyak T. What America can learn from China's use of robots and telemedicine to combat the coronavirus. CNBC. 2020 Mar 18. URL: https://www.cnbc.com/2020/03/18/how-china-is-using-robots-and-telemedicine-to-combat-the-coronavirus.html [accessed 2026-07-10]
49. Ackerman E. iRobot and InTouch Health announce RP-VITA telemedicine robot. IEEE Spectr. 2012 Jul 24. URL: https://spectrum.ieee.org/irobot-and-intouch-health-announce-rpvita-telemedicine-robot [accessed 2026-07-10]
50. Gleiss A. The patient will see you now-towards an understanding of on-demand healthcare. : IEEE Xplore Presented at: 2020 IEEE 22nd Conference on Business Informatics (CBI); Jun 22-24, 2020; Antwerp, Belgium p. 154-161. [doi: 10.1109/CBI49978.2020.00024]

XSL•FO
RenderX

51. Neprash HT, Mulcahy JF, Cross DA, Gaugler JE, Golberstein E, Ganguli I. Association of primary care visit length with potentially inappropriate prescribing. JAMA Health Forum 2023 Mar 3;4(3):e230052. [doi: 10.1001/jamahealthforum.2023.0052] [Medline: 36897582]
52. Janett RS, Yeracaris PP. Electronic medical records in the American health system: challenges and lessons learned. Cien Saude Colet 2020 Mar;25(4):1293-1304. [doi: 10.1590/1413-81232020254.28922019] [Medline: 32267432]
53. Kang'a SG, Muthee VM, Liku N, Too D, Puttkammer N. People, process and technology: strategies for assuring sustainable implementation of EMRs at public-sector health facilities in Kenya. AMIA Annu Symp Proc 2016;2016:677-685. [Medline: 28269864]
54. Samadbeik M, Fatehi F, Braunstein M, et al. Education and Training on Electronic Medical Records (EMRs) for health care professionals and students: a scoping review. Int J Med Inform 2020 Oct;142:104238. [doi: 10.1016/j.ijmedinf.2020.104238] [Medline: 32828034]
55. Cifuentes M, Davis M, Fernald D, Gunn R, Dickinson P, Cohen DJ. Electronic health record challenges, workarounds, and solutions observed in practices integrating behavioral health and primary care. J Am Board Fam Med 2015;28 Suppl 1(Suppl 1):S63-S72. [doi: 10.3122/jabfm.2015.S1.150133] [Medline: 26359473]
56. Miller GJ. Agile problems, challenges, & failures. Presented at: PMI Global Congress Proceedings; Oct 27-29, 2013 URL: https://www.pmi.org/learning/library/agile-problems-challenges-failures-5869 [accessed 2026-07-08]
57. Dooling C, Meshkaty S. Collaborating outsourcing of the IT help desk. Presented at: SIGUCCS '09: Proceedings of the 37th annual ACM SIGUCCS Fell Conference: Communication and Collaboration; Oct 11-14, 2009; St Louis, MO p. 201-206. [doi: 10.1145/1629501.1629538]
58. Middleton B, Bloomrosen M, Dente MA, et al. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013 Jun;20(e1):e2-e8. [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]
59. Derecho KC, Cafino R, Aquino-Cafino SL, et al. Technology adoption of electronic medical records in developing economies: a systematic review on physicians' perspective. Digit Health 2024;10:20552076231224605. [doi: 10.1177/20552076231224605] [Medline: 38222081]
60. Adeniyi AO, Arowoogun JO, Chidi R, Okolo CA, Babawarun O. The impact of electronic health records on patient care and outcomes: a comprehensive review. World J Adv Res Rev 2024;21(2):1446-1455. [doi: 10.30574/wjarr.2024.21.2.0592]
61. Kohli R, Tan SSL. Electronic health records: how can IS researchers contribute to transforming healthcare? MIS Q 2016 Sep 1;40(3):553-573. [doi: 10.25300/MISQ/2016/40.3.02]
62. Oldenburg J, Chase D, Christensen KT, Tritle B. Engage!: Transforming Healthcare Through Digital Patient Engagement: CRC Press; 2020.
63. Hillestad R, Bigelow J, Bower A, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Aff (Millwood) 2005;24(5):1103-1117. [doi: 10.1377/hlthaff.24.5.1103] [Medline: 16162551]
64. Miller AR. Privacy of digital health information. In: Goldfarb A, Tucker CE, editors. The Economics of Privacy: University of Chicago Press; 2024:127-162 URL: https://academic.oup.com/chicago-scholarship-online/book/60280/chapter-abstract/520142493?redirectedFrom=fulltextI [accessed 2026-08-05]
65. Santos-Junior AG, Ferreira AM, Frota OP, et al. Effectiveness of surface cleaning and disinfection in a Brazilian healthcare facility. Open Nurs J 2018;12:36-44. [doi: 10.2174/1874434601812010036] [Medline: 29643951]
66. How to prevent cross-contamination. Infection Control Today. 2007 Aug 1. URL: https://www.infectioncontroltoday.com/view/how-prevent-cross-contamination [accessed 2026-07-08]
67. Kulińska J, Rypicz Ł, Zatońska K. The impact of effective communication on perceptions of patient safety-a prospective study in selected polish hospitals. Int J Environ Res Public Health 2022 Jul 27;19(15):1-11. [doi: 10.3390/ijerph19159174] [Medline: 35954529]
68. Patient-clinician communication in hospitals - communicating for safety at transitions of care: an information sheet for healthcare providers. Australian Commission on Safety and Quality in Healthcare. 2016. URL: https://www.safetyandquality.gov.au/resources/patient-clinician-communication-hospitals-communicating-safety-transitions-care-information-sheet-healthcare-providers [accessed 2026-07-29]
69. Alexander KE, Ogle T, Hoberg H, Linley L, Bradford N. Patient preferences for using technology in communication about symptoms post hospital discharge: BMC Health Services Research; 2021, Vol. 21:1-11. [doi: 10.21203/rs.3.rs-50582/v1]
70. Timmers T, Janssen L, Kool RB, Kremer JA. Educating patients by providing timely information using smartphone and tablet apps: systematic review. J Med Internet Res 2020 Apr 13;22(4):e17342. [doi: 10.2196/17342] [Medline: 32281936]
71. Owens LM, Wilda JJ, Grifka R, Westendorp J, Fletcher JJ. Effect of ambient voice technology, natural language processing, and artificial intelligence on the patient-physician relationship. Appl Clin Inform 2024 Aug;15(4):660-667. [doi: 10.1055/a-2337-4739] [Medline: 38834180]
72. Paul S, Srinivasan K. On-demand health care: a new paradigm for e-health. Presented at: 2009 11th International Conference on e-Health Networking, Applications and Services (Healthcom); Dec 16-18, 2009. [doi: 10.1109/HEALTH.2009.5406209]

XSL•FO
RenderX

73. Pogorzelska-Maziarz M, Rising KL, Gentsch AT, et al. Home healthcare patient, caregiver and provider perspectives on use of unscheduled acute care and the usability and acceptability of on-demand telehealth solutions. Geriatr Nurs 2021;42(5):1029-1034. [doi: 10.1016/j.gerinurse.2021.06.009] [Medline: 34256152]
74. Ferreira JC, Elvas LB, Correia R, Mascarenhas M. Empowering health professionals with digital skills to improve patient care and daily workflows. Healthcare (Basel) 2025 Feb 5;13(3):329. [doi: 10.3390/healthcare13030329] [Medline: 39942518]
75. Sterling R, LeRouge C. On-demand telemedicine as a disruptive health technology: qualitative study exploring emerging business models and strategies among early adopter organizations in the United States. J Med Internet Res 2019 Nov 15;21(11):e14304. [doi: 10.2196/14304] [Medline: 31730038]
76. Terry C, Cain J. The emerging issue of digital empathy. Am J Pharm Educ 2016 May 25;80(4):58. [doi: 10.5688/ajpe80458] [Medline: 27293225]
77. Drew S. Designing for wellbeing in residential healthcare. The Remi Network. 2020 Jul 10. URL: https://www.reminetwork.com/articles/residential-healthcare-wellbeing/ [accessed 2026-07-08]
78. Efthymiou E. Integrating digital and narrative medicine in modern healthcare: a systematic review. Med Educ Online 2025 Dec;30(1):2475979. [doi: 10.1080/10872981.2025.2475979] [Medline: 40327849]
79. Chen LM, Farwell WR, Jha AK. Primary care visit duration and quality: does good care take longer? Arch Intern Med 2009 Nov 9;169(20):1866-1872. [doi: 10.1001/archinternmed.2009.341] [Medline: 19901138]
80. Oermann MH, Templin T. Important attributes of quality health care: consumer perspectives. J Nurs Scholarsh 2000;32(2):167-172. [doi: 10.1111/j.1547-5069.2000.00167.x] [Medline: 10887716]
81. Murgitroyd E, Madurska M, Gonzalez J, Watson A. 3D digital anatomy modelling - practical or pretty? Surgeon 2015 Jun;13(3):177-180. [doi: 10.1016/j.surge.2014.10.007] [Medline: 25593105]
82. Andersen RM, Davidson PL. Improving access to care in America: individual and contextual indicators. In: Andersen RM, Rice TH, Kominiski GF, editors. Changing the US Health Care System: Key Issues in Health Services Policy and Management: Jossey-Bass; 2007:1-30.
83. Geissler J. Sustainable upscaling: the role of digitalization in providing health care and health insurance coverage in developing countries. In: Herberger TA, Dötsch JJ, editors. Digitalization, Digital Transformation and Sustainability in the Global Economy: Risks and Opportunities: Springer; 2021:53-69. [doi: 10.1007/978-3-030-77340-3_5]
84. Bruce A, Hardy I. No longer science fiction, AI and robotics are transforming healthcare. PWC. 2022. URL: https://www.pwc.com/gx/en/industries/healthcare/publications/ai-robotics-new-health/transforming-healthcare.html [accessed 2022-10-24]
85. Saeed SA, Masters RM. Disparities in health care and the digital divide. Curr Psychiatry Rep 2021 Jul 23;23(9):61. [doi: 10.1007/s11920-021-01274-4] [Medline: 34297202]
86. Combes JR, Arespacochaga E. Physician competencies for a 21st century health care system. J Grad Med Educ 2012 Sep;4(3):401-405. [doi: 10.4300/JGME-04-03-33] [Medline: 23997896]
87. Scalese RJ, Obeso VT, Issenberg SB. Simulation technology for skills training and competency assessment in medical education. J Gen Intern Med 2008 Jan;23 Suppl 1(Suppl 1):46-49. [doi: 10.1007/s11606-007-0283-4] [Medline: 18095044]
88. Epstein RM, Hundert EM. Defining and assessing professional competence. JAMA 2002 Jan 9;287(2):226-235. [doi: 10.1001/jama.287.2.226] [Medline: 11779266]
89. Ahmed A, Xi R, Hou M, Shah SA, Hameed S. Harnessing big data analytics for healthcare: a comprehensive review of frameworks, implications, applications, and impacts. IEEE Access 2023 Oct 10;11:112891-112928. [doi: 10.1109/ACCESS.2023.3323574]
90. Good MJD. American Medicine: The Quest for Competence: University of California Press; 2023. URL: https://www.jstor.org/stable/jj.5973110 [accessed 2026-07-23]
91. Kho J, Gillespie N, Martin-Khan M. A systematic scoping review of change management practices used for telemedicine service implementations. BMC Health Serv Res 2020 Sep 1;20(1):815. [doi: 10.1186/s12913-020-05657-w] [Medline: 32873295]

Abbreviations

COPQ: cost of poor quality
COREQ: Consolidated Criteria for Reporting Qualitative Research
EHR: electronic health record
EMR: electronic medical record
ERP: enterprise resource planning
HIPAA: Health Insurance Portability and Accountability Act
HNR: humanoid nurse robot
HR: human resources
IaaS: Infrastructure as a Service
IoT: Internet of Things
IRB: Institutional Review Committee
mHealth: mobile health

XSL•FO
RenderX

NLP: natural language processing
PaaS: Platform as a Service
PACS: picture archiving and communication systems
PCQ: patient care quality
PDA: personal digital assistant
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension
RPA: robotic process automation
SaaS: Software as a Service
SERVQUAL: Service Quality Model
SLR: systematic literature review
SME: subject matter expert

Edited by S Munce; submitted 04.Aug.2025; peer-reviewed by A Gaudino, D Qian, E Rojas; revised version received 31.May.2026; accepted 01.Jun.2026; published 19.Aug.2026.

Please cite as:
Al Khatib I, Awad M, Shamayleh A
Impact of Emerging Technologies on Outpatient Care Quality: Mixed Methods Study
JMIR Rehabil Assist Technol 2026;13:e81847
URL: https://rehab.jmir.org/2026/1/e81847
doi:10.2196/81847

© Inas Al Khatib, Mahmoud Awad, Abdulrahim Shamayleh. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Natural Language Prompts With AI Models for Low-Cost Assistive Software Design: Exploratory Comparative Evaluation

Francesc Antoni Bañuls-Lapuerta[1*], PhD; Vicent Marti-Miralles[1,2*], MSc; Rómulo Jacobo Gónzalez-García[1*], PhD; Gabriel Martínez-Rico[1*], Prof Dr

[1]Campus Capacitas, Valencia Catholic University Saint Vincent Martyr, Carrer de Joaquin Navarro, 37, Burjassot, Spain

[2]Doctorate School, Valencia Catholic University Saint Vincent Martyr, Valencia, Spain

[*]all authors contributed equally

Corresponding Author:
Vicent Marti-Miralles, MSc
Campus Capacitas, Valencia Catholic University Saint Vincent Martyr, Carrer de Joaquin Navarro, 37, Burjassot, Spain

Abstract

Background: This study investigates the capacity of 7 artificial intelligence (AI) models, 5 free and 2 paid, to generate functional software for designing low-cost, personalized assistive products.

Objective: The objective was to determine which models are most effective, accessible, and consistent in supporting nontechnical professionals in developing inclusive digital solutions and to assess the capabilities of commercially available and easy-to-access AI models to generate code from natural language interactions in the shape of a nontechnical assistive technology design process.

Methods: Each AI model was prompted using natural language, without any technical input, to create a Python program that converts an arcade gamepad into an adapted mouse-like controller. Sixteen progressively complex functions were requested through standardized prompts, delivered without additional feedback or correction. Model performance was evaluated based on the number of successfully implemented functions and the average number of prompts required.

Results: Paid models demonstrated markedly superior performance. Gemini Pro (Google) successfully implemented 14 of 16 requested functions with an average of 1.25 (SD 0.45) prompts, while ChatGPT Plus (GPT-5) achieved 11 functions with an average of 1.31 (SD 0.48) prompts. In contrast, free models produced between 0 and 4 functional outcomes, with DeepSeek and Gemini Free ranking the highest within their category. The enhanced outcomes of paid models were linked to improved contextual understanding, greater tolerance for natural language, and reduced conversational drift.

Conclusions: Paid AI models, particularly Gemini Pro and ChatGPT Plus, exhibit strong potential as tools for bridging the gap between health or education professionals and software development. They enable the creation of affordable, user-centered assistive technology without requiring advanced programming skills. Nevertheless, human oversight and foundational literacy in prompt design remain crucial to guarantee functionality, reliability, and ethical use.

(JMIR Rehabil Assist Technol 2026;13:e86786) doi:10.2196/86786

KEYWORDS

artificial intelligence; assistive products; digital accessibility; Gemini; ChatGPT; inclusive software

Introduction

Background

Artificial intelligence (AI) can be defined as the ability of nonhuman systems, machines, or software to simulate functions inherent to human intelligence, such as perceiving, reasoning, learning, planning, and anticipating future situations [1]. The recent advancements in the field have made clear that generative AI will open new scenarios in which human-machine collaboration will enable faster and more adaptive software solutions to meet the needs of people with disabilities [2]. In these contexts, the professional programmer assumes a supervisory role, while users without technical experience can create functional prototypes through intuitive AI-powered interfaces. It is thought that, by 2030, integrated assistants, such as the so-called HyperAssistant, will be capable of accompanying individuals throughout all stages of development, bridging the gap between software conception and implementation [3].

AI is transforming the way software is conceived and produced by allowing nontechnical users to participate actively in programming. Traditionally, this process was restricted to those mastering specific languages and methodologies; however, generative models in the shape of conversational AI systems have reduced the gap between natural language and coding, fostering digital inclusion and democratizing access to digital creation [4,5]. Generative assistants increase productivity and allow users with limited skills to build functional prototypes by

XSL•FO
RenderX

using natural language–oriented programming, where users express requirements in everyday language and AI translates them into executable code, bridging human thought and computational structures. This AI integration in development environments additionally enhances efficiency through real-time suggestions and code autocompletion, while warning against technological overreliance [6-8].

Design and Characteristics of Effective Prompts for Interaction With AI

The dialogue between a user and an AI system depends on how instructions or prompts, as they are known in the AI environment, are formulated. Prompts act as mediators between human intention and automatic language generation, shaping the trajectory of interaction, structuring intent, and bridging human goals with machine logic [9]. The first key element to achieving functional prompts is goal definition. Effective communication with conversational systems requires explicit purposes that limit and contain the semantic context. A prompt without a defined end, such as "tell me about economics," produces generic responses, whereas "summarize in 200 words the effects of remote work on business productivity" provides focus and precision [10].

The second principle is linguistic specificity, referring to the fact that grammatically well-structured prompts enhance coherence. Since AI operates through probabilistic predictions, lexical clarity, thematic delimitation, and details about format or audience reduce ambiguity [11-13]. The internal structure should ideally include 5 essential elements: topic, style, tone, format, and context. It should also contain action verbs (such as analyze, compare, or synthesize) to guide specific cognitive-like operations [14,15]. This level of precision is fundamental, as prompt design directly influences the reasoning processes activated by the model [16].

Using these key principles and unifying different prompt explanations from the literature, we could divide prompts into 6 specific and differentiated types with their own goals (Table 1).

Table . Types of prompts used.

Prompt type	Goal	Example
Zero-shot or decision	Execute a task or produce a direct response without prior examples [17].	Explain the concept of neuroplasticity in simple terms.
Few-shot or iterative	Guide generation using prior examples [12,18].	Example 1: write a scientific abstract. Example 2: create a new one about artificial intelligence.
Chain of thought	Promote logical processes or explanations step-by-step [19].	Explain, step by step, how you arrive at the conclusion about the benefits of automation.
Sequential	Develop an idea or process step by step [11].	First, define the concept, then provide an example, and finally conclude.
Argumentative	Request a reasoned, well-supported position [11].	Argue why automation can be beneficial.
Verification	Review one's own output and detect errors or biases [20].	Review your previous answer and indicate possible inaccuracies.

Communicative Limitations of Nonexpert Users in Interactions With AI

Several studies agree that interactions between users and AI systems present communicative limitations stemming from the spontaneous and imprecise use of natural language. Users without technical training often formulate ambiguous prompts, making it difficult for the system to interpret communicative intent [21]. Additionally, the absence of nonverbal cues and the limited adaptability of artificial discourse exacerbate these differences, revealing that the effectiveness of interaction depends on the precision and structure of the language used [22].

From a qualitative perspective, users tend to treat chatbots as human interlocutors, reproducing communicative patterns based on empathy and reciprocity [23]. Users tend to apply Grice conversational maxims when evaluating virtual assistants. These maxims include quantity (providing the right amount of information), quality (saying only what is true and verifiable), manner (expressing oneself clearly and orderly, avoiding ambiguity), and relevance (keeping responses focused on the topic at hand) [24]. Complementarily, anthropomorphism, or trying to treat AI like a human, seems to induce a sort of relational dissonance that clearly shows a contradiction as users understand AI as nonhuman yet interact with it as if it were a human [25,26].

Generative models do not interpret implicit intentions but rather textual correlations; therefore, effectiveness depends on the specificity of the prompt [12,14]. Many users are unaware of this mechanism and rely on trial and error [27,28]. Furthermore, the absence of mental models about AI processing creates a cognitive gap between intention and response, reinforced by interfaces that simulate naturalness while not processing information cognitively naturally [29,30].

In summary, most users lack the metalinguistic skills required to design effective prompts, and bridging this gap demands a form of communicative literacy oriented toward AI.

AI in Education: Barriers, Assistive Products, and Opportunities for Digital Inclusion

At this point, it is essential to analyze how AI can be integrated into educational and disability support contexts, expanding

XSL•FO
RenderX

opportunities for student participation and learning. This link between technology and inclusion requires examining the barriers that still limit its effective use in classrooms. The analysis of possibilities for students with disabilities within the school setting must address the obstacles that hinder their access and active participation, emphasizing the importance of identifying factors that prevent the functional use of educational materials [31]. Students with disabilities face multiple barriers—including negative attitudes, lack of appropriate resources, and insufficient technological accessibility—which restrict their access to inclusive and equitable education [32].

Since the COVID-19 pandemic, information and communication technology has played a central role both in education and society [33]. However, the lack of technological accessibility remains a major barrier, as current learning materials heavily depend on digital tools. This not only affects curricular access but also the acquisition of digital skills essential for adulthood. Studies and hiring companies, such as iHire, highlight that 76% of people seek employment online, illustrating the importance of digital literacy for future labor inclusion [34,35].

To address these barriers, assistive products (APs) are key tools for ensuring the educational and social participation of students with disabilities. According to ISO 9999:2022, APs are products that optimize functioning and reduce disability, serving as intermediaries between personal abilities and environmental demands. The previous version of this standard (ISO 9999:2016) specified 3 key functions: facilitating participation, supporting or substituting body functions, and preventing limitations or restrictions in participation. APs associated with information and communication technology include both hardware and software, and the ISO 9999:2023 standard emphasizes their mixed nature, explicitly stating “including software” and recognizing their educational value within category e130 of the International Classification of Functioning [36-38].

Society and public administrations are responsible for ensuring that children with disabilities have access to the APs necessary for their personal and social development [39-41]. However, the World Health Organization warns that high costs and technical complexity limit access, particularly in educational contexts [42]. Although technological advances and globalization have partially reduced prices, technical barriers remain in specific contexts [43].

While many education and health professionals are familiar with physical APs, they often lack expertise in software-based solutions, which usually require technical support from IT specialists. This dependency increases costs and delays implementation. In this context, AI represents a major opportunity: its ability to generate and modify code accessibly allows the creation or adaptation of low-cost technological products, narrowing the gap between technical design and real student needs [44]. Although AI-generated software may not reach the technical refinement of professionally supervised development, it offers an effective, economical, and flexible alternative to enhance accessibility and educational participation for students with disabilities, particularly in low-resource or time-constrained contexts.

Methods

Overview

The software development process was carried out iteratively with the assistance of 8 different AI models, 6 of which were free and 2 were paid (Table 2). These models were selected to generate results applicable to contexts where cost may be a limiting factor. The chatbots were asked to program a software solution to convert an arcade fighting-style controller into a computer control device similar to an adapted mouse. Communication with the AIs was conducted using predetermined prompts prior to testing.

The objective was not only to obtain a functional program but also to significantly reduce the feedback provided to each AI and subsequently analyze each model’s ability to solve the problem as efficiently as possible, assessing whether it could produce a complete and functional solution. This approach enabled the comparison of performance, accuracy, and coherence across different platforms within the same experimental context, providing objective data on their real effectiveness in supporting software development through natural language instructions.

Table . AIs used, models, and price.

Company	AI[a]	Model	Price
OpenAI	ChatGPT (Free)	GPT-4.1 mini	N/A[b]
OpenAI	ChatGPT (Pro)	GPT-5	€23 (US $26.36)/month
Google	Gemini (Free)	Gemini 2.5 Pro	N/A
Google	Gemini (Paid)	Gemini 2.5 Pro	€21.99 (US $25.20)/month
Anthropic	Claude (Free)	Sonnet 4.5	N/A
DeepSeek	DeepSeek	DeepSeek V3.2	N/A
Microsoft	Copilot	o3 mini	N/A

[a]AI: artificial intelligence.

[b]N/A: not applicable.

To ensure experimental validity and avoid prior learning or contamination from conversational memory, each evaluation was conducted in entirely new conversations and, in the case of free plans, through newly created accounts. This strategy is

XSL•FO
RenderX

grounded in empirical evidence showing that conversational history can induce bias or interference between tasks, thereby affecting model responses [45]. Similarly, literature on data contamination in language models indicates that any prior exposure to the evaluated information alters the reliability of results [46,47]. Even slight reformulations or lingering contextual information can modify the direction and quality of the generated output [48]. Therefore, initiating each session in a clean environment constitutes a methodologically necessary measure to control carryover bias, prevent contextual leakage, and ensure independence between trials, thereby preserving the external comparability of performance and accuracy across the evaluated models.

Ethical Considerations

The Research Ethics Committee of Valencia Catholic University Saint Vincent Martyr approved the study (UCV/2023-2024/010).

Test Protocol

The first step in initiating the coding process was to define the functions that could be beneficial and link each of them to the prompts that would be sent to the different AIs (Table 3). The definition of these prompts and the design of the conversation were based on two key concepts:

1. Iterative execution: the initial prompt is sent; the generated code is tested for functionality; and if it works successfully, the next prompt is sent with the intention of iteratively adding new features to the designed program.
2. Error handling: no feedback is provided, only the message "it doesn't work, fix it." This approach is based on the understanding that a person without technical qualifications might not comprehend the underlying problems in the code and would likely just ask the AI to fix the errors encountered. Such a person might be able to inform the AI that, for example, the joystick is not working but not describe more complex interaction failures or interpret Python error messages.

Table . Function desired and chosen prompt.

Function	Prompt
1. Move mouse with joystick	I have no knowledge of programming. I have a generic fighting style Game Pad that has a single joystick and several buttons. I need you to code a program in Python that allows the joystick on this device to generate the movements of the computer mouse.
2. Single right and left click	I need you to code left and click buttons like the mouse has.
3. Allow button reassignment	I need you to code in an app that allows the program to reassign keys in the gamepad.
4. Double left click from a single button	I need you to add an extra button that performs a double click in single click.
5. Toggle hold for left click	I need you to add an extra button that holds left click when pressed and releases when pressed again.
6. Scroll page up and down	I need you to add two extra buttons for scrolling up and down.
7. Increase and decrease volume	I need you to add extra buttons for volume up and down.
8. Assign a program to a button	I need you to add an extra button that opens Google Chrome when pressed.
9. Perform an automatic click after 3 seconds of mouse inactivity	Make it so that when the mouse automatically clicks once when standing still for 3 seconds.
10. Allow double-click with a configurable delay	I need you to add a delay that allows two clicks made within 3 seconds to register as a double click.
11. Type the user's name with a button	I need you to add a button that writes the name Pelayo.
12. Pause and unpause	I need you to add a button that toggles between pause and play.
13. Ignore movement during the first second of joystick input	I need you to add a delay that ignores the first seconds of movement in any direction and then starts moving.
14. Automatically send a help email	I need you to add an extra button that automatically sends an email saying "Help" to francesc.banuls@ucv.es.
15. Generate an app that allows modifying sensitivity and delay time for automatic clicking	I need you to code in an app that allows me to change sensitivity of the joystick and how long the mouse waits after being still to automatically click.
16. Create a program and installer	I need you to compile the program made into an app by coding installer for this app.

XSL•FO
RenderX

Based on this premise, and to reduce inconsistencies, it was considered that modern AIs increasingly incorporate more integrated self-feedback mechanisms, independent from human input, and that models should be capable of performing such functions [49]. All testing was conducted by the authors, in the context of an autonomous reference center for disability in Valencia, Spain, during 4 days, from October 10, 2025, to October 14, 2025. Subsequently, all testing was conducted with the approval of the university's ethics committee with code UCV/2023-2024/010.

Additionally, all testing was conducted using a Kubii USB Arcade Controller reference ODARCADE, connected via USB to a Windows 11 computer using Python 3.13 and having installed the libraries pygame and pyautogui.

The selection and design of the prompt constitute a key methodological element in this study, as its structure determines the comparative validity among the different AI models analyzed. Its uniformity follows the principle of instructional consistency, emphasizing that coherence in the formulation of commands is essential to ensure comparability between conversational systems [10]. Complementarily, the structure of the prompt conditions the type of cognitive processing activated by the model; therefore, keeping it constant allows the isolation of the intrinsic effect of each architecture [11]. Minimal wording differences can change both performance and the relative ranking of models, reinforcing the need to use a standardized prompt. Accordingly, the prompts were designed based on 3 theoretical principles [12,28,50].

The first is defining a nontechnical user role, representative of professionals without programming training. Contextualizing the interlocutor's identity allows for adjusting response complexity, and that assigning an explicit role guides the model's generation toward a semantically coherent and functional framework. This role also demands that all models are accessed from their web-found standard versions without adjusting models or using specific more technical or specially designed models for programming. This can be seen in the use of Copilot instead of specific instances coded into GitHub or other programming platforms [10,11].

The second is presenting a sequence of actions formulated in natural language and logical order. Procedural prompts activate step-by-step reasoning that promotes coherent and structured results, facilitating the translation of human descriptions into computational processes (as the de facto interaction mode with chatbots), although such prompts may carry a higher risk of drift [11].

The third is detailing functional requirements, such as cursor control or command execution. Precision in parameters or specific conditions reduces ambiguity and improves consistency of the output, making the prompt a cognitive interface between human language and automated execution [12,28].

These 3 concepts serve as guiding principles to mediate between the natural language of a nontechnical professional and the AIs. To more faithfully emulate the real process of software design by a person without advanced technical knowledge, 2 conditions were established to determine when to terminate the programming attempts.

The first one is that 3 consecutive nonfunctional generations either fail to execute, only partially implement, or lose communication with the gamepad. It is understood that if 3 consecutive iterations cannot fix inherited errors without feedback, it is unlikely that subsequent ones will do so.

Prompts are considered successful when they can correctly implement the full capabilities described in Table 1 for each prompt without compromising existing functionality. They are considered partially successful when they either maintain previous capabilities but only partially implement new functions or when they correctly implement new functions but lose previous ones. Finally, prompts are considered unsuccessful when they either fail to execute or lose communication with the gamepad. Partially successful prompts are marked as such to better understand the performance of different models but are functionally considered unsuccessful as they did not achieve the desired function. As such, 3 unsuccessful or partially successful prompts (with 2 tries each) stop interactions with a model.

The other way in which interactions with models are stopped is when reaching the daily prompt limit in free plans, to better reflect that the prompt limit is one of the most significant constraints of these models.

Goals

The experimental design was developed with the explicit understanding that current AI systems are inherently nondeterministic and that the replication of results will be inconsistent [51]. Consequently, the study objectives reflect this fact:

- General goal 1: compare the effectiveness of the most widely used AI solutions as software programmers for creating personalized assistive products using nontechnical natural language.
  - Specific goal 1: describe the relationship between cost and effectiveness of the models used.
  - Specific goal 2: analyze the ease of use, consistency, and accuracy of each model.

Results

Overview

The 8 AI solutions initially proposed were used to progressively design robust code solutions, incorporating functions of increasing complexity inspired by APs with similar objectives.

The results can be divided according to the cost of the alternatives, distinguishing between those that are free of charge and those requiring a subscription. The latter provided more consistent results and better adaptation to the project's functional requirements, demonstrating a greater understanding of the requested functions. However, it is important to note that these results should be interpreted within the framework of the specific characteristics of the study and that even under identical

XSL•FO
RenderX

conditions, different outcomes could be obtained due to the nondeterministic nature of these technologies.

Paid Alternatives

The subscription-based alternatives delivered consistent results, generating code that can generally be considered successful in meeting the required functions. This category includes ChatGPT Pro, based on the GPT-5 model, and Gemini Pro, which uses Google's advanced Gemini Pro-2.5 model.

Regarding Gemini Pro, the AI solution developed by Google proved to be the most robust option. It successfully implemented 14 out of the 16 required functions within a single application, without any rollback of previously implemented features during the design process and generated functional code with an average of 1.25 (SD 0.45) prompts. Gemini was even partially able to anticipate the implementation of future functions in earlier items: item 2 was successfully integrated within item 1; item 6 was partially included in item 3 by adding a scrolling mode, although without the 2 specific keys intended for that function; and item 15 was partially addressed in item 3 through a joystick sensitivity slider. The only functions not achieved were items 14 and 16. Item 14 could be considered partially achieved, as the AI managed to open the email app and compose a message addressed to the correct recipient but was unable to send it. All partially achieved items are considered not achieved when counting successful items implemented. Regarding item 16, Gemini could not autonomously generate an installable app, and the alternative solutions proposed in Python were also nonfunctional.

One thing to note is that both items (14 and 16) did not work across the board for security reasons, as AIs are not capable of generating programs that have such deep access to your operative system. Seeing how the AIs approached a task they knew beforehand was impossible to fully be able to perform is interesting and shows the AIs' problem-solving skill navigating natural language generated prompts, which may be difficult or impossible to fulfill.

Complementarily, OpenAI's paid version (ChatGPT Pro using GPT-5) also demonstrated solid performance, successfully implementing 11 out of the 16 required functions and producing functional code with an average of 1.31 (SD 0.48) prompts. GPT-5 showed the same limitation as Gemini Pro with respect to sending the email in item 14 and was likewise unable to compile an executable file or provide stable Python-based compilation solutions. Additionally, OpenAI's AI had difficulties implementing joystick controls in item 1, which caused the device to move the cursor only downward and to the right. This error persisted in subsequent iterations, affecting later results such as items 13 and 15, which were only partially achieved due to the lack of precise joystick control. ChatGPT also displayed inconsistency, producing multiple files for different functions instead of integrating them into a single app and rolling back previously implemented features without justification or explicit request.

Free Alternatives

The free alternatives, in contrast to the subscription-based ones, produced notably less consistent results. In some cases, this was due to the limitations of the models themselves, which were unable to meet the functional requirements set for the task. In other cases, restrictions on the number of interactions significantly limited development, constituting a key distinction. It cannot be asserted that greater time investment would have yielded results equivalent to the paid versions, since the possibility of submitting additional prompts was not available. Therefore, in this study, the maximum number of interactions allowed by each model in its free version was established as a methodological limitation. Conversely, the models that failed to generate functional solutions can, for the time being, be considered inoperative for the intended function.

The free AI with the best performance was DeepSeek. This model generated the first 3 items in the initial prompt and even implemented item 2 without it being explicitly requested in item 1. However, the AI was unable to technically complete items 4, 5, and 6, implementing item 5 through a bind function that only allowed the assignment of a single key and caused the program to crash when attempting to add more combinations. This model generated functional items with an average of 1.88 (SD 0.34) prompts, achieving 4 out of the 16 required functions.

The second-best performance can be attributed to GPT Free. The free version of GPT successfully produced a functional prototype on the first attempt for the first 3 prompts but maintained the inconsistency observed in its paid counterpart. From the fourth prompt onward, it began producing nonfunctional code, leading to discontinuation after the seventh attempt. On average, it generated functional prototypes with 1.81 prompts and achieved 3 out of the 16 intended functions.

After GPT, in terms of performance, is Gemini Pro Free. Although it theoretically uses the same model as the paid version, its results differed appreciably. The free version produced correct and functional outputs up to the permitted prompt limit, forcing the process to stop prematurely. It successfully implemented the first 2 items and partially the third, generating a program that ran natively in the browser. Subsequently, time-based restrictions prevented continuation. On average, it produced functional prototypes with 1.88 prompts and achieved 2 out of the 16 complete functions.

The second-to-last worst-performing model is Claude. The Sonnet 4.5 model managed to implement only the first of the planned functions. After failing to execute functions 3, 4, and 5, its use was discontinued. Claude generated functional prototypes with an average of 1.94 (SD 0.25) prompts, achieving 1 out of the 16 proposed functions.

Finally, the individually worst-performing model is Copilot. Using its Deep Think model, Microsoft's AI failed to generate any executable Python files within the first 3 prompts, leading to the termination of testing after the fourth attempt. It did not produce any functional prototypes, with an average of 2 prompts and a total of 0 out of 16 functions achieved.

All models and a color-coded list of the items successfully implemented by each, along with the number of prompts required, are presented in Table 4. Additionally, prompts achieved and average prompts per model are presented in Figure 1.

XSL•FO
RenderX

Table . Functions achieved and prompts needed.

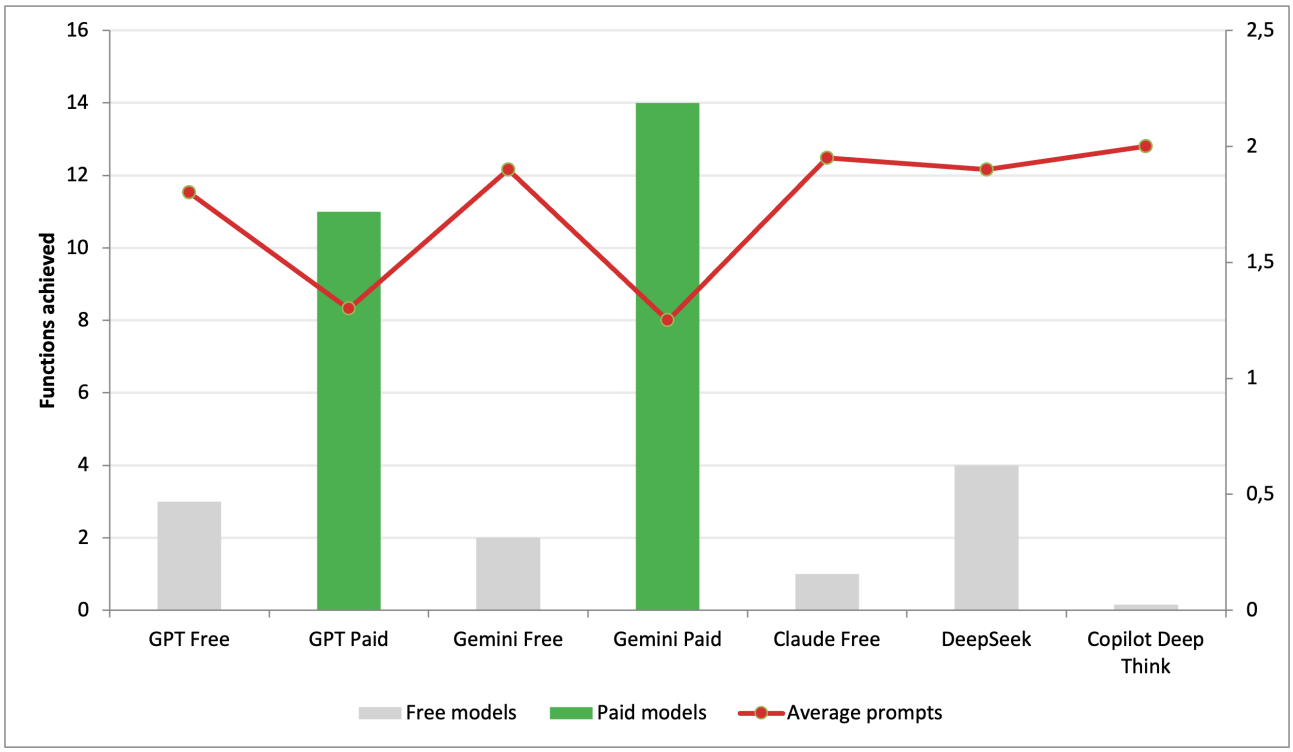
Required function	GPT Free	GPT Paid	Gemini Free	Gemini Paid	Claude Free	DeepSeek	Copilot Deep Think
1	✓[a]	×[b]	✓	✓	✓	✓	×
2	✓	✓	×	✓	×	✓[c]	×
3	✓	✓	✓[c]	×	×	✓	×
4	×	✓	×	✓	×	×	×
5	×	✓	×	✓	×	✓	×
6	×	✓	×	✓[c]	×	×	×
7	×	✓	×	✓	×	×	×
8	×	✓	×	✓	×	×	×
9	×	✓	×	✓	×	×	×
10	×	✓	×	×	×	×	×
11	×	✓	×	✓	×	×	×
12	×	✓	×	✓	×	×	×
13	×	×	×	✓	×	×	×
14	×	×	×	×	×	×	×
15	×	×	×	✓[c]	×	×	×
16	×	×	×	×	×	×	×

[a]Items achieved.

[b]Items not achieved or partially achieved. Partially achieved items are counted as unsuccessful.

[c]Functions already implemented or partially already implemented in previous interactions without direct prompts for their inclusion.

Figure 1. Prompts achieved and average prompts per model.



XSL•FO
RenderX

Discussion

Principal Findings

The results obtained show that the paid AI alternatives produced robust code, achieving 14 out of 16 functions with Gemini Pro and 11 out of 16 functions with ChatGPT Plus, respectively. The free-to-use alternatives, on the other hand, managed to integrate a variable, though consistently lower, number of functions ranging from none to 4 successfully implemented. The discussion of the results obtained can therefore be divided into 3 key areas (Figure 1).

Superiority of Paid AIs

The use of paid AI systems offers several advantages, among which the most significant are access to more advanced models and the elimination or substantial expansion of the number of prompts allowed within a given period. These advantages naturally make paid AI solutions more reliable for performing complex tasks. The main improvements can be summarized in 3 key points.

Paid versions typically operate on newer, more advanced architectures with higher parameter capacity and improved optimization strategies and training based on larger and higher-quality datasets. In literature, larger and better-trained models consistently achieve superior results in both code generation and natural language understanding tasks. For instance, GPT-4 performs substantially better in programming tasks than earlier versions [52].

Due to these technical enhancements, paid models generally demonstrate a stronger grasp of natural language and communicative context, which increases the likelihood of successful interactions in this specific case scenario. They are also better at maintaining coherence across sequential instructions, retaining previously defined functions, and being better at avoiding contradictions and hallucinations. This allows them to integrate multiple features coherently and progressively with fewer rollbacks or loss of content [53-55].

The removal or significant reduction of restrictions on the number of prompts permitted per time interval is particularly relevant when considered alongside the highlighted advantages. Older free models often have weaker memory capabilities and greater difficulty minimizing inconsistencies, rollbacks, and errors across different interaction sessions [56]. Increasing the interaction limits from, for example, 5 per session in the free version of Gemini to 100, or from 10 every 5 hours in GPT-4.1 mini to 160 every 3 hours in GPT-5, not only eases time constraints but also reduces errors and can substantially improve code quality [57,58].

In summary, the differences in model capacity, contextual understanding, memory, and reasoning largely explain why paid AIs may achieve significantly more functional results than their free counterparts. This outcome is expected, as paying for a service naturally entails the expectation of better performance. However, another relevant point of discussion lies in comparing paid AIs with one another and questioning why Google's AI seems to have outperformed GPT in this specific scenario, despite the latter's longer development history and its recent update to version 5.

Differences Between Gemini Pro and GPT-5

Although it is not possible to access the internal technical details of Gemini Pro, or of any major AI model, the existing literature published by the companies behind these models appears to support the results obtained in this specific case scenario, pointing to 3 key aspects that may explain why Gemini stood out as the most robust alternative for generating code from natural language given the restrictions and tasks asked.

On the one hand, Gemini may have better alignment between intention and execution. According to the Google Gemini team, the training of Gemini Pro placed greater emphasis on the correspondence between high-level instructions and structured code generation, giving it an advantage in translating complex requirements into functional code. The model achieved scores above 74% on specialized programming benchmarks [59].

This may also lead to much greater stability and coherence in long sessions. In tasks involving multiple functions (such as the 16 required in this study), the ability to maintain coherence across many iterations is crucial. If Gemini Pro manages internal session states more effectively, it becomes less prone to "forget" previously implemented functions or to inconsistently rename elements [60].

On the other hand, and especially relevant to the study, prompt optimization may be more tolerant of natural language. A distinct training approach regarding the model's handling of natural language results in greater adaptability to intricate instructions, improved contextual understanding, and enhanced problem-solving abilities. This makes Gemini Pro more resilient to inaccuracies or ambiguities in prompts written in nontechnical natural language, whereas ChatGPT Plus may tend to require more carefully formulated instructions to produce reliable outcomes [59,60].

Limitations and Strengths of Free AIs

As previously mentioned, free AI systems may present significant limitations related to the use of simpler and more restricted models, limited numbers of prompts, and, in some cases, the restriction to process multimodal information. These constraints make free models more challenging to use or less adaptable to users' needs. The standardized subscription fee of US $20, as implemented by companies such as OpenAI and Google, remains possibly too expensive for low-resource contexts in many parts of the world.

In recent months, both companies have begun introducing country-based pricing, adjusting subscription costs to local purchasing power. OpenAI (2025) implemented multicurrency billing to reduce conversion costs, and in India, the Plus plan is now offered in rupees, alongside a lower-cost alternative named ChatGPT Go [61]. This strategy echoes *The Economist*'s Big Mac Index, which illustrates how multinational companies adjust prices according to economic context. AI can enhance accessibility in education, but its impact is constrained by the economic barriers preventing access to paid tools [62]. However, if AI design fails to account for income inequalities within a

XSL•FO
RenderX

single country, despite global cost-adjustment strategies, it may reinforce digital exclusion. Thus, exploring the use and functionality of free versions and comparing them with paid alternatives as a potential source of social inequality becomes a social necessity to ensure equitable access to knowledge and accessibility [63,64].

In our test, among the free options, Gemini Free managed to implement 2 functions with relative stability during the first iteration, approaching a third before reaching the interaction limit. Considering the performance of its paid counterpart under fewer restrictions, it is likely that free versions could offer solid, no-cost solutions, albeit requiring a significant time investment, provided the model seems to be able to maintain internal coherence across sessions [59].

In contrast, GPT Free and DeepSeek achieved a greater number of items but lost track of the ongoing programming process, producing unstable code that would require the user to identify errors and explain them to the chatbot. While this may be possible, it is difficult to fully assess how much feedback a nontechnical user may be able to provide, and as such, it is difficult to give certain conclusions about these models. Professionals with intermediate technical knowledge could use these 2 models effectively to develop software solutions, but nontechnical users may be unable to [65,66].

Finally, Claude and Copilot, in their free versions, showed no real ability to generate operational code. In the case of Claude, this is understandable, since its website explicitly states that only the Pro version provides such capabilities [67]. The second model, Copilot, simply appears unable to generate functional, plug-and-play code. This is also due to the several factors, such as copilot being used in the web version (not optimized to do such tasks) without any specific model or implementation chosen [68]. In both case scenarios, using Claude Pro with programming capabilities and a version of Copilot optimized for programming may generate more competent results.

These findings may point to the fact that not all free versions are equivalent: some models are sufficiently robust for basic tasks, while others are less capable with the task chosen, the prompt limit, the version used, and many other factors contributing to their success.

One key limitation to have in mind is that only 1 run has been carried out with each model to mimic a real case scenario of a nontechnical user using these models. Given the nondeterministic nature of these models, more different runs around the same task may produce different results.

Observations and Practical Implications

The observations, limitations, and implications for practice derived from this study can be summarized in 10 key points that integrate the main insights gained throughout the research. Four main points stand out as the implications of the results obtained.

Some AI models, such as ChatGPT (both 4.1 mini and 5), may produce inconsistent results, hallucinating or forgetting its progress. This can be seen by these models including renaming functions without apparent reason, changing project names, generating new programs instead of adding features to the existing one, or rolling back previously correct functions. This suggests that OpenAI models can generate the intended products but require a technical skill level beyond that of nonspecialist professionals, significantly complicating the process of obtaining functional prototypes for nontechnical users in this specific case scenario [66,69].

This behavior is sometimes categorized as AI “drift” and means that, during extended sessions, models may change behavior, forget previously defined functions, or misinterpret prompts that are conceptually equivalent. Maintaining a clear “narrative thread” and consistent prompt formulation is therefore essential. Natural language alone may not suffice, and the risk of accumulating errors that render the code nonfunctional remains high [70,71].

In this specific test, Gemini showed greater consistency interpreting natural-language instructions and required fewer technical reformulations or corrections. It also showed a lower risk of drift. In contexts where users represent the target population, these traits may reduce the barrier to entry and increase the likelihood of producing functional prototypes. The Pro version is notably more efficient and reliable, but the free version can achieve similar outcomes with sufficient time investment if the time between tries does not induce AI drift [72].

While AIs can generate code that is operational and useful, it must be remembered that its efficiency (execution time, memory usage, and design simplicity) is generally suboptimal. For this reason, professional supervision, or ideally full implementation, by a qualified software engineer remains preferable. Studies, such as EffiBench, have shown that AI-generated code tends to be, on average, less efficient than optimized human solutions [73].

The results obtained in this domain (low-cost AP adaptations with 16 functions with no feedback and in a single run) are constrained to the models and prompts used and to the absence of user feedback, maybe reflecting the condition of a nontechnical user. Other models or differently designed prompts might yield different outcomes. The practical implications are as follows:

1. For real-world AP adaptations: Collaboration with software professionals should always be prioritized when budgets allow. If it is not feasible, both paid models tested may be the most effective choice, as they perform better in programming tasks than other paid alternatives. When access to paid tools is possible, they remain the most robust option. However, free models such as Gemini Free, GPT Free, or even DeepSeek can provide functional results in exchange for greater time investment.
2. For low-budget or educational contexts: Free versions, particularly Gemini Free or GPT Free, can serve as valuable tools for initial prototyping, though additional iterations, revisions, and acceptance of functional limitations will be necessary [74].
3. Effective prompt design and problem segmentation: Breaking the task into smaller subfunctions and crafting

XSL•FO
RenderX

clear, stepwise prompts improves the success rate, especially in free models.

4. Continuous monitoring and automated testing: Implementing automated unit tests (eg, by prompting the AI to generate its own test cases) may help detect errors or inconsistencies in the generated code.
5. Caution with long sessions or evolving prompts: Avoid overly long or evolving prompts within a single session and restart the interaction when the model shows signs of drift to preserve consistency.

Conclusions

In conclusion, the comparative analysis of different AI models applied to the generation of functional software for designing personalized APs using natural language reveals clear differences in performance, usability, and consistency. Paid models, such as Gemini Pro and ChatGPT Plus, demonstrated greater efficiency and reliability, achieving 14 and 11 out of 16 required functions, respectively, while free alternatives ranged between 0 and 4. These results confirm that the technical advancements of newer and more sophisticated models, such as expanded context windows, improved multimodal communication, enhanced natural language comprehension, reduced drift, and fewer prompt limitations, directly influence the functional quality of the generated code.

Regarding the relationship between cost and effectiveness, Gemini Pro may stand out as the most balanced option, in this specific case scenario offering higher precision and coherence at a cost equivalent to other paid alternatives. For low-resource contexts, DeepSeek emerges as a viable free alternative that, despite its limitations, can produce acceptable results given sufficient time and technical supervision, while the free versions of Gemini can deliver solid outputs with adequate time investment.

In terms of ease of use and consistency, Gemini demonstrated greater tolerance for nontechnical natural language and more stable performance in long sessions, making it particularly suitable for users without programming knowledge, an essential aspect when the goal is to promote inclusive and low-cost design of APs.

Overall, the findings suggest that AI can serve as an effective bridge between health care or educational professionals and programming, enabling the creation of personalized assistive solutions without the need for advanced software development expertise. Nevertheless, the results also highlight the critical need for collaboration with qualified IT professionals, given the importance of human technical oversight, careful prompt design, and continuous testing to prevent drift and ensure the generation of stable software alternatives.

Future research could focus on exploring how AI models handle problem-solving from visual inputs (eg, images sent to chatbots), their capacity for self-generated prompt design, and the evaluation of emerging or alternative models.

Funding

This study was carried out within the framework of the research project PID2022-142309OB-100 funded by the Ministry of Science and Innovation, State Research Agency and in collaboration with the Capacitas Campus of the Universidad Católica de Valéncia San Vicente Mártir (UCV).

Conflicts of Interest

None declared.

References

1. Xu Y, Liu X, Cao X, et al. Artificial intelligence: a powerful paradigm for scientific research. Innovation (Camb) 2021 Oct 28;2(4):100179. [doi: 10.1016/j.xinn.2021.100179] [Medline: 34877560]
2. Sauvola J, Tarkoma S, Klemettinen M, Riekki J, Doermann D. Future of software development with generative AI. Autom Softw Eng 2024 May;31(1):26. [doi: 10.1007/s10515-024-00426-z]
3. Qiu K, Puccinelli N, Ciniselli M, Di Grazia L. From today's code to tomorrow's symphony: the AI transformation of developer's routine by 2030. ACM Trans Softw Eng Methodol 2025 Jun 30;34(5):1-17. [doi: 10.1145/3709353]
4. Alenezi M, Akour M. AI-driven innovations in software engineering: a review of current practices and future directions. Appl Sci (Basel) 2025;15(3):1344. [doi: 10.3390/app15031344]
5. Yang EW, Waldrup B, Velazquez-Villarreal E. Conversational AI agent for precision oncology: AI-HOPE-WNT integrates clinical and genomic data to investigate WNT pathway dysregulation in colorectal cancer. Front Artif Intell 2025;8:1624797. [doi: 10.3389/frai.2025.1624797] [Medline: 40860720]
6. Yu L. Paradigm shift on coding productivity using GenAI. Presented at: Proceedings of the 29th International Conference on Evaluation and Assessment in Software Engineering (EASE '25); Jun 17-20, 2025. [doi: 10.1145/3756681.3757081]
7. Sergeyuk A, Titov S, Izadi M. In-IDE human-AI experience in the era of large language models; a literature review. Presented at: IDE '24: Proceedings of the 1st ACM/IEEE Workshop on Integrated Development Environments; Apr 20, 2024. [doi: 10.1145/3643796.3648463]
8. Beheshti A. Natural language-oriented programming (NLOP): towards democratizing software creation. Presented at: 2024 IEEE International Conference on Software Services Engineering (SSE); Jul 7-13, 2024. [doi: 10.1109/SSE62657.2024.00047]

XSL•FO
RenderX

9. Dalsgaard P. Thinking through prompting: cognitive mediation in human–AI interaction. 2025 Presented at: Proceedings of the European Conference on Cognitive Ergonomics (ECCE '25); Oct 7-10, 2025. [doi: 10.1145/3746175.3747192]
10. McTear M, Callejas Z, Griol D. The Conversational Interface Talking to Smart Devices: Springer International Publishing; 2016. [doi: 10.1007/978-3-319-32967-3]
11. Morales-Chan M. Explorando el potencial de chat GPT: una clasificación de prompts efectivos para la enseñanza [Report in Spanish]. : Universidad Galileo; 2023 URL: https://biblioteca.galileo.edu/tesario/handle/123456789/1348 [accessed 2026-03-07]
12. Brown TB, Mann B, Ryder N, et al. Language models are few-shot learners. Presented at: Proceedings of the 34rd International Conference on Neural Information Processing Systems (NeurIPS 2020); Dec 6-12, 2020 URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf [accessed 2026-03-07]
13. Serban I, Sordoni A, Lowe R, et al. A hierarchical latent variable encoder-decoder model for generating dialogues. 2017 Presented at: Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI 2017); Feb 4-9, 2017. [doi: 10.1609/aaai.v31i1.10983]
14. Dathathri S, Madotto A, Lan Z, Fung P, Neubig G. Plug and play language models: a simple approach to controlled text generation. Presented at: Findings of the Association for Computational Linguistics: EMNLP 2021; Nov 7-11, 2021. [doi: 10.18653/v1/2021.findings-emnlp.334]
15. Korzynski P, Mazurek G, Krzypkowska P, Kurasinski A. Artificial intelligence prompt engineering as a new digital competence: analysis of generative AI technologies such as ChatGPT. Entrepren Bus Econ Rev 2023;11(3):25-37. [doi: 10.15678/EBER.2023.110302]
16. Kumar H, Musabirov I, Shi J, et al. Exploring the design of prompts for applying GPT-3 based chatbots: a mental wellbeing case study on Mechanical Turk. arXiv. Preprint posted online on Sep 22, 2022. [doi: 10.48550/arXiv.2209.11344]
17. Zhou K, Zhang J, Liu X, Sun M. A systematic survey of prompt engineering in large language models: techniques and applications. arXiv. Preprint posted online on Feb 5, 2024. [doi: 10.48550/arXiv.2402.07927]
18. Fagbohun O, Harrison RM, Dereventsov A. An empirical categorization of prompting techniques for large language models: a practitioner's guide. J Artif Intell Mach Learn Data Sci ;1(4):1-11. [doi: 10.51219/JAIMLD/Oluwole-Fagbohun/15]
19. Wei J, Wang X, Schuurmans D, et al. Chain-of-thought prompting elicits reasoning in large language models. Presented at: Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS 2022); Nov 28 to Dec 9, 2022 URL: https://openreview.net/pdf?id=_VjQlMeSB_J [accessed 2026-03-07]
20. Madaan A, Tandom N, Gupta P. Self-refine: iterative refinement with self-feedback. Presented at: Advances in Neural Information Processing Systems 36; Dec 10-16, 2023 URL: https://proceedings.neurips.cc/paper_files/paper/2023/hash/91edff07232fb1b55a505a9e9f6c0ff3-Abstract-Conference.html [accessed 2026-03-18]
21. Anam RK. Prompt engineering and the effectiveness of large language models in enhancing human productivity. arXiv. Preprint posted online on May 10, 2025. [doi: 10.48550/arXiv.2507.18638]
22. Xu Y, Thomas T, Yu CL, Pan EZ. What makes children perceive or not perceive minds in generative AI? Comput Hum Behav Artif Hum 2025 May;4:100135. [doi: 10.1016/j.chbah.2025.100135]
23. Brandtzaeg PB, Skjuve M, Følstad A. My AI friend: how users of a social chatbot understand their human–AI friendship. Hum Commun Res 2022 Jun 29;48(3):404-429. [doi: 10.1093/hcr/hqac008]
24. Panfili L, Duman S, Nave A, Ridgeway KP, Eversole N, Sarikaya R. Human-AI interactions through a Gricean lens. Proc Ling Soc Amer ;6(1):288. [doi: 10.3765/plsa.v6i1.4971]
25. Zheng Q, Tang Y, Liu Y, Liu W, Huang Y. UX research on conversational human-AI interaction: a literature review of the ACM digital library. 2022 Apr 29 Presented at: CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; Apr 29 to May 5, 2022. [doi: 10.1145/3491102.3501855]
26. Gülay E, Picco E, Glerean E, Coupette C. Relational dissonance in human–AI interactions: the case of knowledge work. arXiv. Preprint posted online on Sep 19, 2025. [doi: 10.48550/arXiv.2509.15836]
27. Kim Y, Lee J, Kim S, Park J, Kim J. Understanding users' dissatisfaction with ChatGPT responses: types, resolving tactics, and the effect of knowledge level. Presented at: Proceedings of the 29th International Conference on Intelligent User Interfaces (IUI 2024); Mar 18-21, 2024. [doi: 10.1145/3640543.3645148]
28. Hernández Caralt M, Sekulić I, Carevic F, et al. Stupid robot, I want to speak to a human!" User frustration detection in task-oriented dialog systems. 2025 Presented at: Proceedings of the 31st International Conference on Computational Linguistics: Industry Track; Jan 19-24, 2025 URL: https://aclanthology.org/2025.coling-industry.23/ [accessed 2026-03-07]
29. Eiband M, Schneider H, Bilandzic M, Fazekas-Con J, Haug M, Hussmann H. Bringing transparency design into practice. Presented at: Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI 2018); Mar 7-11, 2018 URL: https://dl.acm.org/doi/proceedings/10.1145/3172944 [doi: 10.1145/3172944.3172961]
30. Degachi C, Freire SK, Niforatos E, Kortuem G. Understanding mental models of generative conversational search and the effect of interface transparency. arXiv. Preprint posted online on Jun 4, 2025. [doi: 10.48550/arXiv.2506.03807]
31. Jardinez MJ, Natividad LR. The advantages and challenges of inclusive education: striving for equity in the classroom. Shanlax Int J Educ 2024;12(2):57-65. [doi: 10.34293/education.v12i2.7182]
32. Bani Odeh K, Lach LM. Barriers to, and facilitators of, education for children with disabilities worldwide: a descriptive review. Front Public Health 2023;11:1294849. [doi: 10.3389/fpubh.2023.1294849] [Medline: 38292375]

XSL•FO
RenderX

33. Pérez Echeverría MP, Cabellos B, Pozo JI. The use of ICT in classrooms: the effect of the pandemic. Educ Inf Technol 2025 Jul;30(10):14069-14093. [doi: 10.1007/s10639-024-13124-w]
34. Alanazi AS, Benlaria H. Understanding the landscape: assistive technology and work challenges for people with disabilities in Saudi Arabia. Humanit Soc Sci Commun 2024;11(1):1608. [doi: 10.1057/s41599-024-04023-z]
35. The state of online recruiting 2024. iHire. URL: https://www.ihire.com/resourcecenter/employer/pages/the-state-of-online-recruiting-2024 [accessed 2025-10-22]
36. ISO 9999:2022—assistive products for persons with disability: classification and terminology. : International Organization for Standardization (ISO); 2022 URL: https://cdn.standards.iteh.ai/samples/72464/3f3608ed0bff4545bd53c02373f8cdbb/ISO-9999-2022.pdf [accessed 2026-03-07]
37. ISO 9999:2016—assistive products for persons with disability: classification and terminology. : International Organization for Standardization (ISO); 2016 URL: https://cdn.standards.iteh.ai/samples/60547/4694b28419324e45810538491357903c/ISO-9999-2016.pdf [accessed 2026-03-07]
38. International Classification of Functioning, Disability and Health (ICF). : World Health Organization (WHO); 2001 URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2026-03-07]
39. Tomás V. La discapacidad como elemento de discriminación positiva. In: Los Derechos de Los Niños, Responsabilidad de Todos [Book in Spanish]: Universidad de Murcia; 2007:213-218 URL: https://dialnet.unirioja.es/servlet/articulo?codigo=2306106 [accessed 2026-03-07]
40. Vila-Merino ES, Rascón-Gómez T, Calderón-Almendros I. Discapacidad, estigma y sufrimiento en las escuelas. Narrativas emergentes por el derecho a la educación inclusiva [Article in Spanish]. Educ XX1 2024;27(1):353-371. [doi: 10.5944/educxx1.36753]
41. de Beco G. The right to inclusive education according to Article 24 of the UN Convention on the Rights of Persons with Disabilities: background, requirements and (remaining) questions. Neth Q Hum Rights 2014 Sep;32(3):263-287. [doi: 10.1177/016934411403200304]
42. Assistive technology. World Health Organization (WHO). 2024. URL: https://www.who.int/news-room/fact-sheets/detail/assistive-technology [accessed 2025-10-22]
43. Fteiha M, Al-Rashaida M, Elsori D, Khalil A, Al Bustami G. Obstacles for using assistive technology in centres of special needs in the UAE. Disabil Rehabil Assist Technol 2024 Nov;19(8):2934-2944. [doi: 10.1080/17483107.2024.2323698] [Medline: 38436086]
44. Fernández-Batanero JM, Montenegro-Rueda M, Fernández-Cerero J, García-Martínez I. Assistive technology for the inclusion of students with disabilities: a systematic review. Education Tech Research Dev 2022;70(5):1911-1930. [doi: 10.1007/s11423-022-10127-7]
45. Gupta A, Sheth I, Raina V, Gales M, Fritz M. LLM task interference: an initial study on the impact of task-switch in conversational history. Presented at: Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing; Nov 12-16, 2024. [doi: 10.18653/v1/2024.emnlp-main.811]
46. Golchin S, Surdeanu M. Time travel in LLMs: tracing data contamination in large language models. arXiv. Preprint posted online on Aug 16, 2023. [doi: 10.48550/arXiv.2308.08493]
47. Cheng Y, Chang Y, Wu Y. A survey on data contamination for large language models. arXiv. Preprint posted online on Feb 20, 2025. [doi: 10.48550/arXiv.2502.14425]
48. Yang S, Chiang WL, Zheng L, Gonzalez JE, Stoica I. Rethinking benchmark and contamination for language models with rephrased samples. arXiv. Preprint posted online on Nov 8, 2023. [doi: 10.48550/arXiv.2311.04850]
49. Chen Z, Cox D, Gan C, et al. Principle-driven self-alignment of language models from scratch with minimal human supervision. Presented at: 37th Conference on Neural Information Processing Systems (NeurIPS 2023); Dec 10-16, 2023. [doi: 10.52202/075280-0115]
50. Mizrahi M, Kaplan G, Malkin D, Dror R, Shahaf D, Stanovsky G. State of what art? A call for multi-prompt LLM evaluation. Trans Assoc Comput Linguist 2024 Aug;12:933-949. [doi: 10.1162/tacl_a_00681]
51. Ouyang S, Zhang JM, Harman M, Wang M. An empirical study of the non-determinism of ChatGPT in code generation. ACM Trans Softw Eng Methodol 2023;34(2):1-28. [doi: 10.1145/3697010]
52. Moussiades L, Zografos G, Papakostas G. GPT-4 vs. GPT-3.5 as coding assistants. Research Square. Preprint posted online on Feb 7, 2024. [doi: 10.21203/rs.3.rs-3920214/v1]
53. Shuvo UA, Dip SA, Vaskar NR, Al Islam ABMA. Assessing ChatGPT's code generation capabilities with short vs long context programming problems. 2024 Presented at: Proceedings of the 11th International Conference on Networking, Systems, and Security (NSysS 2024); Dec 19-21, 2024. [doi: 10.1145/3704522.3704535]
54. OpenAI. GPT-4 technical report. arXiv. Preprint posted online on Mar 15, 2023. [doi: 10.48550/arXiv.2303.08774]
55. Hou Y, Zhan Z, Zhang R. Benchmarking GPT-5 for biomedical natural language processing. arXiv. Preprint posted online on Aug 28, 2025. [doi: 10.48550/arXiv.2509.04462]
56. Wang C, Sun JV. Unable to forget: proactive lnterference reveals working memory limits in LLMs beyond context length. arXiv. Preprint posted online on Jul 9, 2025. [doi: 10.48550/arXiv.2506.08184]

XSL•FO
RenderX

57. GPT-5.3 and GPT-5.4 in ChatGPT. OpenAI Help Center. 2025. URL: https://help.openai.com/en/articles/11909943-gpt-5-in-chatgpt [accessed 2025-10-20]
58. Gemini Apps limits & upgrades for Google AI subscribers. Gemini Apps Help. 2025. URL: https://support.google.com/gemini/answer/16275805?hl=en [accessed 2025-10-15]
59. Anil R, Borgeaud S, Alayrac JB, et al. Gemini: a family of highly capable multimodal models. : Google DeepMind; 2023 URL: https://storage.googleapis.com/deepmind-media/gemini/gemini_1_report.pdf [accessed 2026-03-07]
60. Alsajri A, Salman HA, Steiti A. Generative models in natural language processing: a comparative study of ChatGPT and Gemini. Babylonian J Artif Intell 2024;2024:134-145. [doi: 10.58496/BJAI/2024/015]
61. OpenAI goes local in India, ChatGPT now available in INR: check ChatGPT Plus, Pro prices here. India Today. 2025. URL: https://www.indiatoday.in/technology/news/story/openai-goes-local-in-india-chatgpt-now-available-in-inr-check-chatgpt-plus-pro-prices-here-2770840-2025-08-13 [accessed 2025-10-22]
62. Melo-López VA, Basantes-Andrade A, Gudiño-Mejía CB, Hernández-Martínez E. The impact of artificial intelligence on inclusive education: a systematic review. Educ Sci 2025;15(5):539. [doi: 10.3390/educsci15050539]
63. Umucu E. Artificial intelligence and health equity for people with disabilities: an integrated framework for disability-inclusive AI design. Inquiry 2025;62:469580251365472. [doi: 10.1177/00469580251365472] [Medline: 40847466]
64. Daepp MIG, Counts S. The emerging generative artificial intelligence divide in the United States. Presented at: Proceedings of the Nineteenth International AAAI Conference on Web and Social Media (ICWSM 2025); Jun 23-26, 2025. [doi: 10.1609/icwsm.v19i1.35825]
65. Laban P, Hayashi H, Zhou Y, Neville J. LLMs get lost in multi-turn conversation. arXiv. Preprint posted online on May 9, 2025. [doi: 10.48550/arXiv.2505.06120]
66. Liu Y, Le-Cong T, Widyasari R, et al. Refining ChatGPT-generated code: characterizing and mitigating code quality issues. ACM Trans Softw Eng Methodol 2024 Jun 30;33(5):1-26. [doi: 10.1145/3643674]
67. Claude AI. 2025. URL: https://claude.ai/upgrade [accessed 2025-10-21]
68. Zhang B, Liang P, Zhou X, Ahmad A, Waseem M. Practices and challenges of using GitHub Copilot: an empirical study. Presented at: The 35th International Conference on Software Engineering and Knowledge Engineering (SEKE 2023); Jul 10-12, 2023. [doi: 10.18293/SEKE2023-077]
69. Bartsch H, Jorgensen O, Rosati D, Hoelscher-Obermaier J, Pfau J. Self-consistency of large language models under ambiguity. Presented at: Proceedings of the 6th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP; Dec 7, 2023. [doi: 10.18653/v1/2023.blackboxnlp-1.7]
70. Donge V, Rossi RA, Lai VD, Yoon DS, Hakkani-Tür D, Bui T. Drift no more? Context equilibria in multi-turn LLM interactions. arXiv. Preprint posted online on Oct 9, 2025. [doi: 10.48550/arXiv.2510.07777]
71. Luo Y, Yang Z, Meng F, Li Y, Zhou J, Zhang Y. An empirical study of catastrophic forgetting in large language models during continual fine-tuning. IEEE Trans Audio Speech Lang Process 2023;33:3786-3776. [doi: 10.1109/TASLPRO.2025.3606231]
72. Comanici G, Bieber E, Schaeckerman M, et al. Gemini 2.5: pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. : Google DeepMind; 2025 URL: https://storage.googleapis.com/deepmind-media/gemini/gemini_v2_5_report.pdf [accessed 2026-03-07]
73. Huang D, Qing Y, Shang W, Cui H, Zhang JM. Effibench: benchmarking the efficiency of automatically generated code. Presented at: Proceedings of the 38th International Conference on Neural Information Processing Systems (NeurIPS 2024); Dec 9-17, 2024 URL: https://papers.nips.cc/paper_files/paper/2024/file/15807b6e09d691fe5e96cdecde6d7b80-Paper-Datasets_and_Benchmarks_Track.pdf [accessed 2026-03-07]
74. Rakotonirina NC, Hamdy M, Campos JA, et al. From tools to teammates: evaluating LLMs in multi-session coding interactions. Presented at: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); Jul 27 to Aug 1, 2025. [doi: 10.18653/v1/2025.acl-long.964]

Abbreviations

AI: artificial intelligence
AP: assistive product

XSL•FO
RenderX

Edited by S Munce; submitted 30.Oct.2025; peer-reviewed by A Elnashar, J Liu, M Al-Agil; accepted 19.Feb.2026; published 24.Mar.2026.

Please cite as:
Bañuls-Lapuerta FA, Marti-Miralles V, Gónzalez-García RJ, Martínez-Rico G
Using Natural Language Prompts With AI Models for Low-Cost Assistive Software Design: Exploratory Comparative Evaluation
JMIR Rehabil Assist Technol 2026;13:e86786
URL: https://rehab.jmir.org/2026/1/e86786
doi:10.2196/86786

© Francesc Antoni Bañuls-Lapuerta, Vicent Marti-Miralles, Rómulo Jacobo Gónzalez-García, Gabriel Martínez-Rico. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Eye-Tracking Technologies for Cognitive Assessment After Acquired Brain Injury: Systematic Review

Andrea Calderone[1], PsyD; Rosaria De Luca[1], MSc, PhD; Francesco Corallo[1], PsyD; Rosalia Calapai[1], PsyD; Alessio Mirabile[1], PsyD; Angelo Quartarone[1], MD, PhD; Alessandro Marco De Nunzio[2,3], PhD; Carmela Casella[4], MD; Rocco Salvatore Calabrò[1], MD, PhD

[1]Department of Neurorehabilitation, IRCCS Centro Neurolesi Bonino-Pulejo, S.S. 113 Via Palermo, C.da Casazza, Messina, Italy

[2]Department of Health, LUNEX University of Applied Sciences, LUNEX University of Applied Sciences, 50 Avenue du Parc des Sports, Differdange, Luxembourg

[3]Luxembourg Health & Sport Sciences Research Institute, 50 Avenue du Parc des Sports, Differdange, Luxembourg

[4]Stroke Unit, University of Messina, Piazza Pugliatti 1, Messina, Italy

Corresponding Author:
Andrea Calderone, PsyD
Department of Neurorehabilitation, IRCCS Centro Neurolesi Bonino-Pulejo, S.S. 113 Via Palermo, C.da Casazza, Messina, Italy

Abstract

Background: Acquired brain injury (ABI) is a heterogeneous umbrella term encompassing traumatic and nontraumatic etiologies and is frequently associated with persistent cognitive dysfunction. Conventional neuropsychological assessment remains central to clinical evaluation, but feasibility and measurement precision may be limited in individuals with motor impairment, aphasia, reduced stamina, or fluctuating arousal. Eye tracking offers an objective, low-burden approach that can quantify gaze behavior during task engagement and may provide complementary process-level markers of cognition.

Objective: This study aimed to systematically synthesize the evidence on eye-tracking paradigms used as a primary approach for cognitive assessment in ABI and to summarize findings by cognitive domain, paradigm, and clinical interpretability.

Methods: We conducted a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020–compliant systematic review and registered the protocol in PROSPERO (CRD420251038768). PubMed, Web of Science, the Cochrane Library, Embase, EBSCOhost, PsycINFO, and Scopus were searched from inception to April 10, 2025. We included peer-reviewed English-language studies enrolling children or adults with ABI in which eye tracking was the primary assessment modality used to quantify at least one cognitive domain or clinically relevant cognitive-communication process. Two reviewers independently screened studies, extracted data, and assessed methodological quality using design-appropriate tools (Risk of Bias 2, Risk of Bias in Non-Randomized Studies of Interventions, Quality Assessment of Diagnostic Accuracy Studies 2, and the Newcastle-Ottawa Scale). A structured narrative synthesis was performed because of heterogeneity in paradigms and outcome definitions.

Results: Twenty-seven studies met the inclusion criteria (N=872 participants; females: n=354 and males: n=518), with most evidence derived from mild traumatic brain injury cohorts, and fewer studies involving stroke, mixed etiologies, and disorders of consciousness. Across domains, antisaccade and related paradigms were commonly associated with differences in inhibitory control and executive function, while predictive tracking, smooth pursuit, and target-blanking paradigms frequently captured alterations in attentional prediction and timing. Virtual reality (VR) free-viewing paradigms identified visuospatial exploration asymmetries in stroke-related neglect, and gaze-based human-computer interface approaches demonstrated above-chance task performance in a subset of patients with disorders of consciousness. Evidence for incremental validity beyond conventional assessment was mixed and often indirect, and safety reporting was uncommon. Overall certainty of evidence was generally low and limited by small sample sizes, cross-sectional designs, and heterogeneity in acquisition procedures, metrics, and analytic pipelines.

Conclusions: Eye tracking shows potential as an adjunctive, process-level approach for quantifying specific cognition-relevant behaviors after ABI, particularly within paradigms targeting inhibitory control and predictive attention. Current evidence is insufficient to support broad diagnostic claims or the routine replacement of conventional neuropsychological assessment. Future research should prioritize harmonized paradigms and reporting standards, external validation of classification models, longitudinal designs, and explicit feasibility and safety reporting to clarify when eye tracking provides incremental clinical value for precision neurorehabilitation.

(JMIR Rehabil Assist Technol 2026;13:e81276) doi:10.2196/81276

XSL•FO
RenderX

KEYWORDS

acquired brain injury; eye-tracking technology; cognitive assessment; oculomotor biomarkers; saccadic eye movements; traumatic brain injury; neurocognitive impairment; neurorehabilitation

Introduction

Background

Acquired brain injury (ABI) is an important health concern, and its definition varies according to whether the medical condition is present at birth [1,2]. It includes all types of brain injury occurring after birth and not due to genetic or congenital pathology [3]. Unlike progressive neurological disorders such as Alzheimer disease or Parkinson disease, ABI is typically induced by an event that disrupts the normal function of the brain and results in changes to neuronal activity that can be observed at the physical, metabolic, or functional level of the nerve cells [1-3]. There are 2 main categories under ABI, which include traumatic brain injury (TBI) and nontraumatic brain injury (nTBI) [4]. Among the most common causes of TBI in older children and adults are assaults, falls, and crashes in motor vehicles, as well as sports injuries and blast injuries, including military blast injuries [5]. nTBI is an umbrella term that includes brain lesions seen in strokes in which blood enters but does not reach brain tissue; infections (eg, meningitis and encephalitis); hypoxia or anoxia, where the brain is deprived of oxygen; brain tumors; exposure to toxins; and drug or alcohol abuse [6]. Epidemiological data show that ABI is a global health problem and one of the major causes of morbidity and mortality [7,8]. More than 20 million new cases of TBI alone are estimated to occur each year worldwide; in fact, some estimates put the overall number of all TBIs, whatever their cause, at between 64 and 74 million annually [9]. TBI can be mild, moderate, or severe based on the score on the Glasgow Coma Scale or the duration of loss of consciousness or posttraumatic amnesia [10,11]. In addition to traditional clinical scales, motor and cognitive functional outcomes are significant limitations when trying to predict long-term recovery following ABI [12]. As many as 65% of individuals with moderate-to-severe TBI continue to experience chronic cognitive impairments, most likely to impact memory (20% - 79%), attention (18% - 54%), processing speed (17% - 19%), and executive functioning (eg, planning, problem solving, and multitasking) during the chronic phase of recovery [13]. In the motor domain, more than one-third of people with severe TBI still present with signs of neuromotor dysfunction (eg, paresis, ataxia, and postural abnormality) two years following brain injury [14,15]. These cognitive disturbances and motor deficits directly impact gait stability, fine motor coordination, and activities of daily living, underscoring the need for precise assessment tools [16].

Rehabilitation after ABI is increasingly characterized by technology-assisted interventions that aim to support cognitive and functional recovery. Recent syntheses highlight growing interest in noninvasive brain stimulation, virtual and extended reality–based rehabilitation, computerized cognitive training, and robot-assisted interventions, while also emphasizing heterogeneity in protocols, target domains, and patient profiles [17-20]. This evolving landscape strengthens the need for outcome measures that are sensitive to change, feasible across clinical settings, and capable of capturing real-time performance during task engagement.

Eye-tracking technology, characterized by excellent temporal and spatial resolution, not only allows for the sensitive detection of subtle oculomotor abnormalities associated with these functional deficits but also can be used as an assistive interface to restore communication for patients lacking volitional motor control [21]. Eye-tracking systems, which convert eye movements into control signals, can also assist with assessment, environmental control, communication, and rehabilitation exercises, enabling smooth transition between assessment and intervention and providing the connection to the applications [22]. Eye-tracking technology enables the measurement of various aspects of eye behavior (eg, eye gaze and pupil size) and has been applied in psychology research to study visual attention and cognitive processing states [23-25]. Often, the technology involves the use of infrared cameras and light sources that track the pupil and corneal reflections, which are then translated into data streams that allow for information about various cognitive states [26,27]. There are various types of eye trackers available, with options designed for controlled laboratory settings, real-world environments, virtual reality (VR) or augmented reality headsets, high-precision head-stabilized systems, and more accessible webcam-based trackers [28-31]. Eye tracking measures key oculomotor parameters such as fixations, saccades, pupillometry, smooth pursuit, vergence, blinks, and other eye movements (gaze paths) [32,33]. In cognitive assessment, eye tracking has been used to screen attention, memory, executive functions, language processing, and reading fluency and has demonstrated promise in the early diagnosis of neurodegenerative diseases [34,35]. Eye-tracking technology has particular implications for cognitive assessment in cases of ABI. Its usefulness has been extended to detect oculomotor dysfunction due to mild traumatic brain injury (mTBI), including concussive injuries [36]. It has also proven useful for determining levels of consciousness in patients with extreme brain injury, as well as for examining spatial biases in visual exploration after stroke [37].

Rationale, Objective, and Theoretical Framework for Investigating Eye Tracking in ABI Cognitive Assessment

As ABI represents a substantial global health problem, the previous overview shows that eye-tracking technology can be a useful and powerful tool to explore cognitive states through oculomotor behavior. Conventional neuropsychological assessments remain central to cognitive evaluation after ABI. Nevertheless, administration and interpretation can be challenging in individuals with motor impairment, aphasia, reduced stamina, or fluctuating arousal, which may limit feasibility and measurement precision [34]. Sensitivity can also be reduced for subtle deficits that are common after mTBI, and repeated testing can be influenced by practice effects. Scoring and interpretation may vary across raters and clinical contexts,

XSL•FO
RenderX

and many instruments provide only a snapshot rather than capturing dynamic cognitive control during continuous interaction with the environment. In contrast, eye tracking offers a potentially less demanding, systematic, and more objective approach to assess some cognitive functions, such as attention, memory, and executive control, through the direct measurement of neural processing correlates manifested in gaze patterns (fixation duration, saccadic patterns, and pupillary responses) [35]. However, although promising applications and research are emerging and interest in the field of ophthalmoplegia continues to grow, the available evidence supporting the implementation of eye tracking for cognitive assessment across the broad spectrum of ABI, including multiple etiologies (TBI and nTBI), levels of severity, and phases of recovery, remains mixed. Importantly, ABI encompasses heterogeneous etiologies and clinical phenotypes, including traumatic injury, stroke, hypoxic-ischemic injury, and disorders of consciousness (DoC), with meaningful variation by severity and phase of recovery. This heterogeneity can influence both the choice of eye-tracking paradigm and the interpretation of findings, underscoring the need to synthesize evidence with attention to ABI subgroup and target cognitive domain. There is a gap in the literature regarding a consolidated synthesis of this evidence. Therefore, this study aimed to identify and synthesize the available literature on the use of eye-tracking technology for cognitive assessment in populations with ABI. This work is based on a theoretical foundation that closely links oculomotor control to cognitive functioning, such that ABI-related alterations are thought to reflect measurable changes in eye movement patterns, which in turn are related to the affected cognitive domains [36,37]. From a neurobiological perspective, oculomotor behavior is governed by distributed control systems that overlap with cognitive networks. Antisaccade paradigms rely on frontostriatal circuitry and executive inhibition, whereas smooth pursuit and predictive tracking recruit cerebellar and parietal contributions that support timing, prediction, and attentional allocation. Fixation stability, gaze variability, and pupil dynamics can index distinct components of cognitive processing rather than a single unitary construct. These mechanistic links support the rationale for using eye tracking as a process-level window into cognition, provided that paradigms and metrics are interpreted within their specific neurocognitive targets [38,39]. This effort should allow for a thorough evaluation of the existing scientific evidence and help guide future researchers and clinicians in the rehabilitation process of this patient population. Table 1 provides an overview of eye-tracking technologies [40-48].

XSL•FO
RenderX

Table . Overview of eye-tracking technologies.

Eye tracking technology	Description	Mechanism of action	Relevance to ABI[a] cognitive assessment	Advantages and limitations	Clinical applications
Tobii 5 (laptop-mounted) [40]	Binocular eye tracker (133 Hz[b]) with head tracking (33 Hz), mounted on a laptop.	Infrared tracking of pupil and corneal reflection; calibrated using Tobii Pro Lab software (Tobii AB).	Captures changes in gaze stability and saccade latency, useful post injury.	Portable, high sampling rate, and limitation may be sensitive to ambient lighting.	Used for fixation and saccadic movement assessment after ABI or mTBI[c].
RightEye System [41]	Computerized eye-tracking system evaluating oculomotor behavior.	Digitally tracks smooth pursuits, saccades, vergence, and reaction times.	Assesses oculomotor deficits and symptom severity in RHI[d] and concussion.	Comprehensive eye metrics, and limitation may be less sensitive to subtle impairments.	Concussion and neurovisual symptom assessment.
BEAM Task Eye Tracker [42]	Automated tracking system integrated with BEAM[e] cognitive tasks.	Monitors saccades in response to visual cognitive tasks on screen.	Correlates with neuropsychological symptoms (eg, depression and PTSD[f]).	Continuous, task-integrated, and limitation may be task-dependent metrics.	Mental health–linked eye tracking in the ABI context.
Tobii T60/T60 [43]	Seventeen-inch monitor with noninvasive eye-tracking unit.	Infrared beams detect gaze location and fixations.	Detects fixation patterns and attention to scenes and grids.	Stable and accurate, and limitations may be fixed setup and screen size.	Visual attention and scene perception in ABI research.
Tobii EyeX [44]	Monitor-mounted tracker (>60 Hz) for fixation-point recording.	Infrared corneal reflection for eye-gaze coordinates.	Differentiates visual fixation across GCS[g] levels.	Affordable, good for basic research, and the limitation may be lower resolution.	Fixation-based cognitive-status screening.
HTC Vive with Pupil Labs [45]	VR[h] headset with integrated eye tracking for immersive testing.	IR[i] eye cameras track gaze with 1° accuracy in virtual environments.	Assesses VR-based gaze symmetry, fixation, and orientation in patients.	Ecological validity, and the limitation may be potential motion sickness and setup cost.	VR cognitive assessment and rehabilitation for ABI/mTBI.
EyeLink 1000 [46]	High-speed eye tracker (up to 1000 Hz) and head-stabilized.	IR beam reflection records eye movements with high temporal precision.	Tracks reaction times and fixation accuracy in cognitive tasks.	High precision and limitation may require head stabilization.	Detailed oculomotor assessment in laboratory settings.
Gazepoint GP3 HD [47]	150 Hz eye tracker with 9-point calibration screen.	Measures saccades and working memory through gaze behavior.	Differentiates performance in working memory and inhibition tasks.	Compact and high frequency, and limitation may miss microsaccades.	Working memory and executive function testing in ABI.
Applied Science Laboratories D6 [48]	High-speed desktop tracker for effort testing.	Records gaze patterns to detect malingering/fake effort.	Validates effort and symptom exaggeration in mTBI.	Reliable classification and limitation may be specialized purpose.	Effort testing and malingering detection.

[a]ABI: acquired brain injury.

[b]HZ: Hertz.

[c]mTBI: mild traumatic brain injury.

[d]RHI: right hemisphere infarction.

[e]BEAM: Bethesda Eye and Attention Measure.

[f]PTSD: posttraumatic stress disorder.

[g]GCS: Glasgow Coma Scale.

[h]VR: virtual reality.

[i]IR: infrared.

Methods

Protocol and Reporting Standard

This systematic review was conducted and reported in line with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidance, with all stages planned to maximize transparency and reproducibility [49]. The review protocol was prospectively registered in PROSPERO (CRD420251038768) to reduce the risk of selective reporting and to provide an auditable record of the prespecified methods [50]. Any deviations from the registered protocol were documented and justified within the paper to preserve methodological transparency [49,50]. Following peer-review feedback, the PROSPERO record was updated on February 26, 2026, to clarify eligibility criteria, including comparator requirements and inclusion of eligible case series, and to specify the design-appropriate risk-of-bias tools used in this review.

XSL•FO
RenderX

Review Question Framework (PICO)

To ensure conceptual clarity and consistent eligibility decisions, we framed the review question using a PICO-informed structure adapted for diagnostic and assessment technologies. The population comprised children and adults with ABI across etiologies, severities, and recovery phases. The index assessment modality was eye tracking, used as the primary approach to quantify cognition or clinically relevant cognitive-communication processes, with paradigms and derived metrics prespecified and mapped to cognitive domains. Comparators were not mandatory and, when available, included healthy controls, ABI subgroups, within-participant conditions (eg, baseline vs follow-up), and/or conventional neuropsychological measures. Outcomes were eye-tracking–derived metrics relevant to cognition (eg, antisaccade latency or error rates for inhibitory control, predictive tracking measures for attentional timing, and fixation or exploration indices for visuospatial processing), along with diagnostic or classification performance indices when explicitly reported. This framework guided study selection, data extraction, and synthesis and reduced the risk of misclassifying eye tracking as an intervention rather than an assessment approach.

Eligibility Criteria and Operational Definitions

We included peer-reviewed, full-text studies published in English that enrolled children or adults with ABI of any etiology, severity, and recovery phase. Eligible etiologies encompassed TBI, stroke, hypoxic-ischemic injury, infectious or inflammatory ABI, and DoC following ABI. Studies were eligible when eye tracking was used as the primary assessment modality to quantify cognitive functioning. Eye tracking was operationally considered "primary" when the paradigm and derived oculomotor metrics were central to the cognitive assessment objective and were used to operationalize at least one cognitive domain or a clinically meaningful cognitive-communication construct. For the purposes of this review, eligible cognitive targets included attention, executive function, memory and working memory, visuospatial exploration and neglect-related processes, language and cognitive-communication processes, social-cognitive paradigms, and gaze-based paradigms aimed at detecting command following, volitional control, or awareness in DoC. We included randomized controlled trials, nonrandomized intervention studies, and observational designs, including cohort, cross-sectional, and case-control studies.

We excluded protocols, conference abstracts, dissertations, reviews, animal studies, and single-case reports. We also excluded studies in which eye tracking was used only for technical device validation, usability testing without cognitive interpretation, or oculomotor characterization without an explicit link to a cognitive target or classification framework. Studies that included mixed neurological samples were eligible only when ABI data were reported separately or could be unambiguously extracted.

Information Sources and Search Strategy

We searched PubMed, Web of Science, the Cochrane Library, Embase, EBSCOhost, PsycINFO, and Scopus from database inception to April 10, 2025. The initial search was performed on March 5, 2025 and updated on April 10, 2025. No publication date restrictions were applied. Searches combined controlled vocabulary terms and free-text keywords related to eye tracking, oculomotor metrics, cognition, and cognitive assessment, and ABI and related etiologies. The strategy was iteratively refined to optimize sensitivity while preserving relevance, and the complete search strategy is reported in Multimedia Appendix 1.

To reduce the risk of missing relevant studies, the reference lists of included studies and of closely related reviews were screened, and forward citation tracking was performed for key sentinel studies identified during full-text screening. Only peer-reviewed full-text papers were included to align with the prespecified scope and to ensure interpretability of methods and results.

Study Selection

Records were exported to a single dataset and deduplicated using multiple identifiers, including title, authorship, year, and digital object identifier when available. Two reviewers (AC and RDL) independently screened titles and abstracts, followed by full-text assessment against eligibility criteria. Before formal screening, a calibration phase was performed to ensure consistent application of criteria. Disagreements were resolved through discussion, with adjudication by a third reviewer (RSC) when needed. Interrater agreement at the abstract and full-text stages was quantified using the Cohen kappa statistic, and agreement was interpreted using established benchmarks to support transparent reporting of selection reliability [51,52]. Reasons for exclusion at the full-text stage were recorded and are available upon request.

Data Extraction and Data Items

Two reviewers (AC and RDL) independently extracted data using a piloted extraction form, with discrepancies resolved by consensus and third-reviewer adjudication (RSC) when necessary. Extracted study-level information included design, setting, recruitment procedures, sample size, participant demographics, ABI etiology, severity indices where reported, and phase of recovery at assessment.

Eye-tracking–specific data were extracted at a granular level to support clinical interpretability and reproducibility. We recorded the eye-tracking platform and device model, sampling rate, calibration procedures, and success criteria when reported, head stabilization conditions, and the testing environment. We extracted paradigm characteristics, including task instructions, stimulus type, trial structure and duration, and any concurrent task demands such as dual tasking or working memory load manipulations. We extracted primary and secondary oculomotor metrics and documented how each metric was operationalized by the original authors. We then mapped each paradigm and metric to the prespecified cognitive target domain using the authors' stated rationale and the task requirements.

Comparator conditions were recorded when present, including healthy controls, ABI subgroups, baseline within-participant comparators, and conventional neuropsychological instruments used as reference measures. We extracted statistical information required for synthesis, including group differences,

XSL•FO
RenderX

within-subject changes, correlation coefficients, regression coefficients, and classification performance metrics such as sensitivity, specificity, and area under the curve when reported. Feasibility and data quality indicators were extracted when available, including calibration failures, exclusion due to inadequate tracking quality, missing data rates, and any preprocessing or artifact handling described by the authors. Safety information was extracted and coded as not reported, explicitly absent, or reported with no adverse events to avoid misclassification of missing safety reporting as evidence of safety.

Risk of Bias Assessment

Risk of bias was assessed independently by 2 reviewers using design-appropriate tools, with disagreements resolved by consensus and third-reviewer adjudication when needed. For randomized controlled trials, we applied Risk of Bias 2 across the standard domains and derived an overall judgment to inform interpretation of intervention effects [53]. For nonrandomized intervention studies, we applied Risk of Bias in Non-Randomized Studies of Interventions, with particular attention to confounding, selection bias, and outcome measurement, which are especially relevant in heterogeneous ABI populations [54].

For studies whose primary purpose was classification or diagnostic discrimination using eye-tracking metrics, we evaluated risk of bias and applicability using Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2), which is specifically structured to address concerns in patient selection, index test conduct and interpretation, reference standards, and flow and timing [55]. For observational cohort and case-control studies that did not meet diagnostic-accuracy criteria, we assessed methodological quality using the Newcastle-Ottawa Scale (NOS) to capture selection, comparability, and outcome or exposure ascertainment features that may drive systematic bias [56]. Risk-of-bias assessments were not used as exclusion criteria, but they were incorporated into synthesis and certainty judgments to maintain transparency in interpretation.

Data Synthesis and Effect Measures

Given the expected heterogeneity in paradigms, metrics, and ABI subgroups, we prespecified a primarily narrative synthesis structured by cognitive domain and eye-tracking paradigm, with additional stratification by etiology and recovery phase where feasible. Narrative synthesis followed established principles for integrating quantitative and qualitative evidence, with emphasis on consistency of direction and robustness across designs rather than vote counting [57].

When quantitative pooling was feasible, we planned meta-analysis only when at least 3 studies used sufficiently comparable paradigms, reported the same outcome metric, and provided adequate summary data. For continuous outcomes, we planned to compute standardized mean differences for between-group comparisons or standardized mean change for within-subject designs, with 95% CIs. Random-effects models were prespecified, given expected clinical and methodological diversity. When meta-analysis was not feasible, we summarized findings descriptively and reported effect sizes as presented by the original studies, clarifying the metric and its interpretation. Diagnostic performance indices such as sensitivity, specificity, and area under the curve were summarized descriptively and were not pooled due to heterogeneity in thresholds, paradigms, and reference standards.

Certainty of Evidence

Certainty of evidence was evaluated using GRADE principles, considering risk of bias, inconsistency, indirectness, imprecision, and publication bias [58,59]. Certainty ratings were planned for each main outcome group aligned with the synthesis structure, with judgments explicitly justified and linked to the corresponding risk-of-bias assessments. Ratings were planned to be presented in a summary of findings table to support transparent translation of the evidence base into clinical interpretability [58,59]. The methodological framework used in this systematic review is summarized in Table 2.

XSL•FO
RenderX

Table . Detailed summary of the systematic review methodology.

Section methodology	Details
Protocol and reporting	• Reporting standard: PRISMA[a] 2020 [49]. • Registration: prospective registration in PROSPERO[b] (CRD420251038768) [50].
Inclusion criteria	• Participants: children and adults with acquired brain injury (ABI) of any etiology, severity, and recovery phase, including traumatic brain injury, stroke, hypoxic-ischemic injury, and disorders of consciousness after ABI. • Technology: studies in which eye tracking was used as the primary assessment modality to quantify cognitive functioning. • Operational definition: eye tracking was considered primary when the paradigm and derived oculomotor metrics were central to the cognitive assessment objective and were used to operationalize at least one cognitive domain or a clinically relevant cognitive-communication process. • Cognitive targets: attention, executive function, memory and working memory, visuospatial exploration and neglect-related processes, language and cognitive-communication processes, social-cognitive paradigms, and gaze-based paradigms aimed at detecting command following, volitional control, or awareness in disorders of consciousness. • Eye-tracking metrics: studies reporting defined oculomotor outcomes (eg, fixations, saccades, smooth pursuit, vergence, or pupillary dynamics) with an explicit cognitive interpretation, including associations with conventional cognitive measures and/or discrimination or classification performance. • Study designs: randomized controlled trials, nonrandomized intervention studies, and observational designs (cohort, cross-sectional, and case-control). • Language: full-text papers published in English.
Exclusion criteria	• Focus of study: studies not centered on cognitive assessment using eye tracking in ABI populations. • Population: nonhuman studies, non-ABI samples, or mixed neurological samples without extractable ABI-specific data. • Technology use: studies using eye tracking only for technical device validation or usability testing without a cognitive framework, or studies describing oculomotor dysfunction without an explicit link to a cognitive target or classification approach. • Publication type: protocols, conference abstracts, dissertations, reviews, and single-case reports or case series.
Review question framework	• Population: individuals with ABI across etiologies, severity levels, and recovery phases, including pediatric and adult samples. • Index assessment: eye-tracking paradigms and derived oculomotor metrics used to quantify cognitive domains or cognitive-communication constructs. • Comparator or reference: healthy controls, ABI subgroups, within-participant baselines, and/or conventional neuropsychological measures when available. • Outcomes: domain-relevant cognitive inferences from eye-tracking metrics, associations with conventional assessments, group discrimination or classification performance (eg, sensitivity, specificity, and area under the curve), and feasibility and safety indicators.
Search period	• Coverage: database inception to April 10, 2025, with no publication date restrictions. • Initial search: March 5, 2025. • Update search: April 10, 2025. • Information sources: PubMed, Web of Science, the Cochrane Library, Embase, EBSCOhost, PsycINFO, and Scopus. Full database-specific strategies are reported in Multimedia Appendix 1.

XSL•FO
RenderX

Section methodology	Details
Study selection	• Deduplication: records were collated and deduplicated before screening. • Screening: 2 reviewers independently screened titles and abstracts, followed by full-text eligibility assessment, with third-reviewer adjudication when needed. A PRISMA 2020 flow diagram documented the selection process [49]. • Reliability: interrater agreement was quantified using the Cohen kappa statistic and interpreted using established benchmarks [51,52].
Data extraction	• Process: 2 reviewers independently extracted data using a piloted extraction form, with discrepancies resolved by consensus and third-reviewer adjudication when required. • Study-level items: design, setting, sample size, participant demographics, ABI etiology, severity indices where reported, and phase of recovery. • Eye-tracking technical items: device model or platform, sampling rate, calibration procedure, and success criteria when reported, head stabilization conditions, and testing environment. • Paradigm and outcomes: task procedures and stimulus characteristics, trial structure and cognitive load, primary eye-movement metrics, and mapping of paradigms and metrics to cognitive targets. • Comparators and statistics: comparator conditions (eg, controls, ABI subgroups, within-participant baselines, and conventional tests) and extractable statistics (group differences, within-subject changes, correlations, regression coefficients, and diagnostic performance indices when reported).
Risk of bias assessment	• Randomized trials: RoB 2[c] [53]. • Nonrandomized interventions: ROBINS-I[d] [54]. • Diagnostic accuracy or classification studies: QUADAS-2[e] [55]. • Observational cohort or case-control studies: Newcastle-Ottawa Scale [56]. • Process: 2 reviewers assessed risk of bias independently, with disagreements resolved by consensus and third-reviewer adjudication when necessary.
Data synthesis	• Primary approach: structured narrative synthesis organized by cognitive domain and eye-tracking paradigm, with stratification by etiology and recovery phase where feasible, following established narrative synthesis principles [57]. • Quantitative synthesis: meta-analysis was planned only when at least 3 studies used sufficiently comparable paradigms and reported the same outcome metric with adequate summary data; random-effects models were prespecified for continuous outcomes. • Diagnostic indices: classification and diagnostic performance metrics were summarized descriptively and were not pooled due to heterogeneity in thresholds, paradigms, and reference standards.
Safety and feasibility	• Safety: adverse events were extracted and coded as not reported, explicitly absent, or reported with no adverse events. • Feasibility: calibration failures, tracking-quality exclusions, and missing data were extracted when reported.
Certainty of evidence	• Approach: certainty of evidence for each main outcome group was evaluated using GRADE[f] principles [58,59]. • Reporting: certainty ratings were planned to be presented in a summary of findings table.

[a]PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

[b]PROSPERO: International Prospective Register of Systematic Reviews.

[c]RoB 2: Risk of Bias 2.

[d]ROBINS-I: Risk of Bias in Non-Randomized Studies of Interventions.

[e]QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies 2.

XSL•FO
RenderX

[f]GRADE: Grading of Recommendations, Assessment, Development, and Evaluations.

Ethical Considerations

As this systematic review involved secondary data analysis from previously published studies, no new ethical approval was required.

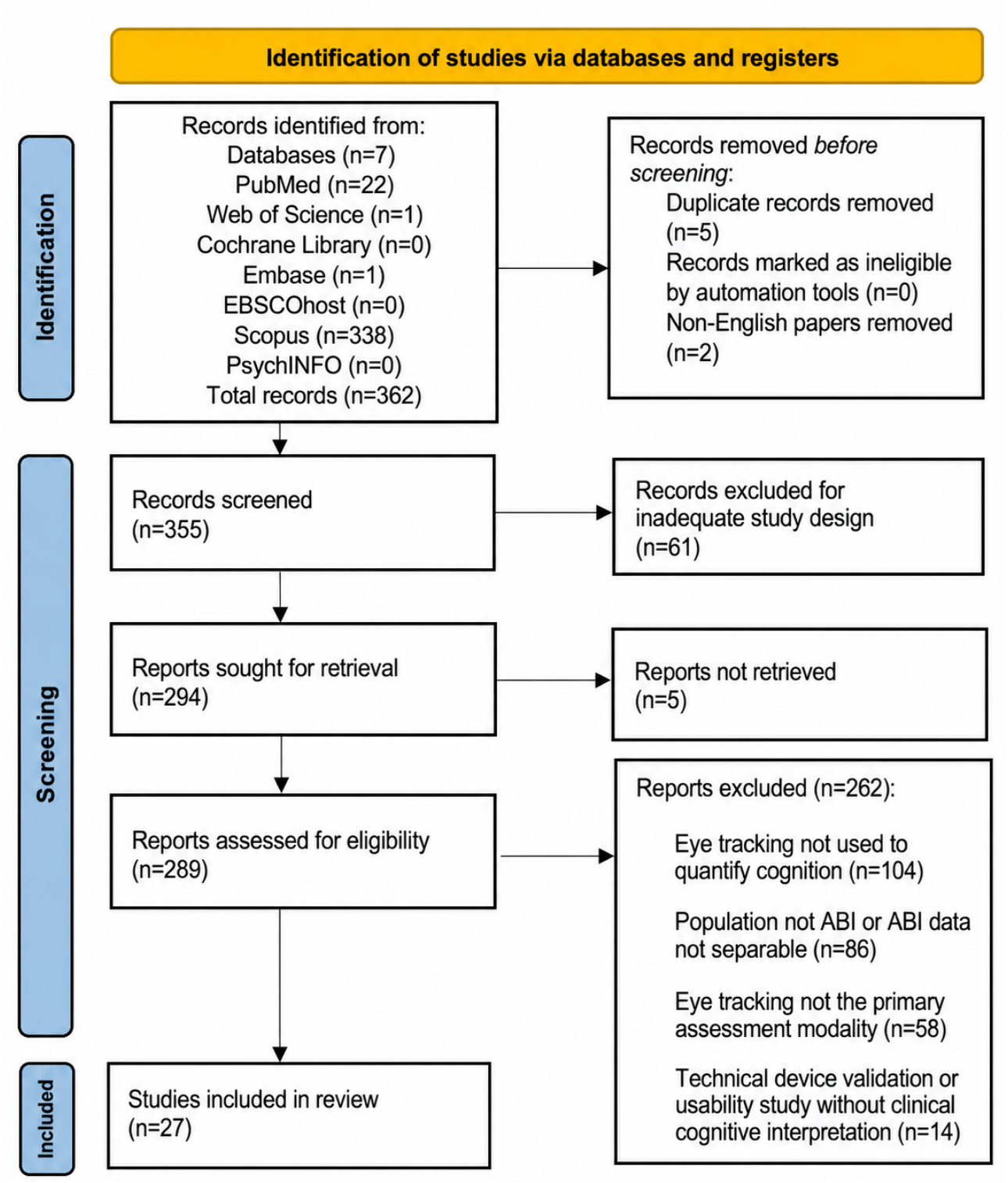
Results

Study Selection

The electronic search across 7 databases identified 362 records. After removal of 5 duplicates and 2 non-English records, 355 records underwent title and abstract screening. At this stage, 61 records were excluded due to inadequate study design, leaving 294 reports sought for retrieval. Despite targeted retrieval efforts, 5 reports could not be obtained, and 289 full-text reports were assessed for eligibility. Of these, 262 reports were excluded for prespecified reasons, most commonly because eye tracking was not used to quantify cognition (104 reports), the population was not ABI or ABI-specific data were not separable (86 reports), eye tracking was not the primary assessment modality (58 reports), or the paper focused on technical and device validation or usability without clinical-cognitive interpretation (14 reports). This process yielded 27 studies that met eligibility criteria and were included in the qualitative synthesis [60-86]. The study selection process is summarized in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram of evaluated studies. ABI: acquired brain injury.



Characteristics of Included Studies and Participants

The included evidence base (refer to Multimedia Appendix 1) was heterogeneous in both clinical populations and methodological approaches, with a strong predominance of TBI cohorts and fewer studies in nontraumatic ABI subgroups. Across included studies, a total of 872 participants were enrolled, including 354 females and 518 males, although sex was not consistently reported or analyzed as an effect modifier across all cohorts [60-86]. Most studies focused on mTBI, concussion, or persistent postconcussive symptoms, often in military or athletic samples [60-62,68-75,78-83,85]. Moderate-to-severe TBI was represented in a smaller subset of studies, generally using case-control paradigms to examine cognitive-communication or higher-order integration processes [63,64,67,76,77]. Stroke was represented primarily through a single visuospatial neglect study using a VR free-viewing paradigm [66], while DoC were represented through gaze-based

XSL•FO
RenderX

paradigms and human-computer interface (HCI) approaches designed to probe command following or awareness in patients with severe impairments [65,84]. One rehabilitation study enrolled adults with mixed etiologies (TBI and stroke) and examined oculomotor rehabilitation for reading-related deficits [86]. A single pediatric study examined mild closed head injury and tracked saccadic outcomes and cognitive performance longitudinally for more than 6 months [80]. Pediatric evidence was limited to one longitudinal study in children with mild closed head injury, which assessed saccadic performance and memory outcomes at acute, 3-month, and 6-month follow-up. This study suggested that developmental factors and age subgroup effects may influence oculomotor trajectories, with a pattern consistent with altered maturation or practice effects rather than a uniform deficit profile across all pediatric participants [80]. Because pediatric data were sparse and paradigms differed from those used in adult cohorts, findings in children should be interpreted cautiously and not generalized to the broader ABI spectrum without additional age-stratified studies using harmonized paradigms and reporting.

Study sizes were typically small to moderate. Median sample size across studies was 45 participants (range 10 - 188), with 15 of 27 studies enrolling fewer than 50 participants and only 2 studies enrolling 100 or more participants [62,85]. Most studies were cross-sectional, with fewer longitudinal designs or repeated-measures components. Longitudinal or baseline-to-follow-up approaches were used in concussion cohorts to relate within-individual changes in eye-tracking metrics to symptom burden [79,83], and one pediatric study implemented repeated assessments at acute, 3-month, and 6-month time points [80]. The evidence base, therefore, offers breadth in paradigms and outcomes, but it remains uneven across etiologies, severity strata, and recovery phases, with limited prospective follow-up outside concussion-focused cohorts.

Results by Cognitive Domain and Clinical Use Case

Given the heterogeneity of paradigms and outcome definitions, no quantitative pooling was performed. Findings are therefore summarized narratively and organized by cognitive domain and clinical application, with attention to consistency across paradigms, ABI subgroups, and the extent to which eye-tracking outcomes were evaluated alongside conventional assessments.

Executive Function and Inhibitory Control

Evidence relevant to executive function and inhibitory control was derived largely from antisaccade and related paradigms, as well as from task batteries embedding oculomotor indices under attentional demands [68,73-77,80,83]. Across adult mTBI and chronic TBI samples, antisaccade paradigms generally demonstrated increased latency and/or elevated error rates relative to controls, consistent with impaired inhibitory control and executive regulation [74,76,83]. In chronic cohorts, executive dysfunction was often most apparent under task conditions requiring suppression of reflexive saccades or sustained inhibitory control, with several studies indicating that oculomotor indices captured impairments not always mirrored by global cognitive screening measures. In one exploratory study integrating antisaccade performance with diffusion measures and Stroop interference, longer antisaccade latencies were associated with greater symptom burden and poorer interference control in acute mTBI and were strongly correlated with white matter integrity markers in the splenium of the corpus callosum within the acute subgroup [83]. These observations support the plausibility that oculomotor inhibition metrics can reflect executive control processes and their neurobiological correlates in at least a subset of mTBI presentations.

Age- and severity-related modulation emerged as important contextual factors. In one study using embedded saccadic measures, older age was associated with slower saccadic response times and reduced inhibitory performance among those with a history of mTBI, suggesting that eye movement–based indices may capture vulnerability patterns that interact with age-related changes in attentional control [75]. Severity-dependent effects were also reported in chronic TBI cohorts, with moderate-to-severe groups typically exhibiting broader impairment profiles than mild TBI groups under similar saccade paradigms [76,77]. These findings are consistent with an injury-severity gradient, but they also highlight that the magnitude and detectability of inhibitory deficits depend on task selection, recording parameters, and analytic definitions.

Working Memory and Cognitive Load Effects

A smaller evidence base evaluated working memory or cognitive load effects through paradigms explicitly manipulating attentional demand, most commonly using n-back designs with embedded oculomotor outcomes [68,73] and related load-dependent paradigms [72]. Across these studies, oculomotor impairment tended to become more apparent as cognitive load increased. In chronic TBI cohorts, saccadic latency and error indices worsened at higher load levels, with greater impairment observed in more severe injury groups and with findings suggesting that load manipulations increase the sensitivity of eye movement indices to cognitive vulnerability [73]. A clinically oriented multimodal assessment incorporating a cognitively demanding paradigm reported that specific saccadic indices contributed to discrimination between chronic mTBI and controls, with conventional neuropsychological measures showing less consistent separation in that setting [68]. Evidence in this domain remains limited by the small number of studies and by task-specific implementations, but the direction of findings is broadly coherent with the notion that cognitive load amplifies executive and attentional demands reflected in oculomotor control.

Attention and Predictive Control

The largest body of evidence relevant to attention, predictive control, and attentional timing involved smooth pursuit, predictive visual tracking, target blanking, and gap paradigms designed to probe anticipatory control [70,72,74,78,79,81,82]. Several studies reported that mTBI or postconcussive symptom groups demonstrated altered pursuit or visual tracking performance, particularly under conditions that required prediction of target trajectories or compensation for missing visual input. In 2 studies using predictable smooth pursuit paradigms, mild TBI groups showed reduced target prediction and increased eye position error and variability relative to controls, with correlations between pursuit-related indices and verbal learning measures related to attention and executive

XSL•FO
RenderX

functioning [81,82]. In a predictive tracking task coupled with magnetoencephalography, deficits were most pronounced under gap conditions, suggesting that anticipatory control may be selectively impaired when the paradigm reduces reliance on continuous sensory feedback and increases top-down prediction demands [72]. Similarly, a visual tracking paradigm used in concussion monitoring showed that within-individual changes in tracking metrics were moderately to strongly correlated with symptom burden when assessed within 2 weeks post injury, whereas simple reaction time metrics showed weaker relationships to symptom scales despite measurable post injury slowing in throughput measures [79]. Together, these findings suggest that predictive timing and sustained attention components of pursuit or tracking paradigms may map more closely to symptom profiles than simple reaction-based outcomes in some concussion contexts.

Not all evidence converged on robust group differences. One study combining diffusion tensor imaging, neurocognitive assessment, and visual tracking reported that the concussion cohort showed largely comparable tracking performance to controls, despite imaging differences and evidence of slowed attention-related reaction time, indicating that tracking metrics may be sensitive in some cohorts but not uniformly abnormal across mTBI presentations [78]. In another study, behavioral eye measures were less discriminatory than imaging-derived indices in separating chronic mTBI from controls, underscoring that the comparative performance of eye tracking depends on the specific paradigm and the selected outcome metric [70]. These mixed findings collectively support the need to interpret attentional tracking outcomes as paradigm- and population-dependent rather than as uniform markers of impairment.

Visuospatial Exploration and Neglect Detection

Evidence for visuospatial exploration was limited but clinically instructive. In a VR free-viewing museum environment, stroke patients with spatial neglect demonstrated lateralized gaze and head-orientation asymmetries relative to controls, and the approach was described as capable of detecting atypical exploration patterns that were not captured consistently by conventional tests in all individuals [66]. This study supports the feasibility and potential incremental utility of immersive, naturalistic paradigms for detecting spatial biases, although generalizability remains limited because the evidence in this domain is currently based on a single study and a single implementation approach.

Cognitive-Communication and Social Cognition

Three studies examined cognitive-communication or socially relevant visual attention patterns through tasks that were not reducible to classic executive function paradigms, but nevertheless quantified cognition through structured gaze behavior [63,64,67]. In a visual-world paradigm assessing online speech–gesture integration, both groups showed facilitation from meaningful gestures, but this effect was attenuated in moderate-to-severe TBI, suggesting reduced integration of multimodal cues during language processing [67]. In severe TBI cohorts evaluating visual attention during interpretation of augmentative and alternative communication displays, grid-based formats were associated with greater visual effort than scene-based formats in both groups, and TBI participants showed less efficient gaze patterns overall, consistent with increased processing demands and reduced efficiency in extracting relevant information [63]. A separate free-viewing study focusing on engagement cues in contextual photographs showed that both groups oriented rapidly to human figures and adjusted fixation patterns according to task-relevant engagement content, with limited between-group differences, suggesting that some aspects of social attention may be relatively preserved depending on task demands and stimulus properties [64]. Taken together, this evidence indicates that eye tracking can quantify differences in processing efficiency and cue integration in cognitive-communication contexts, but the limited number of studies and the task-specific nature of outcomes constrain broad conclusions.

DoC and Gaze-Based Paradigms for Command Following

Two studies focused on DoC populations and used gaze-based paradigms to probe command following or awareness in contexts where conventional assessment is challenged by limited motor output [65,84]. In a gaze-tracking HCI approach designed for DoC, a subset of patients performed above chance on one or more structured tasks, and the authors highlighted the potential for such tools to complement bedside diagnosis and reduce misclassification, including identification of patients who might retain meaningful command-following capacity [84]. A second study combined gaze-tracking tasks with complementary neurophysiologic measures and suggested that multimodal features could relate to clinical status classifications in severe ABI populations [65]. The evidence in this domain remains limited by small samples, variable arousal, and setting-specific feasibility constraints, but the convergence of both studies on above-chance task completion in a subset of patients supports continued evaluation of eye-tracking HCI approaches as adjunctive assessment tools in severe disorders.

Incremental Validity Relative to Conventional Assessment

Direct evidence of incremental validity beyond conventional neuropsychological testing was limited because many studies did not prespecify incremental validity as a primary analytic goal and because neuropsychological comparators varied across cohorts and settings. When conventional measures were included, they were often used as correlates rather than as formal reference standards. Still, several patterns emerged. First, in paradigms tightly linked to attentional prediction and executive inhibition, oculomotor indices often showed associations with symptom burden and executive test performance, suggesting that they can capture clinically relevant variance in domains that are difficult to quantify with brief assessments alone [72,79,83]. Second, some multimodal or embedded-task approaches reported that oculomotor indices contributed meaningfully to group discrimination or classification in settings where conventional cognitive batteries showed limited differentiation, although this was not uniform across platforms and paradigms [68,70]. Third, other studies indicated that symptom-based or vestibular-ocular screening measures may

XSL•FO
RenderX

outperform certain computerized eye-tracking platforms for specific operational goals, emphasizing that incremental value is likely to be context dependent and may vary by hardware, paradigm, and target construct [61]. Collectively, the evidence supports cautious interpretation: eye tracking may add value for specific paradigms and populations, but the current literature does not justify a generalized claim that eye tracking routinely exceeds conventional testing across ABI.

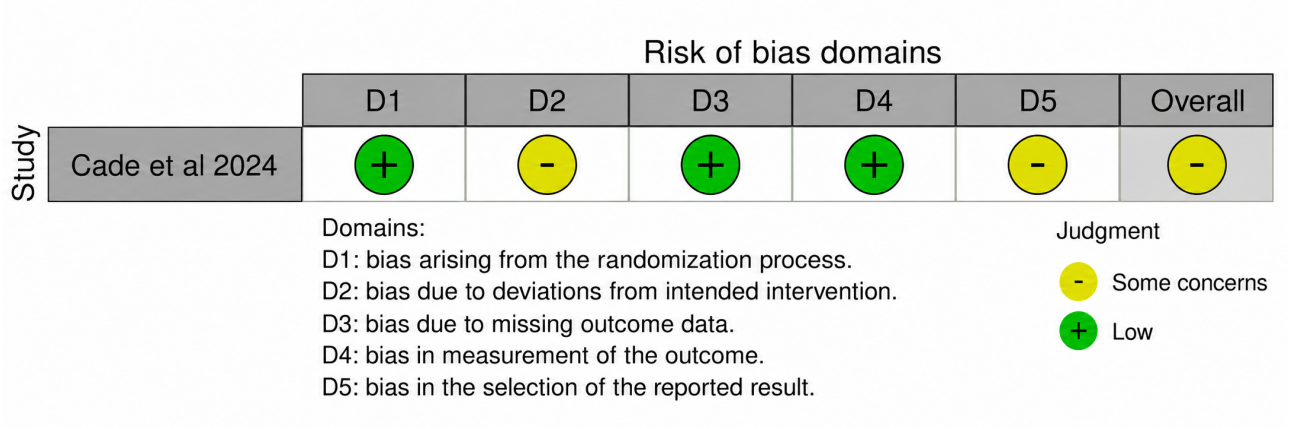
Safety, Feasibility, and Data Quality Reporting

Safety reporting was sparse across the included studies. Only 1 of 27 (4%) studies explicitly reported no harms, while 26 of 27 (96%) studies did not report adverse events in a way that allowed clear classification [60-86]. Feasibility reporting was more common but remained incomplete. Thirteen of 27 (48%) studies provided at least some feasibility or exclusion information related to calibration, equipment failures, incomplete testing, or analytic inclusion, whereas 14 of 27 (52%) studies did not report feasibility indicators beyond general study completion statements. Where reported, feasibility challenges included calibration failure and fatigue limiting completion in a VR neglect assessment [66], equipment-related exclusions in a language-processing study [67], incomplete administration of conventional tests due to physical impairment in a severe TBI cohort [63], and differential analytic sample sizes attributable to unusable eye-tracking records in a prospective concussion cohort [79]. Several paradigms required head stabilization, specialized equipment, or controlled testing conditions, which may limit scalability in routine clinical environments [67,70,74,81,82]. These findings underline that feasibility and access considerations are integral to interpretation and translation, and that standardized reporting of calibration success, data loss, and exclusions remains an important gap.

Methodological Quality and Risk of Bias

Risk-of-bias assessments varied by study design and were interpreted with attention to the appropriate tool for each design. The complete risk-of-bias and methodological quality assessments are provided in Multimedia Appendix 2. The single randomized controlled trial was judged as having some concerns overall, reflecting limitations primarily related to deviations from intended interventions and the transparency of prespecified outcome reporting, while retaining low-risk judgments in several other domains of trial conduct and outcome measurement. Risk of Bias 2 domain-level judgments are provided in Figure 2.

Figure 2. Risk of Bias of included randomized controlled trial studies [60].



Diagnostic and classification studies showed more substantial and heterogeneous concerns. QUADAS-2 assessments, summarized in Table 3, indicated a high risk of bias across all 3 classification studies, primarily driven by nonconsecutive or volunteer sampling strategies and by data-driven selection of predictors and decision thresholds. Applicability concerns were great for Barry and Ettenhofer [71] and Samadani et al [85] because the target condition and/or reference standards were not directly aligned with cognitive-status classification in ABI (performance-validity simulation in the former and computed tomography–defined injury status in the latter). Ettenhofer et al [68] showed lower applicability concerns because chronic mTBI status was established using a standardized identification approach with medical record confirmation, but risk of bias remained high due to in-sample variable selection procedures and exclusions that may inflate classification performance. These considerations suggest that reported classification estimates should be interpreted cautiously, particularly when models and thresholds are derived and evaluated within the same dataset, and external validation is not available [68,71,85].

XSL•FO
RenderX

Table . Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) risk of bias and applicability assessment for diagnostic or classification studies. Judgments follow QUADAS-2 domains. Applicability concerns reflect relevance to the review question (eye-tracking metrics used for cognitive assessment or classification in acquired brain injury).

Study	Patient selection (RoB[a] or applicability)	Index test (RoB or applicability)	Reference standard (RoB or applicability)	Flow and timing (RoB)	Overall (RoB)
Barry and Ettenhofer (2016) [71]	High or High	High or High	High or High	Low	High
Ettenhofer et al (2020) [68]	High or High	High or Low	Low or Low	High	High
Samadani et al (2015) [85]	High or Low	High or High	Low or High	Low	High

[a]RoB: risk of bias.

Observational studies assessed with the NOS generally demonstrated fair methodological quality, with fewer studies achieving good ratings and 1 study rated as poor. Specifically, among the 19 observational studies assessed with NOS, 6 were rated good, 12 were rated fair, and 1 was rated poor, indicating recurring limitations in comparability and exposure or outcome ascertainment despite generally acceptable selection domains in many case-control designs [63,64,66,67,70,72-74,76-83]. NOS ratings are summarized in Table 4. Nonrandomized intervention studies [65,68,69,84-86] were assessed separately for risk of bias using an intervention-focused tool, and the domain-level judgments are provided in Figure 3. Across these designs, confounding and participant selection were common sources of concern, consistent with the predominance of small samples and heterogeneous inclusion criteria.

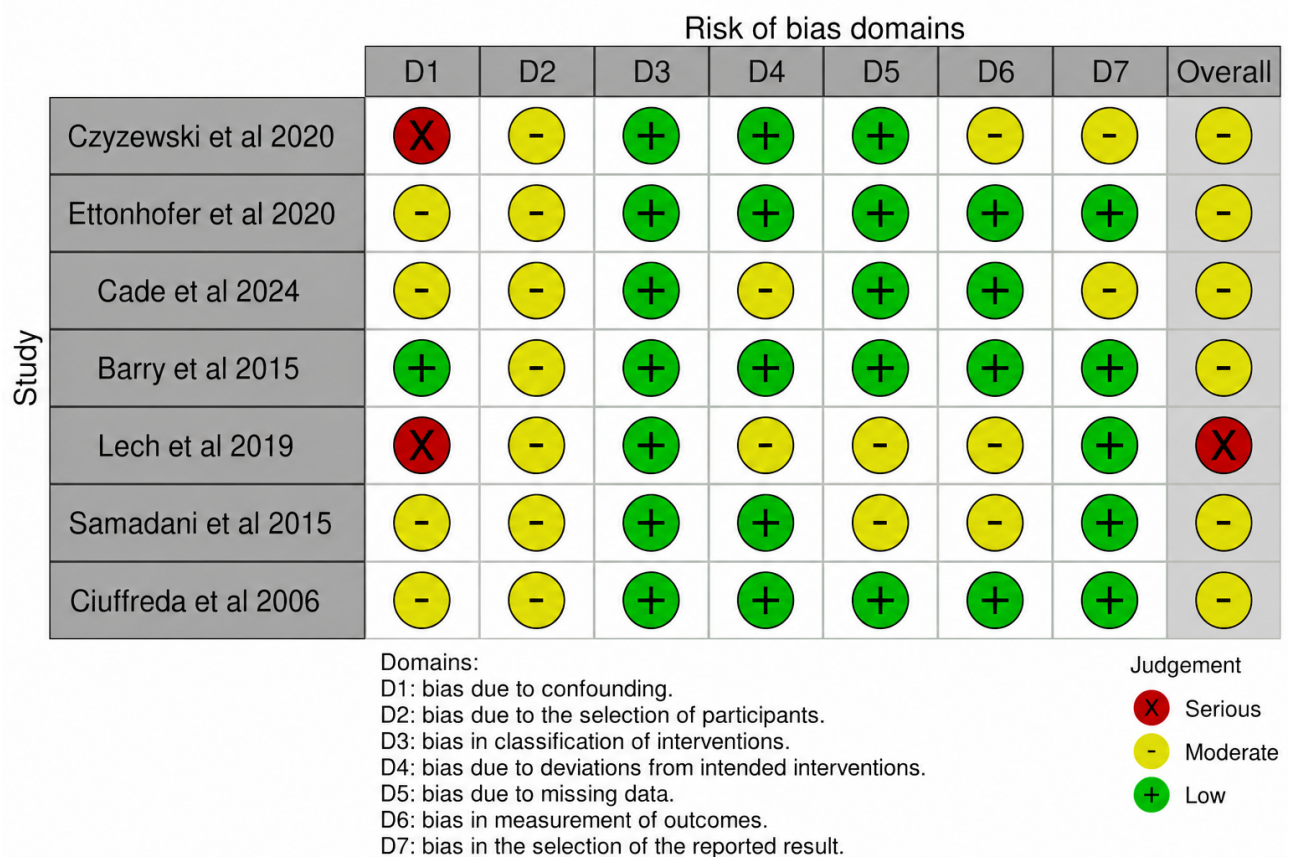
XSL•FO
RenderX

Table . Newcastle-Ottawa Scale (NOS) assessment for observational studies. NOS star ratings were applied in an adapted manner to accommodate nonrandomized observational designs commonly used in eye-tracking research. Overall quality labels are based on total stars (good: 7 - 9, fair: 5 - 6, and poor: <5).

Study	Design	Selection (0 - 4)	Comparability (0 - 2)	Exposure/outcome (0 - 3)	Total (0 - 9)	Overall quality
Astafiev et al (2015) [70]	Case-control	3	2	2	7	Good
Brown et al (2019) [63]	Case-control	3	2	2	7	Good
Clough et al (2024) [67]	Case-control	3	1	2	6	Fair
Diwakar et al (2015) [72]	Case-control	3	1	2	6	Fair
Ettenhofer et al (2018) [73]	Case-control	3	1	2	6	Fair
Heitger et al (2009) [74]	Case-control	3	1	2	6	Fair
Hershaw et al (2017) [75]	Case-control	3	1	2	6	Fair
Hougaard et al (2021) [66]	Case-control	4	1	2	7	Good
Hungerford et al (2023) [62]	Cross-sectional	3	2	2	7	Good
Kontos et al (2024) [61]	Case-control	3	1	2	6	Fair
Kraus et al (2007) [76]	Case-control	3	1	2	6	Fair
Kraus et al (2010) [77]	Case-control	3	1	2	6	Fair
Maruta et al (2016) [78]	Case-control	3	1	2	6	Fair
Maruta et al (2018) [79]	Cohort (single group)	2	0	2	4	Poor
Phillipou et al (2014) [80]	Case-control (longitudinal)	3	1	2	6	Fair
Suh et al (2006) [81]	Case-control	3	2	2	7	Good
Suh et al (2006) [82]	Case-control	3	2	2	7	Good
Thiessen et al (2017) [64]	Case-control	3	1	2	6	Fair
Ting et al (2016) [83]	Case-control	3	1	2	6	Fair

XSL•FO
RenderX

Figure 3. Cochrane Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) [65,68,69,84-86].



Certainty of evidence across outcome domains was generally low, driven by the predominance of nonrandomized designs, small samples, heterogeneity of paradigms and metrics, and limited external validation for classification analyses. Evidence was judged to be very low for most cognitive domains, with low certainty reserved for diagnostic or classification performance outcomes where classification metrics were reported but where risk-of-bias concerns and inconsistency limited confidence. Certainty ratings and primary reasons for downgrading are summarized in Table 5.

XSL•FO
RenderX

Table . Grading of Recommendations Assessment, Development and Evaluation (GRADE) certainty of evidence summary by cognitive domain and clinical use case. Certainty ratings reflect the overall body of evidence for each domain or use case. No quantitative pooling was undertaken; therefore, the summary reflects the direction and consistency of findings reported across studies. Domains may overlap because individual studies often reported multiple paradigms or outcomes.

Outcome domain or clinical use case	Evidence base (design and studies)	Summary of findings (narrative)	Certainty of evidence (GRADE[a])	Main reasons for downgrading
Executive function and inhibitory control (eg, antisaccade performance, inhibition errors, and reaction time variability)	Primarily observational (case-control or cross-sectional and classification studies; 10 studies)	Most studies reported worse inhibitory control or higher variability in ABI[b] or symptomatic mTBI[c] compared with controls, and/or associations with symptom burden; findings were not uniform across paradigms and populations.	Very low	Serious risk of bias (mostly nonrandomized and incomplete reporting), serious inconsistency (heterogeneous paradigms or metrics), and serious imprecision (small samples and limited external validation).
Working memory or cognitive load effects (eg, n-back and dual-task paradigms)	Observational (case-control) with some classification analyses (3 studies)	Across cognitive-load paradigms, oculomotor impairment tended to be more evident under higher load; evidence is based on a few studies and selected samples.	Very low	Serious imprecision (few studies and small samples) and indirectness (task-specific paradigms and selected populations), with additional concerns about risk of bias in classification analyses.
Attention and predictive control (eg, smooth pursuit, visual tracking, predictive pursuit, and target blanking paradigms)	Observational and mixed designs (case-control and longitudinal or pre-post in a minority; 10 studies)	Several studies reported altered pursuit or visual tracking or predictive control in ABI groups, particularly under conditions that increase anticipatory demands; other studies reported smaller or null differences, limiting consistency.	Very low	Serious inconsistency (mixed results across paradigms), serious imprecision (small samples), and serious indirectness due to heterogeneity in paradigms, acquisition settings, and outcome definitions.
Visuospatial exploration or neglect detection (eg, free-viewing VR[d] environments and spatial bias metrics)	Observational (1 study)	VR-based free-viewing eye tracking captured lateralized exploration patterns compatible with neglect phenotypes and may complement conventional tests; evidence is limited to a single study.	Very low	Very serious imprecision (single study) and indirectness (single setting or paradigm and limited generalizability).
Cognitive-communication and social cognition (eg, visual-world paradigms, AAC processing, and context interpretation)	Observational (case-control; 3 studies)	Eye-tracking measures suggested differences in efficiency of information processing and reduced integration of multimodal cues in ABI, but evidence is based on small, task-specific studies.	Very low	Serious imprecision (small samples), indirectness (task- and context-specific outcomes), and risk of bias related to selection and measurement.
Disorders of consciousness (DoC): command following or awareness detection using gaze-based paradigms	Noncomparative case series (2 studies)	A subset of patients achieved above-chance performance in gaze-selection tasks, suggesting potential to support bedside assessment; evidence is limited, and feasibility constraints were variably reported.	Very low	Very serious imprecision (small case series), serious risk of bias (noncomparative designs), and indirectness (variable arousal, assistive adaptations, and setting-specific implementation).
Diagnostic or classification performance of eye-tracking metrics (eg, discrimination of symptomatic mTBI vs controls and performance validity)	Diagnostic or classification studies (3 studies)	Eye-tracking–derived metrics showed discriminatory performance in specific paradigms, but thresholds, reference standards, and model validation procedures varied across studies.	Low	Risk of bias concerns (patient selection, index test conduct, and potential overfitting), inconsistency (different paradigms or thresholds), and imprecision (limited external validation).

[a]GRADE: Grading of Recommendations, Assessment, Development, and Evaluation.

[b]ABI: acquired brain injury.

XSL•FO
RenderX

[c]mTBI: mild traumatic brain injury.

[d]VR: virtual reality.

Discussion

Principal Findings

This systematic review synthesized 27 studies evaluating eye-tracking–derived metrics as a primary approach to quantify cognition or clinically relevant cognitive-communication processes after ABI. The evidence suggests that eye tracking can capture measurable differences in gaze control and visual attention that are plausibly related to executive inhibition, attentional prediction, and cognitive load. Findings were most consistently reported in chronic or persistent mTBI cohorts, while evidence for moderate-to-severe TBI, stroke, hypoxic-ischemic injury, and DoC was comparatively sparse and methodologically variable. Direct evidence of incremental validity over conventional neuropsychological testing was mixed and often indirect because comparator measures and analytic strategies varied across studies and were not consistently designed to test incremental value. Safety and feasibility reporting was inconsistent, limiting firm conclusions about tolerability and implementation at scale. Overall certainty of evidence was low across domains, reflecting small samples, heterogeneity of paradigms and outcome definitions, and limitations in internal and external validity.

Interpretation and Clinical Implications

Eye-tracking outcomes can be interpreted at 2 complementary levels that should be kept conceptually distinct in clinical translation. Some outcomes primarily index sensorimotor oculomotor control, such as conjugate gaze stability, pursuit gain, and basic tracking precision. These can be clinically informative because visual sequelae are common after TBI, and ocular motor changes may reflect injury-related disruptions within the visual system or its supporting networks [87,88]. In contrast, paradigms such as antisaccade tasks, predictive pursuit under blanking or gap conditions, and cognitive load manipulations are more tightly coupled to executive inhibition, attentional prediction, and control processes. These paradigms can provide a process-level window into cognition, but their interpretability depends on careful alignment between task demands, analytic definitions, and the intended clinical construct. This distinction is important because a technically "abnormal" eye-movement pattern does not automatically imply impairment in a higher-order cognitive domain, and conversely, a patient may report cognitive symptoms despite relatively preserved performance on a specific oculomotor task.

From a biomarker perspective, the current literature supports a cautious stance. Eye-tracking metrics show promise as candidate behavioral biomarkers, but the present evidence base is not sufficiently standardized to support broad clinical adoption or claims of routine superiority over conventional assessment. Contemporary biomarker frameworks emphasize the need for rigorous validation, harmonized measurement pipelines, and clinically meaningful thresholds before a candidate measure can be used for diagnostic decision-making or treatment monitoring [89]. In our synthesis, diagnostic and classification analyses were reported in a subset of studies, but many relied on in-sample threshold selection, nonconsecutive sampling, or reference standards that were not directly aligned with cognitive-status classification. This pattern limits confidence in generalizability and reinforces the need for external validation and predefined analytic plans when classification performance is a central claim.

The clearest near-term clinical role for eye tracking may be as an adjunct rather than a replacement. In concussion and persistent symptom cohorts, eye tracking may be helpful when the goal is to quantify specific components of visual attention and control over short time windows, particularly when repeated measures are needed, and practice effects are a concern. Naturalistic and immersive implementations may also offer clinically meaningful advantages because they can capture cognitive-motor performance under conditions closer to real-world demands. VR-based neurocognitive testing that integrates eye tracking is an example of this direction, as it can link gaze behavior to task performance within ecologically relevant contexts [90]. At the same time, practical constraints remain nontrivial. Hardware costs, space requirements, calibration procedures, head stabilization needs, and trained personnel requirements can limit feasibility, particularly outside specialized centers. These barriers are consistent with broader concerns about implementation and scalability for technology-assisted assessment approaches in neurorehabilitation.

Specificity is another recurring concern that should be made explicit. Oculomotor metrics can be influenced by psychiatric symptoms, sleep disturbance, pain, medications, and demographic factors, all of which may be common after TBI and can co-vary with cognitive complaints. These influences do not invalidate eye tracking, but they constrain interpretation unless these covariates are measured and addressed analytically. In vulnerable or socially complex cohorts, comorbidities can be substantial and may contribute to both symptom burden and performance variability, underscoring the importance of careful phenotyping and confounder control in future studies [91,92]. Interpreting eye-tracking changes as cognition-specific without considering these influences risks overattribution and may inflate perceived diagnostic specificity.

Translation to Neurorehabilitation and Future Directions

The results also inform the difference between demonstrated clinical utility and future potential. At present, most studies provide evidence of association, group discrimination, or feasibility rather than evidence that eye tracking improves clinical decisions or outcomes. Translational claims should therefore be framed as emerging. Still, several pathways appear promising. First, eye tracking may support precision rehabilitation by providing objective, high-frequency measures that can be tracked longitudinally and linked to individualized targets. This aligns with broader arguments that eye tracking can serve as a scalable behavioral signal of cognition in

XSL•FO
RenderX

neurological conditions, provided that measurement is standardized and clinically interpretable [93]. Second, methodological work on quantifying and analyzing saccades, pursuit, and fixations offers a foundation for harmonizing metrics across devices and paradigms, which is necessary for multisite comparability and eventual normative datasets [94]. Third, broader clinical applications of eye tracking in medical practice suggest that eye tracking can be integrated into clinical workflows when the technical and interpretive challenges are addressed transparently [22].

A particularly important translational direction is the development of "digital biomarker" programs in which eye tracking is evaluated alongside complementary sources of evidence, including imaging, neurophysiology, and symptom trajectories. Reviews in other neurological conditions emphasize that eye-tracking metrics can be informative but require strong attention to measurement validity, confounding, and reproducibility to be credible as biomarkers [95]. Multimodal rehabilitation technologies, including VR and brain-computer interface concepts, may further enhance ecological validity and enable adaptive assessment and training paradigms, but these approaches require careful validation in ABI populations, especially regarding feasibility and tolerability in symptomatic cohorts [96,97]. In DoC, eye-tracking and HCI approaches may assist in detecting potential command following or residual cognition in select patients, but the literature remains small and sensitive to fluctuations in arousal and implementation constraints. These concerns align with broader discussions on recovery mechanisms and prognostication in DoC and reinforce the need for cautious interpretation of single-session performance metrics [98].

Future studies should prioritize methodological features that directly address the limitations identified in this review. Prospective registration, predefined primary outcomes, standardized reporting of calibration success and data loss, and device-specific acquisition parameters should become routine. External validation of classification thresholds and replication in broader etiologic subgroups are particularly important, given that the existing evidence is heavily weighted toward chronic mTBI and selected populations. In addition, longitudinal designs are needed to clarify whether eye-tracking metrics track recovery trajectories or identify subgroups at risk for persistent impairment. This is especially relevant given growing evidence that moderate-to-severe TBI can show chronic-stage cognitive decline patterns in some cohorts, which underscores the need for sensitive, repeatable measures over extended follow-up [99]. Finally, expansion beyond TBI is warranted. Eye tracking has been used to probe language processing and lexical integration, and emerging work suggests feasibility for remote or online eye tracking in language-related disorders, which may be informative for ABI cognitive-communication assessment if appropriately validated [100,101]. Similarly, additional evidence in prolonged DoC, stroke, and systems-level implementation could support a more balanced etiologic evidence base [102,103].

Strengths and Limitations

This review provides a structured synthesis across cognitive domains and clinical use cases, incorporating design-appropriate risk-of-bias tools and an explicit certainty-of-evidence framework. The broad inclusion criteria allowed capture of a wide range of paradigms, including traditional oculomotor tasks, embedded cognitive-load assessments, VR paradigms, and HCI approaches. These strengths increase coverage and relevance across settings.

Several limitations constrain inference. First, the literature is highly heterogeneous in paradigms, outcome definitions, acquisition procedures, and analytic pipelines, which preclude quantitative pooling and limit the ability to estimate typical effect size ranges with confidence. Second, most studies were cross-sectional and underpowered, increasing vulnerability to imprecision and selective reporting. Third, the evidence base is dominated by chronic mTBI cohorts and selected populations, limiting generalizability to other ABI etiologies and to broader clinical settings. Fourth, safety and feasibility reporting was incomplete, which limits translation conclusions and obscures implementation barriers. Finally, diagnostic and classification claims require caution, as model development and thresholding were often performed within the same dataset, and reference standards varied, creating a risk of overestimated performance.

Conclusion

Eye-tracking technology shows promising potential as an adjunctive approach for quantifying specific components of visual attention and oculomotor control that are plausibly linked to cognitive processes after ABI. The current evidence is most developed for paradigms involving inhibitory control, predictive tracking, and cognitive load in mTBI cohorts, while evidence for stroke, hypoxic-ischemic injury, and DoC remains preliminary and methodologically diverse. At present, the literature does not support broad claims that eye tracking routinely exceeds conventional neuropsychological assessment. Instead, the most defensible interpretation is that eye tracking may add clinically relevant information in paradigm- and population-specific contexts, particularly when standardized acquisition, transparent reporting, and appropriate confounder assessment are in place. Future research should focus on harmonized paradigms and reporting standards, external validation of classification thresholds, longitudinal designs, and explicit feasibility and safety reporting, with a goal of defining when and for whom eye tracking provides incremental clinical value within precision neurorehabilitation pathways.

Acknowledgments

Generative artificial intelligence tools were used only for limited language-related support. Specifically, the latest available version of ChatGPT (OpenAI) was used to assist with grammar refinement, English-language editing, and paper polishing. It was not used to generate data, perform analyses, select studies, assess risk of bias, interpret findings, or verify references and

XSL•FO
RenderX

citations. All content, references, and citations were reviewed and approved by the authors, who remain fully responsible for the accuracy, originality, and integrity of the paper.

Funding

This study was supported by Current Research Funds 2026, Ministry of Health, Italy.

Authors' Contributions

AC, RDL, and FC conceptualized the study. FC, RDL, CC, and AC developed the methodology. Validation was performed by AMDN and RC. RDL, FC, CC, and AMDN conducted the investigation, while RSC, FC, and AQ provided resources. Data curation was carried out by AM and AC. AC, RDL, RC, and FC prepared the original draft of the paper. AC, RDL, and AM contributed to reviewing and editing the paper. Visualization was completed by AMDN, CC, RC, and AM. RSC, AQ, and RDL supervised the study, and project administration was managed by RSC and AQ. Funding acquisition was secured by RSC and AQ.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full search strategies.
[DOCX File, 28 KB - rehab_v13i1e81276_app1.docx]

Multimedia Appendix 2
Graphical abstract.
[PNG File, 1660 KB - rehab_v13i1e81276_app2.png]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 80 KB - rehab_v13i1e81276_app3.pdf]

References

1. Goldman L, Siddiqui EM, Khan A, et al. Understanding acquired brain injury: a review. Biomedicines 2022 Sep 2;10(9):2167. [doi: 10.3390/biomedicines10092167] [Medline: 36140268]
2. Gilmore N, Katz DI, Kiran S. Acquired brain injury in adults: a review of pathophysiology, recovery, and rehabilitation. Perspect ASHA Spec Interest Groups 2021 Aug;6(4):714-727. [doi: 10.1044/2021_persp-21-00013] [Medline: 34746412]
3. Walsh RS, Fortune DG, Gallagher S, Muldoon OT. Acquired brain injury: combining social psychological and neuropsychological perspectives. Health Psychol Rev 2014;8(4):458-472. [doi: 10.1080/17437199.2012.733914] [Medline: 25211210]
4. Capizzi A, Woo J, Verduzco-Gutierrez M. Traumatic brain injury: an overview of epidemiology, pathophysiology, and medical management. Med Clin North Am 2020 Mar;104(2):213-238. [doi: 10.1016/j.mcna.2019.11.001] [Medline: 32035565]
5. Mckee AC, Daneshvar DH. The neuropathology of traumatic brain injury. Handb Clin Neurol 2015;127:45-66. [doi: 10.1016/B978-0-444-52892-6.00004-0] [Medline: 25702209]
6. Irastorza-Valera L, Soria-Gómez E, Benitez JM, Montáns FJ, Saucedo-Mora L. Review of the brain's behaviour after injury and disease for its application in an agent-based model (ABM). Biomimetics (Basel) 2024 Jun 14;9(6):362. [doi: 10.3390/biomimetics9060362] [Medline: 38921242]
7. Karaboue MAA, Ministeri F, Sessa F, et al. Traumatic brain injury as a public health issue: epidemiology, prognostic factors and useful data from forensic practice. Healthcare (Basel) 2024 Nov 13;12(22):2266. [doi: 10.3390/healthcare12222266] [Medline: 39595464]
8. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Care Services. Traumatic Brain Injury: A Roadmap for Accelerating Progress: National Academies Press (US); 2022. URL: https://www.ncbi.nlm.nih.gov/books/NBK580076/ [accessed 2026-05-05]
9. Dewan MC, Rattani A, Gupta S, et al. Estimating the global incidence of traumatic brain injury. J Neurosurg 2018;130(4):1080-1097. [doi: 10.3171/2017.10.JNS17352]
10. Mena JH, Sanchez AI, Rubiano AM, et al. Effect of the modified Glasgow Coma Scale score criteria for mild traumatic brain injury on mortality prediction: comparing classic and modified Glasgow Coma Scale score model scores of 13. J Trauma 2011 Nov;71(5):1185-1192. [doi: 10.1097/TA.0b013e31823321f8] [Medline: 22071923]

XSL•FO
RenderX

11. Cantu RC. Posttraumatic retrograde and anterograde amnesia: pathophysiology and implications in grading and safe return to play. J Athl Train 2001 Sep;36(3):244-248. [Medline: 12937491]
12. Kaurani P, Moreira de Marchi Apolaro AV, Kunchala K, et al. Advances in neurorehabilitation: strategies and outcomes for traumatic brain injury recovery. Cureus 2024 Jun;16(6):e62242. [doi: 10.7759/cureus.62242] [Medline: 39006616]
13. Rabinowitz AR, Levin HS. Cognitive sequelae of traumatic brain injury. Psychiatr Clin North Am 2014 Mar;37(1):1-11. [doi: 10.1016/j.psc.2013.11.004] [Medline: 24529420]
14. Corrigan F, Wee IC, Collins-Praino LE. Chronic motor performance following different traumatic brain injury severity-a systematic review. Front Neurol 2023;14:1180353. [doi: 10.3389/fneur.2023.1180353] [Medline: 37288069]
15. Walker WC, Pickett TC. Motor impairment after severe traumatic brain injury: a longitudinal multicenter study. J Rehabil Res Dev 2007;44(7):975-982. [doi: 10.1682/jrrd.2006.12.0158] [Medline: 18075954]
16. Williams G, Morris ME, Schache A, McCrory PR. Incidence of gait abnormalities after traumatic brain injury. Arch Phys Med Rehabil 2009 Apr;90(4):587-593. [doi: 10.1016/j.apmr.2008.10.013] [Medline: 19345773]
17. Andrei D, Mederle AL, Ghenciu LA, Borza C, Faur AC. Efficacy of neurorehabilitation approaches in traumatic brain injury patients: a comprehensive review. Life (Basel) 2025 Mar 20;15(3):503. [doi: 10.3390/life15030503] [Medline: 40141847]
18. Pisano F, Rizzo G, Bilotta F. Cognitive rehabilitation for patients with acquired brain injury. Curr Opin Anaesthesiol 2025;38(5):591-597. [doi: 10.1097/ACO.0000000000001520]
19. Bulle-Smid L, Keuning W, Van Den Heuvel R, et al. The use of extended reality in rehabilitation for patients with acquired brain injury: a scoping review. Stud Health Technol Inform 2023 Aug 23;306:583-590. [doi: 10.3233/SHTI230682] [Medline: 37638965]
20. Munoz DP, Everling S. Look away: the anti-saccade task and the voluntary control of eye movement. Nat Rev Neurosci 2004 Mar;5(3):218-228. [doi: 10.1038/nrn1345] [Medline: 14976521]
21. Bilbao C, Carrera A, Otin S, Piñero DP. Eye tracking-based characterization of fixations during reading in children with neurodevelopmental disorders. Brain Sci 2024 Jul 26;14(8):750. [doi: 10.3390/brainsci14080750] [Medline: 39199445]
22. Tahri Sqalli M, Aslonov B, Gafurov M, Mukhammadiev N, Sqalli Houssaini Y. Eye tracking technology in medical practice: a perspective on its diverse applications. Front Med Technol 2023;5:1253001. [doi: 10.3389/fmedt.2023.1253001] [Medline: 38045887]
23. Lai ML, Tsai MJ, Yang FY, et al. A review of using eye-tracking technology in exploring learning from 2000 to 2012. Educ Res Rev 2013 Dec;10:90-115. [doi: 10.1016/j.edurev.2013.10.001]
24. Carter BT, Luke SG. Best practices in eye tracking research. Int J Psychophysiol 2020 Sep;155:49-62. [doi: 10.1016/j.ijpsycho.2020.05.010] [Medline: 32504653]
25. Eckstein MK, Guerra-Carrillo B, Miller Singley AT, Bunge SA. Beyond eye gaze: what else can eyetracking reveal about cognition and cognitive development? Dev Cogn Neurosci 2017 Jun;25:69-91. [doi: 10.1016/j.dcn.2016.11.001] [Medline: 27908561]
26. Wilkinson KM, Mitchell T. Eye tracking research to answer questions about augmentative and alternative communication assessment and intervention. Augment Altern Commun 2014 Jun;30(2):106-119. [doi: 10.3109/07434618.2014.904435] [Medline: 24758526]
27. Tangsgaard Hvelplund K. Eye tracking and the translation process: reflections on the analysis and interpretation of eye-tracking data. MonTI Monogr Transl Interpret 2014;ne1:201-223. [doi: 10.6035/MonTI.2014.ne1.6]
28. Adhanom IB, MacNeilage P, Folmer E. Eye tracking in virtual reality: a broad review of applications and challenges. Virtual Real 2023 Jun;27(2):1481-1505. [doi: 10.1007/s10055-022-00738-z] [Medline: 37621305]
29. Zhu L, Chen J, Yang H, et al. Wearable near-eye tracking technologies for health: a review. Bioengineering (Basel) 2024 Jul 22;11(7):738. [doi: 10.3390/bioengineering11070738] [Medline: 39061820]
30. Pastel S, Marlok J, Bandow N, Witte K. Application of eye-tracking systems integrated into immersive virtual reality and possible transfer to the sports sector - a systematic review. Multimed Tools Appl 2023 Jan;82(3):4181-4208. [doi: 10.1007/s11042-022-13474-y]
31. Menéndez González N, Bozkir E. Eye-tracking devices for virtual and augmented reality metaverse environments and their compatibility with the European Union General Data Protection Regulation. Digital Society 2024 Aug;3(2):39. [doi: 10.1007/s44206-024-00128-9]
32. Mahanama B, Jayawardana Y, Rengarajan S, et al. Eye movement and pupil measures: a review. Front Comput Sci 2021;3. [doi: 10.3389/fcomp.2021.733531]
33. Hunfalvay M, Bolte T, Singh A, Greenstein E, Murray NP, Carrick FR. Age-based developmental biomarkers in eye movements: a retrospective analysis using machine learning. Brain Sci 2024 Jul 9;14(7):686. [doi: 10.3390/brainsci14070686] [Medline: 39061426]
34. Tao L, Wang Q, Liu D, Wang J, Zhu Z, Feng L. Eye tracking metrics to screen and assess cognitive impairment in patients with neurological disorders. Neurol Sci 2020 Jul;41(7):1697-1704. [doi: 10.1007/s10072-020-04310-y]
35. Readman MR, Polden M, Gibbs MC, Wareing L, Crawford TJ. The potential of naturalistic eye movement tasks in the diagnosis of Alzheimer's disease: a review. Brain Sci 2021 Nov 12;11(11):1503. [doi: 10.3390/brainsci11111503] [Medline: 34827502]

XSL•FO
RenderX

36. Pearson HC, MacKenzie DE, Oystreck DT, Westwood DA. Using eye-tracking technology to measure cognitive function in mild traumatic brain injury: a scoping review. Journal of Concussion 2023 Jan;7:20597002231194151. [doi: 10.1177/20597002231194151]
37. Zhang J, Kong W, Ma M, Yang X, Li W, Song A. A review of eye-tracking technology and its application in stroke diagnosis and assessment. Measurement (Lond) 2025 Aug;252:117325. [doi: 10.1016/j.measurement.2025.117325]
38. Lencer R, Trillenberg P. Neurophysiology and neuroanatomy of smooth pursuit in humans. Brain Cogn 2008 Dec;68(3):219-228. [doi: 10.1016/j.bandc.2008.08.013] [Medline: 18835076]
39. Kowler E, Rubinstein JF, Santos EM, Wang J. Predictive smooth pursuit eye movements. Annu Rev Vis Sci 2019 Sep 15;5:223-246. [doi: 10.1146/annurev-vision-091718-014901] [Medline: 31283450]
40. Torregrossa W, Torrisi M, De Luca R, et al. Neuropsychological assessment in patients with traumatic brain injury: a comprehensive review with clinical recommendations. Biomedicines 2023 Jul 14;11(7):1991. [doi: 10.3390/biomedicines11071991] [Medline: 37509630]
41. Wolf A, Ueda K. Contribution of eye-tracking to study cognitive impairments among clinical populations. Front Psychol 2021;12:590986. [doi: 10.3389/fpsyg.2021.590986] [Medline: 34163391]
42. Duprez J, Houvenaghel JF, Naudet F, et al. Evaluating cognitive action control using eye-movement analysis: an oculomotor adaptation of the Simon task. Front Hum Neurosci 2016;10:84. [doi: 10.3389/fnhum.2016.00084] [Medline: 26973499]
43. Aljehane SD, Sharif B, Maletic JI. Studying developer eye movements to measure cognitive workload and visual effort for expertise assessment. Proceedings of the ACM on Human-Computer Interaction (PACMHCI) 2023 May 17;7(ETRA):1-18. [doi: 10.1145/3591135]
44. González-Vides L, Hernández-Verdejo JL, Cañadas-Suárez P. Eye tracking in optometry: a systematic review. J Eye Mov Res 2023;16(3):3. [doi: 10.16910/jemr.16.3.3] [Medline: 38111688]
45. McDonald MA, Holdsworth SJ, Danesh-Meyer HV. Eye movements in mild traumatic brain injury: ocular biomarkers. J Eye Mov Res 2022;15(2):4. [doi: 10.16910/jemr.15.2.4] [Medline: 36439911]
46. Leharanger M, Liu P, Vandromme L, Balédent O. Eye tracking post processing to detect visual artifacts and quantify visual attention under cognitive task activity during fMRI. Sensors (Basel) 2024 Jul 29;24(15):4916. [doi: 10.3390/s24154916] [Medline: 39123963]
47. Johansen SA, San Agustin J, Skovsgaard H, Hansen JP, Tall M. Low cost vs. high-end eye tracking for usability testing. Presented at: ACM CHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC p. 1177-1182. [doi: 10.1145/1979742.1979744]
48. Gibaldi A, Vanegas M, Bex PJ, Maiello G. Evaluation of the Tobii EyeX eye tracking controller and Matlab toolkit for research. Behav Res Methods 2017 Jun;49(3):923-946. [doi: 10.3758/s13428-016-0762-9] [Medline: 27401169]
49. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
50. Page MJ, Shamseer L, Tricco AC. Registration of systematic reviews in PROSPERO: 30,000 records and counting. Syst Rev 2018 Feb 20;7(1):32. [doi: 10.1186/s13643-018-0699-4] [Medline: 29463298]
51. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282. [Medline: 23092060]
52. Li M, Gao Q, Yu T. Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters. BMC Cancer 2023;23(1):799. [doi: 10.1186/s12885-023-11325-z]
53. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898. [doi: 10.1136/bmj.l4898]
54. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919. [doi: 10.1136/bmj.i4919]
55. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011 Oct 18;155(8):529-536. [doi: 10.7326/0003-4819-155-8-201110180-00009] [Medline: 22007046]
56. Gualdi-Russo E, Zaccagni L. The Newcastle–Ottawa Scale for assessing the quality of studies in systematic reviews. Publications 2026;14(1):4. [doi: 10.3390/publications14010004]
57. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. J Health Serv Res Policy 2005 Jan;10(1):45-53. [doi: 10.1177/135581960501000110] [Medline: 15667704]
58. Bezerra CT, Grande AJ, Galvão VK, Santos D, Atallah Á, Silva V. Assessment of the strength of recommendation and quality of evidence: GRADE checklist. A descriptive study. Sao Paulo Med J 2022;140(6):829-836. [doi: 10.1590/1516-3180.2022.0043.R1.07042022] [Medline: 36102459]
59. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926. [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]
60. Cade AE, Turnbull PRK. Effect of chiropractic intervention on oculomotor and attentional visual outcomes in young adults with long-term mild traumatic brain injury: a randomized controlled trial. J Manipulative Physiol Ther 2024;47(1-4):1-11. [doi: 10.1016/j.jmpt.2024.08.003] [Medline: 39503641]

XSL•FO
RenderX

61. Kontos AP, Zynda AJ, Minerbi A. Comparison of vestibular/ocular motor screening (VOMS) and computerized eye-tracking to identify exposure to repetitive head impacts. Mil Med 2024 Nov 5;189(11-12):2291-2297. [doi: 10.1093/milmed/usae065] [Medline: 38531077]
62. Hungerford L, Agtarap S, Ettenhofer M. Impact of depression and post-traumatic stress on manual and oculomotor performance in service members with a history of mild TBI. Brain Inj 2023 Jul 3;37(8):680-688. [doi: 10.1080/02699052.2023.2210293] [Medline: 37204183]
63. Brown J, Thiessen A, Freeland T, Brewer CH. Visual processing patterns of adults with traumatic brain injury when viewing image-based grids and visual scenes. Augment Altern Commun 2019 Sep;35(3):229-239. [doi: 10.1080/07434618.2019.1609578] [Medline: 31257929]
64. Thiessen A, Brown J, Beukelman D, Hux K. The effect of human engagement depicted in contextual photographs on the visual attention patterns of adults with traumatic brain injury. J Commun Disord 2017 Sep;69:58-71. [doi: 10.1016/j.jcomdis.2017.07.001] [Medline: 28783543]
65. Czyżewski A, Kurowski A, Odya P, Szczuko P. Multifactor consciousness level assessment of participants with acquired brain injuries employing human-computer interfaces. Biomed Eng Online 2020 Jan 10;19(1):2. [doi: 10.1186/s12938-019-0746-y] [Medline: 31924202]
66. Hougaard BI, Knoche H, Jensen J, Evald L. Spatial neglect midline diagnostics from virtual reality and eye tracking in a free-viewing environment. Front Psychol 2021;12:742445. [doi: 10.3389/fpsyg.2021.742445] [Medline: 34912268]
67. Clough S, Brown-Schmidt S, Cho SJ, Duff MC. Reduced on-line speech gesture integration during multimodal language processing in adults with moderate-severe traumatic brain injury: evidence from eye-tracking. Cortex 2024 Dec;181:26-46. [doi: 10.1016/j.cortex.2024.08.008] [Medline: 39488986]
68. Ettenhofer ML, Gimbel SI, Cordero E. Clinical validation of an optimized multimodal neurocognitive assessment of chronic mild TBI. Ann Clin Transl Neurol 2020 Apr;7(4):507-516. [doi: 10.1002/acn3.51020] [Medline: 32207241]
69. Cade AE, Turnbull PRK. Cervical spine vibration modifies oculomotor function in young adults with traumatic brain injury. J Manipulative Physiol Ther 2024;47(1-4):96-105. [doi: 10.1016/j.jmpt.2024.08.001] [Medline: 39412450]
70. Astafiev SV, Shulman GL, Metcalf NV, et al. Abnormal white matter blood-oxygen-level-dependent signals in chronic mild traumatic brain injury. J Neurotrauma 2015 Aug 15;32(16):1254-1271. [doi: 10.1089/neu.2014.3547] [Medline: 25758167]
71. Barry DM, Ettenhofer ML. Assessment of performance validity using embedded saccadic and manual indices on a continuous performance test. Arch Clin Neuropsychol 2016 Dec 1;31(8):963-975. [doi: 10.1093/arclin/acw070] [Medline: 27625047]
72. Diwakar M, Harrington DL, Maruta J, et al. Filling in the gaps: anticipatory control of eye movements in chronic mild traumatic brain injury. Neuroimage Clin 2015;8:210-223. [doi: 10.1016/j.nicl.2015.04.011] [Medline: 26106545]
73. Ettenhofer ML, Hershaw JN, Engle JR, Hungerford LD. Saccadic impairment in chronic traumatic brain injury: examining the influence of cognitive load and injury severity. Brain Inj 2018 Dec 6;32(13-14):1740-1748. [doi: 10.1080/02699052.2018.1511067]
74. Heitger MH, Jones RD, Macleod AD, Snell DL, Frampton CM, Anderson TJ. Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. Brain 2009 Oct;132(Pt 10):2850-2870. [doi: 10.1093/brain/awp181] [Medline: 19617197]
75. Hershaw JN, Barry DM, Ettenhofer ML. Increased risk for age-related impairment in visual attention associated with mild traumatic brain injury: evidence from saccadic response times. PLoS One 2017;12(2):e0171752. [doi: 10.1371/journal.pone.0171752] [Medline: 28166259]
76. Kraus MF, Little DM, Donnell AJ, Reilly JL, Simonian N, Sweeney JA. Oculomotor function in chronic traumatic brain injury. Cogn Behav Neurol 2007 Sep;20(3):170-178. [doi: 10.1097/WNN.0b013e318142badb] [Medline: 17846516]
77. Kraus MF, Little DM, Wojtowicz SM, Sweeney JA. Procedural learning impairments identified via predictive saccades in chronic traumatic brain injury. Cogn Behav Neurol 2010 Dec;23(4):210-217. [doi: 10.1097/WNN.0b013e3181cefe2e] [Medline: 21150346]
78. Maruta J, Palacios EM, Zimmerman RD, Ghajar J, Mukherjee P. Chronic post-concussion neurocognitive deficits. I. Relationship with white matter integrity. Front Hum Neurosci 2016;10:35. [doi: 10.3389/fnhum.2016.00035] [Medline: 26903842]
79. Maruta J, Spielman LA, Rajashekar U, Ghajar J. Association of visual tracking metrics with post-concussion symptomatology. Front Neurol 2018;9:611. [doi: 10.3389/fneur.2018.00611] [Medline: 30093880]
80. Phillipou A, Douglas J, Krieser D, Ayton L, Abel L. Changes in saccadic eye movement and memory function after mild closed head injury in children. Dev Med Child Neurol 2014 Apr;56(4):337-345. [doi: 10.1111/dmcn.12345] [Medline: 24350895]
81. Suh M, Kolster R, Sarkar R, McCandliss B, Ghajar J. Deficits in predictive smooth pursuit after mild traumatic brain injury. Neurosci Lett 2006 Jun;401(1-2):108-113. [doi: 10.1016/j.neulet.2006.02.074]
82. Suh M, Basu S, Kolster R, Sarkar R, McCandliss B, Ghajar J. Increased oculomotor deficits during target blanking as an indicator of mild traumatic brain injury. Neurosci Lett 2006 Dec;410(3):203-207. [doi: 10.1016/j.neulet.2006.10.001]

XSL•FO
RenderX

83. Ting WKC, Schweizer TA, Topolovec-Vranic J, Cusimano MD. Antisaccadic eye movements are correlated with corpus callosum white matter mean diffusivity, Stroop performance, and symptom burden in mild traumatic brain injury and concussion. Front Neurol 2015;6:271. [doi: 10.3389/fneur.2015.00271] [Medline: 26834693]
84. Lech M, Kucewicz MT, Czyżewski A. Human computer interface for tracking eye movements improves assessment and diagnosis of patients with acquired brain injuries. Front Neurol 2019;10:6. [doi: 10.3389/fneur.2019.00006] [Medline: 30728799]
85. Samadani U, Ritlop R, Reyes M, et al. Eye tracking detects disconjugate eye movements associated with structural traumatic brain injury and concussion. J Neurotrauma 2015 Apr 15;32(8):548-556. [doi: 10.1089/neu.2014.3687] [Medline: 25582436]
86. Ciuffreda KJ, Han Y, Kapoor N, Ficarra AP. Oculomotor rehabilitation for reading in acquired brain injury. NeuroRehabilitation 2006;21(1):9-21. [Medline: 16720933]
87. Rauchman SH, Zubair A, Jacob B, et al. Traumatic brain injury: mechanisms, manifestations, and visual sequelae. Front Neurosci 2023;17:1090672. [doi: 10.3389/fnins.2023.1090672] [Medline: 36908792]
88. Rauchman SH, Albert J, Pinkhasov A, Reiss AB. Mild-to-moderate traumatic brain injury: a review with focus on the visual system. Neurol Int 2022 May 30;14(2):453-470. [doi: 10.3390/neurolint14020038] [Medline: 35736619]
89. Wilde EA, Wanner IB, Kenney K, et al. A framework to advance biomarker development in the diagnosis, outcome prediction, and treatment of traumatic brain injury. J Neurotrauma 2022 Apr;39(7-8):436-457. [doi: 10.1089/neu.2021.0099] [Medline: 35057637]
90. Wilf M, Korakin A, Bahat Y, et al. Using virtual reality-based neurocognitive testing and eye tracking to study naturalistic cognitive-motor performance. Neuropsychologia 2024 Feb;194:108744. [doi: 10.1016/j.neuropsychologia.2023.108744]
91. Topolovec-Vranic J, Schuler A, Gozdzik A, et al. The high burden of traumatic brain injury and comorbidities amongst homeless adults with mental illness. J Psychiatr Res 2017 Apr;87:53-60. [doi: 10.1016/j.jpsychires.2016.12.004]
92. Kumar RG, Ketchum JM, Corrigan JD, Hammond FM, Sevigny M, Dams-O'Connor K. The longitudinal effects of comorbid health burden on functional outcomes for adults with moderate to severe traumatic brain injury. J Head Trauma Rehabil 2020;35(4):E372-E381. [doi: 10.1097/HTR.0000000000000572] [Medline: 32108718]
93. Bueno APA, Sato JR, Hornberger M. Eye tracking – the overlooked method to measure cognition in neurodegeneration? Neuropsychologia 2019 Oct;133:107191. [doi: 10.1016/j.neuropsychologia.2019.107191] [Medline: 31521634]
94. DiCesare CA, Kiefer AW, Nalepka P, Myer GD. Quantification and analysis of saccadic and smooth pursuit eye movements and fixations to detect oculomotor deficits. Behav Res Methods 2017 Feb;49(1):258-266. [doi: 10.3758/s13428-015-0693-x] [Medline: 26705117]
95. Cecchetti S, Duchowski AT, Cavallo M. Eye-tracking metrics as a digital biomarker for neurocognitive disorders in multiple sclerosis: a scoping review. Brain Sci 2025 Jan 31;15(2):149. [doi: 10.3390/brainsci15020149] [Medline: 40002482]
96. Drigas A, Sideraki A. Brain neuroplasticity leveraging virtual reality and brain-computer interface technologies. Sensors (Basel) 2024 Sep 3;24(17):5725. [doi: 10.3390/s24175725] [Medline: 39275636]
97. Georgiev DD, Georgieva I, Gong Z, Nanjappan V, Georgiev GV. Virtual reality for neurorehabilitation and cognitive enhancement. Brain Sci 2021 Feb 11;11(2):221. [doi: 10.3390/brainsci11020221] [Medline: 33670277]
98. Edlow BL, Claassen J, Schiff ND, Greer DM. Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies. Nat Rev Neurol 2021 Mar;17(3):135-156. [doi: 10.1038/s41582-020-00428-x] [Medline: 33318675]
99. Green REA, Dabek MK, Changoor A, Rybkina J, Monette GA, Colella B. Moderate-severe TBI as a progressive disorder: patterns and predictors of cognitive declines in the chronic stages of injury. Neurorehabil Neural Repair 2023 Dec;37(11-12):799-809. [doi: 10.1177/15459683231212861] [Medline: 37990972]
100. Mack JE, Ji W, Thompson CK. Effects of verb meaning on lexical integration in agrammatic aphasia: evidence from eyetracking. J Neurolinguistics 2013 Nov;26(6):619-636. [doi: 10.1016/j.jneuroling.2013.04.002] [Medline: 24092952]
101. van Boxtel WS, Linge M, Manning R, Haven LN, Lee J. Online eye tracking for aphasia: a feasibility study comparing web and lab tracking and implications for clinical use. Brain Behav 2024 Nov;14(11):e70112. [doi: 10.1002/brb3.70112] [Medline: 39469815]
102. Johansson J, Möller M, Franzon K, Stenberg J, Godbolt AK. Eye tracking to support assessment of patients with prolonged disorder of consciousness - a case series. J Rehabil Med 2025 Jan 3;57:jrm41324. [doi: 10.2340/jrm.v57.41324] [Medline: 39749424]
103. French MA, Daley K, Lavezza A, et al. A learning health system infrastructure for precision rehabilitation after stroke. Am J Phys Med Rehabil 2023 Feb 1;102(2S Suppl 1):S56-S60. [doi: 10.1097/PHM.0000000000002138] [Medline: 36634332]

Abbreviations

ABI: acquired brain injury
DoC: disorders of consciousness
HCI: human-computer interface
mTBI: mild traumatic brain injury
NOS: Newcastle-Ottawa Scale
nTBI: nontraumatic brain injury

XSL•FO
RenderX

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies 2
TBI: traumatic brain injury
VR: virtual reality

Edited by S Munce; submitted 25.Jul.2025; peer-reviewed by F Pisano, M Bucur, MM Khan; revised version received 01.Mar.2026; accepted 12.Mar.2026; published 02.Jun.2026.

Please cite as:
Calderone A, De Luca R, Corallo F, Calapai R, Mirabile A, Quartarone A, De Nunzio AM, Casella C, Calabrò RS
Eye-Tracking Technologies for Cognitive Assessment After Acquired Brain Injury: Systematic Review
JMIR Rehabil Assist Technol 2026;13:e81276
URL: https://rehab.jmir.org/2026/1/e81276
doi:10.2196/81276

© Andrea Calderone, Rosaria De Luca, Francesco Corallo, Rosalia Calapai, Alessio Mirabile, Angelo Quartarone, Alessandro Marco De Nunzio, Carmela Casella, Rocco Salvatore Calabrò. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 2.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient-Perceived Factors Influencing Physical Activity Sensor Use in Stroke Prevention and Rehabilitation: Systematic Review of Qualitative Studies Using Thematic Synthesis

Paul Harris[1], MClinRes; Ingrid Maine[1,2], BAppSc (OT), MClinRes

[1]Department of Medical Education, The University of Melbourne, Grattan Street, Parkville, Melbourne, Victoria, Australia

[2]Division of Chronic and Complex Care, Western Health, St Albans, Melbourne, Victoria, Australia

Corresponding Author:
Paul Harris, MClinRes
Department of Medical Education, The University of Melbourne, Grattan Street, Parkville, Melbourne, Victoria, Australia

Abstract

Background: A robust correlation exists between physical activity (PA) and stroke risk reduction, and wearable PA sensors have emerged as promising adjuncts for rehabilitation and risk self-management. However, evidence regarding their long-term efficacy in facilitating sustained behavioral change remains equivocal.

Objective: This study aimed to explore the barriers and facilitators influencing the effective use of PA sensors in older adults at risk of primary or recurrent stroke.

Methods: A systematic search and thematic synthesis of qualitative research was conducted, focusing on PA sensor use among older adults at risk of primary or recurrent stroke. Given the emerging qualitative evidence base, inclusion was extended to proxy populations with analogous cardiovascular risk and functional profiles. Data were analyzed using line-by-line coding of primary text to generate descriptive themes and synthesize overarching analytical constructs.

Results: A search of 6 bibliographic databases (January 2010 to December 2023) identified 18 eligible studies. This systematic review and thematic synthesis revealed predominantly technological (user experience and device attributes) and psychological (motivation) barriers. Key facilitators were psychological (feedback and motivation), technological (user experience), and social/environmental supports. Higher-level analysis revealed a critical interrelationship between effective user engagement and optimally assistive device characteristics.

Conclusions: This review revealed a synergistic user-device interaction driving sustained PA in older adults at risk of primary or recurrent stroke. Future interventions developed in collaboration with patients and informed by the factors identified by this study will improve participation in rehabilitation and functional outcomes in this population.

International Registered Report Identifier (IRRID): RR2-10.1371/journal.pone.0301983

(JMIR Rehabil Assist Technol 2026;13:e86915) doi:10.2196/86915

KEYWORDS

sensors; stroke; physical activity; rehabilitation; qualitative research

Introduction

The terms “physical activity” (PA) and “exercise,” although often used interchangeably, are distinct concepts. PA broadly refers to any bodily movement resulting in energy expenditure, while exercise is more precisely defined as a subcomponent of PA that is “planned, structured, and... done to improve or maintain one or more components of physical fitness” [1]. Within this context, physical fitness encompasses the attributes and abilities an individual possesses or acquires to perform PA or exercise. Physical inactivity, or sedentary behavior, is a significant global risk factor for mortality associated with numerous adverse health outcomes including stroke [2].

As with other cardiovascular diseases, the pathogenesis of stroke is influenced by a range of risk factors, broadly categorized into nonmodifiable, medically modifiable, and lifestyle-modifiable attributes [2]. Nonmodifiable factors include inherent characteristics such as age and genetic predisposition. Medically modifiable factors are amenable to pharmacological or surgical interventions. Lifestyle-modifiable (behavioral) factors are those that can be altered through changes to daily habits and practices. Hypertension, for example, the most significant modifiable risk factor for stroke, can be effectively managed through both medical and lifestyle interventions, including pharmacotherapy, PA, dietary changes, and weight management. Despite the established efficacy of these interventions, one study reported that fewer than 50% of patients had their stroke or cardiovascular disease risk factors appropriately assessed,

XSL•FO
RenderX

treated, or managed [3]. Similarly, a recent audit of stroke rehabilitation services in Australia revealed that a comparable proportion of patients received information on self-management strategies or support programs [3]. PA is a critical, accessible, and cost-effective means of mitigating underlying risk factors, such as hypertension, contributing to the primary prevention of initial stroke events and the secondary prevention of recurrent stroke [4,5].

A growing body of evidence suggests that increased PA confers substantial health benefits across all age groups [6]. Notably, reduced sedentary time demonstrably lowers health risks and mortality, even when accounting for activity levels and established risk factors. Regular PA also mitigates the progression of hypertension and has been linked to improved weight management, reduced adiposity, and a decreased incidence of type 2 diabetes mellitus—all major risk factors for stroke. While increased PA significantly improves walking speed and balance, the evidence regarding the impact on poststroke mortality, dependence, and long-term disability remains insufficient or unclear [7]. A survey conducted in the United States revealed that a significant majority (71%) of older adults, 65 years and older, engage in self-monitoring for health indicators such as weight, diet, or exercise [8]. However, despite the prevalence of digital health tools, only a small minority (12%) utilize PA sensors and companion mobile apps. The predominant method for tracking health measures in this group was pen-and-paper [8]. The limited adoption of technology for health monitoring among this demographic may be attributable to a combination of objective factors such as PA sensor accuracy and subjective considerations such as perceived usability, which may influence engagement levels among older adults.

The need for increasingly sophisticated and persuasive strategies to motivate sustained healthy behavior remains a central challenge for clinicians designing interventions and engineers developing wearable PA sensors. Remarkably, few qualitative studies have thoroughly investigated the use of PA sensors in older adults, specifically exploring reasons for nonadoption and discontinuation. A deeper understanding of these factors is essential for the effective design and implementation of PA sensor-based interventions for stroke prevention and rehabilitation [9]. With that in mind, the current study aimed to explore the barriers and facilitators influencing the effective use of PA sensors in older adults at risk of primary or recurrent stroke.

Methods

Methodological Framework

This systematic review and thematic synthesis was conducted and reported according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Checklist 1) and the JBI (Joanna Briggs Institute) framework for evidence synthesis. Study quality was evaluated using JBI critical appraisal tools. Review methods were established a priori, as detailed in the published protocol [10]. The inclusion of proxy groups, such as individuals with hypertension or age-related motor deficits, is justified by their functional and psychological alignment with stroke survivors. These cohorts experience near-identical ergonomic barriers, such as the manual dexterity required for small device interfaces, and share psychological burdens, including diminished self-efficacy due to device failure, for example. By prioritizing the phenomenology of functional experience over narrow neurological diagnoses, this study applies conceptual transferability to enhance thematic depth where disease-specific evidence is still emerging [11-13]. The primary objective of the study was to identify and synthesize patient-perceived factors acting as either barriers or facilitators, influencing the use of wearable PA sensors in the context of stroke prevention and rehabilitation for older adults.

Inclusion and Exclusion Criteria

Peer-reviewed articles published in English between January 1, 2010, and December 31, 2023, that had a primary focus on wearable PA sensor use were included in the analysis. Studies were required to have a qualitative design or a significant qualitative component. The target population included older stroke survivors and populations with analogous clinical profiles. Eligible participants were characterized by shared cardiovascular risks (eg, atrial fibrillation, hypertension), metabolic conditions (eg, type 2 diabetes, dyslipidemia), and lifestyle-related factors. Inclusion was further defined by stroke-related functional deficits, specifically, impaired physical mobility (gait and balance), cognitive decline (executive dysfunction), and reduced independence in activities of daily living [14]. Study participants were stratified into 3 age cohorts to reflect distinct clinical and life-stage profiles. The 50‐ to 64-year age group (“Young-Old”) was included to capture the critical transition period where stroke risk doubles, and this allowed for an analysis of whether PA sensor–based interventions performed differently in individuals with higher neuroplasticity and higher work/life stress [2,15]. The 65‐ to 84-year group (“Old-Old”) represents the standard geriatric stroke demographic. Finally, those aged 85+ years (“Oldest-Old”) were analyzed separately to ensure the unique frailty and comorbid complexities of the oldest-old did not skew the results of the younger cohorts.

Search Strategy

The methodology for this synthesis followed the prepublished protocol [10]. A high-sensitivity search strategy was conducted across MEDLINE, EMBASE, CINAHL, Cochrane, PsycINFO, and Scopus using MeSH and keywords related to wearable PA sensors and exercise. To maximize recall within an emerging qualitative evidence base, population-specific filters (eg, “stroke”) were omitted; instead, the search was broadened to include all studies of populations with analogous risk and functional profiles. These results were supplemented by manual bibliography scans of included studies. This approach ensured a comprehensive capture of qualitative data despite the paucity of stroke-specific literature.

Screening and Data Extraction

Following the prepublished protocol, publication titles and abstracts were independently screened by two reviewers (IM and PH), with a pilot test conducted to ensure interrater reliability. Disagreements were resolved through consensus or third-party adjudication. Full-text screening and deduplication

XSL•FO
RenderX

were managed via Covidence. A standardized form was used to extract study characteristics (eg, demographics, research design, intervention specifics) and participant data, which were subsequently managed in NVivo. Methodological quality was evaluated using CASP (Critical Appraisal Skills Program) and JBI checklists [16,17]. Quality assessment focused on congruity of perspective, research questions or objectives, methodology, bias, participant voice, ethical standards, analysis, and conclusions. Papers were also evaluated for analytical richness, reflecting qualitative depth and interpretation with respect to the study aims.

Data Analysis

This systematic review and thematic synthesis was conducted in 3 stages, summarized below, and discussed in more detail in the study protocol [10].

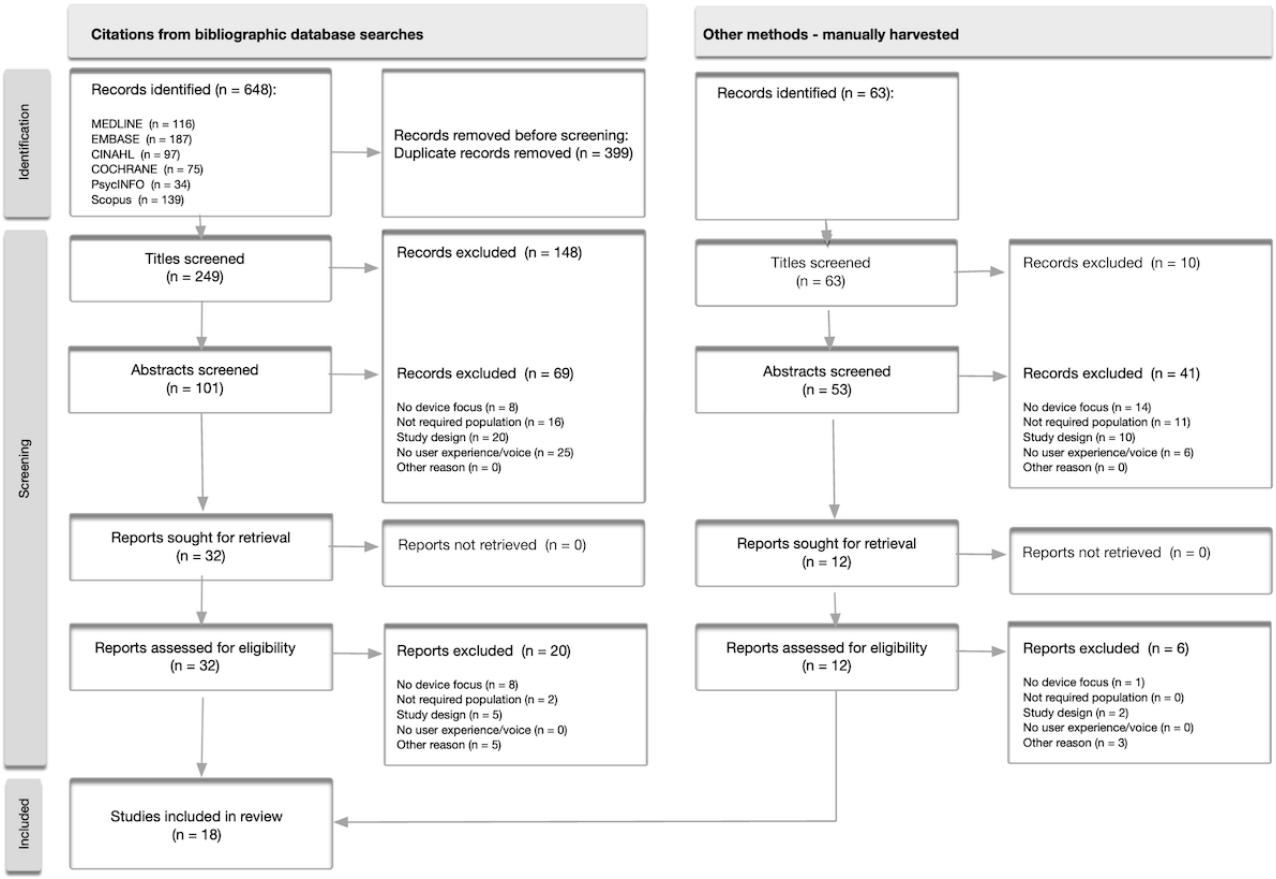
- *Stage 1. Inductive coding of included studies*: Reviewers independently coded the included studies, identifying salient sections and discussing coded sections as a team. A set of a priori codes based on the Person-Environment-Occupation-Performance and Technology Acceptance Model provided a provisional structure, with themes differentiated according to valence under binary (barrier and facilitator) top-level nodes [18,19].
- *Stage 2. Summarize under descriptive themes*: Related codes were aggregated into broader "descriptive themes" that summarized concepts across studies without going beyond the source texts. These higher-level descriptive themes were generated, discussed, and refined iteratively. Coded region frequencies (coding densities) and referenced study counts were analyzed to identify major and minor themes. The use of density metrics is well established in qualitative research, somewhat like matrix-coding queries, and is indicative of theme recurrence for included papers only [20,21].
- *Stage 3. Generate analytical constructs*: Descriptive themes were analyzed in relation to the research questions and objectives. Links between descriptive themes were mapped to generate analytical constructs and visualized diagrammatically to develop a coherent understanding from the evidence base.

Results

Study Selection and Characteristics

A systematic search of 6 electronic bibliographic databases yielded 648 records. An additional 63 records were identified through manual harvesting from other methods (ie, screening citation lists of included studies). After duplicate removal and screening, 44 studies were sought for full-text retrieval. Finally, 18 studies met the inclusion criteria and were included in this systematic review and thematic synthesis. The selection process and search results are summarized in the PRISMA flow diagram (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for systematic reviews.



XSL•FO
RenderX

To categorize heterogeneous participant data across the 18 studies listed in Table 1, a standardized coding system was applied:

- Population (N): [PRX] proxy; [STR] stroke
- Age: [YNG] 50-64; [OLD] 65-84; [ELD] ≥85 years
- Risk factors: [LFS] lifestyle; [MET] metabolic; [VAS] vascular
- Functional deficits: [ADL] activities of daily living; [COG] cognitive; [MOT] motor

XSL•FO
RenderX

Table . Characteristics of included studies.

Studies	Year	Location	Design	Duration	Pop (n)	Age group	Risk profile	Functional profile	Device(s)
Batsis et al [22]	2016	USA[a]	QLT[b]	1 mo	PRX[c] (8)	[OLD[d]]	[LFS[e], MET[f], VAS[g]]	[ADL[h], MOT[i]]	Fitbit
Nguyen et al [23]	2017	AUS[j]	QLT	18 wk	PRX (14)	[YNG[k], OLD]	[LFS, MET, VAS]	[ADL, MOT]	Various
Schlomann [24]	2017	DEU[l]	QLT	1 y	PRX (6)	[YNG, OLD]	[LFS, MET, VAS]	[MOT]	ViFit connect
Hamilton et al [25]	2018	AUS	QLT	6 mo	STR[m] (15)	[YNG, OLD, ELD[n]]	[LFS, MET, VAS]	[ADL, COG[o], MOT]	Fitbit
Johansson et al [26]	2018	SWE[p]	SYS[q], MIX[r]	NR[s]	STR (~43) PRX (54)	[YNG, OLD, ELD]	NR	[ADL, COG, MOT]	Various
Farina et al [27]	2019	GBR[t]	QLT	1 mo	STR (1) PRX (28)	[OLD, ELD]	[LFS, VAS]	[ADL, COG, MOT]	GENEactiv
Kononova et al [28]	2019	USA	QLT	6 wk	PRX (48)	[OLD, ELD]	[LFS, MET, VAS]	[ADL, MOT]	Various
Western et al [29]	2019	GBR	QLT	1 wk	PRX (31)	[YNG, OLD]	[LFS, MET, VAS]	[MOT]	SenseWear Pro
Whelan et al [30]	2019	GBR	MIX	6 wk	PRX (45)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Fitbit, Freestyle
Støve and Larson [31]	2019	DNK[u]	QLT	3 wk	STR (20) PRX (5)	[YNG, OLD]	[LFS, VAS]	[ADL, COG, MOT]	Garmin
Ummels et al [32]	2019	NLD[v]	QLT	2 - 8 wk	PRX (30)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Tracker
Östlind et al [33]	2022	SWE	QLT	12 wk	PRX (18)	[YNG, OLD]	[LFS, MET, VAS]	[MOT]	Fitbit
Gualtieri et al [34]	2016	USA	QLT	14 wk	PRX (10)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Withings Pulse
Mercer et al [35]	2016	CAN[w]	MIX	3+ d	STR (1) PRX (33)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Various
Randriambelonoro et al [36]	2017	CHE[x]	QLT	7 mo	PRX (18)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Fitbit
Ehn et al [37]	2018	SWE	QLT	9 d	PRX (8)	[OLD, ELD]	[VAS, MET, LFS]	[MOT, ADL]	Activite, Jawbone
Takemoto et al [38]	2018	USA	QLT	3 wk	PRX (15)	[YNG, OLD, ELD]	[VAS, MET, LFS]	[MOT, ADL]	Various
Brickwood et al [39]	2020	AUS	MIX	1 y	PRX (20)	[OLD]	[VAS, MET, LFS]	[MOT, ADL]	Jawbone

[a]USA: United States.

[b]QLT: qualitative.

[c]PRX: proxy.

[d]OLD: Old-Old (65–84 y).

[e]LFS: lifestyle.

[f]MET: metabolic.

[g]VAS: vascular.

[h]ADL: activities of daily living.

[i]MOT: motor.

[j]AUS: Australia.

XSL•FO
RenderX

[k]YNG: Young-Old (50–64 y).
[l]DEU: Germany.
[m]STR: stroke survivor.
[n]ELD: Oldest-Old (≥85 y).
[o]COG: cognitive.
[p]SWE: Sweden.
[q]SYS: systematic review.
[r]MIX: mixed methods.
[s]NR: not reported.
[t]GBR: Great Britain.
[u]DNK: Denmark.
[v]NLD: Netherlands.
[w]CAN: Canada.
[x]CHE: Switzerland.

Of the included studies, 1 was a systematic review that included quantitative and qualitative studies, 3 employed mixed methods, and 14 were qualitative. Studies were conducted in North America (n=5), Australia (n=3), Great Britain (n=3), and Europe (n=7). A detailed breakdown of the quality and risk of bias assessment for each study is provided in Table S1 in Multimedia Appendix 1. The characteristics of the included studies with country designations according to International Organization for Standardization 3-letter standards (eg, DEU: Germany and CHE: Switzerland) are summarized in Table 1.

Approximately 78% (n=14) of the studies included individuals with metabolic risks like obesity, prediabetes, or type 2 diabetes. Participants with vascular risks, such as hypertension, heart disease, and high cholesterol, were even more common, appearing in 83% (n=15) of the studies. All 18 studies included participants with lifestyle-related risks, specifically high sedentary behavior and poor dietary choices. Functional deficits further defined these populations: 89% (n=16) of studies reported on participants with motor or physical limitations such as joint pain from osteoarthritis, reduced walking speed, or the use of mobility aids like walkers. Participants with cognitive deficits (eg, memory loss or forgetfulness, impaired task comprehension, and execution) were represented in 4 (22%) studies, and participants with impaired activities of daily living affecting self-care and mobility were represented in 6 (33%) studies.

Specific stroke-related representation was documented in 5 studies. Johansson et al [26] systematically reviewed 24 studies of stroke survivors, while Støve and Larsen [31] included patients with stroke (80%) in active inpatient rehabilitation with hemiparesis. Similarly, Hamilton et al [25] reported a sample where 60% had a neurological primary diagnosis, including stroke, alongside significant motor deficits. Finally, Mercer et al [35] included a participant with stroke-related gait asymmetry (shuffling) that impacted activity sensor accuracy, and Farina et al [27] included a participant with stroke-related vascular dementia but no acute motor symptoms.

Synthesis of Findings

The thematic coding frame in Table 2 was applied to all included studies and reviewed for consistency of interpretation. This process, known as “axial coding,” ensured coded themes remained connected to and understood within the context of the source studies [13].

Table . Low-level theme and subtheme nodes.

Theme	Subtheme
Built environment and technology	Acceptance, accuracy, attitudes, community, customizable, damage, familiarity, functionality, home, influences, privacy, reliability, safety, usability, usefulness, wearability
Cognitive	Attention, awareness, comprehension, language, memory, planning
Natural environment	Safety, terrain, weather
Neurobehavioral	Balance, gait, motor, sensory, vision
Physiological	Endurance, fatigue, flexibility, injury, mobility, strength
Psychological	Accountability, adherence, attainment, autonomy, competence, feedback, goals, habits, mood, motivation, preconceptions, privacy, self-awareness, self-efficacy, trust, well-being
Social support	Caregiver, information, peer, professional, training
Societal and economic	Attitudes, commitments, cost, cultural, norms, strata
Spiritual	Fulfillment, meaningfulness, values

XSL•FO
RenderX

Coding densities informed the generation of the higher-level descriptive (aggregate) themes and subthemes listed in Table 3. Density cutoff points were used to classify major (n≥30), minor (n=10 - 29), and negligible (n<10) themes. Finally, descriptive themes were analyzed with respect to the research questions and objectives of the current study.

Table . Descriptive themes and subthemes.

Theme	Subtheme
Technological	User experience, device attributes, security, other
Psychological	Self-identity, motivation, affective
Neurological	Neurophysiological, cognitive
Support	Social, environment

Densities were also used to visualize descriptive barrier and facilitator themes and subthemes. For readability, descriptive themes in this section have been indicated by quotation marks, and subthemes have been italicized.

Major Barriers

Major barrier themes were largely "technological," more specifically related to user *experience* (24.5% of barrier themes) and *device attributes* (20.1%), and "psychological," related to *motivation* (15.6%). *User experience* themes summarized in Multimedia Appendix 1 (Table S2) included older adults perceiving themselves as not "tech-savvy" but able to use basic functions such as step count. The perceived obtrusiveness of the PA sensor, both aesthetically and functionally, was a recurrent barrier theme. Devices were perceived to be designed for younger people and not older adults due to form factor, usability, and functionality. A lack of customization and the perceived irrelevance of some metrics hindered sustained engagement. Underlying *device attribute* themes referred to metric accuracy, the ability to register low-intensity activities such as standing, slow walking (strolling), or gardening, and to correctly record other activity types like yoga or strength training. The inability of PA sensors to correctly differentiate between similar activities such as sitting versus standing or to accurately track specific movements relevant to stroke recovery (eg, differentiated movements in hemiparesis) were seen as serious limitations. Users wanted more granular and relevant data (eg, time spent in moderate to vigorous PA) and exercise categorization (eg, fat-burning, cardio, peak); cadence and stride length for running and walking; stairs climbed as measures of vertical movement; swim metrics including laps, stroke type, and count; and strength training metrics such as repetition and set counts. *Motivation* themes referred to the "honeymoon period" or novelty effect of PA sensors quickly diminishing. Users lost interest once they felt they fully understood their activity patterns or had achieved their desired goals. Buzzing reminders were seen as annoying, especially when activity change was not possible, such as while driving. Goals that were not established in partnership with a therapist were often perceived as unrealistic, presenting a significant barrier to adherence. PA sensor unreliability or an inability to accurately measure certain activities (eg, nonlinear occupations and nonrepetitive movements) was demotivating. Users expressed a desire for real-time feedback that helped them to understand the impact of their PA level, for example, on blood sugar, rather than just step counts.

Major Facilitators

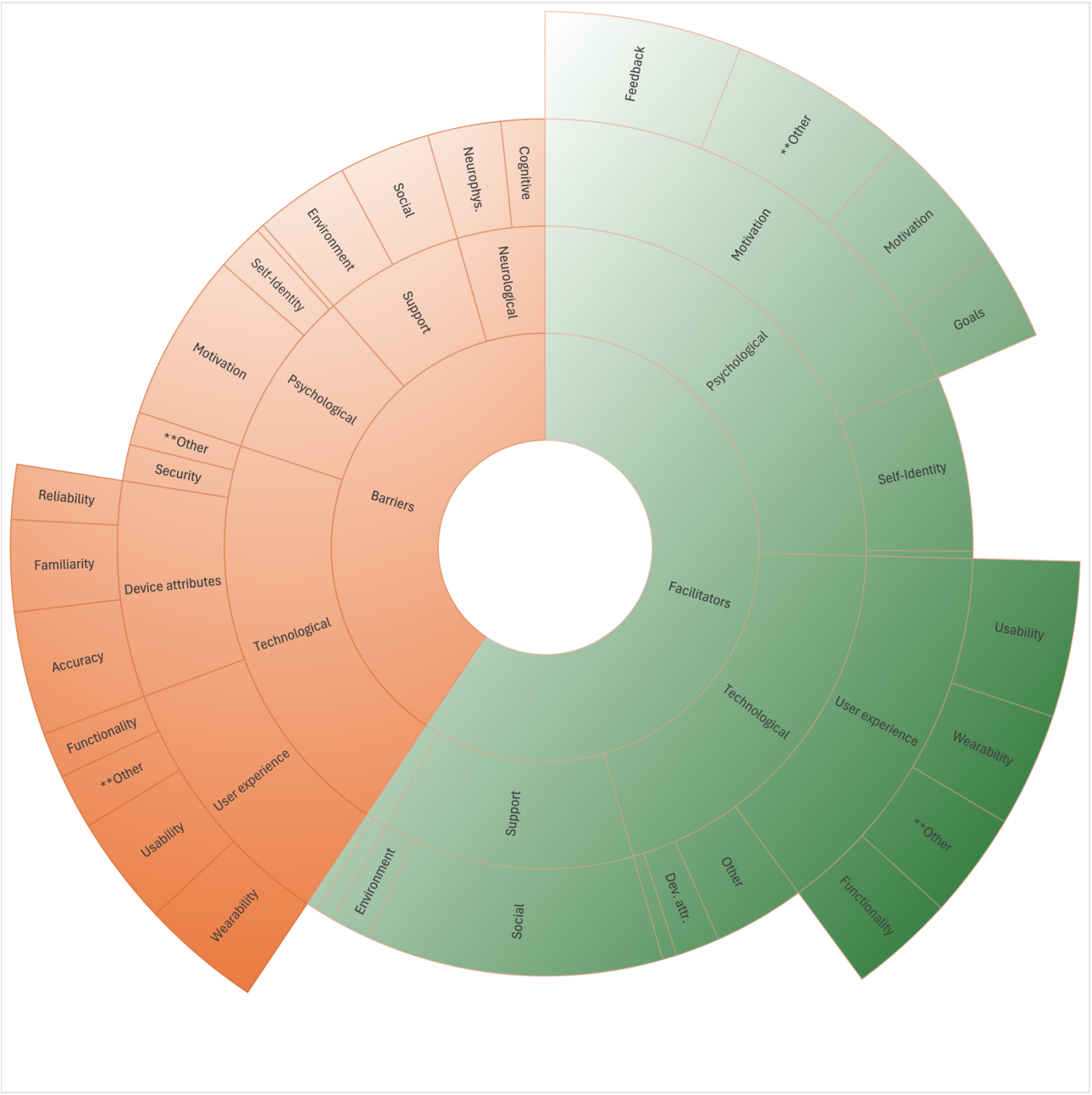
Major facilitator themes were "psychological" related to *motivation* (31.2% of facilitator themes), "technological" related to *user experience* (24.4%), and "support" related to both *social* and *environmental* support (21.4%). Underlying *motivation* themes summarized in Multimedia Appendix 1 (Table S3) included PA sensors that significantly increased self-awareness of PA levels. Achieving goals leads to feelings of inspiration, satisfaction, accountability, and a desire for additional effort. Long-term use was perceived to be strongly linked to intrinsic motivation, driven by personal enjoyment, feelings of accomplishment, and perceived benefits. Features like virtual incentives, nudges, reminders, and other forms of positive encouragement could act as effective extrinsic motivators, especially initially. PA sensors that effectively incorporated behavior change techniques such as self-monitoring, goal setting, and feedback promoted self-efficacy and PA. Individualized and tailored goals, and personalized feedback, for example, based on health status and functional capacity, were seen as critical for sustained motivation. Findings suggested that engagement with PA sensors was often nonlinear, with some users temporarily disengaging and later resuming use rather than abandoning the device entirely. Recognizing these flexible usage patterns is essential for developing interventions that foster long-term adherence. Underlying *user experience* themes included preferences for devices with simple, intuitive interfaces that required minimal user input. Users preferred PA sensors that were comfortable and unobtrusive, which did not interfere with daily activities and routines. Users expressed interest in tracking a more diverse range of activities such as swimming and cycling. The ability to customize goals, feedback type (vibration or visual), frequency of prompts, and data display was critical for effective motivation and sustained use. *Social* support themes related to therapists introducing and setting up the PA sensor, providing clear usage instructions, and interpreting data. This included personalized advice, goal setting, and addressing emotional responses to feedback. Social interaction and group engagement, both face-to-face and online via social networks, were regarded as powerful motivators. Embedding PA sensors into the clinical care process, with therapists actively discussing data and translating it into actionable insights, enhanced the perceived value and utility of the PA sensor. *Self-identity* themes referred to PA sensors significantly enhancing self-awareness of activity levels. Users were often surprised by their baseline level. Objective data helped users clarify real versus perceived activity. Tracking

XSL•FO
RenderX

progress and achieving goals, especially with immediate feedback, directly contributed to increased self-efficacy and confidence in the user's ability to be active and manage their health. For some, using the tracker shifted their self-perception, moving them from a state of pre-contemplation (unmotivated) to maintenance of a healthier lifestyle.

Visualized together in Figure 2, major barriers clustered mainly around *user experience* and *device attribute* aggregates, while facilitators were clustered around *user experience* and *motivation*. Drilling down to the underlying (low-level) themes, it was apparent that major *user experience* barriers were related to *wearability*, *usability*, and *familiarity* in descending order, and major *device attribute* barriers were *accuracy*, *reliability*, and *familiarity* in descending order. For facilitators, which comprised 60% of themes, most were clustered around *motivation*, *user experience*, and *support*. Underlying *motivation*, *feedback* (*motivation* more specifically), and *goals* in descending order were coded most frequently. Major *user experience* descriptive facilitators were *usability*, *wearability*, and *functionality*. Social and environmental *support* facilitators included most notably *professional*, that is, therapist and *caregiver* assistance.

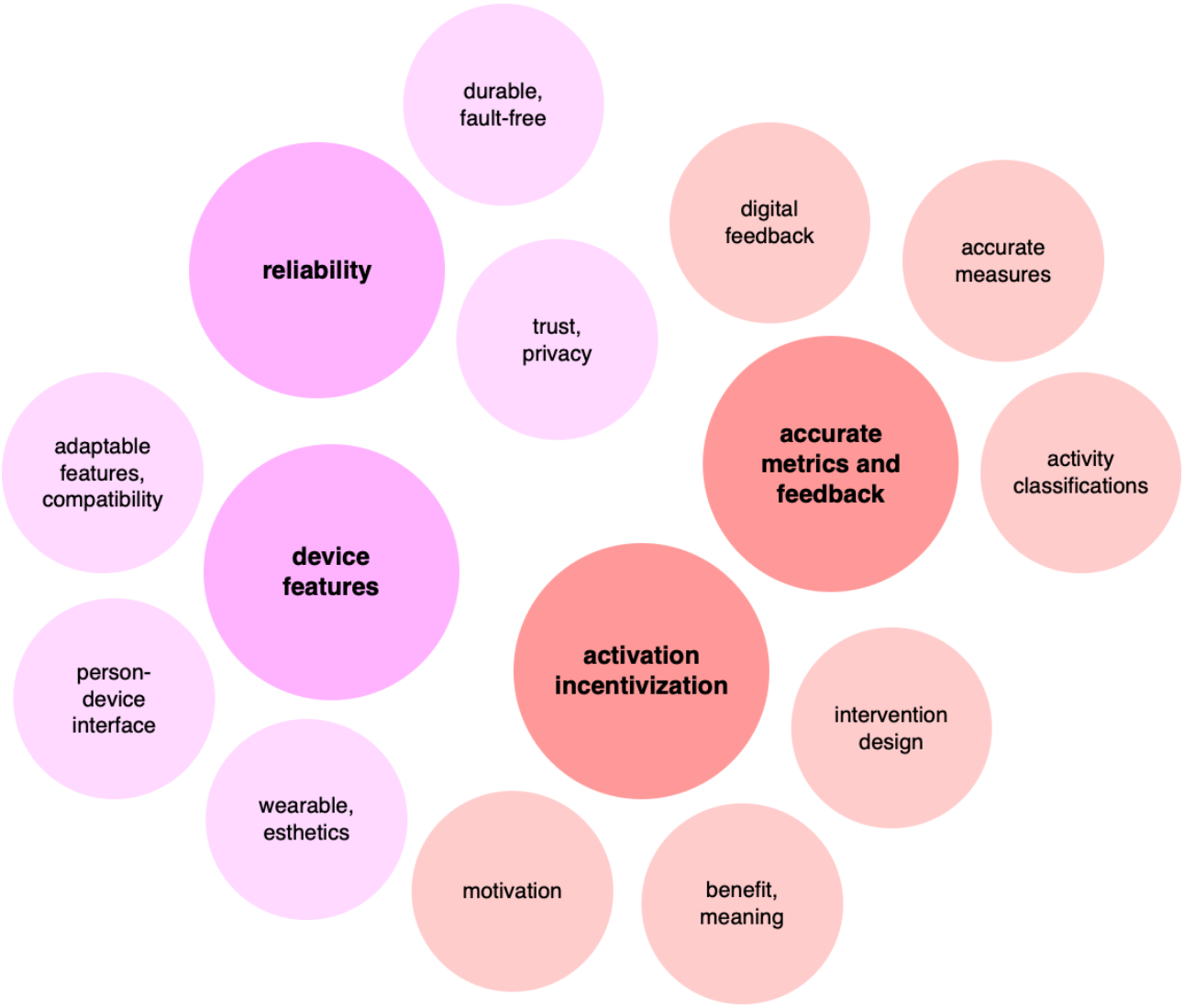
Figure 2. Major barrier/facilitator descriptive and substrate themes. Dev. attr.: device attributes; Neurophys.: neurophysiological



The initial analysis revealed major barriers clustered around device themes: subjective *accuracy*, *wearability*, and *usability* (Figure 3). In this context, *accuracy* referred to the perceived precision of the PA sensor; *wearability* to characteristics that influenced the user's ability and willingness to wear the PA sensor consistently for extended periods (eg, size, fit, esthetics); and *usability* referred to the ease with which users interacted with the PA sensor and companion app.

XSL•FO
RenderX

Figure 3. Major barrier clustering by descriptive theme.



Predominant facilitators were "psychological," that is, *feedback*, *motivation*, *self-awareness*; device-specific; or related to professional support from the physician or therapist (Figure 4).

XSL•FO
RenderX

Figure 4. Major facilitator clustering by descriptive theme.



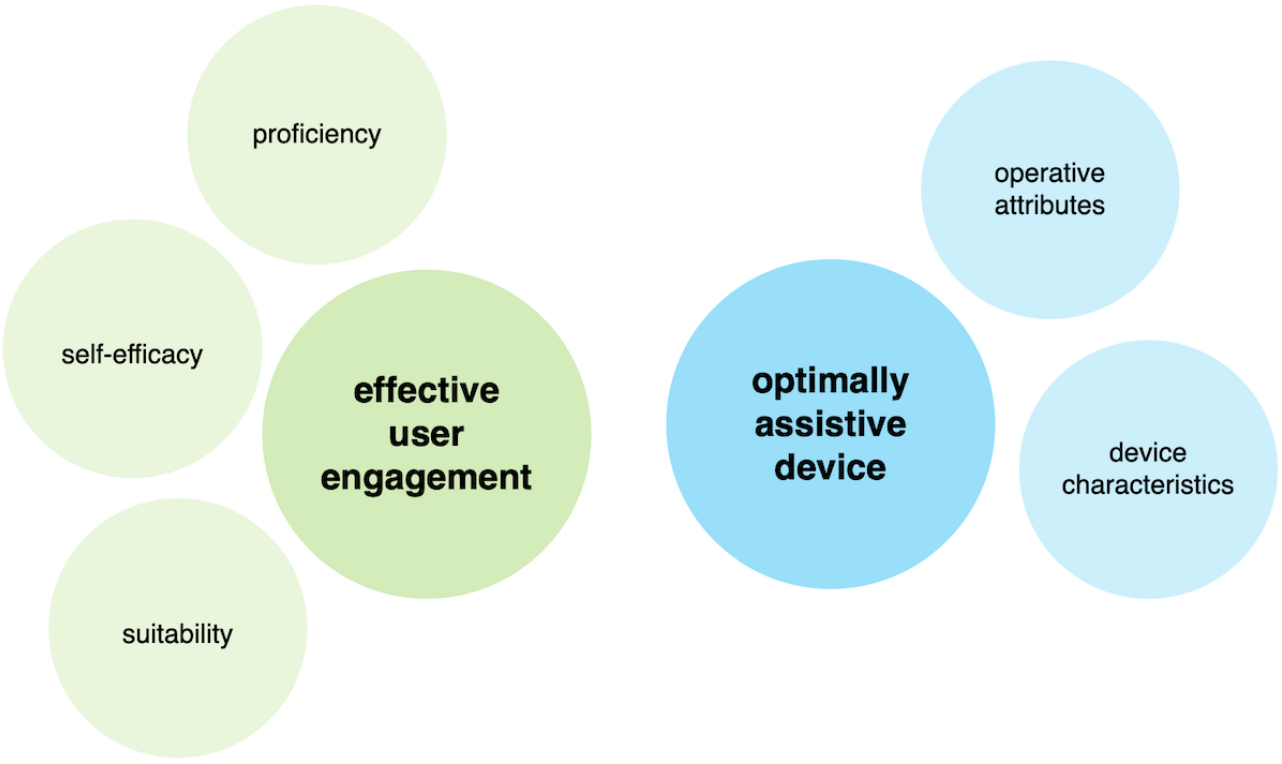
Building on the descriptive themes, higher-level analytical or theory-driven constructs were developed. Two distinct analytical constructs emerged, one person-centric, emphasizing factors related to user experience and engagement, and the other essentially device-centric, focused on aspects attributable to the PA sensor itself. Together, these suggested a synergistic interrelationship between effective user engagement and optimally assistive device factors, highlighting their combined contribution to long-term behavior change (Table 4 and Figure 5).

XSL•FO
RenderX

Table . Analytical constructs, descriptive themes, and subthemes.

Analytical construct and descriptive themes	Subthemes
Effective user engagement	
Proficiency	• Requisite skill(s) and knowledge: *degree of complexity; training and information* • Independent operation: *acceptability; ease of use*
Self-efficacy	• Supportive environment: *cost, commitments, natural environment; professional support; family, peers, friends, carers*
Suitability	• Appropriate for population: *cognitive, sensorimotor ability/impairment; cultural congruence; older adults, chronic conditions* • Thoughts and preconceptions: *values, attitudes; expectations; self-awareness*
Optimally assistive device	
Operative attributes	• Activation, incentivization: *motivation; intervention design; perceived benefit and meaningfulness* • Accurate metrics, feedback: *accurate measures; activity type/classification; useful and appropriate feedback*
Device characteristics	• Reliability: *durable, fault-free; trust, privacy* • Device features: *aesthetics and wearability; person-device interface; adaptable features, compatibility*

Figure 5. High-level analytical constructs.



Effective user engagement, as a "person-centric" analytical construct, represented the extent to which users meaningfully and effectively interacted with PA sensors. Contributing to this overarching construct were 3 key descriptive themes: *Proficiency* referred to the level of skill and technical knowledge required by the user and their capacity to use the PA sensor independently. Inadequate training, lack of technical skills, inability to resolve usage issues independent of external support, and poor instructions were major technical barrier themes evident in 9 of the studies included in the current analysis [25,26,28,30,32,34-37]. When users felt confident in their ability to operate a PA sensor correctly, they were more likely to engage with it. *Self-efficacy* referred to the individual's belief in their capacity or readiness to initiate PA to achieve their goals. This encompassed the ability to integrate PA within personal and broader social/cultural contexts to establish or

XSL•FO
RenderX

participate in enabling social networks, exercise buddies, peers, and family, and to become better informed about their health, the benefits of regular activity, and maintaining long-term self-motivation. An interaction between motivation and adherence was also evident, suggesting an interplay between the analytical themes discussed in this section, that is, how PA sensors might facilitate a sense of self-efficacy or self-empowerment that in turn promoted increased activity. A lack of self-efficacy or belief in one's ability to effectively engage in PA and achieve goals was identified as a significant theme in 8 of the studies [22,25,28-30,33,34,39]. Human support from peers, family, coaches, and clinicians was seen as an essential adjunct that aided increased PA together with PA sensor usage; the absence of these supports was evident in 4 studies as major barriers [24,25,32,36]. *Suitability* referred to the appropriateness and acceptability of the PA sensor within the target population. This reflected how well the PA sensor aligned with the physical, cognitive, and cultural needs of the user.

Optimally assistive device, as a "device-centric" analytical construct, referred to the inherent characteristics of the PA sensor and its capacity to effectively promote increased PA. Contributing to this were 2 primary descriptive themes: *Operative attributes* that referred to the functional performance of the PA sensor, accurately capturing and conveying relevant activity data to the user. This included precision, for example, accurate step counts, correct activity classification, and the clarity, relevance, and appropriateness of feedback provided. *Device characteristics*, on the other hand, referred to the design elements of the PA sensor. This included user-friendliness and ease of use of the PA sensor, functional capacity, that is, support for required or expected functionality, timeliness and appropriateness of information provided, and integration with other devices such as the companion app. Perceived inaccuracies and PA sensor defects were major barriers evident in 8 of the studies [23,27,29,33,36,37,40,41]. Similarly, themes that referred to a perceived lack of PA sensor validation, or characteristics that did not appear to have been designed for the abilities of older persons outlined earlier (eg, form factor, usability, functionality), highlighted broader device characteristic issues. Wearability themes reflected how easily people were able to use the PA sensor in everyday life situations and adapt usage to daily routines.

Discussion

Principal Findings

This study was primarily undertaken to address noted inconsistencies in the long-term efficacy of device-based PA interventions, particularly in older adults at risk of primary or recurrent stroke. Through a review of the qualitative literature, nuanced human and device-related factors influencing sustained engagement in this specific population were explored through a qualitative lens. By synthesizing lived experiences, insights emerged that explained why these inconsistencies occurred, reflecting individual and contextual factors at play. This qualitative depth is crucial for bridging the implementation gap between technological potential and real-world impact. These findings extend the understanding of existing models by demonstrating how device characteristics must be inherently assistive and person-centric to promote sustained engagement, especially in more vulnerable populations.

The preliminary synthesis was underpinned by a binary schema, categorizing overarching descriptive themes under top-level barrier and facilitator nodes. This implicitly assumed influencing factors were mutually exclusive. However, emergent themes within the included studies suggested a more nuanced reality. A similar observation might be made regarding the conceptual dichotomy between human and device factors. Indeed, one systematic review highlighted an interrelationship between these factors, reporting improved adherence in studies with longer monitoring durations [26]. The authors of the review posited that this could be attributed to increased user confidence over time. Consistent with this finding, studies included in the current analysis similarly indicated an interplay between motivation, technical integrity (eg, absence of defects, data accuracy), and adherence factors, suggesting that these elements are not isolated but rather interact synergistically in complex ways to influence outcomes. Studies included in the current analysis consistently indicated an interplay between motivation and adherence to PA.

Numerous theoretical frameworks have attempted to elucidate behavior change mechanisms through constructs such as self-efficacy (perceived confidence), decisional balance (weighing pros and cons), and processes of change (covert and overt activities facilitating progression) [42]. In the context of PA, self-efficacy specifically refers to an individual's confidence in identifying and achieving planned PA goals. This confidence is reinforced through mastery experiences, vicarious learning (ie, observing others' success), persuasion, and performance feedback [40]. While some behavioral goals are characterized by defined end states, others, such as increased PA, necessitate ongoing motivation for sustained maintenance. It has been suggested that individuals are more inclined to adopt healthy behaviors when confronted with a substantial threat and when their actions are perceived as effectively reducing the likelihood or severity of that threat [43]. This applies to patients with stroke concerned about secondary events, as well as individuals at risk of a primary stroke. However, the absence of universally agreed-upon construct measures complicates efforts to isolate which specific intervention components successfully drive behavior change and which do not.

The effectiveness of health care interventions and PA sensors is often complicated by social dynamics, as individuals' interactions can lead to unforeseen outcomes that alter the intended effects of the intervention [44]. For researchers, deciphering the complex evolution and adaptation of behaviors stemming from these interactions is a formidable task. This complexity is arguably amplified in PA sensor–based interventions, given the intricate interplay of human and device factors. The current study revealed prominent device-specific themes, with technical barriers featuring more significantly than psychological barriers related to motivation, support, adherence, or physiological limitations such as sensorimotor impairment.

While PA sensor–based devices offer objective measures of PA, potentially reducing biases inherent in self-report

XSL•FO
RenderX

instruments, themes related to perceived accuracy and trustworthiness were nonetheless evident in this synthesis [28]. The unregulated nature of the consumer market and clinometric quality of consumer-grade PA sensors warrant consideration in this context. A systematic review assessing consumer-grade PA sensors in older adults (mean age 70.2, SD 4.8 y) reported high accuracy for average daily step counts compared to research-grade reference PA sensors [41]. However, the reliability of daily step counts varied considerably across PA sensors, with gait, device placement, and walking speed influencing reliability. Several of the consumer-grade PA sensors contributing to the current thematic synthesis were also evaluated in the aforementioned review. While these findings offer some reassurance regarding accuracy, they simultaneously raise questions about public perceptions of consumer PA sensor accuracy and critical intervention design factors such as training and support. The level of support required for effective PA sensor use also emerged as a recurring major theme in the current study. The degree of intuitiveness and interactivity of the PA sensors may be contributing factors.

The synergistic factors identified in the current study suggest that future interventions should prioritize co-design methodologies, actively involving older patients with stroke in the development and refinement of PA sensors and health care interventions based on these technologies. This would ensure that PA sensors are not only functionally more robust but also intuitively usable, comfortable, and psychologically reinforcing for this specific cohort. For clinicians, these findings underscore the importance of tailored support, including training and goal setting, while addressing individual user concerns regarding PA sensor accuracy and usability when employing these devices in rehabilitation or preventive health care interventions. Future research should focus on developing and rigorously testing co-designed PA sensor–based interventions, employing mixed methods approaches to quantitatively validate the impact of the qualitative insights identified here on sustained levels of PA. Specific areas for investigation might include the optimal frequency and type of professional support, the most effective feedback mechanisms, and ergonomic design features critical in the context of stroke prevention and rehabilitation.

While the hypothesized advantages of PA sensors have garnered considerable scholarly attention, a notable gap exists in the literature regarding their application and, more critically, the lived experience of their use by stroke survivors and older adults at risk of stroke. The current study partially addresses that gap by synthesizing existing qualitative research concerning the utilization of PA sensors within this specific demographic. Through a rigorous qualitative synthesis, both major thematic elements and less predominant, yet illuminating, experiential aspects associated with PA sensor use were identified. The insights gleaned from these findings hold significant implications for the development of more efficacious interventions aimed at mitigating sedentary behavior and fostering increased PA in older adults at risk of primary or recurrent stroke.

Limitations

Despite a comprehensive search, only 2 included studies specifically addressed older stroke survivors: a qualitative study of 15 survivors and a systematic review incorporating one qualitative stroke study and 5 qualitative proxy population studies [25,26]. While limited stroke-specific data might traditionally constrain the generalizability of these findings, the methodology employed prioritized *conceptual transferability* over statistical representativeness [13]. By extending inclusion to proxy populations with analogous risk and functional profiles, this synthesis identified higher-level analytical constructs that are specifically relevant to older adults at risk of primary or recurrent stroke. These overarching themes provide a robust explanatory framework for PA sensor–based interventions that is theoretically relevant to the stroke experience, even as empirical qualitative data for this population are emerging. However, it is acknowledged that the inherent subjectivity of qualitative interpretation means this synthesis may represent one of several possible frameworks. Furthermore, the lack of detailed technical specifications across the diverse range of included PA sensors precluded an analysis of device-specific effects, such as wearability or discrete usability features. Future research should aim to validate these conceptual constructs through larger, stroke-specific cohorts and more granular investigations into device-specific factors.

Conclusion

By pooling existing qualitative research, this study contributes a novel and deeper understanding of the benefits and limitations of wearable PA sensors through the subjective experiential lens of older adults at risk of primary or recurrent stroke. This systematic review and thematic synthesis identified a synergistic interplay between human and device-specific factors that influence engagement and sustained increased PA. However, the precise nature of this relationship and its impact on objective PA levels would require verification through higher levels of evidence. These findings provide a framework for more effective PA sensor–based interventions in stroke prevention and rehabilitation, offering valuable insights for technologists designing PA sensors and therapists developing PA sensor–based health interventions for older stroke cohorts.

Acknowledgments

This study was based on a clinical research thesis written by PH under the supervision of A/Prof Anita Horvath from the Department of Medical Education, University of Melbourne. The coding structure and review process were jointly developed by IM and PH based on the thematic synthesis methodology. Dr Phillis Lau from the Department of General Practice, University of Melbourne, reviewed the methodological rigor of both the thesis and published protocol. Review comments for a preliminary analysis were also provided by Prof Kathleen Gray from the Centre for Digital Transformation of Health, University of Melbourne. All authors have read and approved the final version of the manuscript.

XSL•FO
RenderX

Funding

The authors received no specific funding for this work.

Data Availability

The detailed search strategy, selection process documentation, PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)/JBI (Joanna Briggs Institute) checklists, and Data Management Plan are openly available on Figshare [45].

Authors' Contributions

PH conceived and designed the study, conducted the search and data curation, and performed the primary project administration. PH and IM collaborated on the methodology and formal analysis. PH drafted the original manuscript. IM and PH contributed to the critical review and editing of the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Quality and risk of bias assessments for included studies; barrier themes, coded region frequencies (density), and source study counts; facilitator themes, coded region frequencies (density), and study references; and methodological data.
[DOCX File, 279 KB - rehab_v13i1e86915_app1.docx]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 85 KB - rehab_v13i1e86915_app2.pdf]

References

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131. [Medline: 3920711]
2. Boehme AK, Esenwa C, Elkind MSV. Stroke risk factors, genetics, and prevention. Circ Res 2017 Feb 3;120(3):472-495. [doi: 10.1161/CIRCRESAHA.116.308398] [Medline: 28154098]
3. Cooper C, Gross A, Brinkman C, et al. The impact of wearable motion sensing technology on physical activity in older adults. Exp Gerontol 2018 Oct 2;112:9-19. [doi: 10.1016/j.exger.2018.08.002] [Medline: 30103025]
4. Kelley GA, Kelley KS. Leisure time physical activity reduces the risk for stroke in adults: a reanalysis of a meta-analysis using the inverse-heterogeneity model. Stroke Res Treat 2019;2019:8264502. [doi: 10.1155/2019/8264502] [Medline: 31275539]
5. Lee CD, Folsom AR, Blair SN. Physical activity and stroke risk: a meta-analysis. Stroke 2003 Oct;34(10):2475-2481. [doi: 10.1161/01.STR.0000091843.02517.9D] [Medline: 14500932]
6. Chodzko-Zajko WJ, Proctor DN, Fiatarone Singh MA, et al. Exercise and physical activity for older adults. Med Sci Sports Exerc 2009;41(7):1510-1530. [doi: 10.1249/MSS.0b013e3181a0c95c]
7. Saunders DH, Sanderson M, Hayes S, et al. Physical fitness training for stroke patients. Cochrane Database Syst Rev 2016 Mar 24;3(3):CD003316. [doi: 10.1002/14651858.CD003316.pub6] [Medline: 27010219]
8. Fox S, Duggan M. Tracking for health. : Pew Research Center; 2013 Jan 28 URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/media/Files/Reports/2013/PIP_TrackingforHealth-with-appendix.pdf? [accessed 2026-02-21]
9. Zhang MW, Chew PY, Yeo LL, Ho RC. The untapped potential of smartphone sensors for stroke rehabilitation and after-care. Technol Health Care 2016;24(1):139-143. [doi: 10.3233/THC-151099] [Medline: 26484884]
10. Harris PT, Maine I. Patient-perceived factors influencing physical activity sensor use in stroke prevention and rehabilitation: a thematic synthesis protocol. PLOS ONE 2024;19(4):e0301983. [doi: 10.1371/journal.pone.0301983] [Medline: 38687706]
11. Booth A, Papaioannou D, Sutton A. Systematic Approaches to a Successful Literature Review: SAGE Publications; 2012. URL: https://www.researchgate.net/profile/Andrew-Booth-2/publication/235930866_Systematic_Approaches_to_a_Successful_Literature_Review/links/5da06c7f45851553ff8705fa/Systematic-Approaches-to-a-Successful-Literature-Review.pdf [accessed 2026-02-21]
12. Sandelowski M. “To be of use”: enhancing the utility of qualitative research. Nurs Outlook 1997;45(3):125-132. [doi: 10.1016/s0029-6554(97)90043-9] [Medline: 9210160]
13. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
14. O’Donnell MJ, Chin SL, Rangarajan S, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. Lancet 2016 Aug 20;388(10046):761-775. [doi: 10.1016/S0140-6736(16)30506-2] [Medline: 27431356]

XSL•FO
RenderX

15. Lundebjerg NE, Trucil DE, Hammond EC, Applegate WB. When it comes to older adults, language matters: Journal of the American Geriatrics Society adopts modified American Medical Association style. J Am Geriatr Soc 2017 Jul;65(7):1386-1388. [doi: 10.1111/jgs.14941] [Medline: 28568284]
16. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Healthc 2015 Sep;13(3):132-140. [doi: 10.1097/XEB.0000000000000055] [Medline: 26360830]
17. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc 2015 Sep;13(3):179-187. [doi: 10.1097/XEB.0000000000000062] [Medline: 26262565]
18. Baum C, Christiansen C. The person-environment-occupation-performance (PEOP) model. In: Christiansen CH, Baum CM, Bass JD, editors. Occupational Therapy: Performance, Participation, and Well-Being, 4th edition: SLACK Incorporated; 2015:49-56.
19. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
20. Bassett B. Computer-based analysis of qualitative data: NVIVO. In: Mills AJ, Durepos G, Weibe E, editors. Encyclopedia of Case Study Research: Sage Publications; 2010:193-194. [doi: 10.4135/9781412957397.n71]
21. Dulmage BO, Akintilo L, Welty LJ, Davis MM, Colavincenzo M, Xu S. A qualitative, cross-sectional study of positive and negative comments of residency programs across 9 medical and surgical specialties. Am J Med 2018 Sep;131(9):1130-1134. [doi: 10.1016/j.amjmed.2018.05.019] [Medline: 29908767]
22. Batsis JA, Naslund JA, Gill LE, Masutani RK, Agarwal N, Bartels SJ. Use of a wearable activity device in rural older obese adults: a pilot study. Gerontol Geriatr Med 2016;2:2333721416678076. [doi: 10.1177/2333721416678076] [Medline: 28138502]
23. Nguyen NH, Hadgraft NT, Moore MM, et al. A qualitative evaluation of breast cancer survivors' acceptance of and preferences for consumer wearable technology activity trackers. Support Care Cancer 2017 Nov;25(11):3375-3384. [doi: 10.1007/s00520-017-3756-y] [Medline: 28540402]
24. Schlomann A. A case study on older adults' long-term use of an activity tracker. Gerontechnology 2017 Aug 9;16(2):115-124. [doi: 10.4017/gt.2017.16.2.007.00]
25. Hamilton C, McCluskey A, Hassett L, Killington M, Lovarini M. Patient and therapist experiences of using affordable feedback-based technology in rehabilitation: a qualitative study nested in a randomized controlled trial. Clin Rehabil 2018 Sep;32(9):1258-1270. [doi: 10.1177/0269215518771820] [Medline: 29696990]
26. Johansson D, Malmgren K, Alt Murphy M. Wearable sensors for clinical applications in epilepsy, Parkinson's disease, and stroke: a mixed-methods systematic review. J Neurol 2018 Aug;265(8):1740-1752. [doi: 10.1007/s00415-018-8786-y] [Medline: 29427026]
27. Farina N, Sherlock G, Thomas S, Lowry RG, Banerjee S. Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia. Int J Geriat Psychiatry 2019 Apr;34(4):617-624. [doi: 10.1002/gps.5064]
28. Kononova A, Li L, Kamp K, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. JMIR Mhealth Uhealth 2019 Apr 5;7(4):e9832. [doi: 10.2196/mhealth.9832] [Medline: 30950807]
29. Western MJ, Thompson D, Peacock OJ, Stathi A. The impact of multidimensional physical activity feedback on healthcare practitioners and patients. BJGP Open 2019 Apr;3(1):bjgpopen18X101628. [doi: 10.3399/bjgpopen18X101628] [Medline: 31049409]
30. Whelan ME, Orme MW, Kingsnorth AP, Sherar LB, Denton FL, Esliger DW. Examining the use of glucose and physical activity self-monitoring technologies in individuals at moderate to high risk of developing type 2 diabetes: randomized trial. JMIR Mhealth Uhealth 2019 Oct 28;7(10):e14195. [doi: 10.2196/14195] [Medline: 31661077]
31. Støve MP, Larsen BT. Self-monitoring – usability evaluation of heart rate monitoring using wearable devices in patients with acquired brain injury. Eur J Physiother 2020 Nov 1;22(6):364-372. [doi: 10.1080/21679169.2019.1628300]
32. Ummels D, Beekman E, Moser A, Braun SM, Beurskens AJ. Patients' experiences with commercially available activity trackers embedded in physiotherapy treatment: a qualitative study. Disabil Rehabil 2020 Nov;42(23):3284-3292. [doi: 10.1080/09638288.2019.1590470] [Medline: 30973026]
33. Östlind E, Ekvall Hansson E, Eek F, Stigmar K. Experiences of activity monitoring and perceptions of digital support among working individuals with hip and knee osteoarthritis - a focus group study. BMC Public Health 2022 Aug 30;22(1):1641. [doi: 10.1186/s12889-022-14065-0] [Medline: 36042425]
34. Gualtieri L, Rosenbluth S, Phillips J. Can a free wearable activity tracker change behavior? The impact of trackers on adults in a physician-led wellness group. JMIR Res Protoc 2016 Nov 30;5(4):e237. [doi: 10.2196/resprot.6534] [Medline: 27903490]
35. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7. [doi: 10.2196/mhealth.4225] [Medline: 26818775]
36. Randriambelonoro M, Chen Y, Geissbuhler A, Pu P. Exploring physical activity monitoring devices for diabetic and obese patients. Presented at: UbiComp/ISWC'15 Adjunct: Adjunct Proceedings of the 2015 ACM International Joint Conference

XSL•FO
RenderX

on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers; Sep 7-11, 2015 URL: http://dl.acm.org/citation.cfm?doid=2800835 [accessed 2026-02-21]
37. Ehn M, Eriksson LC, Åkerberg N, Johansson AC. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 1;6(2):e34. [doi: 10.2196/mhealth.8345] [Medline: 29391342]
38. Takemoto M, Lewars B, Hurst S, et al. Participants' perceptions on the use of wearable devices to reduce sitting time: qualitative analysis. JMIR Mhealth Uhealth 2018 Mar 31;6(3):e73. [doi: 10.2196/mhealth.7857] [Medline: 29599105]
39. Brickwood KJ, Williams AD, Watson G, O'Brien J. Older adults' experiences of using a wearable activity tracker with health professional feedback over a 12-month randomised controlled trial. Digit Health 2020;6:2055207620921678. [doi: 10.1177/2055207620921678] [Medline: 32426152]
40. Jones F, Riazi A. Self-efficacy and self-management after stroke: a systematic review. Disabil Rehabil 2011;33(10):797-810. [doi: 10.3109/09638288.2010.511415] [Medline: 20795919]
41. Straiton N, Alharbi M, Bauman A, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]
42. Prochaska JM. The transtheoretical model applied to the community and the workplace. J Health Psychol 2007 Jan;12(1):198-200. [doi: 10.1177/1359105307071754] [Medline: 17158853]
43. Jones CJ, Smith H, Llewellyn C. Evaluating the effectiveness of health belief model interventions in improving adherence: a systematic review. Health Psychol Rev 2014;8(3):253-269. [doi: 10.1080/17437199.2013.802623] [Medline: 25053213]
44. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. BMJ 2001 Sep 15;323(7313):625-628. [doi: 10.1136/bmj.323.7313.625] [Medline: 11557716]
45. Harris P, Maine I. Patient-perceived factors influencing physical activity sensor use in stroke prevention and rehabilitation. Figshare. URL: https://doi.org/10.6084/m9.figshare.c.7992074 [accessed 2026-02-26]

Abbreviations

CASP: Critical Appraisal Skills Program
JBI: Joanna Briggs Institute
PA: physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by S Munce; submitted 31.Oct.2025; peer-reviewed by A Al-Muqarm, H Tariq; revised version received 09.Feb.2026; accepted 10.Feb.2026; published 12.Mar.2026.

Please cite as:
Harris P, Maine I
Patient-Perceived Factors Influencing Physical Activity Sensor Use in Stroke Prevention and Rehabilitation: Systematic Review of Qualitative Studies Using Thematic Synthesis
JMIR Rehabil Assist Technol 2026;13:e86915
URL: https://rehab.jmir.org/2026/1/e86915
doi:10.2196/86915

© Paul Harris, Ingrid Maine. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use of Wearable Devices to Augment Traditional Measurements of Postoperative Outcomes Following Total Joint Arthroplasty: Systematic Review

Robert Raymond Hall III[1], MD; Alexander P Hamel[2], BS; Pamela A Chan[2], BS, MS; Hua Zheng[1], PhD; David C Ayers[1], MD

[1]Department of Orthopedics and Physical Rehabilitation, University of Massachusetts Chan Medical School, 55 Lake Avenue North, Worcester, MA, United States

[2]University of Massachusetts Chan School of Medicine, Worcester, MA, United States

Corresponding Author:
Robert Raymond Hall III, MD
Department of Orthopedics and Physical Rehabilitation, University of Massachusetts Chan Medical School, 55 Lake Avenue North, Worcester, MA, United States

Abstract

Background: Wearable devices enabling remote monitoring by surgeons of their patients have gained prominence around total joint arthroplasty (TJA), offering continuous patient data to identify those not meeting postoperative goals, thereby facilitating timely interventions. While multiple studies highlight the utility of these devices in tracking postoperative progress, a standardized approach to their application is lacking. This review aims to synthesize existing literature on the use of wearable device-tracked activity for monitoring TJA outcomes.

Objective: We examined the current literature to evaluate how wearable devices are used in monitoring and improving patient rehabilitation and outcomes following TJA.

Methods: A systematic review was conducted following Cochrane methodology. A literature search of all available literature was performed in April 2024 and identified 102 studies to undergo full-text review. Systematic reviews, duplicate papers, and theoretical papers were excluded. Ultimately, 35 studies met the selection criteria.

Results: The review revealed that 32 of 35 (91.4%) studies used wearable devices to monitor step counts. A total of 21 (60%) studies incorporated joint-specific patient-reported outcome measures, though the specific measures varied. Further, 9 studies used standardized performance-based outcome measures, which also differed across studies. Finally, 7 (20%) studies collected sleep data; however, the methods and outcomes for sleep measurement were inconsistent among these studies.

Conclusions: Remote monitoring via wearable devices offers a novel approach to tracking outcomes in TJA patients. Although the use of these devices in perioperative care is expanding, significant variability exists in the data reported across studies. Wearable monitoring is often integrated with patient-reported outcome measures and standardized functional assessments, yet the optimal data parameters that best correlate with established outcome metrics remain undefined. Additionally, data collected by wearables has not yet been shown to predict patient recovery or satisfaction. Further research is essential to refine these data parameters and the development of postoperative protocols that leverage wearable devices to enhance patient compliance and improve clinical outcomes.

Trial Registration: PROSPERO CRD420261346230; https://www.crd.york.ac.uk/PROSPERO/view/CRD420261346230

(JMIR Rehabil Assist Technol 2026;13:e84671) doi:10.2196/84671

KEYWORDS

remote monitoring; wearable device; total joint arthroplasty recovery assessment; patient-reported outcome measures; mobility; osteoarthritis

Introduction

Total hip arthroplasty (THA) and total knee arthroplasty (TKA) are currently among the most commonly performed surgeries in the United States, with significant growth expected by 2030. The number of TKAs performed annually is projected to reach 1 million, and the number of THAs is projected to reach 600,000 by 2030 [1]. Despite high patient satisfaction rates, 10% to 20% of TKA patients and 5% to 15% of THA patients are not satisfied after surgery, suggesting that current recovery assessments may not adequately address all aspects of patient care [2,3]. The variability of current methods of assessing total

XSL•FO
RenderX

joint arthroplasty (TJA) recovery exists due to divergent guidelines, institutional policies, difficulty in obtaining data, and physician preferences. Most commonly, assessment of patient recovery has used patient-reported outcome measures (PROMs).

The adoption of wearable technology is on the rise, with 30% - 50% of adults in the United States using devices that track various health metrics [4,5]. Wearables can continuously monitor critical health parameters such as heart rate, physical activity, and sleep during the postoperative recovery phase, providing valuable data for assessing recovery outcomes in TJA patients. Wearable devices provide the opportunity to obtain data detailing the recovery of patients after TJA, improve patient compliance with rehabilitation programs, and enhance patient engagement through real-time feedback [6].

Previous studies demonstrate that wearable devices can enhance physical activity levels and offer insightful data throughout the postoperative recovery period [7-9]. Hameed et al found that remote monitoring of TKA patients improved outcomes in the early postoperative period [10].

Presently, there is a notable absence of standardized methods for interpreting the extensive data generated by wearable devices in the context of TJA outcomes. This systematic review aims to clarify how wearable devices can effectively measure TJA outcomes, focusing on metrics such as sleep quality, physical activity, and step counts. Data from both TKA and THA patients will be included to potentially yield more holistic insights into the application of wearable technology in joint replacement scenarios. The goal of this systematic review is to significantly contribute to ongoing discussions aimed at enhancing patient outcomes and satisfaction following TJA surgery by leveraging the innovative capabilities of wearable technology.

Methods

Search Strategy

A thorough systematic review was conducted according to the Cochrane methodology for systematic reviews as described in the Cochrane Handbook for Systematic Reviews of Interventions [11] (Multimedia Appendix 1). Our trial was registered with PROSPERO (registration number CRD420261346230). A literature search was conducted in Scopus, Medline, CINAHL, PubMed, CENTRAL, and Google Scholar databases in April 2024 using relevant search terms associated with the research question, “Can wearable device data be used to effectively augment postoperative patient-reported outcomes after total joint replacement?” (Multimedia Appendix 2). Literature searches were not limited by date of publication.

Screening and Assessment of Eligibility

A total of 471 papers were identified from the databases or registers used, as well as 7 studies from the reference lists of other papers. Selection criteria included availability in English and full-text accessibility. Systematic reviews and theoretical papers were excluded. After the removal of 100 duplicates, the abstracts of 371 studies were screened for compliance with the criteria. Further, 269 studies were excluded during abstract screening due to a lack of adherence to selection criteria, resulting in 102 studies to undergo full-text review. Finally, 35 studies were identified from the full-text review that satisfied the selection criteria for inclusion in this systematic review.

Data Extraction and Analysis

Data extraction of the 35 studies was performed on Covidence. Data extraction was originally performed independently by 2 reviewers, then reviewed for consensus. Extracted data included information regarding study identification, aims, study design and population, inclusion and exclusion criteria, data collection characteristics, interventions, accelerometer details and data, PROMs and quality of life assessments used, and study end points.

Risk of Bias Assessment

Study quality and bias risk were assessed using Cochrane’s Risk of Bias tool for randomized trials (RoB-2) [12] and the Risk of Bias in Nonrandomized Studies of Interventions (ROBINS-I) tool [13]. RoB-2 includes 5 domains: the randomization process, deviations from the intended intervention, missing outcomes data, measurement of the outcome, and selection of the reported result. ROBINS-I evaluates 7 domains: confounding, participant selection, intervention classification, deviations from intended interventions, missing data, outcome measurement, and selection of reported results. RoB-2 categorized bias as low risk, some concerns, high risk, or no information, while ROBINS-I classified risk as low, moderate, serious, critical, or no information. A total of 2 authors independently assessed the study quality of each paper and resolved any concerns with thorough discussion. Traffic light plots of domains regarding risk of bias were generated using the risk-of-bias visualization tool [14] (Multimedia Appendix 3). Notably, 5 studies were funded by Zimmer Biomet [15-19], in 1 study [20], 360 Knee Systems supplied the wearable device, and in 1 study [21], the lead author was a creator of the MoveUp Application that was used to assess patient recovery.

Study Quality

The quality of the studies was assessed using RoB-2 for the 5 randomized controlled trials (RCTs) and ROBINS-I for the 30 nonrandomized studies. RCT data are presented in Table 1, and nonrandomized data in Table 2. For overall risk of bias, 5 studies were characterized as low risk, 16 studies were ranked moderate risk or some concerns, 11 studies were ranked serious risk, and 3 studies were characterized as not enough information.

XSL•FO
RenderX

Table . Risk of bias assessment for randomized controlled trials using the Cochrane Risk of Bias tool.

Study ID	Bias due to randomization	Bias due to the deviations from intended intervention	Bias due to missing outcomes data	Bias due to outcome measurement	Bias in the selection of the reported result	Overall
Paxton et al, 2018 [22]	High	Low	Low	Low	Low	Some concerns
Van der Walt et al, 2018 [23]	Low	Low	Low	Low	Low	Low
Duong et al, 2023 [24]	Low	Low	Low	Low	Low	Low
Mehta et al, 2020 [25]	No information	Low	Low	Low	Low	Some concerns
Pechstein et al, 2023 [26]	Low	Low	Low	Low	Low	Low

XSL•FO
RenderX

Table . Risk of bias in Nonrandomized Studies of Interventions.

Study ID	Bias due to confounding	Bias due to participant selection	Bias in classification of intervention	Bias due to deviations from intended intervention	Bias due to missing data	Bias in measurements of outcomes	Bias in the selection of the reported result	Overall
Tang et al, 2021 [27]	Moderate	Moderate	Low	Low	Moderate	Low	Low	Moderate
Crizer et al, 2017 [28]	Moderate	Moderate	Low	Low	No Information	Low	Low	No Information
Caliskan et al, 2020 [29]	Moderate	Serious	Low	Low	Low	Low	Low	Serious
Brandes et al, 2011 [30]	Moderate	Serious	Low	Low	Moderate	Low	Low	Serious
Frimpong et al, 2020 [31]	Low	Serious	Low	Low	Serious	Low	Low	Serious
Höll et al, 2018 [32]	Low	Serious	Low	Low	No Information	Low	Low	No Information
Redfern et al, 2023 [15]	Moderate	Moderate	Low	Low	No Information	Low	Low	No Information
Komann et al, 2024 [33]	Low	Low	Low	Low	Serious	Low	Low	Serious
Bolander et al, 2024 [34]	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate
Harding et al, 2014 [35]	Moderate	Moderate	Low	Low	Moderate	Low	Low	Moderate
Bendich et al, 2019 [36]	Moderate	Moderate	Low	Low	Serious	Low	Low	Serious
Twiggs et al, 2018 [20]	Moderate	Moderate	Low	Low	Moderate	Low	Low	Moderate
Bini et al, 2019 [37]	Moderate	Moderate	Low	Low	Serious	Low	Low	Serious
Shah et al, 2019 [38]	Moderate	Low	Low	Low	Moderate	Low	Low	Moderate
Frimpong et al, 2019 [39]	Moderate	Serious	Low	Low	Serious	Low	Low	Serious
Lebleu et al, 2021 [40]	Low	Moderate	Low	Low	Moderate	Low	Low	Moderate
Vitale et al, 2023 [41]	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate
van Diemen et al, 2023 [42]	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate
Gibian et al, 2023 [19]	Moderate	Moderate	Low	Low	Serious	Low	Low	Serious
Christensen et al, 2023 [16]	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate
Sato et al, 2023 [17]	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate
Fallahzadeh et al, 2023 [43]	Low	Moderate	Low	Low	Moderate	Low	Low	Moderate
Karas et al, 2020 [44]	Low	Low	Low	Low	Low	Low	Low	Low
Vitale et al, 2023 [41]	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate

XSL•FO
RenderX

Study ID	Bias due to confounding	Bias due to participant selection	Bias in classification of intervention	Bias due to deviations from intended intervention	Bias due to missing data	Bias in measurements of outcomes	Bias in the selection of the reported result	Overall
Lebleu et al, 2024 [21]	Low	Low	Low	Low	Low	Low	Low	Low
Harada et al, 2024 [45]	Low	Serious	Low	Low	Serious	Low	Low	Serious
Surmacz et al, 2024 [46]	Moderate	Low	Low	Low	Low	Low	Low	Moderate
Lyman et al, 2020 [47]	Low	Moderate	Low	Low	Serious	Low	Low	Serious
Pasqualini et al, 2024 [18]	Moderate	Moderate	Low	Low	Moderate	Low	Low	Moderate
Goeb et al, 2021 [48]	Low	Serious	Low	Low	Moderate	Moderate	Low	Serious

Systematic Review Procedure and Narrative Synthesis Process

The literature identified for review demonstrated considerable variability in variables collected via wearable devices as well as outcomes reported. As such, a meta-analysis was not performed. Rather, a systematic review was performed according to the Synthesis Without Meta-Analysis in systematic reviews guideline [49]. Studies were grouped according to outcomes reported, as demonstrated in Table 3. The results, limitations, and data presentation from our synthesis are also presented in Table 3.

XSL•FO
RenderX

Table . Papers included in the systematic review grouped by measurements of interest.

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
Physical activity and PROMs[c]								
Bendich et al, 2019 [36]	Determine if perioperative patient-generated health data, such as steps and activity levels, are associated with early post-op PROMs.	22	Both	Fitbit Flex Google LLC; wrist)	Steps; physical activity; HR[d]	KOOS[e] or HOOS[f]	• Pre-op:[g] N/A[h] • Post-op[i]: ≤6 weeks	Changes from pre-op to 6-weeks post-op levels steps are associated with improvements in PROMs. Only a weak association between absolute sensor values and PROM values after TJA[j].
Bini et al, 2019 [37]	Predict early post-op TJA outcomes using wearable sensor data and machine learning.	22	Both	Fitbit Flex (wrist)	Steps; physical activity; HR; personalized activity intelligence; cadence; bounce	KOOS or HOOS, VR-12[k]	• Pre-op: ≤4 weeks • Post-op: ≤8 weeks	Moderate correlation between post-op activity levels (steps, minutes active, and metabolic output) and VR-12 PCS[l] scores. No correlation between preoperative steps and PROMs.
Brandes et al, 2011 [30]	Measure change in physical activity and the relationship between objective function data and pain post-TKA.	53	Total knee	SAM[m] step activity monitor (Cyma Corporation; ankle)	Steps	KSS[n], SF-36[o]	• Pre-op: N/A • Post-op: ≤12 mo	Post-op physical activity is influenced by pre-op physical activity, though post-op physical activity at 12 mo is not correlated with KSS or SF-36 scores pre-op or post-op.
Caliskan et al, 2020 [29]	Identify accelerometer-recorded physical activity patterns in patients undergoing TKA in addition to ROM[p], strength, and PROMs.	36	Total knee	ActiCal (Respironics Inc; wrist)	Steps; physical activity; energy expenditure	ROM; muscle strength; KSS, OKS[q]	• Pre-op: ≤1 week • Post-op: ≤6 mo	KSS, OKS, light and moderate physical activity, steps, and daily energy expenditure significantly improved from baseline to 6 months post-op, while knee flexion and extension strength and sedentary behavior did not improve.

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
Christensen et al, 2023 [16]	Describe the trajectory of post-TKA recovery based on PROMs and smartwatch physical activity.	1005	Total knee	Apple Watch (Apple Inc; wrist)	Steps; stairs; gait speed; gait asymmetry	KOOS JR[r], EQ-5D	• Pre-op: N/A • Post-op: ≤12 mo	Perceived knee health and QoL[s] scores increased from pre-op to 12 months post-op while following a maximal recovery trajectory in the first 3 months. Steps per day increased to pre-op levels by 1 month and followed a maximal recovery trajectory in the first 3 months, while stairs per day returned to baseline by 3 months and continued to improve beyond 6 months.
Crizer et al, 2017 [28]	Track daily step count during pre-op and post-op periods and compare with PROMs.	589	Both	Smartphone (N/A)	Steps	LEFS[t]	• Pre-op: ≤4 weeks • Post-op: ≤12 weeks	LEFS scores only weakly correlated with objective step count, with increasingly weaker correlation with time post-op. No clear correlation at 12 weeks post-op suggests PROMs may only provide part of the story. Physical activity recovery is most rapid in THA patients, with most surpassing baseline step counts by 5 weeks compared to 9 in TKA.
Duong et al, 2023 [24]		102	Total knee	Fitbit Inspire (Google LLC; wrist)	Steps; physical activity; sleep		• Pre-op: N/A • Post-op: ≤12 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
	Compare a digital technology package with usual care in reducing pain among participants post-TKA.					NRS[u], PDI[v], Sedentary Behavior Questionnaire, AQoL-8D, Patient Activation Measure, Modified Computer Self-Efficacy Scale, Change in pain intensity (Global Rating of Change)		Patient-reported physical activity (AAS[w]) significantly increases across all participants from baseline (10 days post-op) to 3, 6, and 12 months. Fitbit physical activity measures (ie, steps and sedentary time) do not see the same significant increases. Sleep duration and time in bed remain unchanged from 3 to 6 months, though they significantly decrease from 6 months post-op to 12 months. Digital technology group demonstrates reduced pain intensity and PDI scores and increased physical activity (step count and stepping time).
Fallahzadeh et al, 2023 [43]	Use wearable device data and immune profiling to quantify physical recovery in individuals following THA.	49	Total hip	ActiGraph (Amertis LLC; wrist)	Steps; physical activity; sleep	WOMAC[x], SRS[y]	• Pre-op: ≤5 days • Post-op: ≤40 days	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Activity capacity, sleep disruption, and sedentary activity are most predictive of time after surgery, while light, moderate, and overall activity are less predictive in the model. Objective physical recovery trajectory is significantly associated with SRS, but not WOMAC, suggesting that smartwatch data may be more associated with self-reported fatigue and ability to perform daily activities than pain and overall physical function. Preop activity levels are not associated with post-op recovery speed.
Frimpong et al, 2020 [31]	Assess whether changes in objectively measured physical activity were associated with PROMs before and after TKA.	89	Total knee	activPAL (PAL Technologies Ltd; thigh)	Steps; sleep; sedentary time; standing time; walking time	KOOS, OKS	• Pre-op: ≤2 weeks • Post-op: ≤6 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Despite significant improvements in self-reported perceived knee pain and functional ability (KOOS and OKS) 6 months post-TKA, changes do not correlate with improvements in objectively measured (activPAL) light-intensity and sedentary activity behaviors. Average activity behaviors worsened 6 weeks post-op compared to pre-op, though by 6 months post-TKA, activity behaviors had improved to at or beyond pre-op levels.
Gibian et al, 2023 [19]	Measure patient sleep using a wearable device following TKA and compare with PROMs.	110	Total knee	Fitbit Inspire (wrist)	Sleep	PSQI[z], VAS[aa]	• Pre-op: N/A • Post-op: ≤90 days	Patient-reported sleep quality (very bad sleep) correlated well with 30-day VAS pain score, while sleep duration (monitored or patient-reported) did not correlate with any clinical measure and was not shown to be a useful metric in assessing TKA outcome.
Harada et al, 2024 [45]	Compare objective and subjective measurements of activity levels in patients undergoing THA pre-op and post-op.	42	Total hip	ActivPal3 (PAL Technologies Ltd; thigh)	Steps; upright time; sit-to-stand transitions	OHS[ab], UCLA[ac], FJS-12[ad]	• Pre-op: N/A • Post-op: ≤1 year	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Steps, OHS, and UCLA significantly improve from pre-op to 3 months and 1 year post-op, and age and hip abductor muscle volume are significant predictors of post-op step counts, while pre-op UCLA score does not predict postoperative objective activity data.
Harding et al, 2014 [35]	Measure physical activity pre-op and (6 months) postarthroplasty and examine the proportion of people meeting the American Physical Activity Guidelines.	44	Both	ActiGraph GT1M (Ametris LLC≈; wrist)	Physical activity	NPRS[ae], OHS/OKS, SF-12[af], UCLA	• Pre-op: N/A • Post-op: ≤6 mo	No change in objectively measured physical activity after arthroplasty, despite significant improvements in PROMs post-THA/TKA. There were significant improvements in self-reported pain during activity, physical function, QoL, and in physical activity measured by the UCLA activity score after arthroplasty.
Höll et al, 2018 [32]	Measure activity levels before and after THA.	46	Total hip	StepWatch 3 (Modus Health LLC; ankle)	Steps; physical activity	HHS[ag], WOMAC	• Pre-op: ≤1 week • Post-op: ≤3 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Gait cycles (1 gait cycle =approximately 2 steps) and time in moderate to vigorous physical activity did not increase from pre-op to 6 weeks, but did increase at 3 mo follow-up, while PROMs increased at both follow-up time points. Gait cycles per day did not correlate with PROMs at any time point.
Karas et al, 2020 [44]	Use patient-generated health data from a wearable device to predict recovery trajectories from lower limb surgeries.	196		Fitbit (Google LLC; wrist)	Steps; HR; sleep	Self-reported recovery time	• Pre-op: ≤26 weeks • Post-op: ≤26 weeks	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Daily steps return to baseline in ≈12 weeks post-op, while 95th percentile HR returns to baseline in roughly 4 weeks. Sleep efficiency does not follow a specific trajectory pre-op or post-op, though there are significant reductions in sleep efficiency from baseline up to 24 weeks post-op. Long-term recovery can be accurately predicted on an individual level 1 month after surgery. Further, 12 weeks pre-op and 26 weeks post-op trajectories of daily behavioral measurements (steps sum, HR, and sleep efficiency score) can capture activity changes relative to an individual's baseline.
Komann et al, 2024 [33]	Assess the relationship between postoperative physical activity and PROMs.	55 (TKA)	Total knee	ActiGraph wGT3X (Amertis LLC; wrist)	Steps	BPI[ah], IPO[ai] Questionnaire	• Pre-op: N/A • Post-op: ≤7 days	There is no statistically significant correlation between steps and average pain intensity post-op.
Lebleu et al, 2021 [40]		146	Both	Nokia Go (Nokia Corporation; wrist)	Steps	KOOS or HOOS	• Pre-op: ≤1 week • Post-op: ≤3 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
	Assess physical activity recovery post-op and determine perioperative features that could help predict physical activity recovery at 3 mo with telerehabilitation intervention.							Patients receiving telerehabilitation reached pre-op step count at 7 weeks post-op. Change in step counts post-op at 3, 6, and 11 weeks could be predicted by pre-op step count and post-op length of time on crutches. Further, 3-month step increases are additionally associated with pre-op symptoms level, while early (3 - 6 weeks) step count improvements are associated with pre-op QoL.
Lyman et al, 2020 [47]	Investigate the feasibility of mobile technology to collect daily step data and biweekly PROMs to track TJA recovery: (1) what proportion of patients provided full 6-month follow-up step and PROMs information, (2) whether THA and TKA patients recovered differently, (3) if new patterns of recovery could be discovered, and (4) how well step counts and PROMs scores were correlated.	267	Both	Smartphone accelerometer (N/A)	Steps	NRS, HOOS JR or KOOS JR	• Pre-op:≥2 days • Post-op: ≤6 mo	Recovery was faster by all metrics in THA than in TKA patients, correlations between step counts and PROMs were modest, and 2 distinct clusters of recovery could be described using step counts.
Pasqualini et al, 2024 [18]		2333	Total knee	Apple Watch (Apple Inc; wrist)	Steps; stairs	KOOS JR, NRS	• Pre-op:≥14 days • Post-op: ≤3 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
	Investigate the correlation between pain and KOOS JR with step and stair flight counts in patients undergoing TKA.							Step count is correlated with NRS and KOOS JR, both pre-op and at 1 and 3 mo post-op in TKA patients, while stair flight count is significantly correlated with PROMs pre-op and at 1 month, though no longer associated at 3 months.
Redfern et al, 2023 [15]	Investigate objective measures and PROMs as a function of preoperative PA[aj] levels in TKA patients.	1941	Total knee	Apple Watch (Apple Inc; wrist)	Steps; physical activity	EQ-5D-5L, EQ-VAS, KOOS JR, KSS, NRS	• Pre-op: ≤2 weeks • Post-op: ≤12 mo	Patients in the 2 middle quartiles of wearable device-measured pre-op step counts recovered activity with a modest increase in step counts 1 year after surgery, while those in the highest quartile of activity did not return to preoperative levels at any interval through 1 year. The lowest-activity participants reported greater improvements in joint function, pain, and health-related QoL. Despite this, satisfaction scores did not vary between groups.
Sato et al, 2023 [17]	Identify the trajectory of physical activity and PROMs after THA.	1898	Total hip	Apple Watch (Apple Inc; wrist)	Steps; stairs; walking asymmetry; gait speed	HOOS JR, EQ-5D	• Pre-op: N/A • Post-op: ≤1 year	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								At all post-op time points, THA patients see statistically significant increases in steps, stair flights, walking asymmetry, and gait speed, though only steps and stair flights were clinically significant at 3, 6, and 12 months. Meanwhile, PROM improvements are clinically significant at all time points. Greatest PROM improvements were reached by 1 mo and physical activity by 3 months.
Shah et al, 2019 [38]	Determine the frequency and timeline of wearable device activity data that correlate with early PROMs following TJA.	22	Both	Lumo Lift (Lumo Inc, ND[ak])	Steps; bounce; braking; drop or roll; pelvic rotation or yaw; or pelvic tilt or pitch	VR-12 PCS, HOOS or KOOS	• Pre-op: N/A • Post-op: 42 days	Raw wearable device data predicts 6-week post-op outcomes significantly better than 24-hour summarized data does, and clinical outcomes may be predictable by wearable device data as early as 1 - 2 weeks post-op. Increasing sampling frequency above the standard 24-hour average provided by consumer-grade activity sensors improves the ability of machine learning algorithms to predict 6-week PROMs.
Surmacz et al, 2024 [46]		327	Both			HOOS or KOOS JR		

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
	Determine the association between pain and PROMs or objective mobility data.			Apple Watch (Apple Inc; wrist)	Steps; gait speed; gait asymmetry		• Pre-op: ≤30 days • Post-op: ≤90 days	Pain scores are associated with gait speed at all 3 time points, whereas step count was not significantly associated with pain at any time point.
Tang et al, 2021 [27]	Compare accelerometer-measured physical activity and sleep to PROMs following THA.	41	Total hip	Fitbit Flex (wrist)	Steps; sleep	HOOS JR	• Pre-op: N/A • Post-op: ≤3 mo	No correlation between accelerometry-measured activity level nor sleep and functional PROM scores in the early follow-up period (within 3 months) following THA. Average daily steps and minutes slept pre-op were similar to average daily steps and minutes slept 3 months post-op. Weak negative correlation between average steps pre-op and pre-op HOOS JR. Significant improvement in patients' average daily sleep occurs between 1 - 2 weeks and 1 month post-op.
Twiggs et al, 2018 [20]	Assess pre-op and early post-op physical activity in TKA patients and determine benchmarks for expected activity.	91	Total knee	Fitbit Flex (wrist)	Steps	SF-12, KOOS	• Pre-op: ≤2 weeks • Post-op: ≤6 weeks	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								SF-12 PCS and KOOS ADL were significantly correlated with step counts at all 3 time points. KOOS QoL was significantly associated with post-op 6-week step count, while KOOS pain was significantly associated with post-op day 2 - 4 step count and pre-op step count. Pre-op step count is both strongly significantly correlated to and not statistically different from 6-weeks post-op step count. BMI and SF-12 PCS scores may also be predictive of post-op recovery.
Van der Walt et al, 2018 [23]	Determine if feedback from a wearable device improves activity levels over 6 weeks post-op following TJA.	163	Both	Garmin Vivofit 2 (Garmin Ltd; wrist)	Steps	KOOS, EQ-5D	• Pre-op: ≤2 weeks • Post-op: ≤6 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Feedback from wearable activity trackers increases daily activity in the acute post-op phase and at 6 months. The greatest increase in average daily steps post-op for both groups occurs during the 1 - to 2-week period. Most patients surpass their pre-op step counts after 6 weeks post-op. Patients who received feedback from a commercial wearable device activity tracker with a daily step goal had significantly higher activity levels after hip or knee arthroplasty over 6 weeks and 6 months than those who did not receive feedback.
Vitale et al, 2021 [50]	Assess the objective physical activity and rest-activity daily rhythm in TJA patients.	20	Both	Actiwatch 2 (Amertis LLC; wrist)	Physical activity; RAR[al]	FIM, Barthel Index	• Pre-op: 1 day • Post-op: 10 days	Patients restored and maintained a significant and physiological RAR beginning 2 days post-op. Improvements can be seen in physical activity, BI[am], FIM[an], TUG[ao], pain, and 10MWT[ap] as early as 10 days post-op.
Physical activity and performance-based measures								

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
Paxton et al, 2018 [22]	Evaluate the feasibility of a physical activity feedback intervention for patients after TJA.	45	Both	Fitbit Zip (Google LLC; waist) and GT3X ActiGraph Activity Monitor (wrist) (Amertis LLC)	Steps; physical activity	6MWT, TUG; gait speed (m/s)	• Pre-op: N/A • Post-op: ≤12 weeks	Intervention participants (physical activity feedback) improved daily step count from baseline to 12 weeks by 20% compared to 9% for the control group. 6MWT increased 20% and 7% while TUG improved 1% and 4% for intervention and control groups, respectively. Goal setting needs to be adjusted to set realistic improvements.
Pechstein et al, 2023 [26]	Relate walking endurance with physical activity post-TKA and determine if this relationship differs between above vs below age-normative endurance.	109	Total knee	ActiGraph GT3X (Amertis LLC; waist)	Steps; physical activity	6MWT	• Pre-op: N/A • Post-op: ≤12 mo	Potential dose-response relationship between walking endurance and PA in adults after TKR[aq], which may be present in those below age-normative walking endurance standards but absent in those with restored walking endurance.

Physical activity, PROMs, and performance-based measures

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
Bolander et al, 2024 [34]	Characterize presurgical behavior using wearable devices and compare to post-op.	82	Total knee	TracPatch Duo (TracPatch Health LLC, knee)	Steps; flexion or extension	ROM; VAS	• Pre-op: ≤6 days • Post-op: ≤50 days	Daily step count and flexion return to baseline at 6- to 7-weeks post-op. Extension is regained within 1 week post-op. Pain scores improve past the pre-op baseline by 3 weeks post-op.
Frimpong et al, 2019 [39]	Describe changes in objectively measured physical activity and sedentary behavior in patients undergoing TKA.	79	Total knee	ActiGraph GT3X (wrist)	Steps; physical activity; sleep	ROM, UCLA, WOMAC	• Pre-op: ≤7 days • Post-op: ≤6 mo	While PROMs and knee ROM significantly improve from baseline at both 6 weeks and 6 months post-TKA, steps per day decrease significantly at 6 weeks, then significantly increase beyond baseline between 6 weeks and 6 months. Further, overall sedentary time and time spent in light physical activity improve significantly at 6 months, while moderate to vigorous physical activity does not. Clinically, functional improvements in patients post-TKA may be assessed by objectively measuring changes in low-intensity activity behaviors.
Goeb et al, 2021 [48]		82	Total hip	Letscom (Letscom International Corporation, Ltd; wrist)	Steps		• Pre-op: ≤1 week • Post-op: ≤6 weeks	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
	Assess early outcomes of the posterior THA approach with modified precautions with the aid of wrist-based activity trackers.					Pain medication use, return to work or driving, assistive device use, HOORS JR, LEFS, pain		Patients consistently increased their average daily step count over the first 6 weeks post-op in a nearly linear fashion, with a slight reduction in the rate of improvement in weeks 5 and 6. Improvement in steps significantly correlated with changes in HOOS JR, LEFS, and pain scores over time.
Lebleu et al, 2024 [21]	Examine the incorporation of unsupervised evaluations and continuous monitoring with an activity tracker in the rehabilitation of total joint surgery.	1144	Both	Garmin Vivofit 4 (Garmin Ltd; wrist)	Steps; physical activity	6MWT, OKS, FJS[ar], HOOS or KOOS, UCLA, EQ-5D	• Pre-op: ≤1 week • Post-op: ≤3 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Cadence measurements (ie, P6MC[as] and P1M[at]) may offer more stable indicators of gross motor function compared to the number of steps per day. Patients typically resumed their pre-op activity levels 6 - 10 weeks after undergoing surgery. Age significantly affects steps per day and cadence in TKA/THA. Patients who met MCID[au] for FJS at 3 months surpassed baseline physical activity metrics, particularly light intensity activity, between 25 and 50 days post-op. Patients who did not meet the MCID for FJS were less likely to surpass baseline PA.
Mehta et al, 2020 [25]	Evaluate the effect of activity monitoring and bidirectional text messaging on the rate of discharge to home and clinical outcomes in TJA patients.	242	Both	Withings (Apple Inc, ND)	Steps	TUG, rehospitalizations, ED[av] visits, PT visits, pain scale	• Pre-op: N/A • Post-op: ≤6 weeks	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								No significant increase in step count in those receiving remote monitoring plus gamification and social support compared to remote monitoring alone. Activity monitoring might have encouraged increases in step count, which may have improved the recovery function. Mean increase in daily step counts of 833 was found in intervention and control groups from 2 to 6 weeks post-op.
van Diemen et al, 2023 [42]	Assess the relationship between mitochondrial function at baseline and recovery of mobility following TKA.	30	Total knee	Nokia Go (Nokia Corporation; wrist)	Steps	TUG, quadriceps strength, grip strength, KOOS	• Pre-op: ≤2 weeks • Post-op: ≤6 mo	Pre-op mitochondrial function, grip strength, and steps correlate significantly with post-op step recovery. Post-op activity correlated significantly with KOOS-ADL[aw] and TUG, but not with quadriceps or grip strength. Median recovery of steps was 101 steps per week (32-392 steps). Per 1 step of activity, recovery increases by 0.02 steps/weeks (P=.005).
Sleep								

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
Duong et al, 2023 [24]	Compare a digital technology package with usual care in reducing pain among participants following TKA.	102	Total knee	Fitbit Inspire (wrist)	Steps; physical activity; sleep	NRS, Pain Disability Index, Sedentary Behavior Questionnaire, AQoL-8D, Patient Activation Measure, Modified Computer Self-Efficacy Scale, Change in pain intensity (Global Rating of Change)	• Pre-op: N/A • Post-op: ≤12 mo	While patient-reported PA (AAS) significantly increases across all participants from baseline (10 days post-op) to 3, 6, and 12 months, Fitbit-recorded physical activity measures, such as steps and sedentary time, do not see the same significant increases, although there was no correlation analysis between the outcomes. Sleep duration and time in bed appear to remain unchanged from 3 to 6 months, though they significantly decrease from 6 to 12 months following TKA. The intervention significantly reduced pain intensity, improved the Pain Disability Index score, and increased physical activity (step count and stepping time).
Fallahzadeh et al, 2023 [43]	Use wearable device data and immune profiling to quantify physical recovery in individuals following THA.	49	Total hip	ActiGraph (wrist)	Steps; physical activity; sleep	WOMAC, SRS	• Pre-op: ≤5 days • Post-op: ≤40 days	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Activity capacity, sleep disruption, and sedentary activity are most predictive of time after surgery, while light, moderate, and overall activity are less predictive in the model. Objective physical recovery trajectory is significantly associated with SRS, but not WOMAC, suggesting that smartwatch-derived data may be more associated with self-reported fatigue and ability to perform daily activities than pain and overall physical function. Pre-op activity levels are not associated with postoperative recovery speed.
Frimpong et al, 2020 [31]	Assess whether changes in objectively measured physical activity were associated with PROMs before and after TKA.	89	Total knee	activPAL (thigh)	Steps; sleep; sedentary time; standing time; walking time	KOOS, OKS	• Pre-op: ≤2 weeks • Post-op: ≤6 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								Despite significant improvements in self-reported perceived knee pain and functional ability 6 months post-TKA, these changes do not correlate with improvements in objectively measured (using activPAL) light-intensity and sedentary activity behaviors. On average, activity behaviors worsened 6 weeks after surgery compared to preoperative levels, though by 6 months post-TKA, activity behaviors either were similar to pre-op levels or had improved beyond pre-op levels, such as in the case of step measures.
Gibian et al, 2023 [19]	Measure patient sleep using a wearable device following TKA and compare with PROMs.	110	Total knee	Fitbit Inspire (wrist)	Sleep	PSQI, VAS	• Pre-op: N/A • Post-op: ≤90 days	Patient-reported sleep quality (very bad sleep) correlated well with VAS pain score at 30 days, while sleep duration (monitored or patient-reported) did not correlate with any clinical measure and was not shown to be a useful metric in assessing TKA outcome.
Karas et al, 2020 [44]		196		Fitbit (wrist)	Steps; HR; sleep	Self-reported recovery time	• Pre-op: ≤26 weeks • Post-op: ≤26 weeks	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
	Use patient-generated health data from a wearable device to predict recovery trajectories from lower limb surgeries.							Daily steps return to baseline in ≈12 weeks post-op, while 95th percentile HR returns to baseline in roughly 4 weeks. Sleep efficiency does not follow a specific trajectory pre-op or post-op, though there are significant reductions in sleep efficiency from baseline up to 24 weeks post-op. Long-term recovery can be accurately predicted on an individual level 1 month after surgery. Further, 12 weeks pre-op and 26 weeks post-op trajectories of daily behavioral measurements (steps sum, HR, and sleep efficiency score) can capture activity changes relative to an individual's baseline.
Tang et al, 2021 [27]	Compare accelerometer-measured physical activity and sleep to PROMs following THA.	41	Total hip	Fitbit Flex (wrist)	Steps; sleep	HOOS JR	• Pre-op: N/A • Post-op: ≤3 mo	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								No correlation between accelerometry-measured activity level or sleep and functional PROM scores in the early follow-up period (within 3 months) following TJA (THA). Average daily steps and minutes slept pre-op were statistically like average daily steps and minutes slept 3 months post-op. Weak negative correlation between average steps pre-op and pre-op HOOS JR. Significant improvement in patients' average daily sleep occurs between 1 - 2 weeks and 1 month post-op.
Vitale et al, 2023 [41]	Assess wearable device-derived sleep characteristics and pain scores in patients hospitalized for recovery from TJA.	20	Both	Actiwatch 2 (Amertis LLC; wrist)	Sleep	TUG, 10MWT, PSQI, VAS	• Pre-op: 1 day • Post-op: 10 days	

XSL•FO
RenderX

Study ID	Aim of the study	Participants (n)	TKA?[a] THA?[b] both?	Device (placement)	Accelerometer data collected or assessed	Recovery assessments	Timeline	Summary of findings
								ActiGraph parameters of sleep quantity and timing did not differ from pre-op to 10-day post-op periods in TJA patients, though sleep efficiency and quality drastically decreased the first night after the surgery before improving in the following 9 days. There was a significant negative correlation between pain scores and sleep quality measures (sleep efficiency, immobility time, and fragmentation index).

[a]TKA: total knee arthroplasty.
[b]THA: total hip arthroplasty.
[c]PROM: patient-reported outcomes measure.
[d]HR: heart rate.
[e]KOOS: Knee Injury and Osteoarthritis Outcome Score.
[f]HOOS: Hip Disability and Osteoarthritis Outcome Score.
[g]pre-op: preoperative.
[h]N/A: not available.
[i]post-op: postoperative.
[j]TJA: total joint arthroplasty.
[k]VR-12: Veterans Rand 12-Item Health Survey.
[l]PCS: Physical Composite Score.
[m]SAM: step activity monitor.
[n]KSS: Knee Society Score.
[o]SF-36: 36-Item Short Form Health Survey.
[p]ROM: range of motion.
[q]OKS: Oxford Knee Score.
[r]JR: joint replacement.
[s]QoL: quality of life.
[t]LEFS: Lower Extremity Functional Scale.
[u]NRS: Numeric Pain Rating Scale.
[v]PDI: Pain Disability Index.
[w]AAS: Active Australia Survey.
[x]WOMAC: Western Ontario and McMaster Universities Arthritis Index.
[y]SRS: Surgical Recovery Scale.
[z]PSQI: Pittsburgh Sleep Quality Index.

XSL•FO
RenderX

[aa]VAS: Virtual Analog Scale.
[ab]OHS: Oxford Hip Score.
[ac]UCLA: University of California, Los Angeles Activity Scale.
[ad]FJS-12: Forgotten Joint Score 12.
[ae]NPRS: Numeric Pain Rating Scale.
[af]SF-12: 12-Item Short Form Health Survey.
[ag]HHS: Harris Hip Score.
[ah]BPI: Brief Pain Inventory.
[ai]IPO: International Pain Outcomes Questionnaire.
[aj]PA: physical activity.
[ak]ND: not defined.
[al]RAR: rest activity daily rhythm.
[am]BI: Barthel Index.
[an]FIM: Functional Independence Measure.
[ao]TUG: timed-up-and-go.
[ap]MWT: minute walking test.
[aq]TKR: total knee replacement.
[ar]FJS: Forgotten Joint Socre.
[as]P6MC: peak 6-minute consecutive cadence
[at]P1M: peak 1-minute cadence.
[au]MCID: minimal clinically important difference.
[av]ED: emrgency department.
[aw]KOOS-ADL: Knee Injury and Osteoarthritis Outcome Score-Activities if Daily Living.

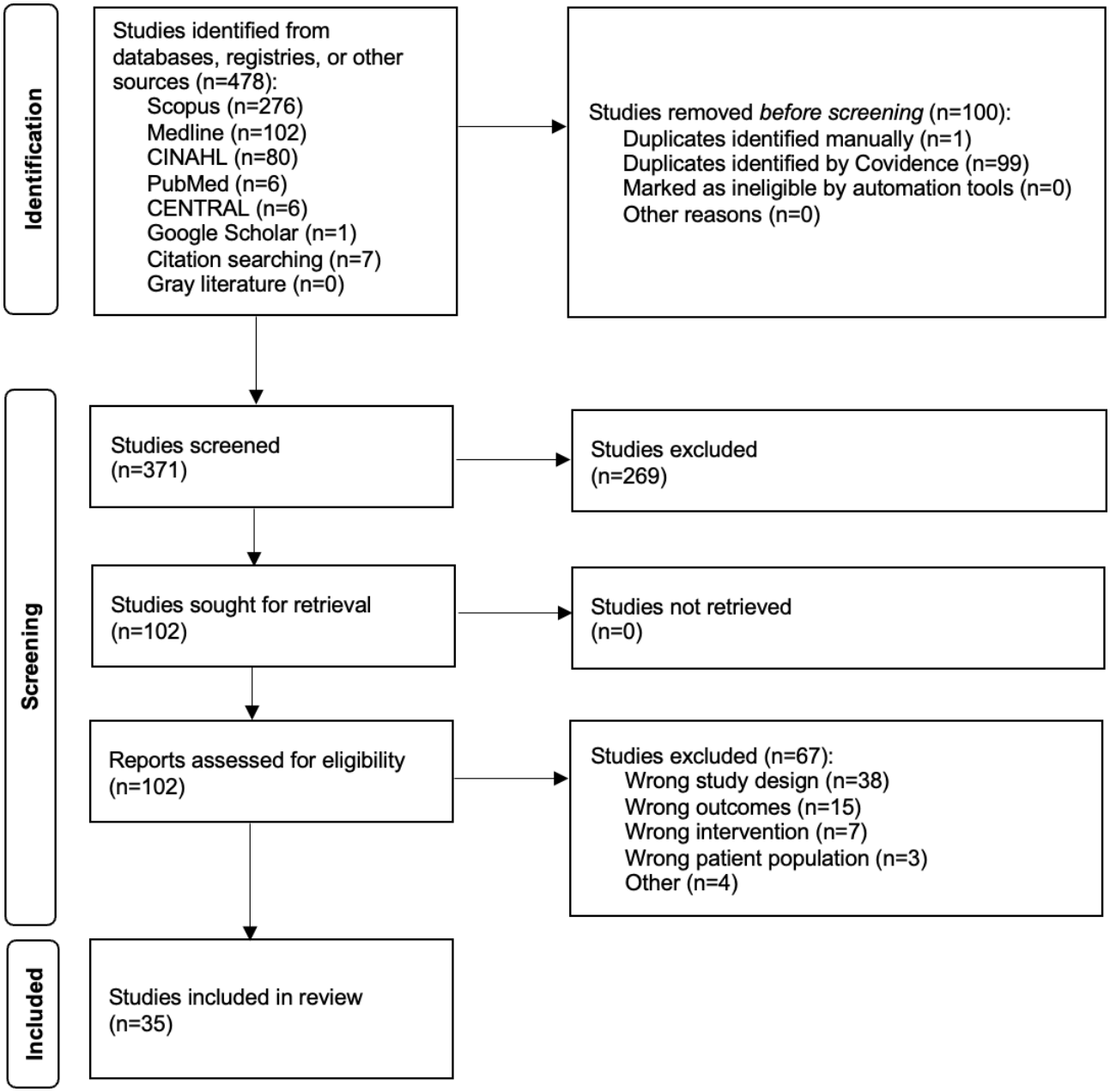
Results

Overview

This systematic review analyzed 35 studies that used wearable devices to assist with patient monitoring in the TJA perioperative period. These papers detail how monitoring with wearables is used in conjunction with PROMs and functional metrics in the recovery of patients undergoing TJA. The final analysis included 35 papers, including 5 RCTs and 30 nonrandomized papers (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA flow diagram. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Physical Activity and Performance-Based Functional Outcomes

Of the 35 papers included in this study, 32 collected step count, making it the most common data point reported (Table 3). Another component of assessing functional outcomes in TJA is performance-based tests, such as the 6-minute walking test (6MWT), range-of-motion, or timed-up-and-go (TUG; Table 3). A total of 9 studies used both wearable device data and standardized performance-based functional outcomes in their assessments (Table 3).

Physical Activity and PROMs

Physical activity and PROMs used were not consistent among studies (Table 3). Joint-specific PROMs, including Knee Society Score, Knee Injury and Osteoarthritis Outcome Score (KOOS), Oxford Knee Score, Hip Disability and Osteoarthritis Outcome Score (HOOS), and Harris Hip Score, were used in 21 studies. Virtual Analog Scale (VAS) was used in 4 studies [15,19,34,41] and Numerical Rating Scale for pain was used in 4 studies [15,18,24,47]. EQ-5D scores were examined in 4 studies [15,16,21,23], and 3 used Western Ontario and McMaster Universities Arthritis Index (WOMAC) [32,39,43]. The Lower Extremity Functional Scale (LEFS) was used in 2 studies [48,51]. The Veterans Rand 12-Item Health Survey was used in 2 studies [37,38], and 1 used SF-36 (36-Item Short Form Health Survey) [30].

Sleep

Sleep as recorded by wearable devices was reported in 7 of the included studies (Table 3). Gibian et al [19] recorded sleep duration and reported a negative correlation with VAS pain scores [35]. A total of 2 additional studies [17,20,24] also report sleep duration as recorded by wearable devices. Frimpong et al [31] use a validated algorithm to report sleep data as recorded by wearables. Karas et al [44] report “sleep efficacy,” which is

XSL•FO
RenderX

the proportion of minutes asleep to the total time in bed. Finally, Fallahzadeh et al [43] and Vitale et al [41] focused on multiple components of sleep, with the former including overall sleep and sleep quality and the latter recording 9 components of sleep.

Rehabilitation Workflow

Interestingly, only 1 study used a novel application to assess patients [28]. This study demonstrated that step counts as recorded in the MoveUp application correlated with LEFS following TJA. However, the authors did not mention how this application can be used to create tailored rehabilitation protocols.

Discussion

Principal Findings

This systematic review found that most of the literature regarding the utility of wearable devices to monitor TJA outcomes includes step count. In addition to recording step count, these devices can record data regarding gait speed and asymmetry as well. Assessment of such metrics allows providers to evaluate gait comprehensively. Nevertheless, few studies used this capability of wearables [16,17,22,46]. Further, there is a large variability in how wearable device data is used in conjunction with other outcome measures in the current literature. Further, while ample studies include PROMs in addition to wearable data, many different PROMs are used. As demonstrated in Table 3, wearable devices have the capability of measuring multiple objective metrics. While this capability has the potential to allow holistic assessment of patient recovery, there is currently no consensus on which metrics are most valuable to record and assess in the perioperative setting. Further, the studies reviewed are not uniform in when patients are assessed, with certain groups analyzing patients days to weeks after surgery, while others continue for 6 to 12 months. The heterogeneity among these studies demonstrates that the ideal manner in which surgeons can use wearable devices for remote monitoring requires further research. We suggest that stratification regarding the reporting of perioperative data obtained by wearables, and that coupling this data with a unified PROM, will maximize the utility of wearable devices as a tool for assessing patient recovery.

The relationship between wearable-derived physical activity and PROMs, such as pain, quality of life, and self-reported function scores, was explored by 21 studies. The most common PROMs reported were joint-specific. Multiple studies found a correlation between change in step count and KOOS, KOOS JR (KOOS-joint replacement), HOOS, or HOOS JR [15,18,20,36,38,42,47,48]. Both the KOOS and HOOS were developed particularly for patients with osteoarthritis and have long been used to track progress following TJA. Such studies support the use of wearable devices as another component of comprehensive monitoring of TJA recovery, particularly since such devices allow continuous monitoring. However, the current literature does not yet suggest a data point obtained via remote wearables that necessitates intervention or that intervention improves KOOS or HOOS.

Contrastingly, some groups found that steps did not correlate with other joint-specific PROMs. Frimpong et al 2020 [31] report that there is no correlation between KOOS and an increase in physical activity. Additionally, Tang et al [27] reported that there is no correlation between accelerometry-measured activity level and HOOS JR within 3 months of TJA. Multiple studies found no correlation between step count and WOMAC [31,43] or UCLA (University of California, Los Angeles Activity Scale) activity index [35,45]. It is unclear why many studies found that step count correlated with KOOS and HOOS, but not WOMAC or UCLA activity index. One suggestion is that HOOS and KOOS have been shown to better detect changes over time, while WOMAC and UCLA are more adapted for more active populations [52,53]. Regarding the UCLA activity index, it is possible to have large increases in step number but not have an increase in activity level on this scale.

In addition to joint-specific PROMs, other measures were used. Notably, 3 studies [28,43,48] found relationships between perioperative physical activity and patients' ability to perform everyday tasks as measured by the LEFS and Surgical Recovery Scale. Others [20,36-38] found a positive correlation between wearable device reported step count and general health and physical functioning, as measured by the Veterans Rand 12-Item Health Survey PCS (Physical Composite Score) or SF-12 (12-Item Short Form Health Survey) PCS scores.

Further, some of the reviewed studies include functional outcome measures. Performance-based tests, such as the 6MWT and TUG, allow standardized assessment of patients. These tests are normally used immediately following surgery, before patient discharge, and then repeated at certain intervals in the clinic. As such, these tests allow serial, objective evaluation of patient recovery. Paxton found that both 6MWT and step count increased 20% from baseline at 12 weeks post-TKA, while TUG improvements are present but less robust [22]. Similarly, Pechstein et al [26] found a dose-response relationship between 6MWT and device-measured physical activity. Improvements in physical activity and TUG can be seen as early as 10 days postoperatively [41], and van Diemen et al [42] found a significant correlation between activity and TUG. Step count and 6MWT both measure endurance in walking, while TUG focuses on mobility. This correlation suggests that data collected by wearable devices might serve as a surrogate for these objective assessments. In contrast to TUG and 6MWT, wearable devices allow assessments to be performed continuously and remotely. As such, patients can be assessed in between clinic visits. When patients deviate from expected progress or demonstrate setbacks, tailored protocols can be developed.

While this review found that most wearable device literature includes step count, many wearable devices have the capacity to monitor additional variables, such as balance, as well. Further investigation could examine how such parameters correlate with TUG and other performance measures.

Further, because wearable devices are used during the preoperative period, they allow providers to obtain a patient baseline before TJA. The collection of this data allows recovery trajectories to be crafted. Patients demonstrate decreases in activity, balance, and gait speed acutely postoperatively, but then return to or surpass their baseline postoperatively. In the acute postoperative setting, knowledge of patient deviation from

XSL•FO
RenderX

baseline can inform discharge protocols to allow patients to return to benchmarks. As patients progress in their recovery, comparison of metrics to preoperative and immediate postoperative data informs the multidisciplinary team about patient recovery. If a patient is not meeting a step count goal, therapy can be tailored toward endurance. Contrastingly, if a patient is struggling with balance, therapy can be adjusted to focus on improving this deficit. As the number of patients who have data collected increases, there is also potential to create expected recovery trajectories for patients based on baseline preoperative data, allowing providers to provide realistic and individualized recovery expectations.

In addition to papers that focus on activity, 7 of the included studies use wearable devices to monitor sleep in the TJA perioperative period. It has been well established that sleep and pain are closely associated, and it has been further shown that higher sleep quality is associated with better acute pain relief following TKA [28,29]. Wearable devices were shown to both increase physical activity on the Active Australia Survey and reduce time in bed and sleep duration [24]. This study did not report on sleep quality as reported by the wearable device. Another study found that both daily steps and minutes slept returned to preoperative levels 3 months post THA. Again, sleep quality as measured by wearable was not reported. Karas et al [44] report a "sleep efficacy score," which is the proportion of minutes asleep to the total time in bed. They find that reductions in sleep efficiency from baseline up to 24 weeks postoperatively, but that daily steps return to baseline 12 weeks after surgery.

Gibian et al [19] found that patient-reported sleep quality, particularly poor-quality sleep as defined by the Pittsburgh Sleep Quality Index, correlated well with VAS pain score at 30 days postoperatively. However, they found no relationship between device-recorded sleep duration and VAS. Further, they found no correlation between overall Pittsburgh Sleep Quality Index scores and device-recorded hours of sleep at 30, 60, or 90 days postoperatively. Finally, 1 study used wearable devices to monitor multiple components of sleep, including overall sleep, sleep onset, and sleep disruption. Sleep disruption increased after surgery, but overall sleep did not. Sleep duration as an isolated metric may lack sensitivity in predicting changes in a patient's recovery timeline following TJA, but fortunately, smartwatches and activity trackers can collect more advanced measurements of sleep. Taken together, the current literature that uses wearable devices to monitor sleep in TJA is heterogeneous. Different groups report an array of variables. This underscores the immense capacity of wearable devices to track multiple domains and also highlights the fact that more uniform outcomes reporting would allow studies to be more easily compared. Further, despite sleep data being collected and reported, current literature does not report on intervening on patients with altered sleep components as reported by wearables.

Notably, 5 of the 35 reviewed studies were sponsored by Zimmer Biomet [15-19], 1 study used a mobile application the lead author created [21], and 1 study used a wearable supplied by 360 Knee Systems [20]. Redfern et al [15] report that patients in the lowest preoperative step count group have lower KOOS JR scores and demonstrate greater postoperative improvement in KOOS JR. Similarly, Pasqualini et al [18] demonstrate that step counts correlate with HOOS JR both preoperatively and at 1- and 3-month follow-up. Further, 2 industry-funded studies [16,17] suggest that PROM surpasses minimal clinically important difference earlier in recovery than activity as measured by wearables. Twiggs et al [20] found a positive correlation between postoperative step count and KOOS activities of daily living subscore. Lebleu et al [21] perform an analysis using the MoveUp application the lead author created. They find that peak cadence improved beyond minimal clinically important difference more quickly than step count in both TJA and TKA recovery. Taken together, the industry-funded studies suggest that PROMs may improve faster than wearable device measurements and that some PROMs correlate with step counts or other measures of physical activity. Further, this research demonstrates the heterogeneity in the types of wearable data reported and the PROMs used among different studies.

The major strength of this study is the comprehensive examination of existing literature. Multiple databases were searched to ensure a comprehensive review of the literature. Further, 2 authors independently extracted data and then collaborated for consensus regarding study eligibility. There was also an agreement between reviewers regarding bias assessment according to the RoB-2 or ROBINS-I tools. This review reports on the vast capabilities of wearable devices, ranging from tracking step counts and sleep to more sophisticated measures. Moreover, we report how wearable device data correlates with validated PROMs and outcome measures, emphasizing their potential to revolutionize postoperative outcome monitoring. Wearable devices also encourage patient participation in rehabilitation protocols and offer patients objective data by which to monitor their postoperative progress. The studies included for review demonstrate high patient participation and compliance.

There are several limitations to this review. As with all reviews, the quality of the reviewed studies is a limitation. Our review included only 5 RCTs, with 30 studies representing nonrandomized trials, and these have lower levels of evidence. As discussed, the heterogeneity among study methods makes direct comparisons difficult. Different wearables have different capabilities. Further, different groups collected data at different time points or had patients wear devices in different locations. While the included studies provided patients with the wearable devices, requiring patients to use a wearable is a source of bias. Patients who are familiar with these devices may have higher technological literacy, socioeconomic status, or ability to care for themselves. As such, certain populations may be excluded from these studies. Additionally, the use of a wearable device on the wrist immediately postoperatively can result in falsely low recordings if patients are using an assistive device such as a walker or cane.

Finally, we propose that remote monitoring via wearable devices provides a practical method for tracking recovery, supporting patient engagement, and detecting deviations from expected progress. Current literature does not agree on the optimal data that wearables provide or when data collection should be performed. Despite this, the studies included in this review demonstrate that variables collected via wearable devices correlate with certain PROMs, pain measures, and objective

XSL•FO
RenderX

postoperative tests. As such, monitoring patients with wearables may allow a more holistic assessment of patient recovery. Further, wearable devices have the potential to allow analysis informed by baseline preoperative activity, which can enhance preoperative counseling and personalize recovery expectations. Lastly, continuous monitoring may also allow early identification of delayed recovery, enabling timely intervention and tailored rehabilitation. Further studies can explore the ideal time in recovery, variables to assess, and the combination of data collected with additional outcome measures that allow wearable devices to most effectively assess outcomes following TJA.

Conclusions

The increasing prevalence of wearable devices presents significant opportunities for advancing postoperative care through continuous health monitoring and data collection. Remote monitoring using these devices allows surgeons to gather extensive and continuous data, which can be crucial for assessing outcomes following TJA. When integrated with a collection of PROMs and objective assessments, data from wearable devices facilitates more comprehensive patient surveillance and enables tailored interventions during the perioperative period. Early monitoring of patients in the postoperative phase enhances the timely identification of adverse events and complications, thereby allowing prompt intervention. Additionally, remote monitoring of patients has been demonstrated to improve compliance with postoperative protocols and instructions. Through rigorous evaluation and standardization of data collected, remote monitoring has the potential to elevate the quality of care and enhance the recovery experience for TJA patients, ultimately leading to improved patient outcomes and satisfaction. However, further research is needed to delineate which specific remote data points are most effective in facilitating the early identification of adverse events and in optimizing patient outcomes and satisfaction after TJA. Future studies can focus on which data provides the most valuable and complete assessment of patient recovery. As our knowledge of how wearables can be best used to assess outcomes increases, there is potential to use data provided to tailor continuous patient assessment and individualized rehabilitation protocols, customized postoperative evaluation, and early detection of deviation from norms in recovery.

Acknowledgments

Generative artificial intelligence was not used at any stage in the preparation of this paper.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Cochrane methodology for systematic reviews was used in the literature search.
[DOCX File, 393 KB - rehab_v13i1e84671_app1.docx]

Multimedia Appendix 2
Terms used in the literature search.
[DOCX File, 15 KB - rehab_v13i1e84671_app2.docx]

Multimedia Appendix 3
Risk of bias visualization tool traffic plots for included randomized control studies, nonrandomized studies, and interventions.
[DOCX File, 11883 KB - rehab_v13i1e84671_app3.docx]

Checklist 1
PRISMA checklist.
[DOCX File, 276 KB - rehab_v13i1e84671_app4.docx]

References

1. Shichman I, Roof M, Askew N, et al. Projections and epidemiology of primary hip and knee arthroplasty in medicare patients to 2040-2060. JB JS Open Access 2023;8(1):e22.00112. [doi: 10.2106/JBJS.OA.22.00112] [Medline: 36864906]
2. Gunaratne R, Pratt DN, Banda J, Fick DP, Khan RJK, Robertson BW. Patient dissatisfaction following total knee arthroplasty: a systematic review of the literature. J Arthroplasty 2017 Dec;32(12):3854-3860. [doi: 10.1016/j.arth.2017.07.021] [Medline: 28844632]

XSL•FO
RenderX

3. Anakwe RE, Jenkins PJ, Moran M. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. J Arthroplasty 2011 Feb;26(2):209-213. [doi: 10.1016/j.arth.2010.03.013] [Medline: 20462736]
4. Xie Z, Jo A, Hong YR. Electronic wearable device and physical activity among US adults: an analysis of 2019 HINTS data. Int J Med Inform 2020 Dec;144:104297. [doi: 10.1016/j.ijmedinf.2020.104297] [Medline: 33059242]
5. Shandhi MMH, Singh K, Janson N, et al. Assessment of ownership of smart devices and the acceptability of digital health data sharing. NPJ Digit Med 2024 Feb 22;7(1):44. [doi: 10.1038/s41746-024-01030-x] [Medline: 38388660]
6. Master H, Bley JA, Coronado RA, et al. Effects of physical activity interventions using wearables to improve objectively-measured and patient-reported outcomes in adults following orthopaedic surgical procedures: a systematic review. PLoS One 2022;17(2):e0263562. [doi: 10.1371/journal.pone.0263562] [Medline: 35167599]
7. Christiansen MB, Thoma LM, Master H, Schmitt LA, Pohlig R, White DK. A physical therapist-administered physical activity intervention after total knee replacement: protocol for a randomized controlled trial. Phys Ther 2018 Jul 1;98(7):578-584. [doi: 10.1093/ptj/pzy037] [Medline: 29608733]
8. Hullfish TJ, Qu F, Stoeckl BD, Gebhard PM, Mauck RL, Baxter JR. Measuring clinically relevant knee motion with a self-calibrated wearable sensor. J Biomech 2019 May 24;89:105-109. [doi: 10.1016/j.jbiomech.2019.04.003] [Medline: 30981425]
9. Babaei N, Hannani N, Dabanloo NJ, Bahadori S. A systematic review of the use of commercial wearable activity trackers for monitoring recovery in individuals undergoing total hip replacement surgery. Cyborg Bionic Syst 2022;2022:9794641. [doi: 10.34133/2022/9794641] [Medline: 36751476]
10. Hameed D, Sodhi N, Dubin J, Schneider A, Barrack RL, Mont MA. Integrating smartphone applications and wearable devices for postoperative rehabilitation in total knee arthroplasty: a critical review. J Arthroplasty 2024 Aug;39(8):2028-2039. [doi: 10.1016/j.arth.2024.02.003]
11. Cumpston M, Li T, Page MJ, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019 Oct 3;10(10):ED000142. [doi: 10.1002/14651858.ED000142] [Medline: 31643080]
12. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
13. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
14. McGuinness LA, Higgins JPT. Risk-of-Bias Visualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. Res Synth Methods 2021 Jan;12(1):55-61. [doi: 10.1002/jrsm.1411] [Medline: 32336025]
15. Redfern RE, Crawford DA, Lombardi AV Jr, et al. Outcomes vary by pre-operative physical activity levels in total knee arthroplasty patients. J Clin Med 2023 Dec 25;13(1):125. [doi: 10.3390/jcm13010125] [Medline: 38202132]
16. Christensen JC, Blackburn BE, Anderson LA, et al. Recovery curve for patient reported outcomes and objective physical activity after primary total knee arthroplasty-a multicenter study using wearable technology. J Arthroplasty 2023 Jun;38(6S):S94-S102. [doi: 10.1016/j.arth.2023.03.060] [Medline: 36996947]
17. Sato EH, Stevenson KL, Blackburn BE, et al. Recovery curves for patient reported outcomes and physical function after total hip arthroplasty. J Arthroplasty 2023 Jul;38(7):S65-S71. [doi: 10.1016/j.arth.2023.04.012]
18. Pasqualini I, Huffman N, Klika A, et al. Stepping up recovery: integrating patient-reported outcome measures and wearable technology for rehabilitation following knee arthroplasty. J Knee Surg 2024 Aug;37(10):757-763. [doi: 10.1055/a-2315-8110]
19. Gibian JT, Bartosiak KA, Lucey BP, Riegler V, King J, Barrack RL. Sleep disturbances following total knee arthroplasty. J Arthroplasty 2023 Jun;38(6):S120-S124. [doi: 10.1016/j.arth.2023.01.066]
20. Twiggs J, Salmon L, Kolos E, Bogue E, Miles B, Roe J. Measurement of physical activity in the pre- and early post-operative period after total knee arthroplasty for osteoarthritis using a Fitbit Flex device. Med Eng Phys 2018 Jan;51:31-40. [doi: 10.1016/j.medengphy.2017.10.007] [Medline: 29117912]
21. Lebleu J, Daniels K, Pauwels A, et al. Incorporating wearable technology for enhanced rehabilitation monitoring after hip and knee replacement. Sensors (Basel) 2024 Feb 10;24(4):1163. [doi: 10.3390/s24041163] [Medline: 38400321]
22. Paxton RJ, Forster JE, Miller MJ, Gerron KL, Stevens-Lapsley JE, Christiansen CL. A feasibility study for improved physical activity after total knee arthroplasty. J Aging Phys Act 2018 Jan 1;26(1):7-13. [doi: 10.1123/japa.2016-0268] [Medline: 28338406]
23. Van der Walt N, Salmon LJ, Gooden B, et al. Feedback from activity trackers improves daily step count after knee and hip arthroplasty: a randomized controlled trial. J Arthroplasty 2018 Nov;33(11):3422-3428. [doi: 10.1016/j.arth.2018.06.024] [Medline: 30017217]
24. Duong V, Robbins SR, Dennis S, Venkatesha V, Ferreira ML, Hunter DJ. Combined digital interventions for pain reduction in patients undergoing knee replacement: a randomized clinical trial. JAMA Netw Open 2023 Sep 5;6(9):e2333172. [doi: 10.1001/jamanetworkopen.2023.33172] [Medline: 37713201]
25. Mehta SJ, Hume E, Troxel AB, et al. Effect of remote monitoring on discharge to home, return to activity, and rehospitalization after hip and knee arthroplasty: a randomized clinical trial. JAMA Netw Open 2020 Dec 1;3(12):e2028328. [doi: 10.1001/jamanetworkopen.2020.28328] [Medline: 33346847]

XSL•FO
RenderX

26. Pechstein AE, Jakiela JT, White DK. Exploring the relation of walking endurance with physical activity after total knee replacement. Musculoskeletal Care 2023 Sep;21(3):939-942. [doi: 10.1002/msc.1758] [Medline: 36916139]
27. Tang A, Behery OA, Singh V, Yeroushalmi D, Davidovitch R, Schwarzkopf R. Do physical activity and sleep correlate with patient-reported outcomes in total hip arthroplasty? J Hip Surg 2021 Jun;05(2):047-054. [doi: 10.1055/s-0041-1729764]
28. Crizer MP, Kazarian GS, Fleischman AN, Lonner JH, Maltenfort MG, Chen AF. Stepping toward objective outcomes: a prospective analysis of step count after total joint arthroplasty. J Arthroplasty 2017 Sep;32(9S):S162-S165. [doi: 10.1016/j.arth.2017.02.058] [Medline: 28343831]
29. Caliskan E, Igdir V, Dogan O, Bicimoglu A. Primary total knee replacement leads to an increase in physical activity but no changes in overall time of sedentary behaviour: a retrospective cohort study using an accelerometer. Int Orthop 2020 Dec;44(12):2597-2602. [doi: 10.1007/s00264-020-04720-9] [Medline: 32654055]
30. Brandes M, Ringling M, Winter C, Hillmann A, Rosenbaum D. Changes in physical activity and health-related quality of life during the first year after total knee arthroplasty. Arthritis Care Res (Hoboken) 2011 Mar;63(3):328-334. [doi: 10.1002/acr.20384] [Medline: 20981812]
31. Frimpong E, van der Jagt DR, Mokete L, et al. Improvements in objectively measured activity behaviors do not correlate with improvements in patient-reported outcome measures following total knee arthroplasty. J Arthroplasty 2020 Mar;35(3):712-719. [doi: 10.1016/j.arth.2019.10.016] [Medline: 31722854]
32. Höll S, Blum A, Gosheger G, Dieckmann R, Winter C, Rosenbaum D. Activity monitor after minimally invasive total hip arthroplasty. J Orthop Surg Res 2018 Jun 15;13(1):148. [doi: 10.1186/s13018-018-0775-4] [Medline: 29907134]
33. Komann M, Dreiling J, Baumbach P, et al. Objectively measured activity is not associated with average pain intensity 1 week after surgery: a cross-sectional study. Eur J Pain 2024 Sep;28(8):1330-1342. [doi: 10.1002/ejp.2260] [Medline: 38450921]
34. Bolander RP, Pierce AG, Mangal RK, Hernandez VH, Stulberg SD. Characterization of the relative change in objective and subjective metrics by baselining patients who have wearable technology before total knee arthroplasty. J Arthroplasty 2024 Aug;39(8S1):S130-S136. [doi: 10.1016/j.arth.2024.04.001] [Medline: 38604282]
35. Harding P, Holland AE, Delany C, Hinman RS. Do activity levels increase after total hip and knee arthroplasty? Clin Orthop Relat Res 2014 May;472(5):1502-1511. [doi: 10.1007/s11999-013-3427-3] [Medline: 24353051]
36. Bendich I, Chung C, Hwang K, et al. Changes in prospectively collected longitudinal patient-generated health data are associated with short-term patient-reported outcomes after total joint arthroplasty: a pilot study. Arthroplast Today 2019 Mar;5(1):61-63. [doi: 10.1016/j.artd.2019.01.005] [Medline: 31020024]
37. Bini SA, Shah RF, Bendich I, Patterson JT, Hwang KM, Zaid MB. Machine learning algorithms can use wearable sensor data to accurately predict six-week patient-reported outcome scores following joint replacement in a prospective trial. J Arthroplasty 2019 Oct;34(10):2242-2247. [doi: 10.1016/j.arth.2019.07.024] [Medline: 31439405]
38. Shah RF, Zaid MB, Bendich I, Hwang KM, Patterson JT, Bini SA. Optimal sampling frequency for wearable sensor data in arthroplasty outcomes research. a prospective observational cohort trial. J Arthroplasty 2019 Oct;34(10):2248-2252. [doi: 10.1016/j.arth.2019.08.001] [Medline: 31445866]
39. Frimpong E, McVeigh JA, van der Jagt D, et al. Light intensity physical activity increases and sedentary behavior decreases following total knee arthroplasty in patients with osteoarthritis. Knee Surg Sports Traumatol Arthrosc 2019 Jul;27(7):2196-2205. [doi: 10.1007/s00167-018-4987-2] [Medline: 29785450]
40. Lebleu J, Poilvache H, Mahaudens P, De Ridder R, Detrembleur C. Predicting physical activity recovery after hip and knee arthroplasty? A longitudinal cohort study. Braz J Phys Ther 2021;25(1):30-39. [doi: 10.1016/j.bjpt.2019.12.002] [Medline: 31874729]
41. Vitale JA, Banfi G, Viganò M, Negrini F. How do patients sleep after orthopaedic surgery? Changes in objective sleep parameters and pain in hospitalized patients undergoing hip and knee arthroplasty. Int Orthop 2023 Aug;47(8):1929-1938. [doi: 10.1007/s00264-023-05862-2] [Medline: 37300562]
42. van Diemen MPJ, Ziagkos D, Kruizinga MD, et al. Mitochondrial function, grip strength, and activity are related to recovery of mobility after a total knee arthroplasty. Clin Transl Sci 2023 Feb;16(2):224-235. [doi: 10.1111/cts.13441] [Medline: 36401590]
43. Fallahzadeh R, Verdonk F, Ganio E, et al. Objective activity parameters track patient-specific physical recovery trajectories after surgery and link with individual preoperative immune states. Ann Surg 2023 Mar 1;277(3):e503-e512. [doi: 10.1097/SLA.0000000000005250] [Medline: 35129529]
44. Karas M, Marinsek N, Goldhahn J, Foschini L, Ramirez E, Clay I. Predicting subjective recovery from lower limb surgery using consumer wearables. Digit Biomark 2020;4(Suppl 1):73-86. [doi: 10.1159/000511531] [Medline: 33442582]
45. Harada S, Hamai S, Shiomoto K, et al. Predictors of physical activity recovery after total hip arthroplasty: a prospective observational study. Int Orthop 2024 Mar;48(3):753-760. [doi: 10.1007/s00264-023-06022-2] [Medline: 37923880]
46. Surmacz K, Ribeiro-Castro AL, Anderson MB, Van Andel D, Redfern RE, Duwelius PJ. A retrospective study on the feasibility of using low-burden patient-reported pain scores to track recovery and outcomes after total joint replacement. Arthroplast Today 2024 Apr;26:101297. [doi: 10.1016/j.artd.2023.101297] [Medline: 38352707]
47. Lyman S, Hidaka C, Fields K, Islam W, Mayman D. Monitoring patient recovery after THA or TKA using mobile technology. HSS J 2020 Dec;16(Suppl 2):358-365. [doi: 10.1007/s11420-019-09746-3] [Medline: 33380968]

XSL•FO
RenderX

48. Goeb YL, Krell EC, Nguyen JT, et al. Early recovery outcomes in patients undergoing total hip arthroplasty through a posterior approach with modified postoperative precautions. J Arthroplasty 2021 Aug;36(8):2817-2822. [doi: 10.1016/j.arth.2021.03.038] [Medline: 33840540]
49. Campbell M, McKenzie JE, Sowden A, et al. Synthesis Without Meta-Analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]
50. Vitale JA, Banfi G, Tivolesi V, Pelosi C, Borghi S, Negrini F. Rest-activity daily rhythm and physical activity levels after hip and knee joint replacement: the role of actigraphy in orthopedic clinical practice. Chronobiol Int 2021 Dec;38(12):1692-1701. [doi: 10.1080/07420528.2021.1939365] [Medline: 34121554]
51. Alattas SA, Smith T, Bhatti M, Wilson-Nunn D, Donell S. Greater pre-operative anxiety, pain and poorer function predict a worse outcome of a total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2017 Nov;25(11):3403-3410. [doi: 10.1007/s00167-016-4314-8] [Medline: 27734110]
52. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998 Aug;28(2):88-96. [doi: 10.2519/jospt.1998.28.2.88] [Medline: 9699158]
53. Davis AM, Perruccio AV, Canizares M, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. Osteoarthr Cartil 2009 Jul;17(7):843-847. [doi: 10.1016/j.joca.2009.01.005]

Abbreviations

HOOS: Hip Disability and Osteoarthritis Outcome Score
JR: joint replacement
KOOS: Knee Injury and Osteoarthritis Outcome Score
LEFS: Lower Extremity Functional Scale
MWT: minute walking test
PCS: Physical Composite Score
PROM: patient-reported outcome measure
RCT: randomized controlled trial
ROB-2: Cochrane's Risk of Bias tool for randomized trials
ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions
SF-12: 12-Item Short Form Health Survey
SF-36: 36-Item Short Form Health Survey
THA: total hip arthroplasty
TJA: total joint arthroplasty
TKA: total knee arthroplasty
TUG: timed-up-and-go
UCLA: University of California, Los Angeles Activity Scale
VAS: Virtual Analog Scale
WOMAC: Western Ontario and McMaster Universities Arthritis Index

Edited by S Munce; submitted 23.Sep.2025; peer-reviewed by D Singhal, Y Ni; revised version received 24.Dec.2025; accepted 26.Jan.2026; published 17.Apr.2026.

Please cite as:
Hall III RR, Hamel AP, Chan PA, Zheng H, Ayers DC
Use of Wearable Devices to Augment Traditional Measurements of Postoperative Outcomes Following Total Joint Arthroplasty: Systematic Review
JMIR Rehabil Assist Technol 2026;13:e84671
URL: https://rehab.jmir.org/2026/1/e84671
doi:10.2196/84671

© Robert Raymond Hall III, Alexander P Hamel, Pamela A Chan, Hua Zheng, David C Ayers. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Use of 3D Printing Technology in Rehabilitation for Adults Living With Neurological Conditions: Scoping Review

Salena Aggerwal[1], BHS; Zahra Lalani[1], BHS; Jennifer O'Neil[1,2], PT, PhD

[1]Faculty of Health Sciences, School of Rehabilitation Sciences, University of Ottawa, 200 Lees Avenue, Ottawa, ON, Canada

[2]Bruyère Health Research Institute, Ottawa, ON, Canada

Corresponding Author:
Jennifer O'Neil, PT, PhD
Faculty of Health Sciences, School of Rehabilitation Sciences, University of Ottawa, 200 Lees Avenue, Ottawa, ON, Canada

Abstract

Background: Neurorehabilitation plays a key role in improving motor recovery for people with neurological conditions. Although 3D printing has emerged as a promising rehabilitation tool, little is known on how it is used for the rehabilitation of adults living with neurological conditions worldwide.

Objective: We aimed to provide a comprehensive overview of 3D printing in neurorehabilitation and precisely explore how it is used to improve motor recovery for adults with neurological conditions living in higher- and lower-middle–income countries.

Methods: We conducted a scoping review following the Joanna Briggs Institute guidelines. After searching 3 databases (MEDLINE, Web of Science, and Nursing and Allied Health Premium), 2 independent reviewers screened and selected English-language studies involving adults (≥18 years) published between 2019 and 2024 to capture the most recent advancements in this field. We extracted relevant information on neurological conditions, motor recovery outcomes, and types of 3D printing and offered a comparative analysis of 3D printing in physical neurorehabilitation from the perspective of national income levels using a modified Joanna Briggs Institute extraction form. We synthesized the findings narratively with tabular support.

Results: After screening 2752 titles and abstracts and 103 (3.7%) full texts, we included 13 (0.5%) studies based on our inclusion criteria. All included studies were conducted in upper-middle–income or high-income countries, and most studies (9/13, 69.2%) focused on stroke, followed by spinal cord injury (2/13, 15.4%), Parkinson disease (1/13, 7.7%), and central nerve disease (1/13, 7.7%). The 3D-printed rehabilitation tools included orthotics (7/13, 53.8% for the upper extremities [UEs]; 3/13, 23.1% for the lower extremities [LEs]), an exoskeleton (1/13, 7.7%; UEs), a modular assistive hand device (1/13, 7.7%; UEs), and an insole (1/13, 7.7%; LEs). In total, 69.2% (9/13) of the studies targeted UE rehabilitation, measured using the Action Research Arm Test, active range of motion, the box and block test, the Fugl-Meyer Assessment, the Modified Ashworth Scale, the manual function test, range of motion, and the Toronto Rehabilitation Institute Hand Function Test, and 30.8% (4/13) targeted LE rehabilitation, measured using the 10-m walk test, anteroposterior ground reaction force analysis, the Barthel index, the Tinetti scale, the RehaWatch system, and the GaitWatch system.

Conclusions: Used as a rehabilitation tool, 3D printing technology has demonstrated significant potential in improving upper and lower motor recovery for people with certain neurological conditions in high-middle–income countries. Future research should explore the implementation feasibility and effectiveness of these technologies across different neurological conditions and income settings, particularly in low- and lower-middle–income countries.

(JMIR Rehabil Assist Technol 2026;13:e81782) doi:10.2196/81782

KEYWORDS

3D printing technology; neurorehabilitation; rehabilitation technology; motor recovery; assistive devices

Introduction

Background

Neurological conditions encompass a broad spectrum of central and peripheral conditions caused by damage to the brain, spinal cord, or peripheral nerves that interfere with brain development or result in impaired motor functions [1]. According to a 2021 report published in *The Lancet Neurology* by the World Health Organization (WHO), over 3 billion people in the world are living with a neurological condition [2,3]. Neurological conditions such as stroke, traumatic brain injury, spinal cord injury, cerebral palsy, Parkinson disease, and cognitive disorders are some of the leading causes of illness and disability worldwide [2,4,5]. Disability-adjusted life years, which measure the overall burden of disease, including disability, sickness, and premature death due to neurological conditions, have increased by 18% since 1990 [1-3]. Moreover, care disparities persist across demographics. Higher-income countries possess greater access to specialized care, with 70 times more health

XSL•FO
RenderX

professionals specializing in neurological conditions per 100,000 individuals compared to lower-middle–income countries [1]. In the United States alone, the financial and public health burden of neurological conditions is estimated to cost US $1.7 trillion annually [1]. Given the global prevalence of neurological conditions, and with over 80% of neurological-related deaths and health losses occurring in low- and middle-income countries (LMICs), there is a need for effective treatment strategies to improve patient health outcomes and reduce associated burdens worldwide [1-7]. Neurorehabilitation is an interdisciplinary clinical process designed to assess disabilities and provide rehabilitation interventions to enhance functions affected by damage to the nervous system [8,9]. The WHO defines neurorehabilitation as enabling individuals to acquire the knowledge, skills, and resources necessary to achieve optimal physical, psychological, social, and economic functioning [2,3,8]. Many interventions, including assistive devices, therapeutic exercises, and cognitive behavioral therapies, are used in the provision of rehabilitation services for individuals with neurological conditions [5,10,11]. For motor recovery, one promising rehabilitation tool is the use of 3D printing technology.

This is an emerging technology that creates 3D objects out of melting materials into thin layers printed layer by layer based on computer-programmed designs [12]. A range of materials can be used, including nylon, biodegradable plastic filament, polylactic acid, and acrylonitrile butadiene styrene plastic [12,13]. Worldwide, 3D printing has the potential to improve access to rehabilitation devices in lower-middle–income countries where traditional manufacturing processes can be both costly and time-consuming [14,15]. Currently, in health care, 3D printing has been widely adopted beyond rehabilitation, including applications in surgical planning, dental restorations, and the production of anatomical models for medical education and preoperative simulation [16]. In the rehabilitation field, 3D printing is being used to produce a variety of highly customized, patient-specific devices, such as orthotics, prosthetics, and assistive tools [6,12,14]. Many of these designs are low cost, customizable, and easy to assemble, making them well suited for resource-limited settings [12,15]. Specific to neurorehabilitation, 3D printing technologies have gained increasing attention for their potential to deliver personalized, cost-effective, and accessible therapeutic solutions [17]. Despite growing interest, there remains limited synthesis of how these technologies are being applied worldwide, particularly in relation to their use in supporting motor recovery in adults living with neurological conditions. As the scope of neurorehabilitation extends across multiple professions (ie, physiotherapy, occupational therapy, audiology, speech-language pathology, social work, and nursing), focusing on motor recovery specific to physiotherapy or physical rehabilitation is a first step toward understanding 3D printing use in neurorehabilitation.

Research Objectives

To address this gap, the objective of our scoping review was to provide a comprehensive overview of how 3D printing technologies have recently been used in neurorehabilitation for adults. We aimed to document and compare their applications across high-income, upper-middle–income, lower-middle–income, and low-income countries and identify the types of assessments used to evaluate motor recovery outcomes when using 3D-printed rehabilitation tools, including standardized clinical scales, functional performance tests, and biomechanical measurements. To our knowledge, this study is the first to offer a comparative analysis of 3D printing in physical neurorehabilitation from the perspective of national income levels.

Methods

We conducted this scoping review in accordance with the Joanna Briggs Institute (JBI) methodology for scoping reviews [17] and aimed to answer the following question: how is 3D printing used to improve motor recovery for adults with neurological conditions in higher- and lower-middle–income countries?

Population, Concept, and Context Framework

We used the population, concept, and context framework as recommended by the JBI guidelines [17]. We included studies involving adult participants (18 years and older) with neurological conditions who received rehabilitation interventions aimed at improving or restoring motor function. Participants under 18 years of age were excluded as pediatric populations have different developmental trajectories and rehabilitation needs, which fall outside the scope of this review [18]. Eligible conditions included acquired brain injuries, neurodegenerative disorders, and traumatic brain or spinal cord injuries [19].

The concept for this review focused on the application of 3D printing technology in improving motor recovery for adults with neurological conditions. Studies exploring the use of 3D printing rehabilitation interventions or 3D-printed tools to enhance motor function in adults within multiple rehabilitation settings were considered. Rehabilitation interventions included physiotherapy or physical rehabilitation strategies designed to enhance motor abilities. Physiotherapy referred to evidence-based treatments delivered by trained professionals to restore movement and function, often through targeted exercises, manual techniques, and assistive devices. Physical rehabilitation strategies more broadly included any structured interventions aimed at improving motor control, strength, or mobility, particularly after neurological injury. Settings included clinical environments (eg, hospitals, rehabilitation centers, and outpatient clinics), home-based rehabilitation programs, and laboratory-based rehabilitation programs. Studies that focused on 3D printing for surgical planning, medical pharmaceutics, bioengineering, or any interventions that did not report clinical or functional motor recovery outcomes (ie, purely descriptive studies) were excluded as they did not reflect the rehabilitation objective of this review [16].

This scoping review considered studies conducted in contexts of higher-income, upper-middle–income, lower-middle–income, and low-income countries as defined by the World Bank based on gross national income per capita for 2023 (Multimedia Appendix 1) [20]. The geographical and socioeconomic context was explored in terms of access to 3D printing technologies as well as how this impacts the effectiveness of these interventions.

XSL•FO
RenderX

No limitations were placed on the gender, ethnicity, or religion of the participants included in the studies.

Developing a Search Strategy

We developed a rigorous search strategy in 3 steps. First, we conducted a preliminary search of MEDLINE (PubMed) to identify articles on the topic and define main keywords. In this preliminary search, we used the keywords "3D printing" and "rehabilitation" to capture as many articles on the topic as possible, thus providing insights into existing literature. We then reviewed keywords contained in titles and abstracts and the indexing terms used to describe potentially included articles (reviewing keywords and indexing terms from relevant titles and abstracts), which was then used to refine our search strategy. Second, we developed a refined search strategy incorporating additional keywords: "physiotherapy," "physical therapy," or "physical rehabilitation" (Multimedia Appendix 2). This refined search strategy incorporated both MeSH (Medical Subject Headings) terms and text word searches in MEDLINE (PubMed) and was adapted using equivalent text word strategies in 2 other databases (Web of Science and Nursing and Allied Health Premium). In this step, a filter for English-language publications was applied across all databases to maintain consistency (Multimedia Appendix 2). Third, the reference lists of all articles included in the review were manually screened for additional relevant papers. This step aimed to capture any pertinent studies that may not have been identified through the database searches.

Study Screening and Evidence Selection

Following the 3-database search, we collated all identified studies and used Mendeley Reference Manager (version 2.51.0; Elsevier) and Covidence (Veritas Health Innovation) for citation management and removal of duplicates. Titles and abstracts were screened by 2 independent reviewers (SA and ZL) for assessment against the inclusion and exclusion criteria. We screened studies published from 2019 onward, with the final search completed on November 21, 2024, to capture the most recent advancements in 3D printing technologies for neurorehabilitation. We included a variety of study designs (ie, randomized controlled trials [RCTs], nonrandomized experimental studies, pilot and feasibility studies, case studies, and systematic reviews) in our search to capture both the effectiveness of the interventions and the practical considerations surrounding their implementation. RCTs and non-RCTs were included to assess potential causal relationships between 3D printing interventions and motor recovery outcomes [21]. Pilot and feasibility studies were eligible if they examined the usability, acceptability, or logistical feasibility of 3D-printed devices in rehabilitation contexts [22]. Case studies were included if they provided detailed examples of individualized application [23]. Systematic reviews were considered if they synthesized primary research aligned with our inclusion criteria [24]. Opinion pieces, editorials, commentaries, and protocols without full-text availability were excluded as they did not contribute to the empirical evidence required for conclusions. When the inclusion criteria were met at the title and abstract screening stage, the full texts were screened through the same 2-reviewer screening method. Any disagreements were resolved through consensus with a third reviewer (JON) when necessary.

Data Extraction

We extracted relevant data from the included full-text studies. We used a data extraction form that was developed based on the JBI guidelines (Multimedia Appendix 3) to ensure completeness of data extraction. The extraction form included (1) general study characteristics, including authors, title, year of publication, date of data collection, study type, country, and study aim; (2) population details, such as sample size, age, gender, and the health condition being addressed; (3) intervention details, including setting (eg, clinical, home based, or laboratory) and duration of the intervention, type of 3D-printed technology used (eg, orthosis or insole), stiffness and type of polymer, neurological condition (eg, stroke or Parkinson disease), and motor area focus (upper extremities [UEs] or lower extremities [LEs]); and (4) outcomes, including the types of assessments used to evaluate motor recovery, such as standardized clinical scales, functional performance tests, and biomechanical measurements.

Data Analysis and Interpretation

We analyzed the data using descriptive qualitative content analysis and synthesized the extracted data from a geographical perspective. Data were systematically mapped based on the objectives of our scoping review to capture study characteristics, (authors, title, and publication year), the type of 3D-printing technology used, the associated neurological condition and rehabilitation goal, outcome measures, and country of implementation. A comparative approach was applied to assess differences between high-income and lower-middle–income countries, documenting disparities in access (as reported through the availability of 3D printing technology, affordability considerations, descriptions of infrastructure, and integration into clinical practice) and effectiveness. Findings were synthesized into thematic categories, providing a comprehensive overview of how 3D-printed technologies have been used to support motor recovery in different health care contexts.

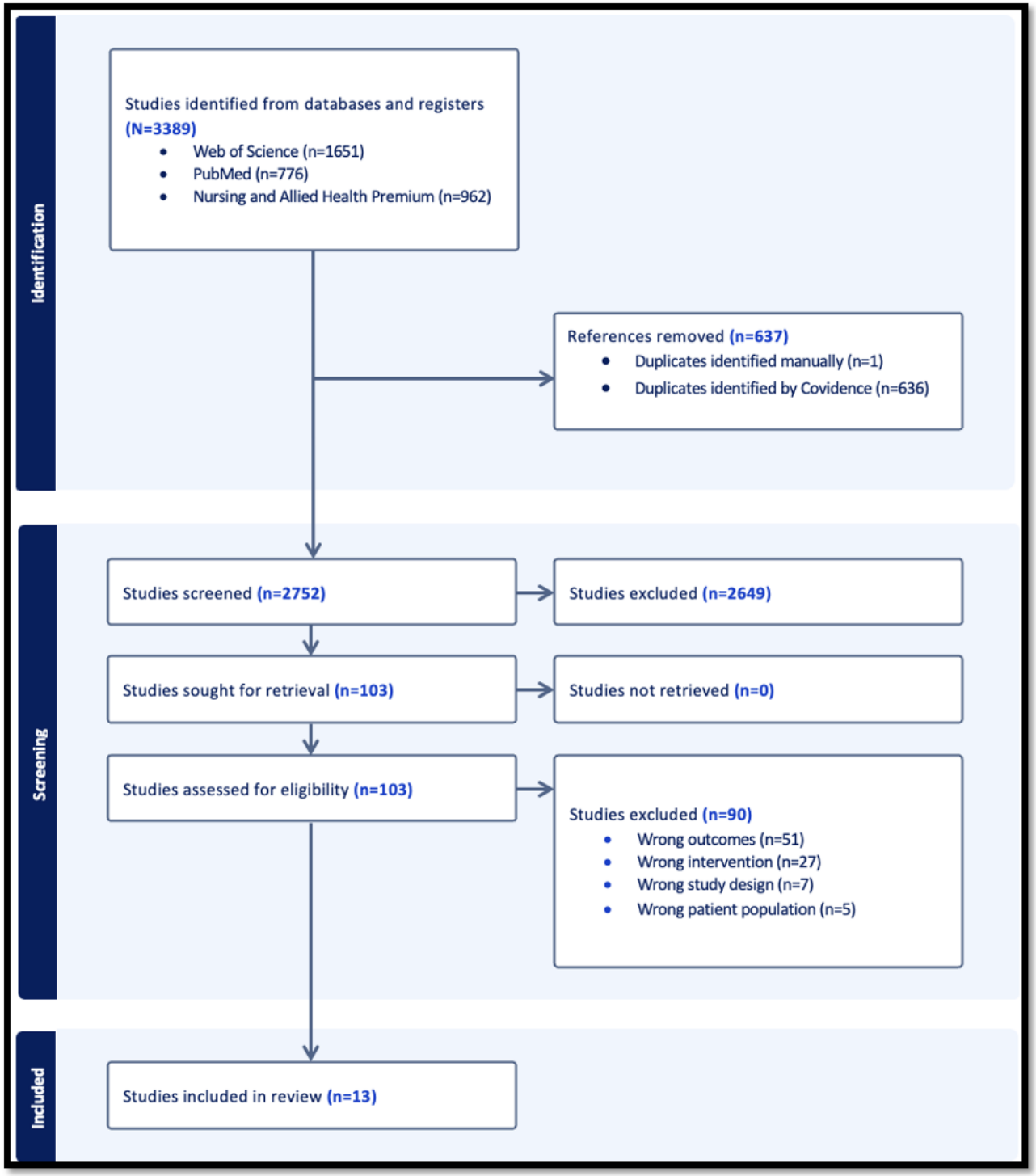
Results

Overview

A total of 2752 titles and abstracts were screened. Of these 2752 articles, after removing 637 (23.1%) duplicates, 103 (3.7%) full-text articles were assessed for eligibility, and 90 (3.3%) were excluded (Figure 1). We included 13 studies [25-37], with sample sizes ranging from single case reports (N=1 individual) to small trials (N=31 individuals) (Table 1). The reported participant ages ranged from 23 to 83 years, aligning with the inclusion criterion of adult participants (18 years and older).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart showing the selection of studies for inclusion in the scoping review.



XSL•FO
RenderX

Table . Technologies used in neurological rehabilitation for upper- and lower-extremity impairments, including the country in which the technology was implemented, the associated neurological condition, the type of technology, 3D printing materials and processes, the outcome summary, and the outcome measures used to assess effectiveness.

Study and country	Neurological condition	Assistive device	3D printing details (device; stiffness and type of polymer) and process	Outcome summary	Outcome measures
Upper-extremity impairments					
Chen et al [25], 2022, China	Stroke	Hand splint	PLA[a]; 3D printing method and process not specified	Improved grip and lateral pinch strength; no significant change in palmar pinch and gross movement	ARAT[b] and pinch force tests[c]
Demeco et al [26], 2023, country not specified	Stroke	Hand splint	Not specified	Improved upper-limb ROM[d], hand dexterity, and pinch strength (palmar and lateral); reduced spasticity	ROM, BBT[e], pinch force tests, FMA[f], and MAS[g]
Dudley et al [27], 2021, United States	Stroke	Exoskeleton	PLA+1% copper additive (PLACTIVE); FDM[h] (Ultimaker 2+ Extended); infill: 35% - 40%; layer height: 0.15 - 0.25 mm; bed: 50 °C; shell: 0.8 mm; print speed: 60 - 100 mm/s	Improved finger flexion and extension, grasp function, and hand dexterity	FMA and BBT
Huber et al [28], 2023, United States	Stroke	Hand splint	Variable-stiffness TPU[i] powder; SLS[j] 3D printing; printer and parameters not reported	Improved hand dexterity and pincer force; no significant change in pincer aperture	Pincer force test, pincer aperture test, and BBT
Kuo et al [29], 2024, China	Neurological central nerve injury	Modular assistive hand technology device	TPU filament (1.75 mm); FDM (Original Prusa i3 MK3S+); CAD[k] designed (SolidWorks, DesignSpark Mechanical, and Meshmixer)	Improved shoulder AROM[l] and functional hand tasks; no significant change in hand dexterity or grip strength	AROM, BBT, Jamar dynamometer, and functional hand tasks
Yeh et al [30], 2023, China	Spinal cord injury	Hand splint	Material not specified; FDM (FlashForge Finder); layer height: 0.2 mm; infill: 15%; speed: 60 - 80 mm/s; print temperature: 200 °C; nonprinted components: silicone finger cots; acrylic linkages (3 mm) laser cut	Improved pinch force and hand dexterity	Pinch force test and BBT
Toth et al [31], 2020, Hungary	Stroke	Hand splint	PA[m] 2200 (nylon) framework; TPU sheets; nitinol SMA[n]+Kapton heating elements; SLS (EOS Formiga P 110); 110 µm layers; FDM; integrated heating elements	Improved pencil holding, handle and tool grasp, vertical handle grasp, cup holding, door handle operation, and device assembly; no significant improvement in eraser holding	MFT[o]

XSL•FO
RenderX

Study and country	Neurological condition	Assistive device	3D printing details (device; stiffness and type of polymer) and process	Outcome summary	Outcome measures
Yang et al [32], 2021, China	Stroke	Hand splint	ABS[p] filament; FDM (UP Box); layer thickness: 0.2 mm; infill: 20%; nozzle temperature: 230 °C; print speed: 60 mm/s	Improved upper-extremity motor function; reduced spasticity	MAS and FMA
Yoo et al [33], 2019, South Korea	Spinal cord injury	Hand splint	PLA; FDM (Moment 2); CAD designed (SolidWorks); post-print heat reshaping; nonprinted components: self-adhesive padding, Velcro, and nylon thread	Improved grasp, lift, and object manipulation	TRI-HFT[q]
Lower-extremity impairments					
Brognara et al [34], 2020, Spain	Parkinson disease	Insole	Material not specified; 3D foot scan (3D Sense); CAD designed (Rhinoceros 3D); FDM (Delta WASP 4070); custom insole with 2 blunted cones (plantar stimulation)	Improved balance; no significant change in daily activities	Barthel index and Tinetti scale
Hsu et al [35], 2020, China	Stroke	Ankle-foot orthoses	Nylon (4611); 3D scan (Structure Sensor; OBJ[r]→STL[s]); FDM (Ultimaker); approximately 3-mm wall thickness	Improved gait parameters; no significant change in walking speed	RehaWatch system and 10MWT[t]
Kwon et al [36], 2019, China	Stroke	Ankle-foot orthoses	TPU filament; VeroBlackPlus (rigid components); FDM (Cubicon Single Plus); PolyJet (Objet 30 Prime)	Improved paretic propulsion; no significant change in walking speed	Anteroposterior ground reaction force analysis and 10MWT

XSL•FO
RenderX

Study and country	Neurological condition	Assistive device	3D printing details (device; stiffness and type of polymer) and process	Outcome summary	Outcome measures
Liu et al [37], 2019, South Korea	Stroke	Ankle-foot orthoses	PA-12; 3D scan (EinScan-Pro); model modification (Geomagic Studio); MJF[u]	Increased gait velocity and stride length; no significant changes in cadence and step symmetry	GaitWatch system

[a]PLA: polylactic acid.

[b]ARAT: Action Research Arm Test.

[c]"Pinch force test" is used as a standardized term across studies measuring pinch strength except for the study by Huber et al [28], which specifically used "pincer force test" based on a defined protocol.

[d]ROM: range of motion.

[e]BBT: box and block test.

[f]FMA: Fugl-Meyer Assessment.

[g]MAS: Modified Ashworth Scale.

[h]FDM: fused deposition modeling.

[i]TPU: thermoplastic polyurethane.

[j]SLS: selective laser sintering.

[k]CAD: computer-aided design.

[l]AROM: active ROM.

[m]PA: polyamide.

[n]SMA: shape-memory alloy.

[o]MFT: manual function test.

[p]ABS: acrylonitrile butadiene styrene.

[q]TRI-HFT: Toronto Rehabilitation Institute Hand Function Test.

[r]OBJ: Wavefront Technologies file format.

[s]STL: stereolithography.

[t]10MWT: 10-m walk test.

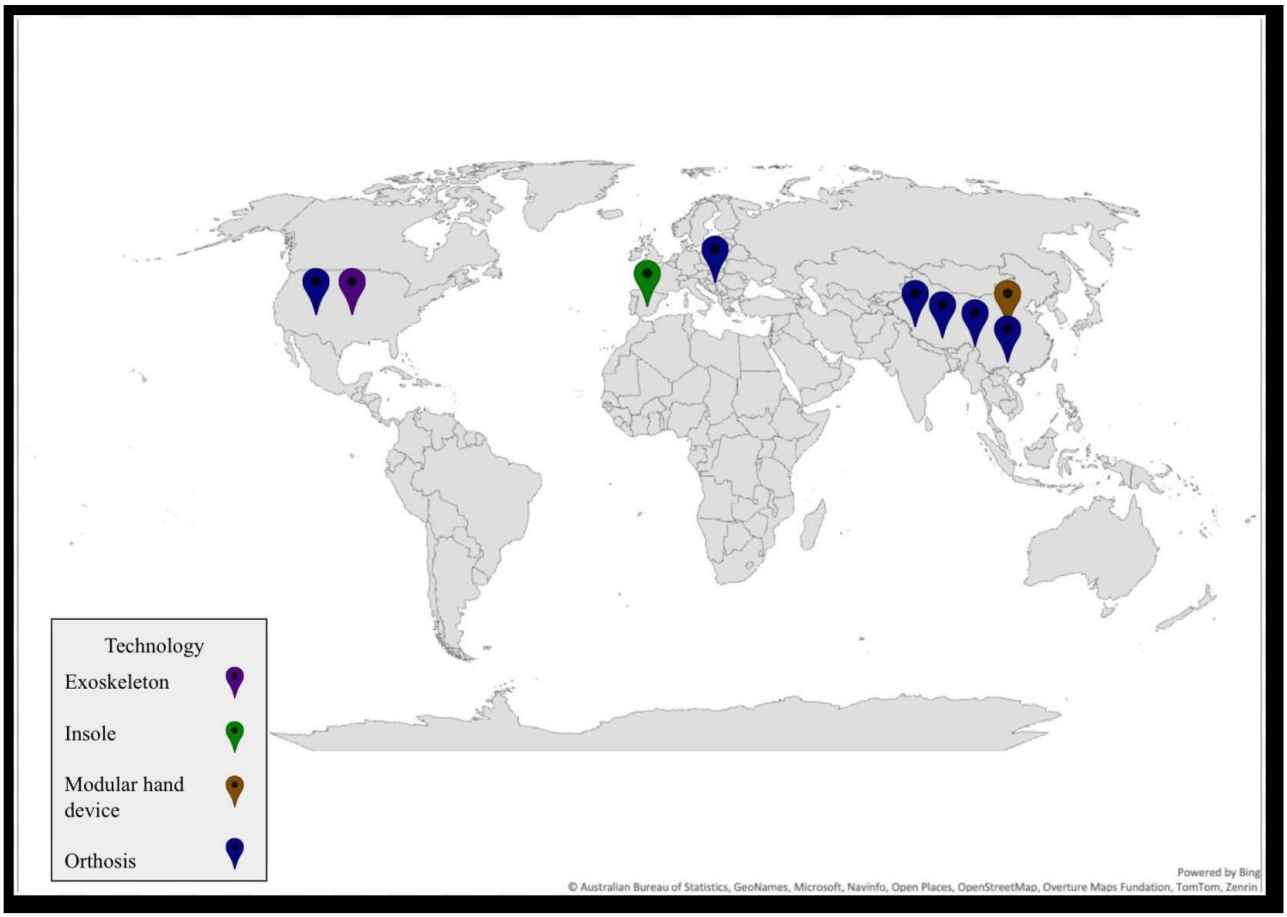
[u]MJF: multijet fusion.

Overview of 3D-Printed Rehabilitation Tools Used in Neurorehabilitation per Country

All included studies [25-37] were conducted in upper-middle–income or high-income countries (ie, China, South Korea, the United States, Spain, and Hungary; Figure 2). While 4 different types of neurorehabilitation tools were designed using 3D printing—orthotics (10/13, 76.9%) [25,26,28,30-33,35-37], an exoskeleton (1/13, 7.7%) [27], a modular assistive hand technology device (1/13, 7.7%) [29], and an insole (1/13, 7.7%) [34] (Multimedia Appendix 4)—most included studies (8/13, 61.5%) primarily focused on orthoses for stroke rehabilitation [25,26,28,30-32,35-37].

XSL•FO
RenderX

Figure 2. Geographical distribution of 3D-printed rehabilitation tools by country and technology type. Map created using Microsoft Excel and Google Drawings.



In total, 69.2% (9/13) of the included studies focused on rehabilitation of UE function, and 30.8% (4/13) focused on LE rehabilitation [25,27-33]. For UE rehabilitation, 5 of the studies used 3D printing technology to design hand splints for stroke-related impairments [25,26,28,31,32], 2 of the studies focused on hand splints for individuals with spinal cord injury [30,33], 1 study included a 3D-printed modular assistive hand device for patients with central nerve injuries [29], and 1 study investigated the use of a 3D-printed upper-limb exoskeleton to support hand rehabilitation for individuals after stroke [27]. For LE rehabilitation, 3 of the studies explored the use of 3D-printed ankle-foot orthoses for gait impairment rehabilitation after stroke [35-37], and 1 study used a 3D-printed insole for people living with Parkinson disease [34].

Types of Assessments

Across studies, outcome measures were predominantly used in pretest-posttest designs to assess changes in motor performance before and after the application of 3D-printed rehabilitation tools. The most frequently used standardized outcome measures for UEs were the pinch force test (4/13, 30.8%) [25,26,28,30] and the box and block test (5/13, 38.5%) [26-30]. For LEs, a commonly used outcome measure was the 10-m walk test (10MWT; 2/13, 15.4%) [35,36].

Discussion

Principal Findings

To our knowledge, this scoping review is the first to explore the application of 3D printing rehabilitation interventions for adults with neurological conditions from a geographical perspective. We selected a geographical perspective and the World Bank national income classification to reflect the socioeconomic and structural determinants of adoption that may influence whether 3D-printed neurorehabilitation devices can be implemented across international health systems. Our geographical analysis showed that rehabilitation research on 3D printing technologies is primarily concentrated in higher- and upper-middle–income countries (ie, China, the United States, South Korea, Spain, and Hungary). This distribution reflects broader economic trends and access to infrastructure and materials. In contrast to this technological growth in higher-income countries, rehabilitation-specific data remain limited in lower- and lower-middle–income countries. No studies were identified in our review from lower-middle– or low-income countries, highlighting the current gap in implementation within resource-constrained health systems.

As outlined in the WHO *Global strategy on digital health 2020 - 2025*, such disparities reflect systemic barriers in low-resource settings, including underdeveloped infrastructure, limited digital capacity, and inadequate financing models that constrain the adoption of digital health technologies [38].

XSL•FO
RenderX

Potential barriers also include the cost of medical-grade 3D printers, which can be substantial depending on material capabilities; infrastructure limitations such as unreliable electricity, insufficient internet connectivity, and shortages of trained personnel [14,39]; sociocultural factors such as potential stigma associated with 3D-printed assistive technologies; and if rehabilitation is not widely used in some regions, this may further affect the uptake of 3D-printed devices for rehabilitation [14,40]. The technical compatibility of 3D printing in lower-middle– or low-income countries may also be limited by gaps in the broader fabrication ecosystem required for clinical device production. This includes unreliable access to printing materials (eg, filament or resin), supplementary assistive device components (eg, straps or Velcro and padding or foam), and the software and digital infrastructure needed for the design process [12,14,40]. Even when 3D printers are available, limited capacity for postprocessing (eg, curing and edge finishing), routine maintenance (eg, calibration and replacement of wear parts), and training of clinicians may hinder sustained implementation in resource-constrained settings [12,14,40]. While barriers must be considered, investigating the use of emerging technologies in LMIC settings is crucial to understanding adoption, adaptation, and implementation in various and diverse contexts. The feasibility and long-term sustainability of 3D printing in low-resource settings remain critical areas for future investigation [14,40].

3D Printing Technology in Neurorehabilitation

Integrating 3D printing technologies into neurorehabilitation represents a paradigm shift in delivering innovative, personalized, clinically relevant interventions for individuals with neurological conditions. Our scoping review synthesized evidence on the application of novel technology, including 3D-printed orthotics, exoskeletons, modular assistive hand devices, and insoles, supporting motor rehabilitation for individuals who have experienced a stroke or spinal cord injury or been diagnosed with Parkinson disease or central nerve injury. We recognize that these are not representative of all neurological conditions, and we consider that our findings suggest a potential shift characterized by the ability to use emerging technologies such as 3D printing to rapidly design and produce patient-specific rehabilitation tools tailored to individual needs and functional impairments [12]. Among the technologies reviewed, orthotic devices emerged as the most widely studied, primarily addressing lower motor impairments for persons who experienced a stroke. This focus is particularly important in both upper-middle– and lower-income countries as stroke remains a leading cause of morbidity and mortality worldwide. Gait impairments after stroke are extensively researched and prevalent in neurological rehabilitation, so investigating cost- and resource-effective tools tailored to these impairments is imperative [41]. Unlike conventional rehabilitation tools, 3D-printed solutions may offer greater flexibility and adaptive adjustments based on the patient, faster production timelines, and affordability [12,14]. Therefore, 3D-printed rehabilitation tools may improve fit, comfort, and usability to enhance motor recovery outcomes and support greater independence in daily activities. Their application across clinical and home-based settings could further broaden access to care, particularly in LMICs. Given this emphasis and the growing potential of 3D printing in neurorehabilitation, it is essential that future research not only continue to refine orthotic solutions but also explore a broader range of technologies and document their effectiveness within neurorehabilitation contexts.

Our findings offer a comprehensive overview of measurements used to evaluate motor recovery outcomes (ie, standardized clinical, functional, and biomechanical measurements). Gold-standard rehabilitation assessments are those that are widely validated, clinically reliable, and consistently used across diverse populations to measure functional outcomes with high sensitivity and specificity. Results from our scoping review identified 4 gold-standard outcome measures being used for assessing the impact of 3D printing technologies on UE rehabilitation: the Fugl-Meyer Assessment (FMA), the Action Research Arm Test (ARAT), the Jamar dynamometer, and the Barthel index. For LE rehabilitation, the 10MWT was the only gold-standard outcome measure identified. The frequent use of the FMA across our included studies reflects its widespread acceptance as an outcome measure, particularly for people after stroke. Its strong psychometric properties, such as high interrater reliability, responsiveness to change, and internal consistency, make it a clinically relevant tool for capturing motor function in the UEs [42-44]. Within the context of 3D printing interventions, the FMA's emphasis on voluntary movement patterns, coordination, and reflex activity aligns well with the types of impairments that these technologies aim to address [42-44]. Its application in studies examining customized orthoses and exoskeletons suggests that researchers prioritize outcome measures that can accommodate the complexity of stroke-related motor deficits and individualized recovery trajectories [42-45].

The Jamar dynamometer was used in 7.7% (1/13) of the studies [29] and serves as a recognized gold-standard tool for objectively quantifying grip strength. In neurorehabilitation, grip strength is a meaningful proxy for upper-limb recovery and is often correlated with broader indicators of functional independence and quality of life in individuals with neurological impairments [46,47]. The Jamar dynamometer is particularly important for evaluating 3D printing interventions as it can provide precise, quantifiable feedback on how assistive devices, such as modular assistive hand devices, can enhance grip function and hand strength for adults with neurological conditions [46,47]. These objective data strengthen the clinical evidence supporting the functional efficacy of 3D-printed devices and enable more informed design refinements for 3D-printed technologies. In resource-limited settings, where comprehensive neurological assessments may not be feasible, dynamometers can offer practical and scalable solutions for tracking fine motor outcomes in response to 3D printing interventions [48].

The Barthel index was used in 7.7% (1/13) of the studies [34] as a gold-standard tool and captures independence in essential activities of daily living, including feeding, grooming, bathing, mobility, and toileting [43,49]. Its relevance lies in its ability to assess the effectiveness of 3D-printed insoles by monitoring changes in self-care and mobility over time, thereby capturing the progressive nature of degenerative neurological conditions such as Parkinson disease [50]. Its ease of administration, low

XSL•FO
RenderX

cost, and global familiarity make it especially well suited for integration into 3D printing rehabilitation protocols in settings with limited infrastructure or personnel [49].

The ARAT was reported in 7.7% (1/13) of the studies [25] as a gold-standard measure to evaluate upper-limb performance across 4 domains: grip, grasp, pinch, and gross movement [51]. The ARAT has widely excellent interrater reliability, high test-retest reliability, and good construct validity, particularly for people after stroke [51]. By assessing both fine and gross components of upper-limb function, it is well suited for evaluating 3D-printed devices targeting grasp and release, dexterity, and functional hand use [25,51].

The 10MWT was used in 15.4% (2/13) of the studies [35,36] as a gold-standard measure of short-distance walking performance. Its relevance to 3D-printed ankle-foot orthoses lies in its ability to quantify clinically meaningful changes in gait capacity through walking speed, a core functional indicator of mobility and community ambulation after neurological injury [52,53]. The 10MWT is particularly well suited for integration into 3D printing rehabilitation protocols in resource-limited settings because it is rapid to administer and requires minimal equipment and training for clinicians and patients [52,53].

The systematic use of gold-standard rehabilitation assessments within studies investigating 3D printing interventions is particularly significant as they provide an important framework for objectively evaluating the efficacy of these technologies, ensuring that observed improvements in motor function are both measurable and clinically meaningful across various neurological populations. Therefore, by aligning 3D printing research with gold-standard tools, clinicians and researchers can ensure more accurate, comparable, and clinically meaningful assessments of motor recovery and functional gains.

Strengths and Limitations

A strength of this scoping review lies in its rigorous methodology, which included a comprehensive search strategy, clear inclusion criteria, and a structured data extraction form. By synthesizing evidence across diverse neurological conditions and outcome measures, this review provides a detailed understanding of how 3D-printed technologies contribute to motor recovery.

Although previous research has demonstrated improvements in motor recovery outcomes through 3D-printed rehabilitation technologies, several limitations remain. Studies included in this review had small sample sizes and heterogeneous intervention protocols, limiting the ability to draw generalizable conclusions. Additionally, the search strategy may have inadvertently excluded relevant studies by not expanding the keywords to include other discipline-specific terminology (eg, “occupational therapy,” “audiology,” and “nursing”). Not indexing these terms potentially narrowed the scope of the findings. Moreover, restricting inclusion to English-language publications may have introduced a language bias and limited the generalizability of the results.

All eligible studies were conducted in high- or middle-income countries, which limited comparisons across income groups and may have contributed to heterogeneity as the effectiveness and feasibility of 3D-printed neurorehabilitation devices may differ in infrastructure-constrained settings. Printer type, printing materials, and fabrication processes were inconsistently reported across the included studies. We reported available details descriptively (Table 1) but were unable to synthesize fabrication methods across studies or provide implementation guidance.

Conclusions

This scoping review highlights the growing interest in 3D-printed rehabilitation technologies within neurorehabilitation, particularly in the development of orthoses for motor recovery [12,14]. While preliminary findings suggest that 3D printing may offer a flexible and cost-effective means of delivering personalized interventions, the current evidence base remains limited by small sample sizes and heterogeneous study designs [14]. To strengthen the field, future research should prioritize larger and more diverse populations; standardized protocols; and consistent use of validated, clinically meaningful outcome measures [54]. These efforts are essential to better assess the functional impact and real-world applicability of 3D-printed technologies in neurorehabilitation [54].

Acknowledgments

The authors acknowledge the University of Ottawa for providing the institutional resources and academic environment that supported this research.

Funding

This research received no external funding. The authors did not receive any financial support for the conduct of this study.

Data Availability

The data extraction form used in this study is provided in Multimedia Appendix 3. The data that support the findings of this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: SA, ZL, JON
Data curation: SA, ZL
Formal analysis: SA, JON

XSL•FO
RenderX

Investigation: SA, ZL
Methodology: SA, ZL, JON
Project administration: SA, JON
Supervision: JON
Validation: JON
Visualization: SA
Writing—original draft: SA
Writing—review and editing: SA, ZL, JON

Conflicts of Interest

None declared.

Multimedia Appendix 1
Classification of countries by income level (2023).
[DOCX File, 18 KB - rehab_v13i1e81782_app1.docx]

Multimedia Appendix 2
Search query.
[PDF File, 89 KB - rehab_v13i1e81782_app2.pdf]

Multimedia Appendix 3
Modified Joanna Briggs Institute data extraction form.
[PDF File, 128 KB - rehab_v13i1e81782_app3.pdf]

Multimedia Appendix 4
Distribution of 3D-printed rehabilitation tools by health condition and motor areas.
[PNG File, 69 KB - rehab_v13i1e81782_app4.png]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 329 KB - rehab_v13i1e81782_app5.pdf]

References

1. Steinmetz J, Seeher K, Schiess NEA. Global, regional, and national burden of disorders affecting the nervous system, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Neurol 2024 Apr;23(4):344-381. [doi: 10.1016/S1474-4422(24)00038-3] [Medline: 38493795]
2. WHO Media Team. Over 1 in 3 people affected by neurological conditions, the leading cause of illness and disability worldwide. World Health Organization. 2024. URL: https://www.who.int/news/item/14-03-2024-over-1-in-3-people-affected-by-neurological-conditions--the-leading-cause-of-illness-and-disability-worldwide [accessed 2026-04-19]
3. The Lancet Neurology. Sex, gender, and the cost of neurological disorders. Lancet Neurol 2023 May;22(5):367. [doi: 10.1016/S1474-4422(23)00120-5] [Medline: 37059496]
4. Abuín-Porras V, Martinez-Perez C, Romero-Morales C, et al. Citation network study on the use of new technologies in neurorehabilitation. Int J Environ Res Public Health 2022;19(1):26. [doi: 10.3390/ijerph19010026] [Medline: 35010288]
5. Tamburin S, Smania N, Saltuari L, Hoemberg V, Sandrini G. Editorial: New advances in neurorehabilitation. Front Neurol 2019;10:1090. [doi: 10.3389/fneur.2019.01090] [Medline: 31681155]
6. Oña ED, Cano-de la Cuerda R, Sánchez-Herrera P, Balaguer C, Jardón A. A review of robotics in neurorehabilitation: towards an automated process for upper limb. J Healthc Eng 2018;2018:9758939. [doi: 10.1155/2018/9758939] [Medline: 29707189]
7. Dinon T, Caimmi M, Chiavenna A, et al. DUALarm: an open-source and 3D-printable device for upper limb neurorehabilitation. J Rehabil Assist Technol Eng 2018;5:2055668317749989. [doi: 10.1177/2055668317749989] [Medline: 31191920]
8. Burke S, McGettrick G, Foley K, Manikandan M, Barry S. The 2019 neuro-rehabilitation implementation framework in Ireland: challenges for implementation and the implications for people with brain injuries. Health Policy 2020 Mar;124(3):225-230. [doi: 10.1016/j.healthpol.2019.12.018] [Medline: 31964508]
9. Iosa M, Verrelli CM, Gentile AE, Ruggieri M, Polizzi A. Gaming technology for pediatric neurorehabilitation: a systematic review. Front Pediatr 2022;10:775356. [doi: 10.3389/fped.2022.775356] [Medline: 35155305]

XSL•FO
RenderX

10. Khan F, Amatya B, Galea MP, Gonzenbach R, Kesselring J. Neurorehabilitation: applied neuroplasticity. J Neurol 2017 Mar;264(3):603-615. [doi: 10.1007/s00415-016-8307-9] [Medline: 27778158]
11. Schneider WN, Wong TM. Cognitive and behavioral disorders in neurorehabilitation. Continuum (Minneap Minn) 2011 Jun;17(3 Neurorehabilitation):462-470. [doi: 10.1212/01.CON.0000399067.69567.f2] [Medline: 22810862]
12. Ishengoma FR, Mtaho AB. 3D printing: developing countries perspectives. Int J Comput Appl 2014;104(11):30-34. [doi: 10.5120/18249-9329]
13. Shahrubudin N, Lee TC, Ramlan R. An overview on 3D printing technology: technological, materials, and applications. Procedia Manuf 2019;35:1286-1296. [doi: 10.1016/j.promfg.2019.06.089]
14. Abbady HE, Klinkenberg ET, de Moel L, et al. 3D-printed prostheses in developing countries: a systematic review. Prosthet Orthot Int 2022 Feb 1;46(1):19-30. [doi: 10.1097/PXR.0000000000000057] [Medline: 34772868]
15. Dally C, Johnson D, Canon M, Ritter S, Mehta K. Characteristics of a 3D-printed prosthetic hand for use in developing countries. Presented at: 2015 IEEE Global Humanitarian Technology Conference (GHTC); Oct 8-11, 2015. [doi: 10.1109/GHTC.2015.7343956]
16. Dodziuk H. Applications of 3D printing in healthcare. Kardiochir Torakochirurgia Pol 2016 Sep;13(3):283-293. [doi: 10.5114/kitp.2016.62625] [Medline: 27785150]
17. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2024. [doi: 10.46658/JBIMES-24-01]
18. Kaufman JN, Lahey S, Slomine BS. Pediatric rehabilitation psychology: rehabilitating a moving target. Rehabil Psychol 2017 Aug;62(3):223-226. [doi: 10.1037/rep0000169] [Medline: 28836805]
19. Carr JH, Shepherd RB. Neurological Rehabilitation: Optimizing Motor Performance: Churchill Livingstone; 2010.
20. Roser M, Rohenkohl B, Arriagada P, Hasell J, Ritchie H, Ortiz-Ospina E. World Bank income groups. Our World in Data. 2025. URL: https://ourworldindata.org/grapher/world-bank-income-groups [accessed 2026-04-19]
21. Verde PE, Ohmann C. Combining randomized and non-randomized evidence in clinical research: a review of methods and applications. Res Synth Methods 2015 Mar;6(1):45-62. [doi: 10.1002/jrsm.1122] [Medline: 26035469]
22. Lawson DO, Mellor K, Eddy S, et al. Pilot and feasibility studies in rehabilitation research: a review and educational primer for the physiatrist researcher. Am J Phys Med Rehabil 2022 Apr 1;101(4):372-383. [doi: 10.1097/PHM.0000000000001797] [Medline: 34091466]
23. Rowley J. Using case studies in research. Manag Res News 2002;25(1):16-27. [doi: 10.1108/01409170210782990]
24. Whitlock EP, Lin JS, Chou R, Shekelle P, Robinson KA. Using existing systematic reviews in complex systematic reviews. Ann Intern Med 2008 May 20;148(10):776-782. [doi: 10.7326/0003-4819-148-10-200805200-00010] [Medline: 18490690]
25. Chen ZH, Yang YL, Lin KW, Sun PC, Chen CS. Functional assessment of 3D-printed multifunction assistive hand device for chronic stroke patients. IEEE Trans Neural Syst Rehabil Eng 2022;30:1261-1266. [doi: 10.1109/TNSRE.2022.3173034] [Medline: 35533167]
26. Demeco A, Foresti R, Frizziero A, et al. The upper limb orthosis in the rehabilitation of stroke patients: the role of 3D printing. Bioengineering (Basel) 2023 Oct 27;10(11):1256. [doi: 10.3390/bioengineering10111256] [Medline: 38002380]
27. Dudley DR, Knarr BA, Siu KC, Peck J, Ricks B, Zuniga JM. Testing of a 3D printed hand exoskeleton for an individual with stroke: a case study. Disabil Rehabil Assist Technol 2021 Feb;16(2):209-213. [doi: 10.1080/17483107.2019.1646823] [Medline: 31385727]
28. Huber J, Slone S, Bazrgari B. An evaluation of 3D printable elastics for post stroke dynamic hand bracing: a pilot study. Assist Technol 2023 Nov 2;35(6):513-522. [doi: 10.1080/10400435.2023.2177774] [Medline: 36780423]
29. Kuo FL, Wu YS, Kuo TY, Lee YS, Huang SW, Lee HC. Effects of 3D-printed assistive device on daily life function in patients with neurological impairment: a pilot study. Disabil Rehabil Assist Technol 2024 May;19(4):1648-1656. [doi: 10.1080/17483107.2023.2227222] [Medline: 37377434]
30. Yeh PC, Chen CH, Chen CS. Using a 3D-printed hand orthosis to improve three-jaw chuck hand function in individuals with cervical spinal cord injury: a feasibility study. IEEE Trans Neural Syst Rehabil Eng 2023;31:2552-2559. [doi: 10.1109/TNSRE.2023.3273300] [Medline: 37159308]
31. Toth L, Schiffer A, Nyitrai M, Pentek A, Told R, Maroti P. Developing an anti-spastic orthosis for daily home-use of stroke patients using smart memory alloys and 3D printing technologies. Mater Des 2020 Oct;195:109029. [doi: 10.1016/j.matdes.2020.109029]
32. Yang YS, Tseng CH, Fang WC, Han IW, Huang SC. Effectiveness of a new 3D-printed dynamic hand-wrist splint on hand motor function and spasticity in chronic stroke patients. J Clin Med 2021 Sep 30;10(19):4549. [doi: 10.3390/jcm10194549] [Medline: 34640564]
33. Yoo HJ, Lee S, Kim J, Park C, Lee B. Development of 3D-printed myoelectric hand orthosis for patients with spinal cord injury. J Neuroeng Rehabil 2019 Dec 30;16(1):162. [doi: 10.1186/s12984-019-0633-6] [Medline: 31888695]
34. Brognara L, Navarro-Flores E, Iachemet L, Serra-Catalá N, Cauli O. Beneficial effect of foot plantar stimulation in gait parameters in individuals with Parkinson's disease. Brain Sci 2020 Jan 27;10(2):69. [doi: 10.3390/brainsci10020069] [Medline: 32012779]

XSL•FO
RenderX

35. Hsu CY, Wu CM, Huang CC, Shie HH, Tsai YS. Feasibility and potential effects of robot-assisted passive range of motion training in combination with conventional rehabilitation on hand function in patients with chronic stroke. J Rehabil Med 2022 Aug 26;54:jrm00323. [doi: 10.2340/jrm.v54.1407] [Medline: 35925030]
36. Kwon J, Park JH, Ku S, Jeong Y, Paik NJ, Park YL. A soft wearable robotic ankle-foot-orthosis for post-stroke patients. IEEE Robot Autom Lett 2019;4(3):2547-2552. [doi: 10.1109/LRA.2019.2908491]
37. Liu Z, Zhang P, Yan M, Xie YM, Huang GZ. Additive manufacturing of specific ankle-foot orthoses for persons after stroke: a preliminary study based on gait analysis data. Math Biosci Eng 2019 Sep 9;16(6):8134-8143. [doi: 10.3934/mbe.2019410] [Medline: 31698659]
38. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2026-04-20]
39. Ashraf M, Choudhary N, Kamboh UA, et al. Early experience with patient-specific low-cost 3D-printed polymethylmethacrylate cranioplasty implants in a lower-middle-income-country: technical note and economic analysis. Surg Neurol Int 2022;13:270. [doi: 10.25259/SNI_250_2022] [Medline: 35855120]
40. Woodson T, Alcantara JT, do Nascimento MS. Is 3D printing an inclusive innovation?: an examination of 3D printing in Brazil. Technovation 2019;80:54-62. [doi: 10.1016/j.technovation.2018.12.001]
41. Nakipoğlu Yüzer GF, Koyuncu E, Çam P, Özgirgin N. The regularity of orthosis use and the reasons for disuse in stroke patients. Int J Rehabil Res 2018 Sep;41(3):270-275. [doi: 10.1097/MRR.0000000000000299] [Medline: 29939857]
42. Liz L, da Silva TG, Michaelsen SM. Validity, reliability, and measurement error of the remote Fugl-Meyer Assessment by videoconferencing: Tele-FMA. Phys Ther 2023 Aug 1;103(8):pzad054. [doi: 10.1093/ptj/pzad054] [Medline: 37255324]
43. Quinn TJ, Langhorne P, Stott DJ. Barthel index for stroke trials: development, properties, and application. Stroke 2011 Apr;42(4):1146-1151. [doi: 10.1161/STROKEAHA.110.598540] [Medline: 21372310]
44. Kumar A, Schmeler MR, Karmarkar AM, et al. Test-retest reliability of the functional mobility assessment (FMA): a pilot study. Disabil Rehabil Assist Technol 2013 May;8(3):213-219. [doi: 10.3109/17483107.2012.688240] [Medline: 22612721]
45. Khramtsov DM, Stoyanov O, Hruzevskyi O, Shaevchuk HY. Complex neurorehabilitation of post-stroke patients. Med Sci Ukraine 2022;18(3):52-59. [doi: 10.32345/2664-4738.3.2022.08]
46. Akter S. Measurement of hand grip strength by digital Jamar hand dynamometer: a study on CRP staffs & BHPI students [Dissertation]. : University of Dhaka; 2018 URL: http://library.crp-bangladesh.org:8080/xmlui/handle/123456789/520 [accessed 2026-04-20]
47. Bertrand AM, Fournier K, Wick Brasey MG, Kaiser ML, Frischknecht R, Diserens K. Reliability of maximal grip strength measurements and grip strength recovery following a stroke. J Hand Ther 2015;28(4):356-362. [doi: 10.1016/j.jht.2015.04.004] [Medline: 26206167]
48. Sánchez-Aranda L, Fernández-Ortega J, Martín-Fuentes I, et al. Reliability and criterion validity of a low-cost handgrip dynamometer: the Camry. medRxiv. Preprint posted online on Jun 25, 2024. [doi: 10.1101/2024.06.25.24309304]
49. Hobart JC, Thompson AJ. The five item Barthel index. J Neurol Neurosurg Psychiatry 2001 Aug;71(2):225-230. [doi: 10.1136/jnnp.71.2.225] [Medline: 11459898]
50. Taghizadeh G, Martinez-Martin P, Meimandi M, et al. Barthel Index and modified Rankin Scale: psychometric properties during medication phases in idiopathic Parkinson disease. Ann Phys Rehabil Med 2020 Nov;63(6):500-504. [doi: 10.1016/j.rehab.2019.08.006] [Medline: 31816448]
51. Sakamoto D, Hamaguchi T, Nakayama Y, Abo M. Reliability and validity of a newly developed Action Research Arm Test for upper limb function assessment in patients with stroke: a comparison with the conventional version. PLoS One 2026;21(3):e0334199. [doi: 10.1371/journal.pone.0334199] [Medline: 41875181]
52. Tasseel-Ponche S, Delafontaine A, Godefroy O, et al. Walking speed at the acute and subacute stroke stage: a descriptive meta-analysis. Front Neurol 2022;13:989622. [doi: 10.3389/fneur.2022.989622] [Medline: 36226075]
53. Peters DM, Fritz SL, Krotish DE. Assessing the reliability and validity of a shorter walk test compared with the 10-Meter Walk Test for measurements of gait speed in healthy, older adults. J Geriatr Phys Ther 2013;36(1):24-30. [doi: 10.1519/JPT.0b013e318248e20d] [Medline: 22415358]
54. Desai N. Addressing the gaps in assessment and treatment of the upper extremity in rehabilitation of individuals with neurological conditions [PhD thesis]. : University of Toronto; 2021 URL: https://utoronto.scholaris.ca/server/api/core/bitstreams/7da1c076-ef0a-4b08-88f0-03c44634e232/content [accessed 2026-04-20]

Abbreviations

10MWT: 10-m walk test
ARAT: Action Research Arm Test
FMA: Fugl-Meyer Assessment
JBI: Joanna Briggs Institute
LE: lower extremity
LMIC: low- and middle-income country
MeSH: Medical Subject Headings

XSL•FO
RenderX

RCT: randomized controlled trial
UE: upper extremity
WHO: World Health Organization

Edited by S Munce; submitted 05.Aug.2025; peer-reviewed by H Wang, J Fielder; revised version received 27.Feb.2026; accepted 23.Mar.2026; published 06.May.2026.

Please cite as:
Aggerwal S, Lalani Z, O'Neil J
The Use of 3D Printing Technology in Rehabilitation for Adults Living With Neurological Conditions: Scoping Review
JMIR Rehabil Assist Technol 2026;13:e81782
URL: https://rehab.jmir.org/2026/1/e81782
doi:10.2196/81782

© Salena Aggerwal, Zahra Lalani, Jennifer O'Neil. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 6.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

State-of-the-Art Knee Exoskeleton for Gait Assistance and Rehabilitation: Systematic Review

Dilnoza Karibzhanova[1], MSc, MD; Aibek Niyetkaliyev[2], PhD; Amna R Khawaja[3], DPT, MSc; Aray Zhaisanbek[2], BSc; Zeeshan[1], DPT; Ardana Utepbergen[2], BSc; Prashant K Jamwal[2], PhD; Kassymbek Ozhikenov[4], PhD; Aiman Ozhikenova[4,5], PhD; Chingiz Alimbayev[4], PhD; Zhadyra Alimbayeva[4], PhD; Zhanar Bigaliyeva[4], PhD; Lashin Bazarbay[4], PhD; Serik Didar[4], BSc

[1]School of Medicine, Nazarbayev University, Kabanbay Batyr 53 Street, Astana, Kazakhstan

[2]School of Engineering and Digital Sciences, Nazarbayev University, Astana, Kazakhstan

[3]Department of Physical Therapy, Rehabilitation Science and Athletic Training, University of Kansas Medical Center, Kansas City, KS, United States

[4]Department of Robotics and Technical Means of Automation, Satbayev University, Almaty, Kazakhstan

[5]Qazaq Youth Science Hub LLP, Almaty, Kazakhstan

Corresponding Author:
Dilnoza Karibzhanova, MSc, MD
School of Medicine, Nazarbayev University, Kabanbay Batyr 53 Street, Astana, Kazakhstan

Abstract

Background: Knee exoskeletons represent a significant advancement in wearable robotic technology, developed to enhance gait assistance and facilitate rehabilitation processes. These devices are increasingly used in clinical and research environments worldwide. Nonetheless, the existing empirical evidence supporting their biomechanical and functional outcomes lacks both breadth and clarity. Further investigation is essential to elucidate the specific effects of knee exoskeletons on gait dynamics and rehabilitation efficacy.

Objective: This study aims to systematically review the clinical and biomechanical evidence on knee exoskeletons used for gait assistance and rehabilitation.

Methods: A systematic search of databases—PubMed, Web of Science, IEEE Xplore, and Scopus—was conducted to identify eligible studies published between January 2015 and March 2025. Studies were included if they investigated knee exoskeleton systems in human participants for gait assistance and rehabilitation. Randomized controlled trials, observational studies, and experimental investigations were eligible. Exclusion criteria were multijoint exoskeletons, nonhuman studies, and simulation-based analyses. Two reviewers independently extracted study characteristics, participant demographics, device features, intervention protocols, and clinical and biomechanical outcomes. Risk of bias was assessed using RoB-2 (Cochrane Risk of Bias 2) for randomized controlled trials and ROBINS-I (Risk of Bias in Nonrandomized Studies of Interventions) for nonrandomized studies.

Results: Thirty-two studies met the inclusion criteria. Most investigations involved healthy adults in feasibility or pilot settings, while a smaller subset included individuals with stroke, spinal cord injury, or orthopedic conditions or pediatric movement disorders. Across studies, knee exoskeletons demonstrated short-term improvements in spatiotemporal gait parameters, including increased gait speed and step length, improved knee joint kinematics, and reductions in muscular effort or metabolic demand. However, substantial heterogeneity in device design, outcome metrics, and intervention duration limits comparability across studies. Sample sizes were typically small, and follow-up outcomes were rarely reported, restricting the interpretation of the long-term clinical efficacy of these devices.

Conclusions: Knee exoskeletons demonstrate promising short-term functional and biomechanical benefits and are increasingly feasible within structured rehabilitation settings. While current evidence is limited by small sample sizes and brief follow-up periods, emerging data suggest potential for enhancing task-specific gait training. Larger, well-designed trials are needed to determine sustained clinical impact and integration into routine rehabilitation practice.

Trial Registration: PROSPERO CRD420261285256; https://www.crd.york.ac.uk/PROSPERO/view/CRD420261285256

(JMIR Rehabil Assist Technol 2026;13:e94916) doi:10.2196/94916

KEYWORDS

knee exoskeleton; gait rehabilitation; wearable robotics; stroke; spinal cord injury; orthopedic rehabilitation

XSL•FO
RenderX

Introduction

Gait disorders and knee joint dysfunction are the most common consequences of neurological and orthopedic diseases, including postoperative conditions such as total knee replacement [1]. Neurological gait disorders are commonly associated with stroke, spinal cord injury, cerebral palsy, and degenerative conditions. At the same time, orthopedic pathologies such as osteoarthritis, joint degeneration, and skeletal deformities of the lower extremities also contribute to gait disorders and progressive disability [2,3]. Knee disorders impose a substantial and growing global burden across all populations. Osteoarthritis of the knee is especially prevalent, affecting approximately 4% - 5% of people worldwide (365 million in 2019) [4] and, by some estimates, about 16% of adults aged >40 years (654 million) [5]. Traumatic knee injuries (eg, sports or accident injuries) also contribute significant morbidity, especially in younger and athletic populations, often precipitating earlier osteoarthritis [6]. Even after surgical treatment (eg, total knee arthroplasty), many patients remain disabled: roughly 15% - 30% report persistent pain or functional limitation postoperatively [7,8]. Economic costs are correspondingly high: one systematic review found mean annual costs of €11,100 (€1=US $1.16 as of March 10, 2026) per knee/hip patient with osteoarthritis [9], and national expenditures reflect this trend. For example, osteoarthritis-related costs in Germany rose from €8.6 billion in 2015 to €12.1 billion in 2020 [10]. These impairments significantly affect mobility, functional independence, and quality of life across adult and pediatric populations.

The knee joint plays a central role in shock absorption, load transfer, and movement during walking and functional loads [11]. Several technology-enhanced modalities have been investigated as adjuncts to conventional knee rehabilitation, except for robotic exoskeletons. Virtual reality–based training has demonstrated clinically meaningful improvements in pain, strength, and function. In patients undergoing anterior cruciate ligament reconstruction, a meta-analysis of 9 randomized controlled trials reported significant reductions in pain and improvements in knee function, quadriceps strength, and knee flexion range of motion compared with standard therapy [12]. Similarly, pooled evidence from studies of rehabilitation in total knee arthroplasty (TKA) showed modest but significant improvements in pain, balance, and knee-specific function with virtual reality interventions [13]. Telerehabilitation and app-guided programs have also demonstrated outcomes comparable to face-to-face therapy, including improved pain relief, knee flexion, and quadriceps strength after TKA [14]. Wearable biofeedback systems further support functional recovery, with faster Timed Up and Go performance and higher patient satisfaction reported following TKA rehabilitation [15].

Despite these benefits, most nonrobotic technologies rely on feedback-based strategies and do not provide direct mechanical assistance to the knee joint during functional tasks.

Wearable knee exoskeletons offer a complementary approach by delivering task-specific, assist-as-needed mechanical support during gait, enabling high-repetition training while reducing physical demand and discomfort [16,17]. Exoskeleton-assisted rehabilitation has been associated with improved gait performance, muscle activation, and functional independence in neurological and orthopedic populations [18,19]. Emerging evidence also suggests that robotic gait training may promote central neuroplastic adaptations, supporting recovery beyond peripheral biomechanics [20]. In addition, knee exoskeletons are increasingly explored for fall prevention, with integrated recovery assistance shown to improve swing-limb kinematics and reduce fall risk in mobility-impaired populations [21]. This highlights the growing use of exoskeletons as both therapeutic and protective devices.

A wide range of technology-based interventions, including virtual reality, telerehabilitation, neuromuscular stimulation, and sensor-driven feedback systems, have been implemented in orthopedic and neurological knee rehabilitation [22-26]. Among these approaches, robotic knee exoskeletons have attracted increasing interest due to their capacity to provide biomechanical support and task-specific training. Despite growing evidence of their potential benefits, substantial gaps remain regarding consistency, affordability, and comparative effectiveness across technologies. With a focus on clinical outcomes, implementation challenges, and integration into rehabilitation care pathways, this systematic review examines knee exoskeletons for gait assistance and rehabilitation, with particular attention to their clinical applications and mechanical components.

This systematic review aimed to critically evaluate knee exoskeletons used for gait assistance and rehabilitation. The objectives were to identify the types of knee exoskeleton systems investigated, the clinical conditions and patient populations in which these devices have been applied, and the reported clinical and biomechanical outcomes associated with their use. In addition, the review identifies key gaps in the existing literature and proposes directions for future research to support the safe, effective, and scalable adoption of advanced rehabilitation technologies.

Methods

Protocol and Registration

This systematic review was registered on PROSPERO (International Prospective Register of Systematic Reviews) under registration number CRD420261285256. The review was conducted and reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines (Checklist 1). Minor methodological changes were made during full-text screening to align the risk-of-bias assessment tools with the final included study designs. These refinements did not affect the overall objectives, eligibility criteria, or review conclusions.

Search Strategy

A comprehensive literature search was conducted across 4 electronic databases: PubMed, Scopus, Web of Science, and IEEE Xplore. The search covered studies published between January 2015 and March 2025, with the final search performed in March 2025. These databases were selected to ensure

XSL•FO
RenderX

coverage of biomedical, rehabilitation, engineering, robotics, and wearable technology literature.

To improve sensitivity and minimize the risk of missing relevant studies, the search strategy combined keywords and synonyms related to exoskeleton technologies and the knee joint. These included terms such as “exoskeleton,” “robotic exoskeleton,” “wearable robot,” “powered orthosis,” and “robotic orthosis,” combined with “knee” or “knee joint,” and terms related to human or clinical testing (eg, “human,” “patient,” “clinical,” “trial,” and “experiment”). Studies that focus solely on simulation, musculoskeletal modeling, finite element analysis, or in silico approaches were excluded at the search level where possible.

Database-specific Boolean search strings were developed and adapted for each platform. The full reproducible search strategies for all databases are provided in Multimedia Appendix 1.

The search was limited to English-language publications due to feasibility constraints in screening and data extraction. Gray literature, trial registries, and industry- or manufacturer-reported studies were excluded because this review focused on peer-reviewed scientific evidence. These decisions may introduce language and publication bias, and this is acknowledged in the limitations.

Eligibility Criteria

Inclusion Criteria

Studies were included if they met the following criteria:

- Investigated knee-focused exoskeletons or robotic knee orthoses designed for gait assistance, rehabilitation, or functional lower-limb activities.
- Involved human participants, including healthy individuals, pediatric populations, or clinical populations with neurological or orthopedic impairments.
- Reported original experimental, observational, feasibility, or clinical trial data and were published as peer-reviewed full-text articles.
- Were published between 2015 and 2025.
- Were available in English.

Exclusion Criteria

Studies were excluded if they:

- Articles, systematic reviews, meta-analyses, editorials, protocols, conference abstracts, or opinion papers were reviewed.
- Investigated multijoint exoskeleton (hip-knee, knee-ankle, or full lower-limb exoskeletons) where knee-specific outcomes could not be isolated.
- Focused solely on simulation, finite element modeling, musculoskeletal modeling, or in silico analysis without human validation.
- Investigated devices not related to knee function or gait rehabilitation (eg, upper-limb exoskeletons).
- Lacked sufficient methodological or outcome data for extraction and evaluation.

Study Process

Following the initial database search, 71 duplicate records were removed, resulting in 2418 unique articles. Study selection was performed in multiple stages.

First, an AI-assisted abstract screening was applied to articles retrieved from Web of Science, IEEE Xplore, PubMed, and Scopus. Screening was performed using the GPT-4 large language model accessed through the OpenAI API. The model was not additionally trained on domain-specific datasets and relied on its pretrained language understanding capabilities. The filtering prompt and classification code were developed by reviewer AZ and applied prior to manual screening (provided in Multimedia Appendix 2). The AI-assisted screening used a prompt-based binary classification approach in which each abstract was evaluated to determine whether the study focused exclusively on knee exoskeletons. The model was instructed to retain studies by default and to exclude articles only when the abstract explicitly involved joints other than the knee (eg, hip, ankle, foot, or multijoint lower-limb systems). In ambiguous cases, studies were retained for further manual evaluation. This conservative filtering strategy was intentionally designed to prioritize sensitivity and minimize false exclusions.

Of the 2418 unique records, 1089 were excluded solely through AI-assisted abstract screening. To validate the performance of the AI-assisted filtering process, a random subset of AI-excluded records was manually reviewed by the authors. This validation confirmed that the majority of excluded studies involved multijoint or non–knee-focused exoskeleton systems. After AI-assisted screening, 71 articles from Web of Science, 21 from IEEE Xplore, 29 from PubMed, and 997 from Scopus remained eligible for subsequent manual screening.

Second, articles retained after AI-assisted screening underwent manual review to confirm eligibility and ensure that the studies focused exclusively on knee exoskeletons and represented original research rather than review articles. Manual screening was performed independently by 4 reviewers (DK, Zeeshan, AN, and ARK). Following AI-assisted and manual screening, 1166 articles were retained for full-text assessment. Full-text screening further evaluated eligibility based on the predefined inclusion and exclusion criteria, resulting in 112 studies selected for data extraction.

Data Extraction and Quality Assessment

Data extraction was performed for all eligible studies following full-text screening. For each study, the following information was systematically extracted and organized into structured tables: author name and DOI, article title, database source, device name, weight, mechanism type, control system, actuators, range of motion, clinical testing details, and patient type. These data formed the basis for qualitative and comparative analysis of knee exoskeleton technologies and applications.

Risk of Bias Assessment

The risk of bias was assessed based on the design and methodological characteristics of the included studies. This approach was adopted because the reviewed literature included randomized clinical trials, nonrandomized clinical studies,

XSL•FO
RenderX

experimental feasibility studies involving healthy participants, and technical validation studies. Risk of bias was assessed according to the study design: randomized controlled trials were evaluated using the RoB-2 tool. Nonrandomized clinical studies involving patient populations were assessed using the ROBINS-I (Risk of Bias in Nonrandomized Studies of Interventions) tool, which evaluates bias due to confounding, participant selection, intervention classification, deviations from intended interventions, missing data, outcome measurement, and selective reporting. Experimental and feasibility studies conducted in healthy participants were assessed using a modified Downs and Black checklist, as these studies typically focus on technical feasibility, biomechanical performance, usability, or short-term functional outcomes rather than clinical efficacy. Risk-of-bias assessment was independently performed by 3 reviewers (DK, PKJ, and Zeeshan), and disagreements were resolved through discussion. Nonhuman validation studies were not assessed using clinical risk-of-bias tools, as they do not evaluate health outcomes, and were analyzed descriptively with emphasis on device architecture, actuation strategy, sensing configuration, control method, and technical validation outcomes.

Risk-of-bias assessments were incorporated into the narrative synthesis, with greater emphasis placed on findings from studies demonstrating lower risk of bias or higher methodological quality.

Results

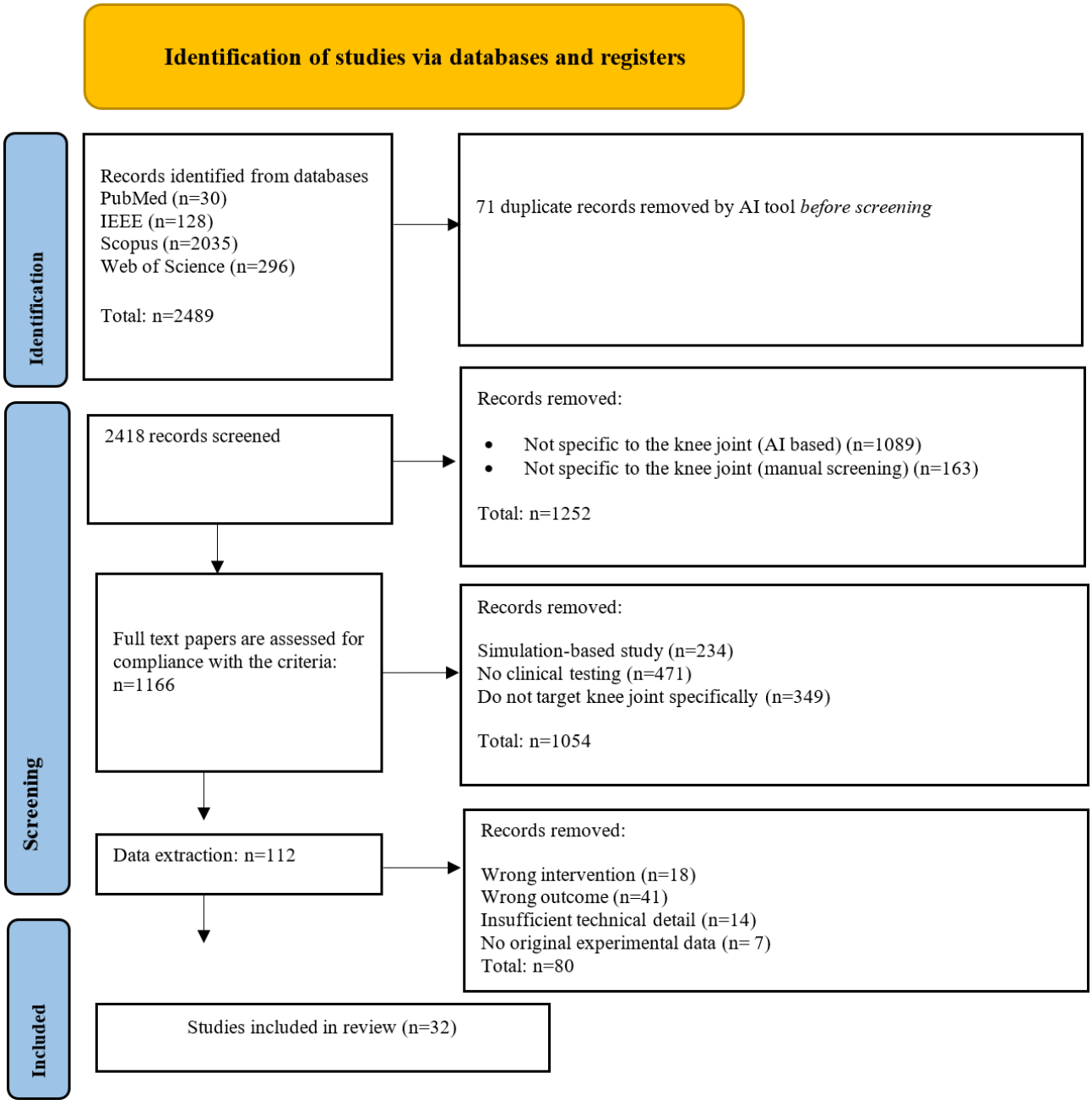
Study Selection

A total of 32 articles met all inclusion criteria and were included in the final review.

Figure 1 illustrates the overall study selection process, detailing the reasons for excluding certain articles and their corresponding numbers.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart on the selection of papers.



To reflect the multidisciplinary nature of this field and to ensure clarity in synthesizing the evidence, the Results section was divided into two complementary parts. The first part focuses on the engineering perspective, summarizing design concepts, actuation strategies, and control approaches that enable the functionality of knee exoskeletons (Table 1). The second part addresses the clinical perspective, highlighting patient populations, rehabilitation goals, and functional outcomes associated with the use of knee exostheses (Table 1). This structured approach highlights the importance of integrating engineering advances with clinically relevant outcomes to ensure effective rehabilitation.

XSL•FO
RenderX

Table . Clinical and engineering characteristics of knee exoskeleton systems.

Author and year	Study design	Device name	Population		Sensor modality	Control architecture	Actuation type	Maximum assistive torque (Nm)
			Group	Sample size				
Pugliese et al, 2025 [27]	Experimental study	Wearable POF[a] knee-sleeve kinematics estimator	Healthy individuals	31	Polymer optical fiber	Predictive learning-based control	Sensor only	—[b]
Barquín-Santos et al, 2024 [28]	Nonrandomized clinical trial	MAK[c] robotic device	Postsurgical	6	Joint encoder and torque sensing	Trajectory tracking control	Motor-driven with gearbox	10 to 30
Puyuelo-Quintana et al, 2020 [29]	Cross-sectional study	Portable lower-limb exoskeleton (prototype)	Stroke and Multiple Sclerosis	5	Joint encoders and foot pressure sensors	Gait-phase-based control	Motor-driven with transmission	10 to 25
Monteiro et al, 2024 [30]	Experimental study	Knee exoskeleton with HIL[d]-optimized assistance (name not specified)	Healthy individuals	5	Joint encoder and gait event sensing	Assist as needed adaptive control	Motor-driven	5 to 20
Dai et al, 2024 [31]	Experimental study	Rigid-soft hybrid knee exoskeleton (name not specified)	Healthy individuals	5	Foot pressure sensors, encoders, and inertial sensors	Gait-phase-based control	Motor-driven cable transmission	10 to 30
Zhou et al, 2023 [32]	Experimental study	Passive knee-assisted exoskeleton (spring mechanism)	Healthy individual	1	—	Passive control	Passive	—
Kittisares et al, 2023 [33]	Single-case feasibility study	Dual 4-bar linkage knee exoskeleton	Healthy individual	1	Hydraulic pressure sensing and joint angle sensing	Gait-phase-based control	Hydraulic artificial muscle	15 to 40
Wang et al, 2021 [34]	Experimental study	Adaptive knee exoskeleton with buffering function	Healthy individual	1	Joint angle sensor and motor encoder	Trajectory tracking control	Motor-driven	10 to 30
Kim et al, 2015 [35]	Experimental study	Knee exoskeleton prototype	Healthy individual	1	Foot pressure sensors and knee torque sensors	Gait-phase-based control	Motor-driven with gearbox	5 to 20
Villa-Parra et al, 2018 [36]	Experimental study	sEMG[e]-controlled knee exoskeleton	Healthy individuals	12	Surface electromyography and joint encoder	Bio-signal proportional control	Motor-driven	5 to 20
Chen et al, 2018 [37]	Benchtop validation study	Wearable pediatric robotic knee exoskeleton	No human participants	—	Joint encoders, foot pressure sensors, and inertial sensors	Gait-phase-based control	Motor-driven with transmission	5 to 15
Wang et al, 2018 [38]	Benchtop validation study	Lightweight backdrivable knee exoskeleton	No human participants	—	Joint encoder and torque sensing	Force and impedance control	Motor-driven low-ratio transmission	5 to 20

XSL•FO
RenderX

Author and year	Study design	Device name	Population		Sensor modality	Control architecture	Actuation type	Maximum assistive torque (Nm)
			Group	Sample size				
Jammeli et al, 2021 [39]	Experimental study	Knee rehab exoskeleton with explicit MPC[f]	Healthy individuals	3	Joint encoder and velocity estimation	Predictive learning-based control	Motor-driven	10 to 30
Park et al, 2020 [40]	Experimental study	Hinge-free tethered knee exosuit	Healthy individuals	6	Cable tension sensing and gait-phase detection	Gait-phase-based control	Motor-driven tendon system	5 to 15
Mosconi et al, 2024 [41]	Experimental study	Knee exoskeleton interaction testbed	Healthy individual	1	Joint encoder and torque sensor	Force and torque control	Motor-driven	5 to 30
Rodríguez-Fernández et al, 2022 [42]	Randomized, crossover clinical trial	Knee-powered exoskeleton and knee-ankle-foot orthoses	Spinal cord injury	10	Joint encoders and gait-phase sensing	Gait-phase-based control	Motor-driven	10 to 40
Kittisares et al, 2020 [43]	Experimental study	Four-bar linkage + HAM[g] knee support device	Healthy individual	1	Hydraulic pressure sensing and angle measurement	Low-level control	Hydraulic artificial muscle	15 to 40
Lora-Millan et al, 2023 [44]	Nonrandomized clinical trial	REFLEX knee exoskeleton	Poststroke hemiparetic	7	Joint encoders, inertial sensors, and foot pressure sensors	Assist-as-needed adaptive control	Motor-driven	10 to 30
Velasco-Guillen et al, 2024 [45]	Experimental study	Torque-assisted knee exoskeleton with elastic fault compensation	Healthy individual	1	Joint encoder and elastic torque sensing	Compliance and fault management	Motor-driven series elastic actuator	10 to 30
Zhang et al, 2024 [46]	Experimental study	Unilateral active knee exoskeleton with intention recognition	Stroke patients with hemiplegic paralysis	3	Inertial sensors, foot pressure sensors, and joint encoders	Predictive learning-based control	Motor-driven	5 to 25
Olinski, 2025 [47]	Experimental study	Knee joint exoskeleton prototypes	Healthy individuals	2	Kinematic measurement systems	Low-level control	Motor-driven or passive	10 to 30
Zhang et al, 2023 [48]	Experimental study	Neuromusculoskeletal model-informed machine learning–based control of a knee exoskeleton	Healthy individuals	8	Kinematic sensing	Predictive learning-based control	Motor-driven	5 to 25
Hidayah et al, 2021 [49]	Experimental study	SpringExo is a coil-spring design that aids in knee extension	Healthy individuals	7	External motion capture and force measurement	Passive control	Passive elastic element	—
Eveld et al, 2023 [21]	Experimental study	Stumble-recovery knee exoskeleton	Healthy individuals	3	Inertial sensors, joint encoder, and foot sensors	Gait-phase-based control	Motor-driven	15 to 40

XSL•FO
RenderX

Author and year	Study design	Device name	Population		Sensor modality	Control architecture	Actuation type	Maximum assistive torque (Nm)
			Group	Sample size				
Shideler et al, 2020 [50]	Experimental study	Pediatric hybrid exoskeleton + noninvasive Neuromuscular Electrical Stimulation	A teenager with bilateral spastic CP[h]	1	Gait-phase sensing and stimulation control	Bio-signal proportional control	Electrical stimulation and motor-driven	5 to 15
Lerner et al, 2017 [51]	Nonrandomized clinical trial	Lower-extremity exoskeleton for crouch gait	Children with CP	7	Joint encoders and gait event sensing	Gait-phase-based control	Motor-driven	5 to 15
Ben-David et al, 2022 [52]	Experimental study	Passive knee exoskeleton for jumping	Healthy individuals	10	—	Passive control	Passive	—
Goffredo et al, 2022 [53]	Observational study	Exoskeleton used for community ambulation (name not specified)	Healthy individuals	5	Device encoders	Gait-phase-based control	Motor-driven	10 to 40
Sarkisian et al, 2021 [54]	Experimental study	Powered knee exoskeleton with self-aligning mechanism	Healthy individuals	14	Joint encoder and interaction force sensing	Force and impedance control	Motor-driven	10 to 30
Wang et al, 2023 [55]	Experimental study	Knee flexion-assisted human-exoskeleton system	Healthy individuals	7	Joint encoder and gait-phase sensing	Gait-phase-based control	Motor-driven	5 to 20
Wang et al, 2023 [56]	Experimental study	Rigid-soft hybrid knee exoskeleton	Healthy individuals	6	Joint encoder inertial sensors and foot pressure sensors	Gait-phase-based control	Motor-driven cable transmission	10 to 30
de Miguel Fernández et al, 2023 [57]	Nonrandomized clinical trial	Knee exoskeleton for adapted assistance/resistance training	Poststroke individuals	6	Joint encoder inertial sensors and foot pressure sensors	Force and impedance control	Motor-driven	10 to 30

[a]POF: polymer optical fiber.

[b]Not applicable.

[c]MAK: Marsi active knee.

[d]HIL: human-in-the-loop.

[e]sEMG: surface electromyography.

[f]MPC: model predictive control.

[g]HAM: hydraulic artificial muscle.

[h]CP: cerebral palsy.

Part 1: Engineering Perspective

Actuation Systems

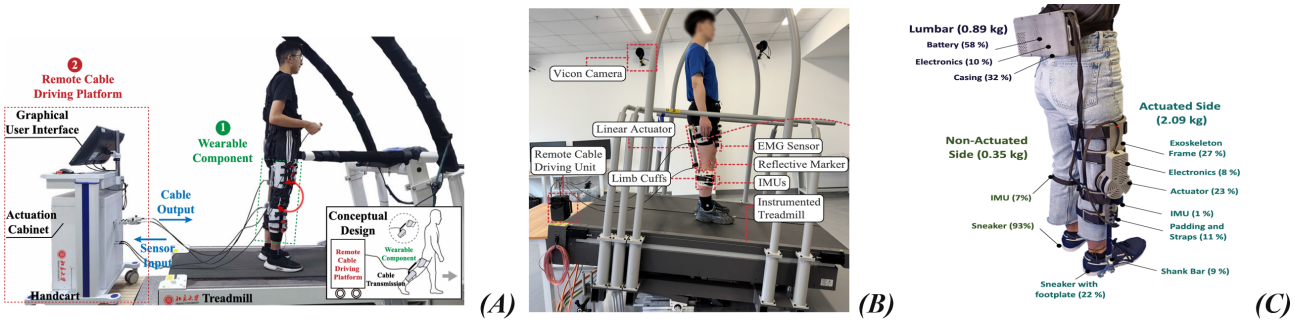
The choice of actuator in knee exoskeletons is constrained by the trade-off between high-fidelity torque control and the penalties of increased mass, complexity, and reduced portability. Three prevailing actuation paradigms can be identified within this design space: motor-driven rigid joints, compliant or distributed actuation systems, and passive elastic mechanisms.

The most common solution has been electric motor-driven actuation for rehabilitation-focused devices and gait-assist, one of which is shown in Figure 2A. In such systems, direct current or brushless direct current motors are typically used, with gear reductions used to provide repeatable, controllable knee torques, which can be used in closed-loop torque or impedance control

XSL•FO
RenderX

structures [29,34,38,42,56]. It provides accurate command tracking and can be easily combined with sensor-based feedback required for safe interactions during rehabilitation activities. In the reviewed literature, assistive torques typically range from 5 to 30 Nm, which is consistent with the mechanical loads of level walking and early rehabilitation [29,31,36,44,56]. Stair-climbing or perturbation-recovery devices are designed for more challenging tasks that require much higher torque capability, with peak assistive torques of nearly 40 Nm [21,33]. They usually accept larger actuators, higher transmission gears, and external power. To dampen interaction forces in high-impedance motorized joints, several papers have added compliance to the actuation chain, either in series-elastic or mechanically compliant form. This was done to enhance shock resistance and the accuracy of torque regulation [45,54].

Figure 2. Representatives of different actuation architectures in knee exoskeletons: (A) a rigid motor-driven powered exoskeleton using electric motor actuation with gearbox transmission for precise torque control [56], (B) a rigid-soft hybrid exoskeleton combining a rigid frame with compliant elements to improve comfort and reduce joint misalignment [31], and (C) a passive knee exoskeleton using coil-spring mechanisms to store and return elastic energy during gait without external power [52]. EMG: electromyography; IMU: inertial measurement unit.



By contrast, a second design focuses on interaction quality and wearability with compliant or distributed actuation. Examples of such solutions include hybrid rigid-soft exoskeletons and hinge-free textile-based exosuits, which redistribute mass closer to the interface, as shown in Figure 2B [31]. The objects of these systems maintain natural knee kinematics and reduce discomfort associated with joint misalignment. Although compliant actuation is less disruptive and less damaging, it presents engineering challenges, including transmission and friction losses and a more complex calibration process.

The third paradigm comprises passive knee exoskeletons that store and recover elastic energy without external power or closed-loop control. Assistance during certain stages of motion is provided by coil springs and compliant mechanisms, as shown in Figure 2C, thereby reducing device mass and simplifying the mechanism [32,49,52]. These systems have the benefit of being portable and low-complexity, but do not offer flexibility across users and tasks.

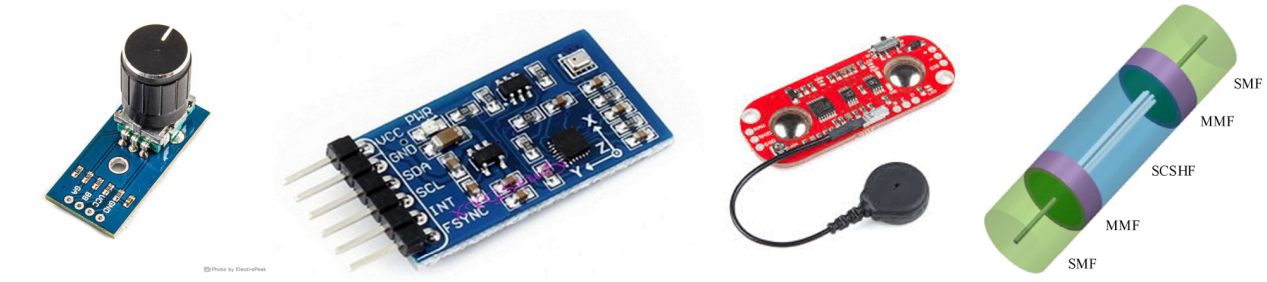
The smaller group of papers investigated hydraulic artificial muscles as an alternate actuation concept, also driven by their high force density and natural compliance, especially when a high torque is needed in large knee flexion angles [33,43]. Given these merits, fluidic actuation systems introduce considerable complexity in pressure and power control. They can only be applied to tethered concept prototypes, thereby hampering translational scalability.

Sensor Design and Engineering Integration

All powered knee exoskeletons have joint-level kinematic sensing as the basic sensing layer. Rotary encoders, as shown in Figure 3A, in the actuation or transmission system can provide high-resolution knee angle measurements needed to control closed-loop torque, knee position, or impedance [34,38,39,54]. In engineering terms, encoder-based sensing operates in a high-bandwidth, low-latency range, but its accuracy also depends on mechanical alignment. Various studies underscore that encoder measurements reflect device joint motion rather than the actual motion of the anatomy, making them susceptible to joint-axis misalignment and soft-tissue deformation.

XSL•FO
RenderX

Figure 3. Examples of sensor modalities integrated into knee rehabilitation exoskeletons: (A) rotary encoder mounted at the joint axis for high-resolution knee angle measurement [58], (B) inertial measurement unit (IMU) attached to the leg segment for gait-phase detection and swing-stance classification [59], (C) surface electromyography (sEMG) electrodes placed over the quadriceps for real-time motor intention estimation [60], and (D) polymer optical fiber (POF) embedded in a textile sleeve for nonrigid kinematic sensing [61]. MMF: multimode fiber; SCSHF: single-core six-hole microstructure optical fiber; SMF: single-mode fiber.



(A) Rotary encoder *(B) IMU sensor* *(C) EMG sensor* *(D) Optic fiber*

Most systems use gait-phase detectors based on inertial measurement units (IMUs), shown in Figure 3B, and plantar pressure sensors to provide task-oriented support, particularly while walking. IMUs provide the segment's angular velocity and orientation, thereby supporting the detection of swing-stance transitions across different walking speeds [29,31,46,56]. Plantar pressure sensors or foot switches provide temporally accurate event information, for example, heel strike and toe-off, and are often used with IMU data to enhance robustness and minimize false state changes [35,46]. The current success of multimodal gait sensing has been a companion to recognition throughout the field due to the limited reliability of single-sensor modalities across conditions, particularly for impaired gait patterns.

A smaller body of related work uses surface electromyography (EMG) to infer purpose and control assistance in real time [36,48,50]. In engineering terms, EMG, shown in Figure 3C, provides an entirely new control signal that occurs before joint motion and, in theory, may enable predictive assistance. There are some practical difficulties identified with the implementation of EMG, such as electrode placement sensitivity, signal variability due to fatigue or perspiration, and susceptibility to motion artifacts. These properties require complex signal-processing pipelines and frequent recalibration, thereby restricting the use of EMG-based sensing to a laboratory setting.

The third way is an emerging technique in nonrigid, textile-based sensing technologies to overcome the comfort and alignment constraints of solid sensor attachment. The use of polymer optical fiber-based strain sensors, shown in Figure 3D, in wearable sleeves indicates that it is possible to measure knee kinematics without placing mechanical limits on the joint [27]. Although these methods make wearable objects less sensitive to alignment and motion tracking, they are currently used predominantly for motion detection rather than for closed-loop actuation.

Control Strategies and Algorithms

The control strategies for knee exoskeletons have been designed to balance accuracy in assistance timing, stability in interactions, and computational viability against real-time requirements. In the literature reviewed, control architectures are always hierarchical, with high-level task or gait-state estimation.

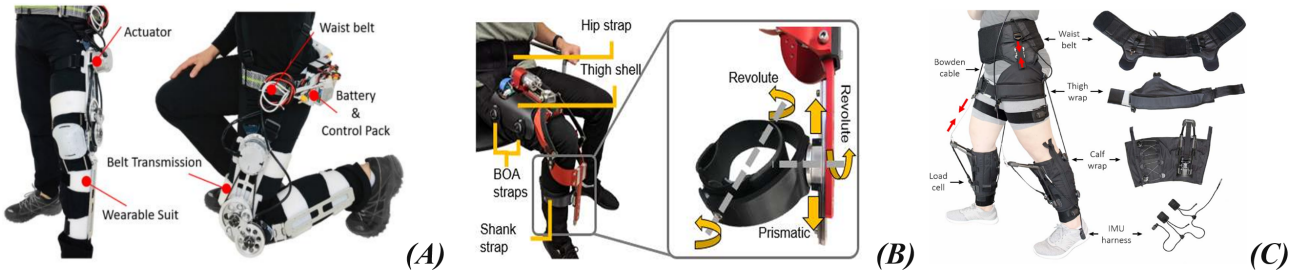
A finite-state machine (FSM)–controlled walking aid is most often implemented for gait-phase-driven control. FSMs allow deterministic, phase-specific assistance modulation in stance and swing, and conform control output to the gait cyclic pattern, with minimal computational cost [29,31,42,46,48]. The method is quite appealing for rehabilitation because the duration of assistance can be predicted. However, FSM-based control is very sensitive to both the correctness and precision of state transitions; misclassifying the gait phase will cause assistance at the wrong time and result in poor-quality interaction.

Some studies use assist-as-needed control strategies, in which the level of assistance is regulated based on the user's interaction forces or task execution [37,39,57]. From an engineering perspective, these solutions are based on the localization of control complexity. That is, a way of state timing, which necessitates stable low-level controllers and robust sensing. Although assist-as-needed paradigms provide enhanced customization, they are mostly done in a heuristic manner, and few formal reports on stability indicators have been described in the studies reviewed.

Knee exoskeletons are dominated by low-level torque- and impedance-based control. Torque monitoring enables direct regulation of assistive moments and is applicable to systems with built-in torque sensors or series compliance [34,45]. Instead, we use impedance control, which is commonly used to reduce interaction forces caused by joint misalignment and to maintain back drivability, particularly in comfort-based designs as depicted in Figure 4A,B [38,54]. As illustrated in Figure 4C, the hinge-free exosuit redistributes actuation forces through Bowden cable routing anchored to textile thigh and calf wraps, with a load cell and IMU harness enabling force and kinematic sensing without imposing any rigid constraint on the knee joint axis. Though it has less control bandwidth, it offers the benefit of perceived transparency, at the cost of achievable responsiveness in high-torque or high-speed processes.

XSL•FO
RenderX

Figure 4. Representatives of different structural design and alignment strategies in knee exoskeletons: (A) a lightweight, back drivable knee exoskeleton featuring a low-ratio transmission for improved transparency and reduced distal inertia [38], (B) a powered knee exoskeleton with a self-aligning mechanism designed to accommodate anatomical variation and reduce parasitic joint forces [54], and (C) a hinge-free tethered exosuit that replaces rigid linkages with textile-based cable routing, preserving natural knee kinematics [40]. BOA: BOA system straps; IMU: inertial measurement unit.



More complex control techniques, such as explicit model predictive control, human-in-the-loop optimization, and machine-learning-based controllers, are reported in fewer studies [30,39,48]. The same strategies are expected to enhance adaptability by incorporating predictive models and iterative optimization. They have shown promising results in controlled experiments but are only practical when model-dependent, computationally expensive, and require large amounts of calibration or training data.

Engineering Limitations and Challenges

Although substantial advances have been made in developing knee exoskeletons, the reviewed articles indicate a range of inherent engineering issues that limit their effectiveness, wearability, and translational appropriateness. Such restrictions are not limited to designs but are repeated as systems of the subsystems of actuation, sensing, control, and human-robot interface.

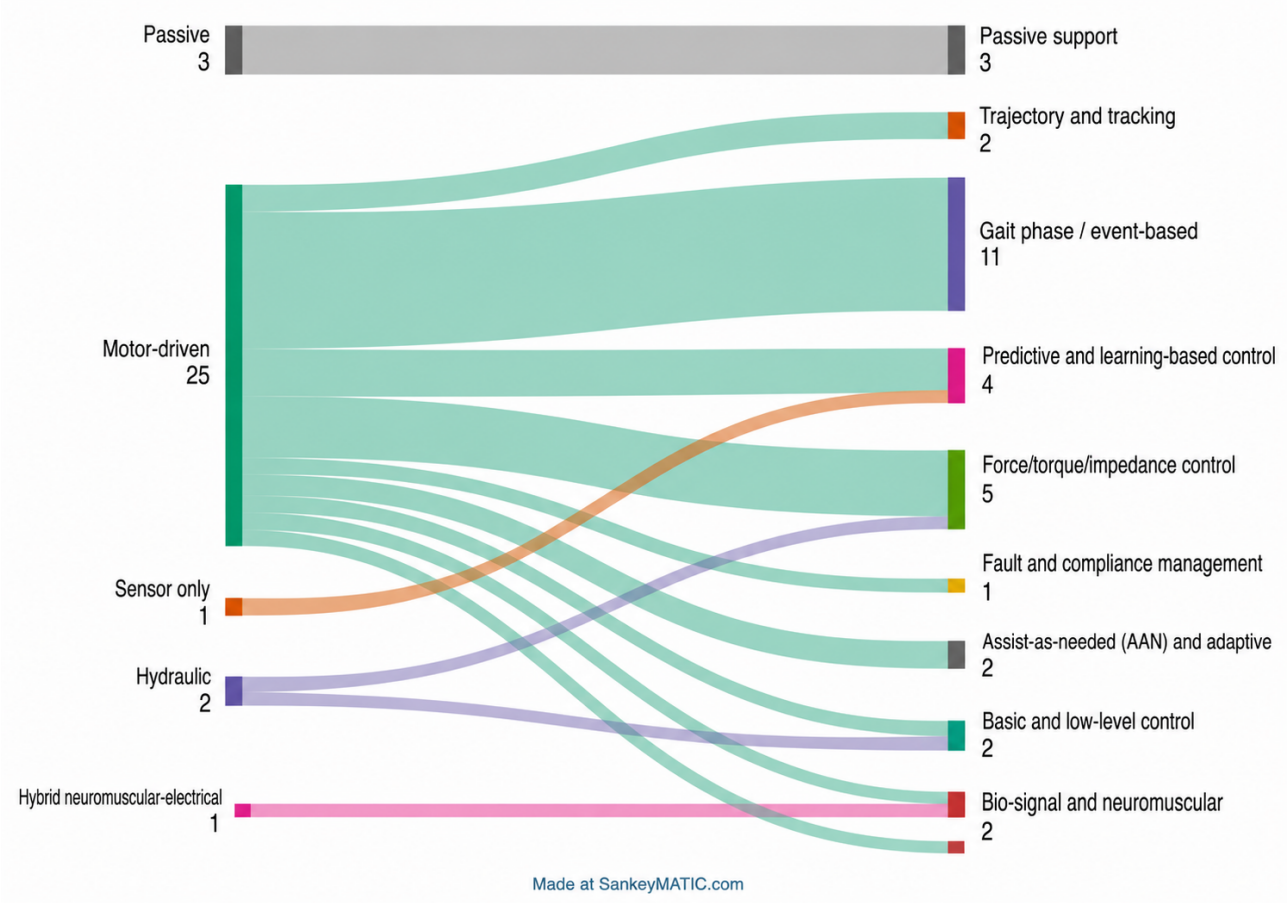
One of the major limitations is the mechanical misalignment between the exoskeleton joint and the anatomical knee axis, which increases discomfort and reduces the efficacy of assistance. In systems with tuned joint-level sensory and torque control, the error is due to soft-tissue deformity and anatomical differences in residual limbs. Multiple attempts to reduce these effects include self-aligning or compliant interfaces, but these solutions introduce additional mechanical complexity and mass and represent an open-ended trade-off between alignment strength and structural simplicity [54].

Another prevailing challenge is the mass and inertia of devices. Exoskeletons with motorized knees usually require greater mass and battery capacity, leading to higher distal inertia and reduced user acceptance. To deliver mass in proximity and enhance wearability, hybrid rigorous-soft architecture and remote actuation strategies are currently being explored, but these methods have their own problems: transmission friction, hysteresis, and sensitivity to environmental conditions make repeatable torque delivery difficult to achieve [31,40,56].

Among other observations, one major finding is that the choice of actuation implicitly limits the entire system architecture. Figure 5 demonstrates that motor-driven systems prevail in the field, and the most diverse control strategies are supported. This is an indication of the predictability and controllability of the electric motors, as well as an indicator that designers are willing to trade the added mass, inertia, and alignment sensitivity to gain the ability to predict the torque generation. Other actuation ideas, such as compliant, passive, or hydraulic actuation, are much less common because they introduce new issues related to controllability, scalability, or power control [34,42]. This implies that the design of knee exoskeletons is constrained by the challenge of achieving both high torque authority and wearability within a single system.

XSL•FO
RenderX

Figure 5. Sankey diagram illustrating the relationship between actuation mechanisms and control architectures across the included knee exoskeleton studies. Arrow width is proportional to the number of studies using each combination. Motor-driven actuation is linked to the widest range of control strategies, reflecting its prevalence and versatility in the reviewed literature. Diagram created using SankeyMATIC [62].



Limitations about sensing also limit the reliability of control. Encoder-based joint sensing quantifies device motion rather than anatomy motion and is thus sensitive to alignment errors, whereas IMU-based gait phase estimation is prone to errors in abnormal gait patterns. All these concerns are part of the problem of the absence of standard benchmarks for sensing accuracy, latency, and durability under realistic operating conditions.

These sensing and mechanical constraints are closely linked to control challenges. FSM-based gait controllers offer robustness and simplicity but lack flexibility when gait patterns deviate from nominal cycles. On the other hand, adaptive and learning-based controllers impose additional computational and modeling constraints that are likely to compromise real-time reliability and safety. It is worth noting that there is limited research on formal stability assessment or long-term assessment, and there remains an open question regarding the performance of controllers when fatigued, affected by sensor drift, or used over long periods of the day [30,39,48].

Lastly, the absence of engineering standards impedes meaningful comparisons across studies. Measures such as energy efficiency, long-term functionality, and durability are not commonly addressed, even though they directly affect practical use. In the absence of regular evaluation of these factors, it remains difficult to estimate the effectiveness of these design trade-offs in nonexperimental settings.

To conclude, engineering issues include energy efficiency and long-term operability. Battery capacity, power consumption, and thermal management are rarely measured in standardized units, even though they directly influence portability and user compliance.

Risk of Bias Quality

Risk of bias and methodological quality were assessed according to study design using the Cochrane RoB-2 tool [63] for randomized controlled trials, ROBINS-I [64] for nonrandomized clinical studies, and a modified Downs and Black checklist [65] for experimental and feasibility studies (Figures 6 and 7, Table 2). Among the nonrandomized clinical studies assessed using ROBINS-I, most were rated as having a moderate risk of bias [28,29,44,51,57] (Figure 6). The most common sources of bias were confounding, participant selection, and lack of blinding during intervention delivery and outcome assessment. One study was classified as having a serious risk of bias due to substantial methodological limitations and potential confounding [33]. The randomized crossover clinical trial by Rodríguez-Fernández et al [42] was rated as having an overall low risk of bias according to the RoB-2 tool (Figure 7). However, some concerns were identified regarding deviations from intended interventions, as participant blinding was not feasible in wearable exoskeleton studies.

XSL•FO
RenderX

Figure 6. ROBINS-I traffic light plot for nonrandomized clinical studies. Risk of bias assessment of nonrandomized clinical studies using the ROBINS-I tool [28,29,33,44,51,57]. Created using the Robvis app. ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions.

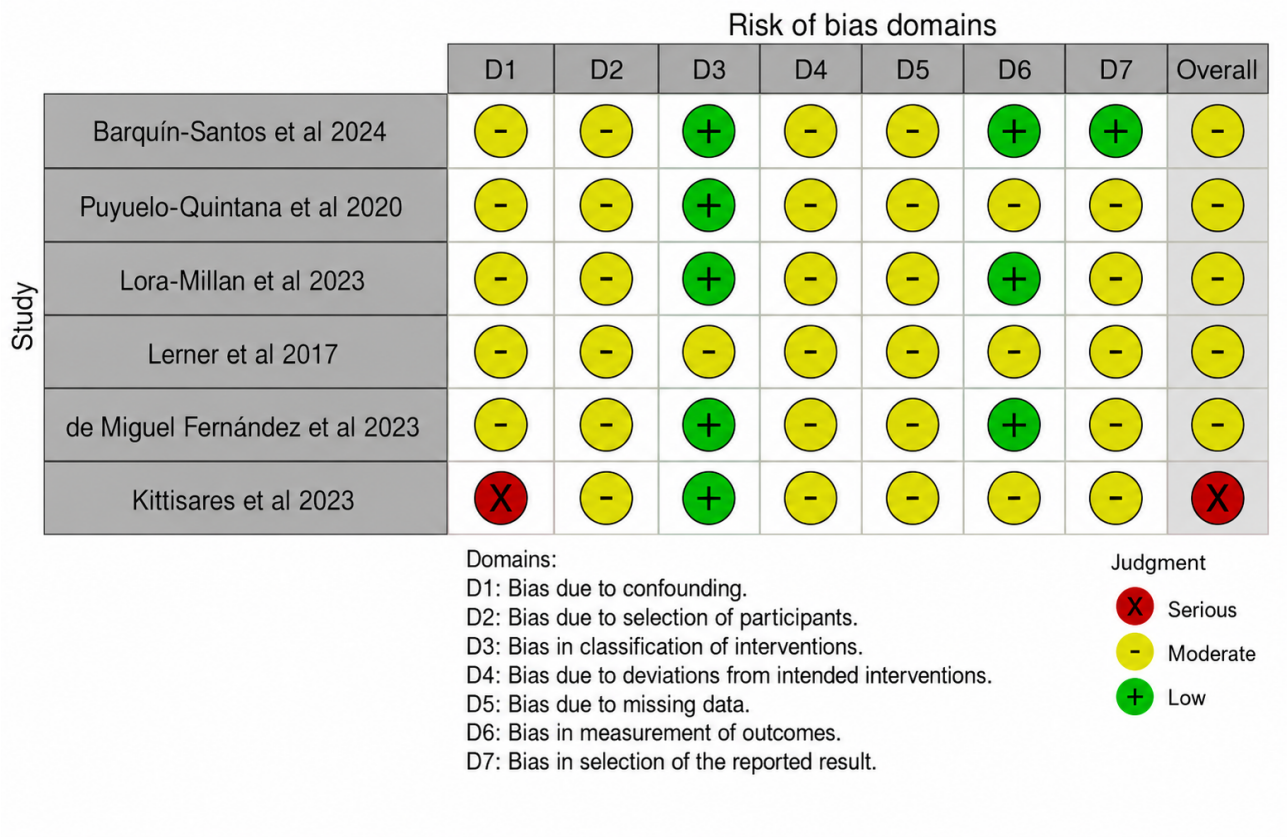
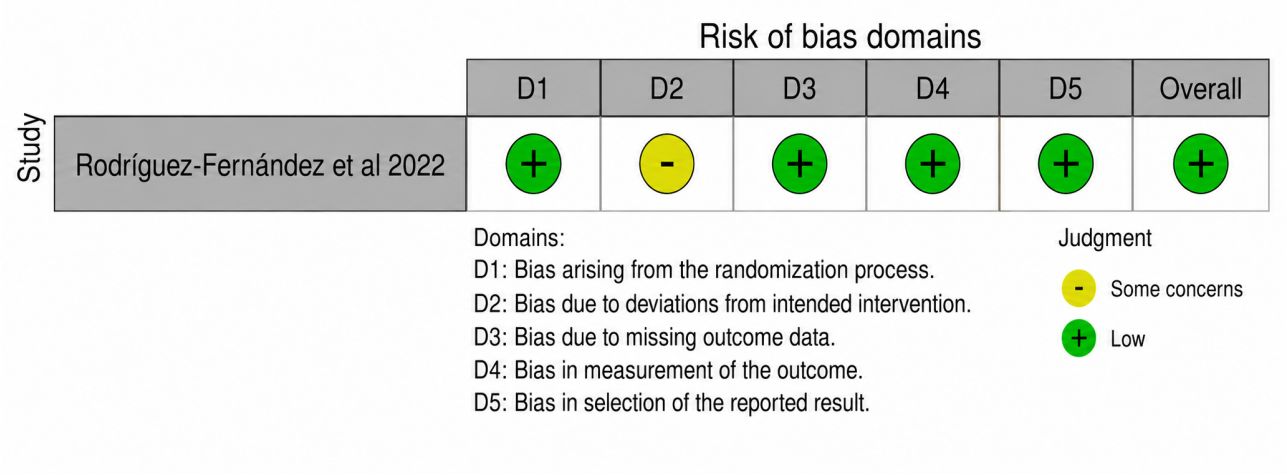


Figure 7. RoB-2 traffic light plot for randomized controlled trials. Risk of bias assessment of the randomized crossover clinical trial using the Cochrane Risk of Bias 2 (RoB-2) tool [42]. Created using the Robvis app.



XSL•FO
RenderX

Table . Methodological quality assessment of experimental and feasibility studies using the modified Downs and Black checklist. Scoring was conducted using the 27-item modified Downs and Black checklist, encompassing the domains of reporting quality, external validity, internal validity (bias and confounding), and statistical power [65].

Study	Reporting	External validity	Bias	Confounding	Power	Total	Quality
Pugliese et al, 2025 [27]	8	1	4	3	0	16	Fair
Monteiro et al, 2024 [30]	9	1	5	4	0	19	Good
Dai et al, 2024 [31]	9	1	5	4	0	19	Good
Villa-Parra et al, 2018 [36]	8	1	4	3	0	16	Fair
Jammeli et al, 2021 [39]	8	1	4	3	0	16	Fair
Park et al, 2020 [40]	8	1	4	3	0	16	Fair
Mosconi et al, 2024 [41]	7	1	4	3	0	15	Fair
Velasco-Guillen et al, 2024 [45]	8	1	4	3	0	16	Fair
Olinski et al, 2025 [47]	7	1	4	3	0	15	Fair
Zhang et al, 2023 [48]	8	1	4	3	0	16	Fair
Hidayah et al, 2021 [49]	8	1	4	3	0	16	Fair
Eveld et al, 2023 [21]	8	1	5	3	0	17	Fair
Ben-David et al, 2022 [52]	9	1	5	4	0	19	Good
Sarkisian et al, 2021 [54]	8	1	4	3	0	16	Fair
Wang et al, 2023 [55]	8	1	4	3	0	16	Fair
Wang et al, 2023 [56]	8	1	4	3	0	16	Fair
Kittisares et al, 2023 [33]	8	1	4	3	0	16	Fair
Wang et al, 2021 [34]	8	1	4	3	0	16	Fair
Kim et al, 2015 [35]	8	1	4	3	0	16	Fair
Zhou et al, 2023 [32]	8	1	4	3	0	16	Fair
Kittisares et al, 2020 [43]	8	1	4	3	0	16	Fair
Zhang et al, 2024 [46]	8	1	4	3	0	16	Fair
Goffredo et al, 2022 [53]	7	1	4	3	0	15	Fair
Shideler et al, 2020 [50]	8	1	4	3	0	16	Fair

XSL•FO
RenderX

Experimental and feasibility studies evaluated using the modified Downs and Black checklist generally demonstrated fair-to-good methodological quality. Total scores ranged from 15 to 19 points (Table 2). Three studies achieved "good" (19 points) quality ratings [30,31,52], whereas the remaining studies were classified as "fair." Overall, studies consistently provided detailed descriptions of exoskeleton architecture, control strategies, sensor integration, and outcome measures. However, important methodological limitations were identified. External validity was often limited because most studies used small convenience samples, healthy participants, or single-subject feasibility designs, thereby reducing generalizability to broader clinical populations. Randomization and blinding were generally not feasible. In addition, none of the included experimental studies reported formal power calculations or sample-size justification, indicating a persistent risk of bias in the statistical power domain.

Two studies involving benchtop or no human-validation experiments were not assessed using formal clinical risk-of-bias tools because they did not evaluate patient outcomes or clinical interventions [37,38]. Instead, these studies were analyzed descriptively with emphasis on technical reporting quality, actuator validation, sensor integration, and controller performance. Both studies provided detailed engineering descriptions and reproducible experimental protocols; however, their findings remain limited to laboratory validation settings and cannot be generalized to clinical rehabilitation outcomes.

Part 2: Clinical Perspective

Study Populations and Clinical Settings

This section summarizes the clinical, biomechanical, and user-oriented outcomes of knee exoskeletons for gait improvement and rehabilitation. The presented studies span orthopedic, neurological, pediatric, and healthy populations and investigate both powered and passive knee exoskeleton systems across laboratory, clinical, and real-world settings.

Among the included studies, 22 studies were conducted in healthy adult participants, primarily to evaluate feasibility, safety, biomechanical effects, and control strategies. Clinical populations were investigated in 8 studies: neurological populations represented the largest clinical subgroup and included individuals with stroke, multiple sclerosis, spinal cord injury, and hemiplegic paralysis [29,42,44,46,57]. Orthopedic applications were limited to patients in the early postoperative phase after total knee arthroplasty [28], whereas pediatric populations included children and adolescents with cerebral palsy, crouch gait, or knee-extension deficits [50,51]. In addition, two studies were conducted without human participants [37,38].

Neurological Rehabilitation Outcomes

Knee exoskeletons were most frequently used in neurological gait rehabilitation. In stroke and multiple sclerosis populations, portable, powered knee exoskeletons have been shown to be feasible and safe, with participants able to complete overground walking tasks without major adverse events [29]. Improvements were reported in spatiotemporal gait parameters and walking performance, although responses varied across individuals and control strategies.

In a randomized crossover clinical trial involving individuals with spinal cord injury, a knee-powered exoskeleton demonstrated improved walking efficiency compared with conventional knee-ankle-foot orthoses, highlighting the potential of active knee assistance to reduce ambulatory effort [42]. Poststroke clinical trials have reported improvements in gait control and adaptability when assistance levels are adjusted to task demands or individual user characteristics, particularly during the stance and swing phases of gait [44,57]. Swing phase–specific assistance was associated with increased knee flexion and improved foot clearance during walking in patients with hemiplegic stroke [17].

Orthopedic and Postsurgical Applications

Orthopedic applications were primarily represented by a clinical trial involving patients undergoing early rehabilitation following total knee arthroplasty. In this study, the integration of a robotic knee device within 48 hours postsurgery was reported to be safe and well-tolerated [28]. Participants successfully completed assisted exercises, including passive mobilization, sit-to-stand transitions, and gait training, with no significant adverse events reported. Patient satisfaction was high, and clinical outcomes were comparable to those achieved with conventional rehabilitation protocols, suggesting that knee exoskeletons may serve as a feasible adjunct during early postoperative recovery. However, despite these promising preliminary findings, validated clinical evidence for orthopedic and postsurgical applications of knee exoskeletons remains extremely limited. Most of the currently available studies are feasibility-focused, involve small sample sizes, and lack long-term follow-up or controlled, comparative designs. In addition, there remains insufficient data on other orthopedic conditions, indicating a significant gap in the current literature and a need for larger clinical trials for specific conditions.

Pediatric Rehabilitation and Gait Correction

Pediatric studies primarily addressed pathological gait patterns, particularly crouch gait and impaired knee extension in children and adolescents with cerebral palsy. Powered knee exoskeletons were associated with clinically meaningful improvements in knee extension during stance and reductions in crouch gait severity [51]. In a case-based study, combining mechanical knee assistance with noninvasive neuromuscular electrical stimulation led to greater improvements in knee extension than exoskeleton assistance alone, highlighting the potential benefits of multimodal rehabilitation strategies in pediatric populations [50].

Functional Tasks and Outcome Domains

Across the included studies, level walking was the most assessed functional task, evaluated in both healthy and clinical populations [29,42]. Additional tasks included stair ascent [33], sit-to-stand transitions [28], squatting [49], stumble recovery [21], jumping [52], and community ambulation in outdoor environments [53].

Clinical outcome measures clustered into four main domains:

XSL•FO
RenderX

1. Gait performance and biomechanics, including walking speed, knee range of motion, joint moments, and stance- and swing-phase characteristics [29,31,51].
2. Neuromuscular outcomes, such as muscle activation amplitude, coordination, and interaction torque, reflecting reductions in muscular demand during assisted walking [36,41,50].
3. Physiological and effort-related outcomes, including metabolic cost and fatigue, with several studies reporting reduced energetic demand during exoskeleton-assisted gait [30,40,49].
4. User-centered outcomes, including comfort, safety, stability, alignment, and perceived trust, are particularly emphasized in studies addressing interface design and migration resistance [45,54].

Relationship Between Engineering Design and Clinical Outcomes

Clinical and biomechanical outcomes were frequently associated with specific engineering design characteristics across the included studies, as shown in Table 3. Motor-driven exoskeletons with gait-phase-based control were among the most commonly investigated systems in neurological rehabilitation and were associated with improvements in gait coordination, swing-phase assistance, and walking support in individuals with stroke and spinal cord injury [29,42,44,57]. Similarly, assist-as-needed and adaptive control strategies were developed to modulate assistance in response to user performance and task demands, thereby supporting individualized rehabilitation approaches [48,57].

Table . Relationship between engineering design features and reported clinical or biomechanical outcomes.

Engineering feature or technical configuration	Reported clinical or biomechanical outcome	Representative studies
Gait-phase-based motor control	Improved gait timing, coordination, and assistive synchronization	Kim et al, 2015 [35]; Wang et al, 2023 [55]
Assist-as-needed and adaptive control	Personalized rehabilitation and task-specific gait assistance	Zhang et al, 2023 [48], de Miguel Fernández et al, 2023 [57]
Passive elastic actuation	Reduced muscular effort and assistance during repetitive movement	Hidayah et al, 2021 [49], Ben-David et al, 2022 [52]
High-torque motor-driven systems	Improved support during rehabilitation and pathological gait assistance	Kittisares et al, 2020 [43], Rodríguez-Fernández et al, 2022 [42]
Hybrid rigid-soft structures	Reduced joint misalignment and improved comfort during walking	Wang et al, 2023 [56]
Self-aligning and polycentric knee mechanisms	Improved biomechanical alignment and preservation of physiological knee kinematics	Sarkisian et al, 2021 [54], Olinski et al, 2025 [47]
EMG[a]-driven intention recognition	Improved synchronization between user intention and robotic assistance	Villa-Parra et al, 2018 [36], Zhang et al, 2023 [48]
NMES/FES[b]-triggered robotic assistance	Increased voluntary muscle activation and enhanced motor relearning potential	Shideler et al, 2020 [50]
Assistance/resistance multimodal training	Improved peak knee flexion and gait endurance after stroke rehabilitation	de Miguel Fernández et al, 2023 [57]
IMU[c]-based outdoor gait monitoring	Improved evaluation of community ambulation and ecological walking behavior	Goffredo et al, 2022 [53]
Cable-driven lightweight exoskeleton systems	Improved comfort and reduced restriction of natural gait motion	Wang et al, 2023 [56]
Adaptive oscillator-based gait prediction	Improved real-time gait synchronization and assistive timing	Wang et al, 2023 [55]

[a]EMG: electromyography.

[b]NMES/FES: neuromuscular electrical stimulation/functional electrical stimulation.

[c]IMU: inertial measurement unit.

Passive and compliant exoskeleton systems have been commonly investigated for repetitive movement assistance and the reduction of muscular effort in healthy individuals [40,49,52]. However, these systems generally provided lower torque output and limited adaptability for pathological gait correction. In contrast, higher-torque motorized systems (10 - 40 Nm) were more frequently used for stair climbing, perturbation recovery, and poststroke rehabilitation tasks that require greater movement support and stabilization [21,33,43].

Several studies also emphasized the importance of wearable comfort, joint alignment, and interface compliance for user acceptance and stability during assisted gait [45,54]. Hybrid rigid-soft architectures, self-aligning mechanisms, and hinge-free exosuit designs were specifically developed to reduce joint misalignment and preserve more physiological gait kinematics while maintaining assistive functionality [31,40,54,56]. Overall, the reviewed studies demonstrated that rehabilitation outcomes were influenced not only by torque generation capacity but also

XSL•FO
RenderX

by sensing modalities, control strategies, biomechanical alignment, and human-robot interaction design.

Discussion

Bridging Engineering Innovation and Clinical Reality

Over the last 10 years, significant progress has been made in knee exoskeleton research, particularly in actuator design, sensor integration, and control system development [38,39,48,54,56]. From an engineering perspective, the domain has evolved significantly in complexity. Nevertheless, a closer look at the literature reveals that clinical validation has not kept up with this advancement. The risk-of-bias assessment indicates that current evidence supporting knee exoskeletons is methodologically heterogeneous. Although engineering reporting and technical characterization are generally robust, clinical evidence remains constrained by small sample sizes and the moderate risk of bias. Therefore, the strength and generalizability of the current evidence regarding long-term clinical efficacy should be interpreted with caution.

Many published studies continue to be conducted in laboratory settings, primarily involving healthy subjects rather than those with neurological or orthopedic conditions [29,32,45,48-50,55]. Although these investigations are essential for confirming feasibility and safety, they fail to fully address the intricacies of rehabilitation in practice.

Motor-driven exoskeletons remain the leading technology in this area due to their ability to accurately adjust assistive torque and their effective integration with control frameworks that use torque and impedance principles [31,34,42,54,56]. In structured walking activities, these systems can consistently provide assistance that varies with the gait cycle phase. Nevertheless, achieving greater torque capability often requires larger actuators and transmission mechanisms, thereby increasing distal mass and inertia [31,40,42,54]. These mechanical additions may influence natural gait mechanics and metabolic demand, yet these effects are not consistently quantified. In rehabilitation settings, the trade-off between torque capacity and wearability becomes particularly relevant, as long-term use depends on comfort and tolerance rather than peak output.

Passive and hybrid systems have emerged as attempts to address these limitations by emphasizing transparency and reduced mechanical burden [32,40,49,52]. Passive elastic devices have demonstrated measurable biomechanical effects during controlled tasks such as squatting and jumping [49,52]. However, their assistance profiles are inherently task-specific and less adaptable to the variability observed in pathological gait. Hybrid rigid-soft architectures seek to preserve torque capability while improving interface compliance and load distribution [31,54,56]. Although conceptually promising, comparative data demonstrating superior long-term outcomes are still limited. Taken together, the current evidence suggests that future progress may depend less on maximizing mechanical output and more on refining the balance between assistance magnitude, comfort, and adaptability to individual impairment.

The review demonstrated that rehabilitation outcomes were strongly influenced by engineering characteristics such as actuation type, control architecture, and interface compliance, highlighting the importance of biomechanically aligned and task-specific exoskeleton design.

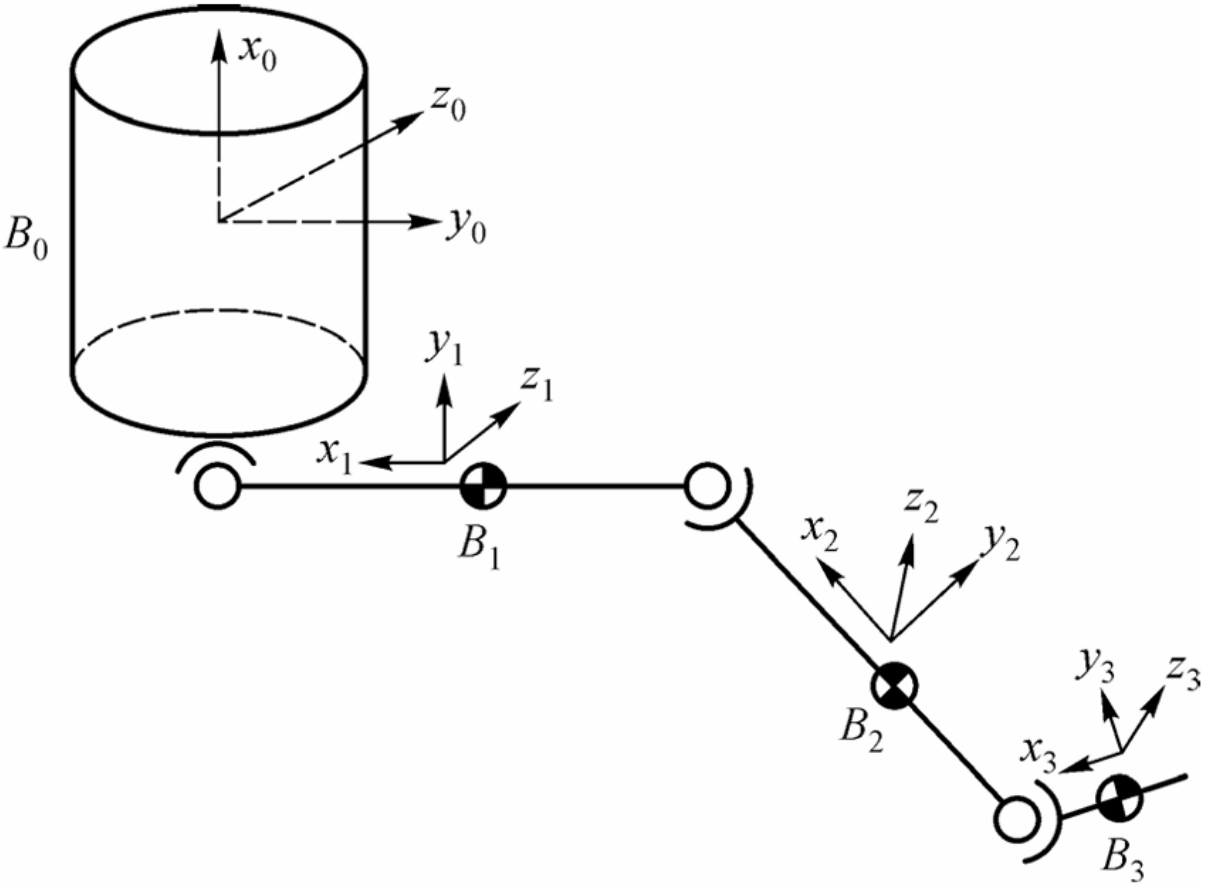
Knee Biomechanics

The observed results in clinical and biomechanical parameters can be directly interpreted using fundamental knee biomechanics. When walking on a flat surface, knee biomechanics are characterized by coordinated changes in joint kinematics and kinetics, particularly during the initial stance phase, in which knee flexion and extension moments regulate shock absorption and stability [66,67]. During daily activities, tibiofemoral forces can reach 2 - 4 times body weight during level walking and up to 5 times body weight during stair descent, highlighting the substantial mechanical demands placed on the joint [68,69]. These biomechanical requirements emphasize the need to develop an exoskeleton adaptable to various tasks.

Figure 8 shows a biomechanical model of the human lower limb, representing the femur and tibia as rigid bodies interconnected by ligamentous and contact elements, used to estimate knee joint kinematics and loading. This framework provides a physiological basis for designing knee exoskeletons that deliver assistive torques aligned with natural joint motion and internal force distributions.

XSL•FO
RenderX

Figure 8. Mechanical model of human lower limbs. The figure illustrates the mechanical representation of the human lower limb modeled as a kinematic chain of rigid segments (pelvis, thigh, leg, and foot), each assigned a local coordinate system to define joint orientations and interactions [68].



Clinical data show that altered knee mechanics lead to pain and functional disorders. Although the peak moment of external knee adduction shows a weak overall correlation with knee pain in osteoarthritis, dynamic anomalies, such as varus deformity, are associated with pain, underscoring the importance of frontal-plane stability [70]. After total knee replacement, knee extension moments decrease in patients, and joint load increases during downhill walking, indicating continued biomechanical vulnerability in the early stages of rehabilitation [71]. Similarly, patients after anterior cruciate ligament reconstruction exhibit altered kinematics during rehabilitation [72]. Overall, knee exoskeletons must integrate accurate anatomical sensing, adaptive control, and compliant actuation to modulate joint moments without disrupting natural movement, underscoring the critical need to align engineering design with clinical biomechanics.

Human-Robot Interface: A Determinant of Clinical Viability

Across studies, the quality of the human-robot interface consistently influences system performance, even when not explicitly analyzed. Mechanical misalignment between the anatomical knee axis and the exoskeleton joint can introduce parasitic torques and increase soft tissue stress [40,42,43,51,54]. Several groups have proposed self-aligning mechanisms, polycentric linkages, and improved suspension strategies to address this issue [33,42,54]. These refinements have been associated with improved user comfort and perceived naturalness of movement.

Despite these advances, many clinical investigations do not systematically report alignment errors, device migration, or long-duration wear tolerance [31,44,52,57]. This omission is clinically meaningful. Rehabilitation depends on repetition and sustained engagement; even minor discomfort or instability may alter neuromuscular activation patterns or reduce adherence over time. Emerging work examining embodiment and user perception suggests that when assistance feels mechanically coherent with voluntary movement, users integrate the device more effectively [44,57]. For this reason, interface optimization should be considered central to rehabilitation efficacy rather than an accessory design feature.

Control Strategies: Reliability vs Personalization

Control architecture remains a defining feature of knee exoskeleton systems. FSM-based controllers remain widely used in clinical contexts for their deterministic phase segmentation and predictable assistance timing [21,31,33,38,39,41,45-47,51]. In patient populations where safety is paramount, this reliability provides practical advantages.

XSL•FO
RenderX

More advanced strategies, including impedance control [38,54], model predictive control [39], human-in-the-loop optimization [30], EMG-driven assistance [36,50], and neuromusculoskeletal model-informed machine learning [48], have demonstrated promising personalization capabilities in controlled environments. For example, human-in-the-loop optimization has reduced metabolic costs in healthy participants [30], and EMG-based systems enable modulation of assistance based on muscle activation signals [36,50]. However, these approaches are rarely evaluated in larger or clinically heterogeneous cohorts [44,57].

Furthermore, few studies report performance under fatigue, uneven terrain, prolonged daily use, or sensor drift [30,39,48]. Without such validation, it is difficult to determine whether algorithmic complexity yields meaningful clinical benefit. At present, clinical implementation appears to favor robustness and predictability over adaptive sophistication.

Assistance Timing and Implications for Motor Learning

Across the reviewed literature, phase-specific assistance consistently contributes to functional improvement. Systems that target swing-phase knee flexion or stance-phase extension tend to improve toe clearance, knee extension, and gait symmetry while preserving voluntary muscle contribution [46,50,51,53]. This observation suggests that the timing of assistance may be more critical than continuous torque magnitude.

In contrast, trajectory-enforced systems may risk reducing active neuromuscular participation [42,44]. Since rehabilitation aims to facilitate motor relearning rather than to replicate passive movements, this distinction is important. Studies in pediatric cerebral palsy populations demonstrate that appropriately timed extension assistance can reduce crouch gait severity without suppressing voluntary engagement [50,51], supporting the concept of deficit-targeted assistance.

However, most outcome assessments are performed during active device use [31,44,57]. Postintervention follow-up is uncommon, making it difficult to determine whether improvements represent true motor adaptation or temporary mechanical augmentation. Longitudinal retention studies remain necessary to clarify this distinction.

Clinical Interpretation and Functional Relevance

Short-term improvements in knee kinematics, gait symmetry, and metabolic efficiency were consistently reported across studies. However, these findings were predominantly derived from studies conducted in healthy participants under controlled laboratory conditions [30,31,40,49].

Evidence from clinical populations, including individuals with stroke, spinal cord injury, and cerebral palsy, remains comparatively limited and heterogeneous [29,42,44,51,57]. While some studies reported improvements in gait parameters and walking efficiency, these findings are based on small sample sizes, short-duration interventions, and variable study designs. Therefore, the strength of evidence supporting clinical effectiveness is substantially lower than that observed in healthy populations.

Importantly, most reported improvements occurred during active exoskeleton use, with limited assessment of retention or transfer of gains after device removal [31,44,57]. This raises uncertainty about whether the observed improvement represents true motor rehabilitation or temporary performance augmentation driven by mechanical assistance.

Short-term improvements in knee extension, step symmetry, and metabolic efficiency have been reported in stroke, spinal cord injury, pediatric cerebral palsy, and postsurgical populations [28,31,36,38,42,44,46,50,51]. In spinal cord injury, powered assistance reduced ambulatory effort compared to orthotic support [42]. Early postoperative implementation after total knee arthroplasty has been reported to be feasible and safe [28].

Nevertheless, most studies involve small cohorts and brief intervention periods [29,44,57]. Outcome measures frequently emphasize kinematic and biomechanical metrics, whereas participation-level outcomes and quality-of-life indicators are less consistently reported [35,37,53]. While biomechanical improvements confirm device functionality, sustained functional independence and integration into daily activities remain the ultimate rehabilitation goals.

The available evidence therefore supports short-term performance enhancement, but durable rehabilitation benefit has not yet been conclusively demonstrated. Larger randomized trials with extended follow-up are required to determine whether knee exoskeleton–assisted training yields meaningful, lasting functional gains.

State of the Field and Translational Challenges

Collectively, the literature indicates that knee exoskeleton technology has achieved a high degree of mechanical and algorithmic sophistication [39,48,54,56]. However, clinical validation remains comparatively early in development [31,44,57]. The field appears technologically advanced yet still consolidating in terms of rehabilitation evidence.

Future progress will likely depend on closer integration between biomechanical analysis, interface refinement, control robustness, and longitudinal clinical evaluation. Demonstrating sustainable functional outcomes in real-world contexts will ultimately determine whether knee exoskeletons transition from promising prototypes to established rehabilitation tools.

Study Limitations

This review is limited by substantial heterogeneity in device design, control strategies, populations, and outcome measures, which precluded quantitative meta-analysis. Many included studies were exploratory or feasibility-focused and frequently involved healthy participants, limiting generalizability to clinical populations. Reporting of mechanical specifications, adverse events, and long-term follow-up was inconsistent, constraining cross-study comparison. Additionally, most studies assessed outcomes only during device use, providing limited insight into retention or neuroplastic effects.

XSL•FO
RenderX

Furthermore, a substantial proportion of included studies were conducted in healthy participants or involved early-stage feasibility or prototype validation, limiting the direct clinical applicability of the findings.

Future Research Directions

Future research should more explicitly align mechanical design and control strategies with specific rehabilitation objectives. Mechanical parameters such as torque capacity, range of motion, and assistance timing should be justified based on task demands and population needs rather than maximized by default. Adaptive and uncertainty-aware control approaches that show promise in healthy users should be systematically evaluated in clinical populations with attention to safety, usability, and interindividual variability.

Most importantly, future studies should move beyond short-term feasibility testing and controlled laboratory demonstrations. Well-designed, multicenter randomized controlled trials are needed, with larger, clinically representative samples, predefined primary outcomes, and standardized intervention protocols. Longer follow-up periods should be included to determine whether improvements persist after device removal and whether observed gains reflect true rehabilitation effects rather than temporary device-assisted performance.

Future trials should include patient-important outcomes, not only biomechanical parameters. These should include functional independence, walking, endurance, fall risk, participation in daily activities, quality of life, and user satisfaction. Greater standardization of outcome measures that integrate biomechanical, functional, and participation-level domains would strengthen cross-study comparisons and improve clinical interpretation. This would help determine whether knee exoskeletons can provide meaningful rehabilitation benefits beyond short-term improvements observed during assisted use.

Conclusions

In summary, knee exoskeleton devices demonstrated consistent short-term improvements in biomechanical and performance-related outcomes, particularly under controlled laboratory conditions [30,31,40,49]. However, the current evidence is largely derived from small-scale studies, frequently involving healthy participants or short-duration feasibility trials [29,44,57].

Evidence from clinical populations remains limited and heterogeneous, with insufficient high-quality data to support routine clinical implementation in physical therapy and rehabilitation practice. While preliminary findings suggest potential benefits for gait rehabilitation, these results should be interpreted cautiously due to methodological limitations, including small sample sizes, lack of long-term follow-up, and variability in outcome measures [29,42,44,57].

Future research should prioritize well-designed, multicenter randomized controlled trials with clearly defined primary outcomes, longer intervention and follow-up periods, and inclusion of patient-centered outcomes such as functional independence, fall risk, and participation. These factors are critical for establishing the true clinical value and translational potential of knee exoskeleton-assisted rehabilitation.

Funding

This research was funded by the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant BR24992820).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 17 KB - rehab_v13i1e94916_app1.docx]

Multimedia Appendix 2
Filtering prompt and classification code.
[DOCX File, 17 KB - rehab_v13i1e94916_app2.docx]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 165 KB - rehab_v13i1e94916_app3.pdf]

References

1. Mahlknecht P, Kiechl S, Bloem BR, et al. Prevalence and burden of gait disorders in elderly men and women aged 60-97 years: a population-based study. PLoS One 2013 Jul 24;8(7):e69627. [doi: 10.1371/journal.pone.0069627] [Medline: 23894511]
2. Pirker W, Katzenschlager R. Gait disorders in adults and the elderly: a clinical guide. Wien Klin Wochenschr 2017 Feb;129(3-4):81-95. [doi: 10.1007/s00508-016-1096-4] [Medline: 27770207]

XSL•FO
RenderX

3. Ali F, Padilla H, Blazek AM, Barnard L, Kaufman KR. Gait analysis in neurologic disorders: methodology, applications, and clinical considerations. Neurology 2025 Oct 21;105(8):e214154. [doi: 10.1212/WNL.0000000000214154] [Medline: 40971743]
4. Li E, Tan J, Xu K, Pan Y, Xu P. Global burden and socioeconomic impact of knee osteoarthritis: a comprehensive analysis. Front Med (Lausanne) 2024 May 16;11:1323091. [doi: 10.3389/fmed.2024.1323091] [Medline: 38818397]
5. GBD 2021 Osteoarthritis Collaborators. Global, regional, and national burden of osteoarthritis, 1990-2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol 2023 Aug 21;5(9):e508-e522. [doi: 10.1016/S2665-9913(23)00163-7] [Medline: 37675071]
6. Ouyang Y, Dai M. Global, regional, and national burden of knee osteoarthritis: findings from the Global Burden of Disease study 2021 and projections to 2045. J Orthop Surg Res 2025 Aug 13;20(1):766. [doi: 10.1186/s13018-025-06140-0] [Medline: 40804406]
7. Berghmans DDP, Lenssen AF, Emans PJ, de Bie RA. Functions, disabilities and perceived health in the first year after total knee arthroplasty; a prospective cohort study. BMC Musculoskelet Disord 2018 Jul 25;19(1):250. [doi: 10.1186/s12891-018-2159-7] [Medline: 30045710]
8. Gao J, Xing D, Dong S, Lin J. The primary total knee arthroplasty: a global analysis. J Orthop Surg Res 2020 May 26;15(1):190. [doi: 10.1186/s13018-020-01707-5] [Medline: 32456654]
9. Salmon JH, Rat AC, Sellam J, et al. Economic impact of lower-limb osteoarthritis worldwide: a systematic review of cost-of-illness studies. Osteoarthr Cartil 2016 Sep;24(9):1500-1508. [doi: 10.1016/j.joca.2016.03.012] [Medline: 27034093]
10. Platzer H, Gantz S, Färber B, Moradi B. Rising direct medical costs of osteoarthritis in Germany from 2002 to 2020: trends, demographic drivers, and public health considerations. Front Public Health 2025 Nov 12;13:1627510. [doi: 10.3389/fpubh.2025.1627510] [Medline: 41312231]
11. Flandry F, Hommel G. Normal anatomy and biomechanics of the knee. Sports Med Arthrosc Rev 2011 Jun;19(2):82-92. [doi: 10.1097/JSA.0b013e318210c0aa] [Medline: 21540705]
12. Cortés-Pérez I, Desdentado-Guillem JM, Camacho-Delgado MS, Del Rocío Ibancos-Losada M, Obrero-Gaitán E, Lomas-Vega R. Virtual reality-based therapy after anterior cruciate ligament injury effectively reduces pain and improves knee function, movement patterns, and dynamic balance: a systematic review and meta-analysis. Knee Surg Sports Traumatol Arthrosc 2025 May;33(5):1736-1753. [doi: 10.1002/ksa.12477] [Medline: 39302094]
13. Shaheen AAM, Algabbani MF, Alshahrani MM, et al. Efficacy of virtual reality-based rehabilitation following total knee replacement: a systematic review and meta-analysis of randomized control trials. Bull Fac Phys Ther 2025 Mar 5;30(1). [doi: 10.1186/s43161-024-00246-y]
14. Liu X, Yang G, Xie W, et al. Efficacy of telerehabilitation for total knee arthroplasty: a meta-analysis based on randomized controlled trials combined with a bibliometric study. J Orthop Surg Res 2024 Dec 26;19(1):874. [doi: 10.1186/s13018-024-05381-9] [Medline: 39726029]
15. Correia FD, Nogueira A, Magalhães I, et al. Home-based rehabilitation with a novel digital biofeedback system versus conventional in-person rehabilitation after total knee replacement: a feasibility study. Sci Rep 2018 Jul 26;8(1):11299. [doi: 10.1038/s41598-018-29668-0] [Medline: 30050087]
16. Morris L, Diteesawat RS, Rahman N, Turton A, Cramp M, Rossiter J. The state-of-the-art of soft robotics to assist mobility: a review of physiotherapist and patient identified limitations of current lower-limb exoskeletons and the potential soft-robotic solutions. J Neuroeng Rehabil 2023 Jan 30;20(1):18. [doi: 10.1186/s12984-022-01122-3] [Medline: 36717869]
17. Zhang Y, Zhao W, Wan C, et al. Exoskeleton rehabilitation robot training for balance and lower limb function in sub-acute stroke patients: a pilot, randomized controlled trial. J Neuroeng Rehabil 2024 Jun 8;21(1):98. [doi: 10.1186/s12984-024-01391-0] [Medline: 38851703]
18. Ma L, Marshall PJ, Wright WG. The impact of external and internal focus of attention on visual dependence and EEG alpha oscillations during postural control. J Neuroeng Rehabil 2022 Jul 26;19(1):81. [doi: 10.1186/s12984-022-01059-7] [Medline: 35883085]
19. Travasso C. India meets target to reduce tuberculosis burden. Br Med J 2015 Nov 17;351:h6183. [doi: 10.1136/bmj.h6183]
20. Fan T, Zheng P, Zhang X, et al. Effects of exoskeleton rehabilitation robot training on neuroplasticity and lower limb motor function in patients with stroke. BMC Neurol 2025 May 3;25(1):193. [doi: 10.1186/s12883-025-04203-7] [Medline: 40319228]
21. Eveld ME, King ST, Zelik KE, Goldfarb M. Efficacy of stumble recovery assistance in a knee exoskeleton for individuals with simulated mobility impairment: a pilot study. Wearable Technol 2023 Sep 6;4:e22. [doi: 10.1017/wtc.2023.17] [Medline: 38510587]
22. Maeda T, Sasaki E, Kasai T, et al. Therapeutic effect of knee extension exercise with single-joint hybrid assistive limb following total knee arthroplasty: a prospective, randomized controlled trial. Sci Rep 2024 Feb 16;14(1):3889. [doi: 10.1038/s41598-024-53891-7] [Medline: 38365932]
23. Soma Y, Mutsuzaki H, Yoshioka T, et al. Single-joint hybrid assistive limb in knee rehabilitation after ACL reconstruction: an open-label feasibility and safety trial. Prog Rehabil Med 2022 Jul 21;7:20220036. [doi: 10.2490/prm.20220036] [Medline: 35935452]

XSL•FO
RenderX

24. Lee D, Lee S, Young AJ. AI-driven universal lower-limb exoskeleton system for community ambulation. Sci Adv 2024 Dec 20;10(51):eadq0288. [doi: 10.1126/sciadv.adq0288] [Medline: 39693442]
25. Xu Q, Li J, Jiang N, et al. A study on the design of knee exoskeleton rehabilitation based on the RFPBS model. Biomimetics (Basel) 2024 Jul 5;9(7):410. [doi: 10.3390/biomimetics9070410] [Medline: 39056851]
26. Wu K, Pan HH, Lin CH. Robotic exoskeletons and total knee arthroplasty: the future of knee rehabilitation and replacement – a meta-analysis. Medicine (Baltimore) 2024 Apr 26;103(17):e37876. [doi: 10.1097/MD.0000000000037876] [Medline: 38669435]
27. Pugliese BL, Angelucci A, Parisi F, et al. Development of a wearable sleeve-based system combining polymer optical fiber sensors and an LSTM network for estimating knee kinematics. IEEE Trans Neural Syst Rehabil Eng 2025;33:728-738. [doi: 10.1109/TNSRE.2025.3540708] [Medline: 40031675]
28. Barquín‐Santos E, Cumplido‐Trasmonte C, Gor‐García‐Fogeda MD, et al. Early implementation of MAK robotic device in total knee arthroplasty rehabilitation: a proof‐of‐concept study. Physiotherapy Res Intl 2024 Oct;29(4):e2134. [doi: 10.1002/pri.2134] [Medline: 39295196]
29. Puyuelo-Quintana G, Cano-de-la-Cuerda R, Plaza-Flores A, et al. A new lower limb portable exoskeleton for gait assistance in neurological patients: a proof of concept study. J Neuroeng Rehabil 2020 May 6;17(1):60. [doi: 10.1186/s12984-020-00690-6] [Medline: 32375815]
30. Monteiro S, Figueiredo J, Fonseca P, Vilas-Boas JP, Santos CP. Human-in-the-loop optimization of knee exoskeleton assistance for minimizing user's metabolic and muscular effort. Sensors (Basel) 2024 May 22;24(11):3305. [doi: 10.3390/s24113305] [Medline: 38894101]
31. Dai X, Zhou Z, Wang Z, et al. Reducing knee joint loads during stance phase with a rigid-soft hybrid exoskeleton. IEEE Trans Neural Syst Rehabil Eng 2024;32:4164-4173. [doi: 10.1109/TNSRE.2024.3498044] [Medline: 40030323]
32. Zhou Y, Wang X, Tang X, Ji X, Gao Z. Design and analysis of a passive knee assisted exoskeleton. Adv Mech Eng 2023 Nov;15(11). [doi: 10.1177/16878132231202578]
33. Kittisares S, Ide T, Nabae H, Suzumori K. Ergonomic dual four-bar linkage knee exoskeleton for stair ascent assistance. Front Robot AI 2023 Dec 6;10:1285520. [doi: 10.3389/frobt.2023.1285520] [Medline: 38124902]
34. Wang Y, Zhang W, Shi D, Geng Y. Design and control of an adaptive knee joint exoskeleton mechanism with buffering function. Sensors 2021 Dec 15;21(24):8390. [doi: 10.3390/s21248390] [Medline: 34960484]
35. Kim JH, Shim M, Ahn DH, et al. Design of a knee exoskeleton using foot pressure and knee torque sensors. Int J Adv Robot Syst 2015 Aug 1;12(8):112. [doi: 10.5772/60782]
36. Villa-Parra AC, Delisle-Rodriguez D, Botelho T, et al. Control of a robotic knee exoskeleton for assistance and rehabilitation based on motion intention from sEMG. Res Biomed Eng 2018;34(3):198-210. [doi: 10.1590/2446-4740.07417]
37. Chen J, Hochstein J, Kim C, Damiano D, Bulea T. Design advancements toward a wearable pediatric robotic knee exoskeleton for overground gait rehabilitation. In: 2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob): IEEE; 2018. [doi: 10.1109/BIOROB.2018.8487195]
38. Wang J, Li X, Huang TH, et al. Comfort-centered design of a lightweight and backdrivable knee exoskeleton. IEEE Robot Autom Lett 2018;3(4):4265-4272. [doi: 10.1109/LRA.2018.2864352]
39. Jammeli I, Chemori A, Moon H, Elloumi S, Mohammed S. An assistive explicit model predictive control framework for a knee rehabilitation exoskeleton. IEEE/ASME Trans Mechatron 2022;27(5):3636-3647. [doi: 10.1109/TMECH.2021.3126674]
40. Park EJ, Akbas T, Eckert-Erdheim A, et al. A hinge-free, non-restrictive, lightweight tethered exosuit for knee extension assistance during walking. IEEE Trans Med Robot Bionics 2020;2(2):165-175. [doi: 10.1109/tmrb.2020.2989321] [Medline: 33748694]
41. Mosconi D, Moreno Y, Siqueira A. Exploring human-exoskeleton interaction dynamics: an in-depth analysis of knee flexion-extension performance across varied robot assistance-resistance configurations. Sensors (Basel) 2024 Apr 21;24(8):2645. [doi: 10.3390/s24082645] [Medline: 38676262]
42. Rodríguez-Fernández A, Lobo-Prat J, Tarragó R, et al. Comparing walking with knee-ankle-foot orthoses and a knee-powered exoskeleton after spinal cord injury: a randomized, crossover clinical trial. Sci Rep 2022 Nov 9;12(1):19150. [doi: 10.1038/s41598-022-23556-4] [Medline: 36351989]
43. Kittisares S, Nabae H, Endo G, Suzumori K, Sakurai R. Design of knee support device based on four-bar linkage and hydraulic artificial muscle. Robomech J 2020 Mar 27;7:16. [doi: 10.1186/s40648-020-00165-2]
44. Lora-Millan JS, Sanchez-Cuesta FJ, Romero JP, Moreno JC, Rocon E. Robotic exoskeleton embodiment in post-stroke hemiparetic patients: an experimental study about the integration of the assistance provided by the REFLEX knee exoskeleton. Sci Rep 2023 Dec 21;13(1):22908. [doi: 10.1038/s41598-023-50387-8] [Medline: 38129592]
45. Velasco-Guillen RJ, Bliek A, Font-Llagunes JM, Vanderborght B, Beckerle P. Compensating elastic faults in a torque-assisted knee exoskeleton: functional evaluation and user perception study. J Neuroeng Rehabil 2024 Dec 28;21(1):230. [doi: 10.1186/s12984-024-01531-6] [Medline: 39732683]
46. Zhang Z, Cai X, Zhang M, Chen W, Chen Y, Wang P. Real-time gait intention recognition for active control of unilateral knee exoskeleton. Appl Bionics Biomech 2024;2024(1):9426782. [doi: 10.1155/2024/9426782] [Medline: 39568574]

XSL•FO
RenderX

47. Olinski M. Design and experimental characterization of developed human knee joint exoskeleton prototypes. Machines 2025 Jan 18;13(1):70. [doi: 10.3390/machines13010070]
48. Zhang L, Zhang X, Zhu X, Wang R, Gutierrez-Farewik EM. Neuromusculoskeletal model-informed machine learning-based control of a knee exoskeleton with uncertainties quantification. Front Neurosci 2023;17:1254088. [doi: 10.3389/fnins.2023.1254088]
49. Hidayah R, Sui D, Wade KA, Chang BC, Agrawal S. Passive knee exoskeletons in functional tasks: biomechanical effects of a SpringExo coil-spring on squats. Wearable Technol 2021;2:e7. [doi: 10.1017/wtc.2021.6] [Medline: 38486637]
50. Shideler BL, Bulea TC, Chen J, Stanley CJ, Gravunder AJ, Damiano DL. Toward a hybrid exoskeleton for crouch gait in children with cerebral palsy: neuromuscular electrical stimulation for improved knee extension. J Neuroeng Rehabil 2020 Sep 3;17(1):121. [doi: 10.1186/s12984-020-00738-7] [Medline: 32883297]
51. Lerner ZF, Damiano DL, Bulea TC. A lower-extremity exoskeleton improves knee extension in children with crouch gait from cerebral palsy. Sci Transl Med 2017 Aug 23;9(404):eaam9145. [doi: 10.1126/scitranslmed.aam9145] [Medline: 28835518]
52. Ben-David C, Ostraich B, Riemer R. Passive knee exoskeleton increases vertical jump height. IEEE Trans Neural Syst Rehabil Eng 2022;30:1796-1805. [doi: 10.1109/TNSRE.2022.3187056] [Medline: 35776830]
53. Goffredo M, Romano P, Infarinato F, et al. Kinematic analysis of exoskeleton-assisted community ambulation: an observational study in outdoor real-life scenarios. Sensors (Basel) 2022 Jun 16;22(12):4533. [doi: 10.3390/s22124533] [Medline: 35746315]
54. Sarkisian SV, Ishmael MK, Lenzi T. Self-aligning mechanism improves comfort and performance with a powered knee exoskeleton. IEEE Trans Neural Syst Rehabil Eng 2021;29:629-640. [doi: 10.1109/TNSRE.2021.3064463] [Medline: 33684041]
55. Wang Z, Yang C, Zhang S, et al. Knee flexion-assisted method for human-exoskeleton system. IEEE Trans Neural Syst Rehabil Eng 2023;31:2800-2808. [doi: 10.1109/TNSRE.2023.3287867] [Medline: 37339043]
56. Wang Z, Zhou Z, Ruan L, Duan X, Wang Q. Mechatronic design and control of a rigid-soft hybrid knee exoskeleton for gait intervention. IEEE/ASME Trans Mechatron 2023;28(5):2553-2564. [doi: 10.1109/TMECH.2023.3245810]
57. de Miguel Fernandez J, Rey-Prieto M, Rio MSD, et al. Adapted assistance and resistance training with a knee exoskeleton after stroke. IEEE Trans Neural Syst Rehabil Eng 2023;31:3265-3274. [doi: 10.1109/TNSRE.2023.3303777] [Medline: 37556332]
58. CJMCU-111 rotary encoder switch module. ElectroPeak Inc. URL: https://electropeak.com/cjmcu-111-rotary-encoder-switch-module [accessed 2026-07-04]
59. IMU sensor working and its applications. ElProCus – Electronics Projects Focus. 2018. URL: https://www.elprocus.com/imu-sensor-working-applications [accessed 2026-07-01]
60. MyoWare muscle sensor kit. SparkFun Electronics. URL: https://learn.sparkfun.com/tutorials/myoware-muscle-sensor-kit/all [accessed 2026-07-01]
61. Li Y, Yu C, Lu P. An optical fiber sensor for axial strain, curvature, and temperature measurement based on single-core six-hole optical fiber. Sensors 2022;22(4):1666. [doi: 10.3390/s22041666]
62. SankeyMATIC: Build a Sankey Diagram. URL: https://sankeymatic.com/build [accessed 2026-07-01]
63. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928. [doi: 10.1136/bmj.d5928] [Medline: 22008217]
64. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
65. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998 Jun;52(6):377-384. [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
66. Hassebrock JD, Gulbrandsen MT, Asprey WL, Makovicka JL, Chhabra A. Knee ligament anatomy and biomechanics. Sports Med Arthrosc Rev 2020 Sep;28(3):80-86. [doi: 10.1097/JSA.0000000000000279] [Medline: 32740458]
67. Chhabra A, Elliott CC, Miller MD. Normal anatomy and biomechanics of the knee. Sports Med Arthrosc 2001 Jul;9(3):166-177. [doi: 10.1097/00132585-200107000-00002]
68. Madeti BK, Chalamalasetti SR, Bolla Pragada S. Biomechanics of knee joint — a review. Front Mech Eng 2015 Jun;10(2):176-186. [doi: 10.1007/s11465-014-0306-x]
69. Hyodo K, Masuda T, Aizawa J, Jinno T, Morita S. Hip, knee, and ankle kinematics during activities of daily living: a cross-sectional study. Braz J Phys Ther 2017;21(3):159-166. [doi: 10.1016/j.bjpt.2017.03.012] [Medline: 28473283]
70. Hutchison L, Grayson J, Hiller C, D'Souza N, Kobayashi S, Simic M. Relationship between knee biomechanics and pain in people with knee osteoarthritis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2023 Jun;75(6):1351-1361. [doi: 10.1002/acr.25001] [Medline: 35997473]
71. Wen C, Cates HE, Weinhandl JT, Crouter SE, Zhang S. Knee biomechanics of patients with total knee replacement during downhill walking on different slopes. J Sport Health Sci 2022 Jan;11(1):50-57. [doi: 10.1016/j.jshs.2021.01.009] [Medline: 33540108]

XSL•FO
RenderX

72. Murrell-Smith ZN, Alabdullah MM, Zhang F, Jennings LM, Astill SL, Liu A. Knee biomechanics during rehabilitation exercise in individuals with and without anterior cruciate ligament reconstruction: a systematic review. Clin Biomech (Bristol) 2025 Jun;126:106559. [doi: 10.1016/j.clinbiomech.2025.106559] [Medline: 40393328]

Abbreviations:

EMG: electromyography
FSM: finite-state machine
IMU: inertial measurement unit
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RoB-2: Cochrane Risk of Bias 2
ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions
TKA: total knee arthroplasty

Edited by A Scano; submitted 10.Mar.2026; peer-reviewed by A Torad, Y Jin; revised version received 11.May.2026; accepted 19.May.2026; published 04.Aug.2026.

Please cite as:
Karibzhanova D, Niyetkaliyev A, Khawaja AR, Zhaisanbek A, Zeeshan , Utepbergen A, Jamwal PK, Ozhikenov K, Ozhikenova A, Alimbayev C, Alimbayeva Z, Bigaliyeva Z, Bazarbay L, Didar S
State-of-the-Art Knee Exoskeleton for Gait Assistance and Rehabilitation: Systematic Review
JMIR Rehabil Assist Technol 2026;13:e94916
URL: https://rehab.jmir.org/2026/1/e94916
doi:10.2196/94916

© Dilnoza Karibzhanova, Aibek Niyetkaliyev, Amna R Khawaja, Aray Zhaisanbek, Zeeshan, Ardana Utepbergen, Prashant K Jamwal, Kassymbek Ozhikenov, Aiman Ozhikenova, Chingiz Alimbayev, Zhadyra Alimbayeva, Zhanar Bigaliyeva, Lashin Bazarbay, Serik Didar. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 4.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Affective Computing in Serious Games for Physical Rehabilitation: Scoping Review

María del Pilar Beristain-Colorado[1*], PhD; Patricia Batres-Mendoza[2*], PhD; Erick Israel Guerra-Hernández[2*], PhD; José Luis Cano-Pérez[3*], PhD; Christian Perezcampos-Mayoral[3*], PhD; Marciano Vargas-Treviño[4*], PhD; Jaime Gutiérrez-Gutiérrez[3*], PhD; Jorge Fernando Ambros-Antemate[1*], PhD

[1]Laboratorio de Computación Centrada en el Humano aplicada a la salud y a la educación, Facultad de Sistemas Biológicos e Innovación Tecnológica, Universidad Autónoma Benito Juárez de Oaxaca, Av. Universidad S/N. Ex-Hacienda 5 Señores, Oaxaca de Juárez, Oaxaca, Mexico

[2]Laboratorio de Inteligencia Artificial, Robótica y Control, Facultad de Sistemas Biológicos e Innovación Tecnológica, Universidad Autónoma Benito Juárez de Oaxaca, Oaxaca de Juárez, Oaxaca, Mexico

[3]Laboratorio de Fibras Ópticas y Sensores, Facultad de Sistemas Biológicos e Innovación Tecnológica, Universidad Autónoma Benito Juárez de Oaxaca, Oaxaca de Juárez, Oaxaca, Mexico

[4]Facultad de Sistemas Biológicos e Innovación Tecnológica, Universidad Autónoma Benito Juárez de Oaxaca, Oaxaca de Juárez, Oaxaca, Mexico

[*]all authors contributed equally

Corresponding Author:
Jorge Fernando Ambros-Antemate, PhD
Laboratorio de Computación Centrada en el Humano aplicada a la salud y a la educación, Facultad de Sistemas Biológicos e Innovación Tecnológica, Universidad Autónoma Benito Juárez de Oaxaca, Av. Universidad S/N. Ex-Hacienda 5 Señores, Oaxaca de Juárez, Oaxaca, Mexico

Abstract

Background: Serious games have become an alternative support for traditional physical therapy. However, many of these games do not address the emotional needs of patients. People with disabilities often experience emotions such as sadness, frustration, and even anger, which can create a barrier to their rehabilitation treatment.

Objective: This review presents a comprehensive overview of technologies, techniques, and methods in affective computing as applied to serious games for physical rehabilitation and identifies key gaps to guide future research in this field.

Methods: A scoping review was conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, using the databases PubMed, ScienceDirect, IEEE Xplore, ACM Digital Library, PEDro, Springer, and Google Scholar.

Results: The initial search yielded 5293 records, of which 9 papers met the inclusion criteria. Data were systematically extracted from these papers based on predefined research questions. Notably, engagement, tiredness, and pain were the most identified emotions, reported in 4 of 9 (50%) studies. Only 3 studies applied theoretical frameworks for emotion classification. Facial expression analysis and gesture recognition were the most frequently used affective computing techniques, yet only 2 studies implemented adaptive gameplay based on affective feedback.

Conclusions: The integration of affective computing into serious games represents a promising approach for detecting affective states in patients undergoing rehabilitation. However, the limited number of primary studies, methodological limitations, and potential selection and reference-standard biases limit the reliability and generalizability of the current findings. Future research should prioritize rigorous multicenter designs, standardized evaluation protocols, and multidisciplinary collaboration. Developing these areas is essential to establishing clinical effectiveness.

(JMIR Rehabil Assist Technol 2026;13:e81344) doi:10.2196/81344

KEYWORDS

affective computing; serious game; physical rehabilitation; rehabilitation technology; emotions

Introduction

Overview of Serious Games for Physical Rehabilitation

Physical rehabilitation is a therapeutic approach aimed at improving the quality of life for individuals who have lost mobility in one or more body parts. Rehabilitation typically involves a series of repetitive exercises using the affected limb to gradually regain mobility. However, the repetitive nature of these exercises often becomes monotonous, leading to patient boredom and, in many cases, therapy discontinuation [1].

To address these challenges, alternative interventions such as serious games have been developed. Serious games are video

XSL•FO
RenderX

games designed to be both entertaining and educational [2]. They have great potential for training, motivation, and entertainment. One of their key features is the ability to simulate reality in a virtual environment, enhancing user immersion and concentration. These types of games have been used in various fields, including health, education, politics, and engineering. Specifically, serious games designed for physical rehabilitation aim to provide a more motivating environment, enhance treatment adherence, and make repetitive exercises more engaging. They provide a different context within therapeutic interventions to help patients reacquire, recover, and maintain physical capabilities, thereby improving quality of life [3]. Applications include motor recovery following stroke, gait rehabilitation, Parkinson disease, multiple sclerosis, orthopedic recovery, and physical stimulation for individuals with cerebral palsy [4,5]. These interventions have gained considerable attention among researchers and health care professionals due to their potential to address complex rehabilitation needs through innovative technologies [6].

A wide range of serious games has been developed for physical rehabilitation [7,8], which have shown positive results compared to conventional rehabilitation therapies. For instance, these games enhance the quality, enjoyment, and effectiveness of rehabilitation by providing an engaging environment and varying levels of difficulty based on injury severity. They can also overcome logistical limitations by enabling patients to engage in therapy remotely, attending clinical sessions only when necessary [5,9]. However, literature reviews show that most existing platforms focus on motor metrics, such as range of motion and kinematics, and often overlook the patients' affective state, including frustration from pain and feelings of sadness, depression, or anxiety [9,10]. Consequently, there are significant gaps in achieving an optimal balance between game challenge and emotional engagement.

One approach to improving interaction between patients and serious games is affective computing, a type of human-computer interaction that enables systems to recognize and adapt to users' affective states [11]. Evidence suggests that dynamic adaptation based on affective states in serious games can enhance rehabilitation outcomes by improving motivation, performance, and treatment adherence [12]. Additionally, affective computing offers real-time diagnostic capabilities that allow for personalized adjustments in gameplay in response to detected emotional cues [13].

Theoretical Background

Affective Computing

Picard [11] introduced the term "affective computing," defining it as "computing that relates to, arises from, or deliberately influences emotion." This definition highlights 3 key elements: (1) recognition, for detecting and classifying affective states; (2) expression, for displaying affective states through interaction media such as avatars; and (3) influencing affect, by using emotional models to guide behavior and decision-making.

Affective computing integrates input recognition to create an affective loop, a dynamic process in which the user's emotional states and the system's responses continuously interact through perception, interpretation, and emotional feedback [14]. Ideally, the system should not only detect or model emotions but also actively influence the user's affective experience, forming a closed loop in which the user's emotional state shapes the system's behavior and, in turn, the user's emotional experience.

Affective computing enables a more interactive and personalized experience by considering users' affective states. Several studies have focused on patient care. For example, Maj et al [15] developed a system that enhances care for individuals with chronic conditions or those in rehabilitation by continuously monitoring their affective state. This system also provides real-time affective feedback and virtual companionship through a personal assistant module, offering emotional support and helping to reduce loneliness and isolation. The benefits of affective computing have been demonstrated in digital mental health interventions [16,17], adaptive systems that adjust difficulty based on users' cognitive or emotional states [18], human-robot interactions with real-time emotion recognition and task adaptation [19], and adaptive systems with real-time feedback [20]. When combined with serious games, affective computing allows adaptive systems to modify game parameters in response to users' emotions, such as reducing difficulty during stress or sadness and increasing it when users are motivated or engaged.

Emotion Recognition Techniques

According to Hurst et al [21], affective recognition techniques can be categorized into invisible and visible inputs. Invisible inputs detect physiological signals imperceptible to the human eye, such as electrodermal activity or heart rate variability. Conversely, visible inputs include observable behaviors such as facial expressions and body movements.

Techniques based on invisible inputs generally use sensors that are placed on a part of the body. However, these techniques may present some challenges, such as mild discomfort for the individual when sensors are attached to the body. The most commonly used techniques are based on electrocardiogram and electroencephalogram (EEG). The electrocardiogram is an electrical tracing of the heart and is recorded noninvasively by electrodes placed on the body. EEG is a test that measures the brain's electrical activity by placing electrodes, which are small metal discs, on the scalp.

Techniques based on visible inputs do not require the user to be in direct contact with a sensor. Examples include facial expression analysis, user behavior monitoring, and gesture recognition. Facial expression techniques gather information about affective states from facial cues, even at rest. As a result, many studies have explored the link between facial movements and affective states [22]. User behavior techniques infer emotional states based on interaction with the system, including task performance, time spent on tasks, number of clicks, and mouse movement [23]. Gesture recognition techniques infer emotions from motor responses such as body movement, hand movement, or finger pressure [24].

Emotions are difficult to identify. According to Gonçalves et al [25], emotions are triggered by interrelated changes in 5 components: cognitive appraisals, behavioral tendencies, motor

XSL•FO
RenderX

expressions, physiological reactions, and subjective feelings. To detect emotions efficiently, it is necessary to use a combination of tests that include as many components as possible. Therefore, no single emotion recognition technique is sufficient on its own.

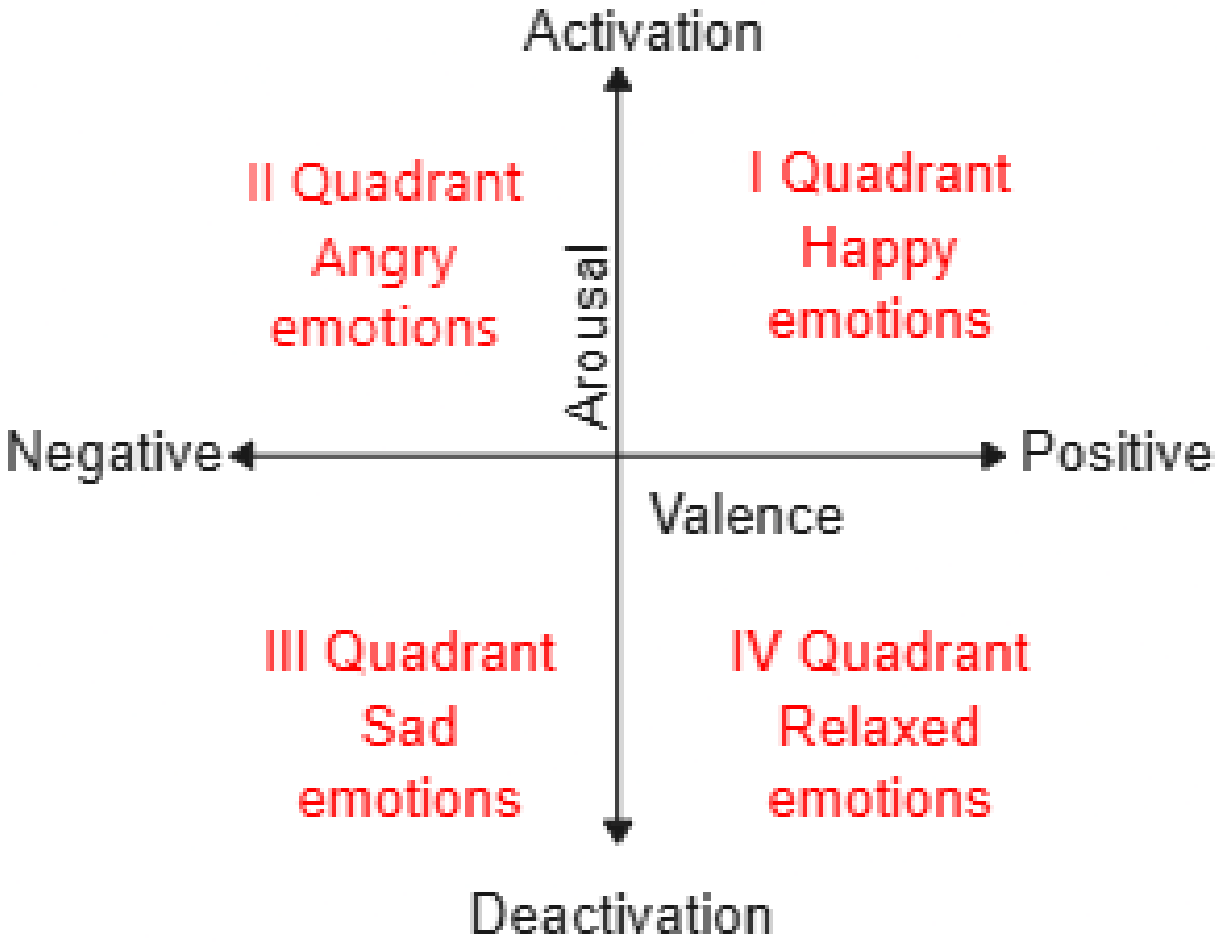
Models for Classifying Affective States

The classification of affective states varies across studies; however, 2 principal models are predominantly used in affective computing: the discrete emotion model and the dimensional emotion model [26].

One discrete model is that proposed by Ekman [27], which conceptualizes emotions as discrete, measurable, and linked to physiological responses. Ekman identified 6 basic emotions considered to be universal: happiness, sadness, anger, fear, surprise, and disgust. In contrast, Plutchik [28] proposed a model comprising 8 basic emotions: joy, fear, surprise, trust, sadness, anger, disgust, and anticipation, organized in a wheel-like structure. This model suggests that more intense emotions are located at the center of the wheel, while less intense emotions are found at the periphery.

Dimensional models were introduced to overcome the limitations of discrete models by acknowledging the complexity and fluidity of emotional experiences [29-32]. For instance, Russell [30] developed the circumplex model based on valence and arousal to represent complex emotions, in which emotions are classified within quadrants. The model proposes a 2-axis plane, where the x-axis represents valence, indicating the degree of pleasure or displeasure (positive or negative), and the y-axis represents arousal, symbolizing the level of activation or deactivation. The first quadrant represents emotions associated with happiness, the second encompasses emotions related to anger, the third includes sadness, and the fourth is associated with calm emotions. The circumplex model is illustrated in Figure 1.

Figure 1. Illustration of the valence-arousal model proposed by Russell [30].



Related Works

Recent advances in affective computing focus on developing techniques to detect affective states and integrating real-time affective feedback within adaptive systems. This is especially relevant in digital health, human-computer interaction, and intelligent rehabilitation. For instance, one study shows that emotion-sensitive adaptive approaches can promote task persistence and enhance in-game performance [33]. Another review finds that machine learning models support real-time personalization of psychological treatments by including behavioral and affective signals [34]. A systematic review by Leong et al [35] focused on facial emotion recognition, evaluating classification techniques and technologies used,

XSL•FO
RenderX

while emphasizing the increasing use of deep learning. Despite its contributions, the review is limited to 2 emotion recognition techniques and does not examine adaptive mechanisms within the software.

These studies do not focus specifically on physical rehabilitation, where emotional engagement is essential for treatment adherence. Therefore, there is a critical need to explore how affective computing can support empathetic interactions in rehabilitation contexts, through either the adaptation of gameplay elements or the application of novel emotion detection techniques.

Objective

This paper presents a scoping review of affective computing technologies, techniques, and methods used in serious games for physical rehabilitation, following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. It offers a comprehensive overview of the current research and highlights key gaps for future developments in this field.

Research Questions

To assess the current landscape of research in affective computing within serious games for physical rehabilitation, the following research questions were formulated:

RQ1: What theoretical approaches to emotion were applied in the studies, and which emotions were recognized?

RQ2: What techniques, including sensors and algorithms, were used for emotion recognition?

RQ3: Which game elements were adapted in response to affective state detection?

RQ4: What specific hardware was used to support the rehabilitation process?

Methods

Study Design

Given the objective of providing a broad overview of the literature and the diverse nature of the research questions, a scoping review was determined to be the most appropriate methodological approach. Scoping reviews are particularly suitable when the volume of publications or the heterogeneity of the topic precludes more focused methodologies such as systematic reviews [36].

This study adhered to the reporting standards outlined in the PRISMA protocols extension for scoping reviews [37] (Checklist 1).

Inclusion and Exclusion Criteria

Due to the limited available literature on this particular topic and to ensure the inclusion of all relevant contributions, studies with varied designs and outcome measures were considered. The inclusion criteria were as follows: papers published in English, regardless of publication date; studies that used affective computing within a serious game for physical rehabilitation; and studies that involved at least 1 patient in the evaluation process. Participants in the included studies were required to have a condition necessitating physical rehabilitation.

Review papers (eg, systematic reviews, scoping reviews, and literature reviews) were excluded to focus on primary empirical evidence.

Information Sources and Search Strategy

A comprehensive literature search was carried out in March 2024 and updated in February 2025 across the following databases: Google Scholar, PubMed, ScienceDirect, IEEE Xplore, ACM Digital Library, SpringerLink, and PEDro (Physiotherapy Evidence Database). The search strategy used the following keywords: “(affective OR emotional) AND serious (game OR games) AND rehabilitation.” The number of results retrieved from each database is presented in Table 1.

Table . Databases used.

Database	Results, n
Google Scholar	4370
PubMed	38
ScienceDirect	453
IEEE Xplore	5
ACM Digital Library	41
SpringerLink	354
PEDro	32

Selection of Sources of Evidence

Three authors independently reviewed the records obtained through the search. Duplicate entries were removed using Mendeley Desktop (version 2.122.0) Reference Manager. Studies that met the inclusion criteria based on the title and abstract were retrieved in full text for further evaluation. Any disagreements among the reviewers were resolved by a fourth reviewer, who made the final decision.

Data Charting Process and Data Items

Once the primary papers were identified, each was analyzed by at least 2 reviewers. These reviewers were responsible for extracting essential information from each study. Subsequently,

XSL•FO
RenderX

the reviewers met to reach a consensus on the extracted data. The following data items were identified per paper: study, citation, affective computing technique, sensor or device used to detect affective states, population characteristics (age and sex), type of disease, type of rehabilitation, game elements adapted in response to affective states, affective states detected, machine learning algorithms or software used for affective state detection, and theoretical approaches to affective states.

Assessment of the Quality of the Included Studies

The methodological quality of the selected studies was evaluated using the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool [38], which assesses 5 domains: sample selection, index test, reference standard, flow, and timing (Multimedia Appendix 1). Risk of bias and overall study quality were reviewed by 2 authors. Any disagreements in the assessment process were resolved through discussion and consensus within the research team. Question 1 in domain 3 was modified to require at least 2 tests to validate the target condition. The question was modified because, although there is no gold standard for validating emotions, Gonçalves et al [25] suggest that 5 components should be considered: cognitive appraisals, behavioral tendencies, motor expression, physiological reactions, and subjective feelings. Although validating the emotion using the 5 components would be more effective, it is also more complex. For this reason, it was expected that at least 2 tests would be used.

Synthesis of Results

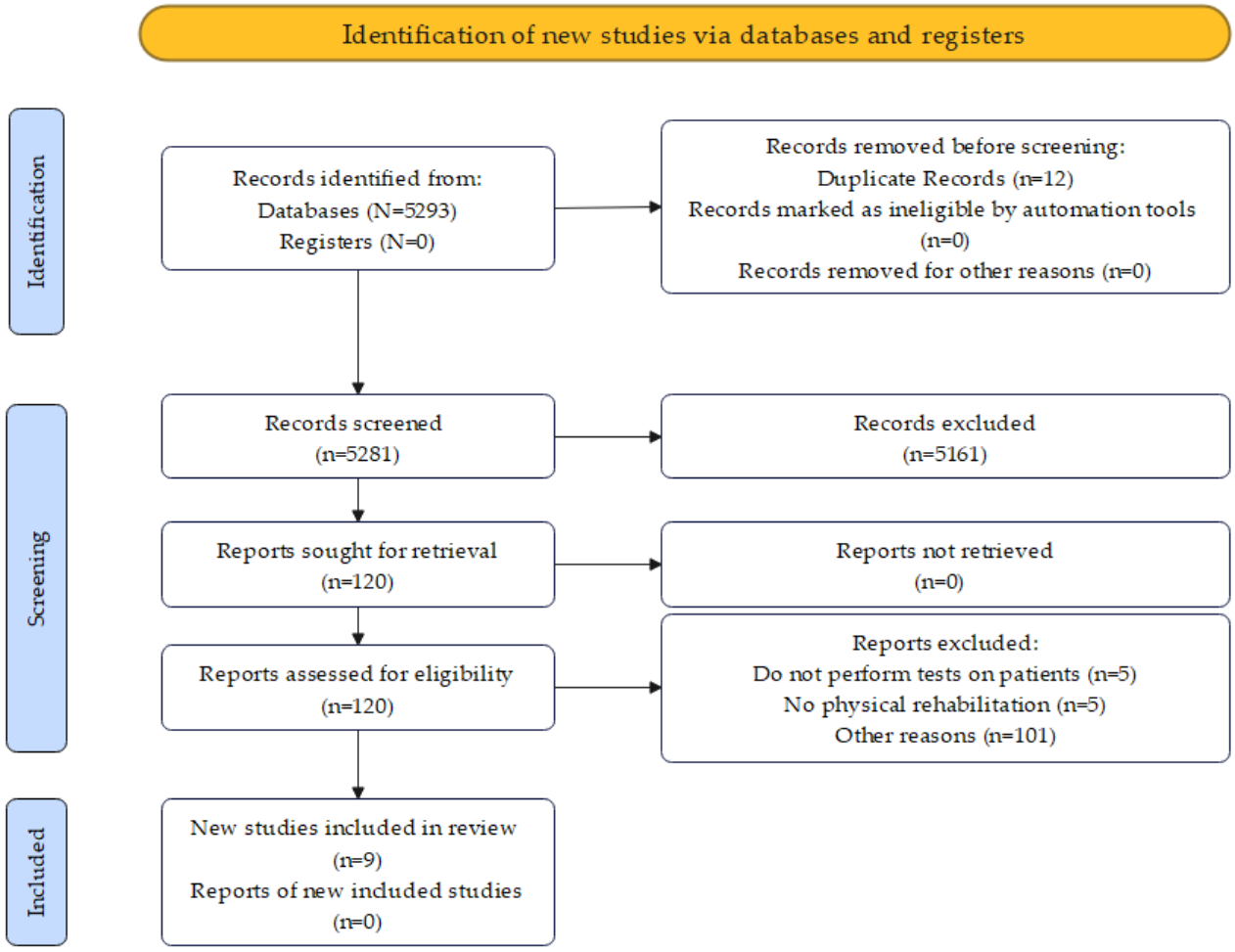
A formal narrative synthesis was conducted to evaluate the influence of affective computing on the rehabilitation process.

Results

Selection and Characteristics of Sources of Evidence

A total of 5293 records were identified through database searching. After removing 12 duplicate records, 5281 records were screened based on title and abstract, of which 5161 were excluded. As a result of this screening process, the full texts of 120 reports were sought for retrieval, all of which were successfully obtained and assessed for eligibility. Of these, 111 reports were excluded for various reasons, including interventions involving healthy individuals, games not involving physical rehabilitation, or review papers. Finally, 9 studies were included in the systematic review (Figure 2).

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram illustrating the selection process of the included studies.



XSL•FO
RenderX

Results From Individual Sources of Evidence

Analysis of the Research Questions

RQ1: What theoretical approaches to emotion were applied in the studies, and which emotions were recognized?

As shown in Table 2, 3 of the analyzed studies used the dimensional model [39-41], while the others did not specify the emotion model applied. The authors noted that affective states were classified as discrete (eg, pain and fatigue), enabling researchers to address these states as events requiring immediate intervention by therapists or the rehabilitation systems. Additionally, psychophysiological states were identified as having a direct impact on patients' motor performance.

Overall, the studies address affective states with direct clinical relevance to patient progress: (1) pain and anxiety are the main barriers to motor recovery, as they may lead patients to discontinue therapy; (2) physical and psychological fatigue are critical to monitor to prevent overexertion and frustration; and (3) engagement and motivation are essential for long-term treatment adherence. Studies such as those by Rivas et al [42,43] also highlight the interdependence among these states. For example, pain often coexists with anxiety, while engagement rarely occurs alongside extreme fatigue.

RQ2: What techniques, including sensors and algorithms, were used for emotion recognition?

Analysis of the 9 studies indicates that affective computing in virtual rehabilitation uses sensor-based techniques, artificial intelligence (AI) algorithms and models, and data integration to assess patients' affective states.

Sensor-Based Techniques

The reviewed studies classify sensors into 3 primary categories based on their nature and degree of intrusiveness.

- Movement-based and force-based systems: A notable trend in the field is the adoption of low-cost devices that integrate 3D position and pressure sensors to capture hand location and grip strength [42-46]. This nonintrusive approach allows affective states to be inferred directly from rehabilitation platform data, eliminating the need for additional sensors on the patient.
- Physiological and neurophysiological signals: Neurophysiological signals, especially EEG, are used to capture brain activity for the direct interpretation of affective states. Some approaches classify basic emotional valence (positive or negative) [41], while others use wearable devices to monitor complex states such as engagement, boredom, or frustration. These methods are also applied in areas such as gait rehabilitation [40].
- Computer vision and posture capture: The application of visual sensors has evolved from specific limb tracking, such as hand position analysis by Funabashi et al [47], to more robust multimodal approaches. Recent studies integrate conventional webcams for simultaneous analysis of facial expressions and body posture [42-45]. This integration enables a more comprehensive, contextually informed interpretation of the user's affective state.

XSL•FO
RenderX

Table . Summary of the studies included in the review.

Study	Affective computing technique	Sensor used	Type of study	Elements adapted in video game	Patient population	Age and sex	Disease	Rehabilitation	Affective states detected	Machine learning algorithms or software used to detect affective states	Emotion theoretical approaches
Funabashi et al [47]	Facial expressions	Webcam	Proof-of-concept	Not available	4	Not available	Stroke	Upper limb	Engagement, valence	Affectiva SDK[a]	Not available
Rivas et al [44]	Facial expressions, gesture recognition	Webcam, sensor to obtain 3D coordinates, and pressure sensor	Feasibility pilot	Not available	2	Not available	Stroke	Upper limb	Tiredness, tension, pain, and satisfaction	Linear support vector machines	Not available
Tironi et al [39][b]	User behavior	Microsoft Kinect	Proof-of-concept	Facial expressions, movements, and gestures of the avatar	1	Not available	Posture problem	Postural	Valence and arousal	Patient's game performances and textual inputs	Core affect, Russell [31]
Rodriguez et al [40]	Electroencephalogram	Emotiv Insight, Affect-Driven Self-Adaptive System (ADAS)	Preliminary study	Not available	8	Range 5 - 13 y; 4 females and 4 males	Gait-related impairment	Lower limb	Engagement, boredom, excitement, frustration, and meditation	Emotive SDK, ADAS software	Pleasure-arousal-dominance dimensional model, Mehrabian [29]
Rivas et al [45]	Facial expressions, gesture recognition	Webcam, sensor to obtain 3D coordinates, and pressure sensor	Feasibility study	Not available	5	Range 41 - 67 y; 4 females and 1 male	Stroke	Upper limb	Tiredness, anxiety, pain, and motivation	Fusion Semi-Naive Bayesian (FSNB) classifier, Multiresolution Semi-Naive Bayesian (MSNB) classifier	Not available
Hou and Sourina [41][b]	Electroencephalogram	Emotiv Insight	Proof-of-concept	Background music, visual adaptation, game difficulty	1	69 years; gender not specified	Stroke	Upper limb	Positive and negative	Real-time fractal dimension-based valence-level recognition algorithm	Valence-arousal-dominance emotion model, Russell and Mehrabian [32]
Rivas et al [46]	Gesture recognition	Webcam, sensor to obtain 3D coordinates, and pressure sensor	Feasibility study	Not available	5	Range 41 - 67 y; 4 females and 1 male	Stroke	Upper limb	Tiredness, anxiety, pain, and engagement	MSNB classifier	Not available

XSL•FO
RenderX

Study	Affective computing technique	Sensor used	Type of study	Elements adapted in video game	Patient population	Age and sex	Disease	Rehabilitation	Affective states detected	Machine learning algorithms or software used to detect affective states	Emotion theoretical approaches
Rivas et al [42]	Facial expressions, gesture recognition	Webcam, sensor to obtain 3D coordinates, and pressure sensor	Methodological and validation study	Not available	11	Range 41 - 76 y; 4 females and 7 males	Stroke	Upper limb	Tiredness, anxiety, pain, and engagement	FSNB classifier, Circular Classifier Chain (CCC)	Not available
Rivas et al [43]	Facial expressions, gesture recognition	Webcam, sensor to obtain 3D coordinates and pressure sensor	Methodological and validation study	Not available	5	Range 41 - 67 y; 4 females and 1 male	Stroke	Upper limb	Tiredness, anxiety, pain, and engagement	Semi-Naive Bayesian classifier, FSNB classifier, CCC	Not available

[a]SDK: software development kit.

[b]Games perform dynamic adaptation.

Algorithms and AI Models

The evolution of algorithms in this field has progressed from simple classifiers to architectures capable of addressing the complexity and interdependence of affective states: (1) probabilistic and Bayesian models, such as Semi-Naive Bayesian classifiers, have been used by Rivas et al [42,43,45,46] to support patients who have experienced stroke during upper-limb rehabilitation exercises; (2) dependency models, including multilabel approaches such as Circular Classifier Chains, have been implemented to simultaneously recognize multiple affective states (eg, pain, anxiety, and fatigue) [42,43]. These algorithms are critical because they account for the correlation among affective states; (3) fuzzy logic has been used in the Emphatic Virtual Caregiver system developed by Tironi et al [39], which applies fuzzy inference to process patient performance and affective state, thereby enabling smooth, natural transitions in the virtual caregiver's responses; (4) traditional machine learning techniques, such as support vector machines [44], have been documented in preliminary stages for classifying affective states, including tiredness, tension, pain, and satisfaction, based on descriptors of 3D hand movement and finger pressure.

Data Integration and Fusion

A trend in the papers is the adoption of multiple techniques for detecting affective states. Rather than relying on a single sensor, the systems proposed by Rivas et al [42-45] combine data from limb movements, facial expressions, and finger pressure. This integration permits the fusion algorithm to function as a higher-level layer, weighting information from each source to generate a more robust final emotional label that is resilient to noise and variability among different motion-sensing devices.

RQ3: Which game elements were adapted in response to affective state detection?

This research question seeks to determine which elements of serious games are adapted or modified in response to users' affective states. Four categories of adaptation were identified: auditory, visual, difficulty, and no adaptation.

Auditory Adaptation

Auditory adaptation is primarily used to regulate negative emotions and increase user engagement in serious games for rehabilitation, generally through background music and sound effects. This approach is frequently implemented when users display negative emotions despite correct task performance, indicating that auditory feedback serves as a compensatory mechanism to improve mood without changing the task structure, as observed in Hou and Sourina [41].

Visual Adaptation

Visual adaptation is a flexible approach to addressing users' affective states by modifying graphical user interface elements, such as avatars, environments, lighting, and in-game objects. These adaptations occur at expressive and environmental levels, involving modifications in facial expressions, gestures, object configurations, color schemes, and visual complexity.

At the expressive level, systems dynamically adjust avatar behavior to reflect users' emotional states, enabling empathetic responses through facial expressions and body movements, as implemented in Tironi et al [39]. At the environmental level, visual properties of the game—such as object shapes, colors, and scene configuration—are modified in response to users' affective states, particularly when negative emotions are detected despite adequate performance, as reported by Hou and Sourina [41].

Overall, visual adaptation works as both a reactive mechanism and a tool for modulating the user's emotional experience. This dual capacity highlights its potential to reflect and influence affective states in rehabilitation contexts.

XSL•FO
RenderX

Difficulty Adaptation

Difficulty adaptation is based on flow theory [48], which highlights the need to balance task challenge with the user's ability to maintain engagement. This balance is especially important in rehabilitation, where tasks must be adjusted to avoid both frustration and disengagement.

Hou and Sourina [41] found that task difficulty increases when users perform successfully and show positive affect. If negative affect is present, task difficulty remains unchanged. This asymmetry indicates a conservative adaptation strategy, prioritizing performance-driven progression while minimizing risks linked to lowering task demands.

No Adaptation

Unlike studies that incorporate adaptive mechanisms, some investigations do not adjust game elements in response to emotional recognition. In these instances, affective computing serves primarily analytical functions and does not directly influence system dynamics.

For example, Funabashi et al [47] evaluate user engagement and affective responses without altering gameplay. Similarly, Rodriguez et al [40] use electroencephalographic signals to identify emotional states, but do not use this information to modify game behavior.

Other studies use various machine learning algorithms for emotion recognition, but these data are not used to adjust game elements in real time. However, the authors agree that identifying states such as pain, anxiety, or fatigue aims to personalize rehabilitation tasks.

Specifically, Rivas et al [42-46] focus on improving the accuracy of affective state inference and propose involving adaptive mechanisms in future research. These mechanisms include adjusting difficulty levels and modifying the virtual environment to maintain a safe, motivating, and needs-based therapeutic context for patients.

The results show that affective computing in rehabilitation remains in a validation phase focused on identifying affective states. Analysis of the studies indicates that researchers have prioritized accurate identification of patients' affective states, such as stress, fatigue, or anxiety, before delegating clinical decision-making to software. The limited implementation of adaptation implies a lack of consensus regarding appropriate system responses to detected emotions, such as whether the system should halt or adjust task difficulty when negative emotions are identified.

RQ4: What specific hardware was used to support the rehabilitation process?

Analysis of the 9 included studies identifies a diverse technological infrastructure designed to support therapeutic exercise and enhance interactivity. Hardware used in physical rehabilitation can be categorized into 3 functional groups: nonintrusive motion capture sensors, haptic devices, and visualization and immersion hardware.

Nonintrusive Motion Capture Sensors

The reviewed studies demonstrate the use of nonintrusive technologies to capture movement during rehabilitation, allowing natural user interaction without invasive wearable devices. Rivas et al [42-46] use 3D position sensors to record therapeutic movements and supplement this data with webcams to capture facial expressions and gestures, enabling richer analysis through multiple data sources.

Similarly, Funabashi et al [47] use the Microsoft Kinect for Xbox One, which uses RGB-D (red, green, blue, and depth) technology to track and evaluate body movement. This device supports rehabilitation therapies by enabling patients to manipulate virtual objects with hand movements, fostering intuitive interaction within gamified therapeutic settings.

Collectively, these approaches demonstrate a trend toward the incorporation of accessible and nonintrusive sensors that enable accurate motion capture and integration with additional modalities. This integration supports a more comprehensive assessment of patients' performance in rehabilitation settings.

Haptic Devices

The literature highlights the use of specialized hardware to facilitate user interaction with rehabilitation environments. These technologies vary according to therapeutic objectives, ranging from high-precision haptic devices such as the Novint Falcon for controlling serious games [41] to advanced robotic assistance like the *Lokomat*. The *Lokomat* enables lower-limb rehabilitation by integrating body-weight support and robotic legs with a treadmill [40]. Additionally, customized interfaces have been developed, including sensorized levers designed to quantify grip strength and finger pressure during exercise [42-45].

Visualization and Immersion Hardware

Advanced visualization systems, such as the *Oculus Go*, enhance immersion and realism in rehabilitation environments. As documented by Rodriguez et al [40], these devices provide high-fidelity 3D visuals, increasing patients' sense of presence, which is critical for engagement and motivation during therapy.

Synthesis of Results

The affective states detected across the studies included engagement, valence, tension, pain, satisfaction, happiness, sadness, anger, surprise, boredom, excitement, frustration, and meditation. The most frequently recognized emotions were engagement, pain, tiredness, and anxiety. Only 3 studies used theoretical frameworks of emotion. The most commonly used affective computing techniques were facial expression analysis and gesture recognition, while user behavior analysis was the least used technique. Merely 2 studies adapted serious game elements based on the affective states of users. The adaptive elements identified included modifications to background music, visual interface parameters, game difficulty, and avatar facial expressions, movements, and gestures.

Quality Assessment

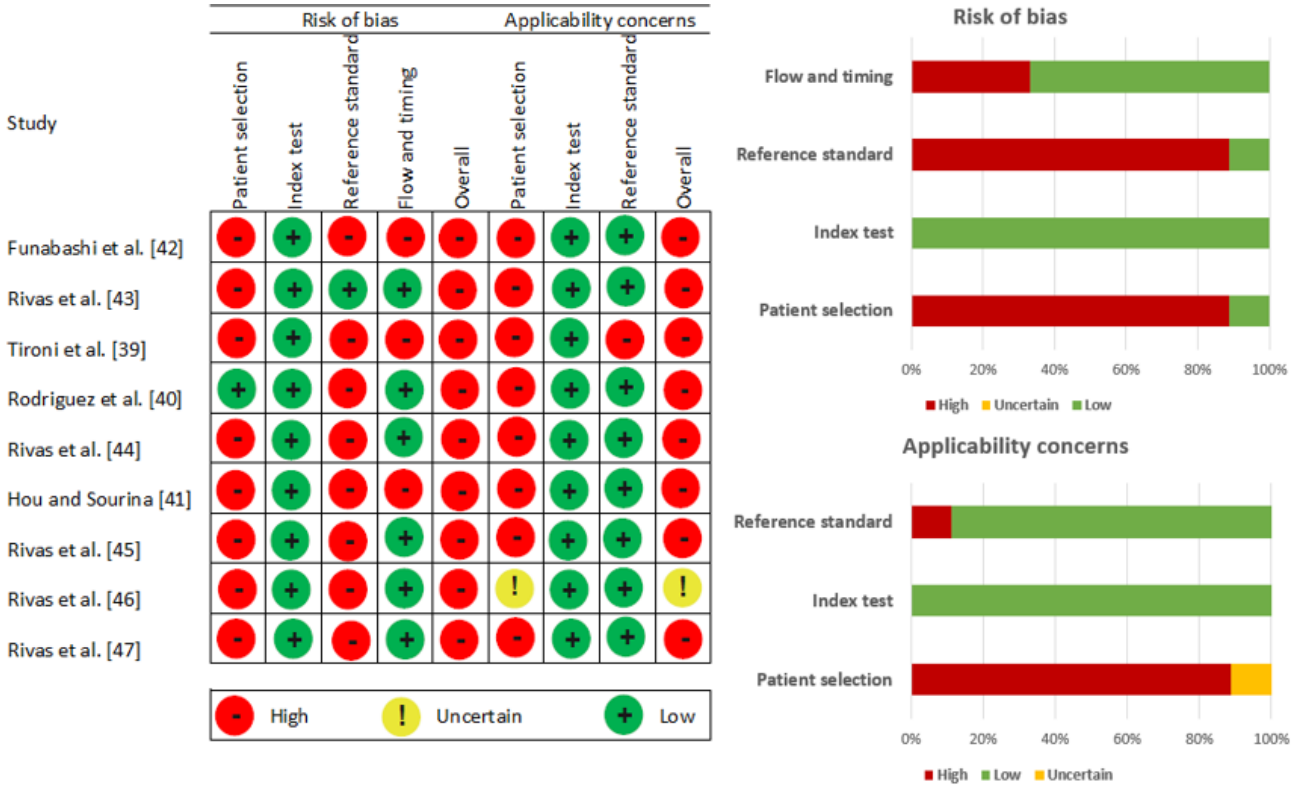
Quality assessment using the QUADAS-2 tool indicated that 88% (8/9) of the included studies exhibited a high risk of bias

XSL•FO
RenderX

in the patient selection category, mainly due to insufficient descriptions of recruitment procedures. Neither random nor consecutive recruitment was reported, which substantially increases the risk of selection bias. In the index test category, all studies (n=9, 100%) clearly defined the algorithms or techniques used for detecting affective states. For the reference standard, although no gold standard exists in affective computing, it is recommended to use multiple emotion recognition methods to improve accuracy [49]. However, 88% (8/9) of the studies did not use appropriate methods to assess reliability or validate their findings. In the flow and timing domain, 33% (3/9) of the studies demonstrated a high risk of bias due to the lack of explanation of the number of sessions and the failure to report participant dropout rates.

Regarding applicability, 88% (8/9) of the studies showed a high risk of bias in patient selection due to small sample sizes, limiting the generalizability of the findings. Additionally, patient recruitment from a single center further restricts the representativeness of the results for the wider patient population with varying severity levels. All studies presented low applicability concerns for the index test, as the techniques or algorithms used were clearly described. For the reference standard, 88% (8/9) of the studies exhibited low applicability concerns because they relied on clinical experts to validate emotions or used at least 2 emotion identification techniques, thereby ensuring that the system measures clinically relevant outcomes. Figure 3 illustrates the quality assessment outcomes.

Figure 3. Risk of bias assessment of the included studies using the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool [39-47].



Discussion

Implications for Clinical Rehabilitation

Given the critical role that affective states play in the rehabilitation process, this scoping review provides a comprehensive examination of the application of affective computing techniques within serious games designed for physical rehabilitation.

Studies involving healthy participants were excluded to ensure the analysis was relevant to clinical contexts. Patients frequently experience emotions such as anxiety, depression, and fear during rehabilitation. Thus, it is crucial to assess whether recognizing and adapting to these affective states can improve therapeutic outcomes. As a result, only 9 studies were included in the final synthesis. This highlights an important opportunity for future research to design studies involving patient populations where emotional state monitoring is integrated into the rehabilitation process.

Analysis of the 9 studies shows clear differences among techniques, each with specific implications for rehabilitation: (1) gestures and motion analysis, used by Rivas et al [42-46] and Funabashi et al [47], represent the predominant technique. The main advantage is its nonintrusive nature, which uses sensors already required for therapy to measure variables such as hand position and finger pressure. This characteristic makes it suitable for routine clinical practice, as it does not increase the sensory burden on patients, although its inference capability may be less direct than that of other techniques; (2) facial expression analysis, applied in multimodal approaches by Rivas et al [42-46], helps detect states such as satisfaction or pain. However, its efficacy can be compromised when patient expressions are altered, for example, throughout intense physical exercise, which may lead to misinterpretation by monitoring algorithms; (3) EEG, as implemented by Hou and Sourina [41]

XSL•FO
RenderX

and Rodriguez et al [40], directly measures brain activity and allows precise detection of affective states such as emotional valence. However, EEG used in serious games for physical rehabilitation is limited because electrodes can be intrusive for patients, and the technique is highly sensitive to noise from muscle movements. This sensitivity may limit its use during dynamic exercise in clinical settings.

A common limitation across the reviewed literature is the small sample size, which raises concerns regarding the generalizability and reliability of the reported findings. Future studies should aim for larger, more representative samples. Methodological guidelines, such as those proposed by Fritz et al [50], should be used for sample size calculation. Additionally, the reasons for participant dropout in physical rehabilitation interventions should be thoroughly documented, as this information can help inform and improve the design of future studies.

Research involving human participants is one of the most highly regulated areas in scientific inquiry, particularly regarding ethical standards. Data obtained from affective computing studies may be used for various secondary purposes, including video labeling, database inclusion, or training AI models. Consequently, adherence to ethical frameworks, such as those established by the American Psychological Association [51], is essential. Among the studies reviewed, only Rodriguez et al [40] explicitly reported compliance with the Declaration of Helsinki, received ethics committee approval, ensured data protection, and obtained explicit consent for video and photographic recordings. In contrast, several studies [42-44,46] merely acknowledged consent for video recordings, and others did not mention compliance with ethical guidelines at all, raising serious concerns regarding the ethical implications of affective computing applications. Clinical data privacy, especially in pediatric populations, must be a priority. Ensuring data confidentiality in storage, transmission, and usage is essential.

Regarding emotion theory, there is a lack of standardization, which makes direct comparisons between systems difficult; most studies did not justify their choice of emotion model. Among those that did [39-41], the dimensional approach to emotion was the most frequently adopted. This approach appears sufficient for many serious game applications, as determining whether a patient's affective state is generally positive or negative may be adequate for making gameplay adaptations such as altering difficulty or modifying environmental elements.

Another issue identified is the classification of pain as an emotion, which was reported in 50% (4/8) of the studies [42-46]. However, existing literature [52,53] suggests that emotions like depression, anxiety, and anger often emerge as psychological responses to pain rather than being emotions themselves. Therefore, classifying pain as an emotion may be conceptually inaccurate, highlighting the need for a more nuanced and consistent emotional taxonomy in future research.

However, detecting pain and fatigue (physical and psychological) is essential in rehabilitation, as these are major barriers to motor recovery. Pain may cause anxiety and fear of movement [43,46]. Nonintrusive pain detection allows systems to adjust challenges in real time, prevent secondary injuries, and maintain patient safety, which can improve engagement. By detecting patient fatigue, the system can recommend breaks, adjust exercises, or lower the difficulty to maintain high motor quality and enhance the rehabilitation process.

Clinical applicability is a further consideration when translating research into practice. When selecting data acquisition devices, factors such as cost, precision, portability, patient comfort, and mobility should be considered. Some devices, such as standard webcams [42-47], are easily accessible and cost-effective, while others, like the Emotiv Insight [40,41], require a more substantial investment.

Patient characteristics should also inform the choice of emotion recognition techniques. For instance, in pediatric rehabilitation, noninvasive methods should be prioritized to accommodate children's natural restlessness and their need for freedom of movement. In studies using facial expression analysis, factors such as ambient lighting, camera specifications, and the distance between the patient and the device can influence outcomes. However, these conditions were not reported in the included studies, and such factors must be considered when selecting or developing software libraries or AI algorithms.

An emerging trend identified is the integration of machine learning techniques. Findings suggest that combining multiple machine learning methods or modifying existing classifiers, such as the one proposed by Rivas et al [45], can enhance emotion recognition accuracy in virtual rehabilitation settings. These findings point to promising future directions for the integration of advanced AI methods in affective computing research.

Due to the variability in study designs and small sample sizes, the findings of this scoping review must be interpreted with caution. While affective computing appears to be a feasible strategy for enabling emotion-sensitive adaptations in serious games, the evidence for its clinical efficacy remains inconclusive. Accordingly, future research should adopt more rigorous methodologies, including larger patient samples, standardized protocols, and validated outcome measures, such as the Fugl-Meyer Assessment or Berg Balance Scale, to better assess the therapeutic impact.

Zhao et al [49] suggest that using multiple affective state recognition techniques can improve classification accuracy and the effectiveness of affective adaptations. However, only 4 of the studies [42-45] used more than 1 recognition method, with gesture and facial expression analysis being the most common combination. It is therefore recommended that future studies use at least 2 complementary techniques, selected based on their noninvasiveness and their ability to preserve patient mobility during rehabilitation.

User behavior analysis was the least explored approach, despite its potential for nonintrusive emotion detection. In this review, only Tironi et al [39] incorporated this technique, and its effectiveness remains to be validated. Future research should consider the application and validation of this approach in clinical rehabilitation settings.

The field remains underdeveloped, with most studies focused on technical validation or proof-of-concept, rather than rigorous experimental designs. The lack of randomized controlled trials

XSL•FO
RenderX

is a significant limitation, preventing affective computing from being established as an evidence-based intervention. As a result, these technologies are still considered experimental. Future research should use multicenter designs with statistically significant samples and control groups to validate the long-term clinical value of affective computing.

Analysis of the 9 studies highlights a major limitation in translating affective state recognition into functional application: adaptive approaches are infrequently implemented, with only 2 out of 9 studies [39,41] using a functional closed-loop system that detects emotion and adapts therapy accordingly. Three primary barriers account for this limitation: (1) clinical safety, as incorrect adaptations based on false positives may delay patient progress or compromise treatment effectiveness, for example, if physical effort is misinterpreted as pain and exercise intensity is reduced; (2) technical complexity, since real-time integration of patient signal processing with game engine logic requires considerable coordination among components such as game mechanics and hardware; and (3) the absence of theoretical intervention models, as current psychological models define affective states but do not provide protocols for automatic responses in affect-based rehabilitation systems [13]. As a result, researchers frequently use open-loop systems, in which software monitors the patient, but a human expert, such as the therapist, ultimately adjusts the therapy.

The application of the QUADAS-2 quality assessment tool revealed a high risk of bias in study participant selection, as most studies have small sample sizes. A high risk of bias was also found in the reference standard, since emotions were labeled through expert observation of video recordings. Subjectivity raises concerns that the emotion-detection algorithm, or index test, may simply replicate the biases of human observers. A moderate risk was observed regarding flow and timing, as several studies did not specify the testing interval. Considerable concerns also exist regarding the applicability of patient selection, since studies are often conducted in controlled research laboratories that may not accurately reflect real clinical practice, given patient variability as well as environmental factors such as movement and external noise.

The combination of serious games with affective computing facilitates the identification of states that may hinder rehabilitation, such as pain, fatigue, or anxiety. Failure to deal with these barriers can result in treatment abandonment or injuries due to overexertion in the short or medium term. Nonintrusive detection of pain or fatigue has direct implications for patient safety. In conventional clinical practice, therapists rely on patient self-reports, which may be biased by a desire to please health care professionals, communication difficulties, or embarrassment. Automated detection systems can support health care professionals in making more informed rehabilitation decisions. Affective serious games may serve as objective safety monitors, enabling the system to reduce challenge intensity before patients reach risk thresholds. Additionally, the inclusion of empathetic virtual agents offers a clinical strategy to address loneliness or depression associated with disability. Furthermore, integrating affective computing improves the feasibility of home-based therapy systems by delivering health care providers and family members with alerts regarding patient progress and emotional health.

Limitations

This scoping review presents several limitations. Although systematic search strategies were used to identify relevant studies across major electronic databases using carefully selected keywords, terminological heterogeneity within the literature may have resulted in the inadvertent omission of pertinent studies.

A major limitation of this review is the small number of included studies (n=9), which limits the ability to draw robust generalizations about the implementation of affective computing in serious games for physical rehabilitation. As a result, the identified trends should be viewed as preliminary, indicating that the integration of clinical rehabilitation and affective computing is still in an early stage of development.

The exclusion of gray literature further limits the scope of this review. By focusing exclusively on peer-reviewed studies, the review ensures a consistent standard of quality, coherence, and reliability. In contrast, gray literature often lacks external validation and standardized reporting structures, which increases the risk of methodological errors and interferes with the assessment of bias, sample representativeness, and the robustness of conclusions.

Another limitation is the absence of searches in languages other than English. All included studies are published in English in high-impact journals or specialized conferences, such as IEEE. This approach may exclude relevant research in other languages and obscure relevant regional developments.

Finally, the findings of this review should be interpreted in light of the methodological limitations of the included studies. First, most studies use small sample sizes, typically 1 to 11 participants. Second, there is also significant selection bias, as many studies rely on samples from specific institutions and often do not clearly describe the patient selection process. Third, most studies focus only on inferring and classifying affective states without implementing adaptive mechanisms to complete the functional loop.

Conclusions

This scoping review presents an updated overview of the state of research on the application of affective computing in serious games for physical rehabilitation. The studies analyzed identify various emotional states in patients undergoing rehabilitation through serious games. These papers use machine learning algorithms or commercial software development kits for emotion detection through at least 1 modality, such as facial expressions, EEG, gesture recognition, or user behavior.

The study shows that applying affective computing techniques to serious games for physical rehabilitation offers a promising complement to traditional rehabilitation approaches for various types of disabilities, incorporating patients' affective states into the design and adaptation of these games. While affective classification methods have advanced, closing the affective loop through automated interventions remains rare in serious game-based rehabilitation systems. Moreover, the limited

XSL•FO
RenderX

availability of primary sources weakens the conclusions regarding adherence and functional performance. Consequently, there is a strong need for multicenter studies with rigorous methodologies to develop systems with a greater impact on rehabilitation.

Therefore, the current evidence base for affective computing in rehabilitation through serious games is technically advanced, yet methodologically fragile. Existing studies remain in a model exploration phase, demonstrating the feasibility of emotion detection but lacking reliability at scale due to selection biases and uncertainties regarding real-life applicability.

Future research should consider the following key aspects: (1) methodological design, including a clear description of the theoretical approaches to emotion, patient selection criteria and characteristics, and a detailed description of the index tests and reference tests used; (2) the adaptation of serious game elements based on the patient's affective state; (3) use of validated motor function assessment scales to evaluate improvements in the rehabilitation treatment; and (4) promotion of multidisciplinary collaboration among experts in affective computing, physical therapists, and mental health professionals such as psychologists or psychiatrists.

We hope this study serves as a foundation for future investigations seeking to advance physical rehabilitation by integrating affective computing technologies.

Acknowledgments

The authors declare that no generative artificial intelligence tools were used in the creation, drafting, or editing of this manuscript.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questions of the QUADAS-2 tool.
[PDF File, 35 KB - rehab_v13i1e81344_app1.pdf]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 196 KB - rehab_v13i1e81344_app2.pdf]

References

1. Jack K, McLean SM, Moffett JK, Gardiner E. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228. [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
2. Michael D, Chen SL. Serious Games: Games That Educate, Train, and Inform: Cengage Learning PTR; 2005.
3. Argenton L, Schek E, Mantovani F. Serious games as positive technologies. In: Shumaker R, Lackey S, editors. Virtual, Augmented and Mixed Reality Applications of Virtual and Augmented Reality: Springer; 2014:169-177. [doi: 10.1007/978-3-319-07464-1_16]
4. Laamarti F, Eid M, El Saddik A. An overview of serious games. Int J Comput Games Technol 2014;2014:358152. [doi: 10.1155/2014/358152]
5. Mubin O, Alnajjar F, Al Mahmud A, Jishtu N, Alsinglawi B. Exploring serious games for stroke rehabilitation: a scoping review. Disabil Rehabil Assist Technol 2022 Feb;17(2):159-165. [doi: 10.1080/17483107.2020.1768309] [Medline: 32508187]
6. Abd-Alrazaq A, Abuelezz I, Hassan A, et al. Artificial intelligence-driven serious games in health care: scoping review. JMIR Serious Games 2022 Nov 29;10(4):e39840. [doi: 10.2196/39840] [Medline: 36445731]
7. Bonnechère B. Serious Games in Physical Rehabilitation: From Theory to Practice: Springer International Publishing; 2018. [doi: 10.1007/978-3-319-66122-3]
8. Vieira C, Ferreira da Silva Pais-Vieira C, Novais J, Perrotta A. Serious game design and clinical improvement in physical rehabilitation: systematic review. JMIR Serious Games 2021 Sep 23;9(3):e20066. [doi: 10.2196/20066] [Medline: 34554102]
9. Bernava G, Nucita A, Iannizzotto G, Caprì T, Fabio RA. Proteo: a framework for serious games in telerehabilitation. Appl Sci 2021;11(13):5935. [doi: 10.3390/app11135935]
10. Koutsiana E, Ladakis I, Fotopoulos D, Chytas A, Kilintzis V, Chouvarda I. Serious gaming technology in upper extremity rehabilitation: scoping review. JMIR Serious Games 2020 Dec 11;8(4):e19071. [doi: 10.2196/19071] [Medline: 33306029]
11. Picard RW. Affective Computing: The MIT Press; 1997.

XSL•FO
RenderX

12. Hocine N, Gouaïch A, Cerri SA, Mottet D, Froger J, Laffont I. Adaptation in serious games for upper-limb rehabilitation: an approach to improve training outcomes. User Model User-Adap Inter 2015 Mar;25(1):65-98. [doi: 10.1007/s11257-015-9154-6]
13. Wang K, Wei L, Liu Y, et al. Affective computing for healthcare: recent trends, applications, challenges, and beyond. In: Huang X, Mao Q, editors. Emotional Intelligence: Springer; 2025:3-19. [doi: 10.1007/978-981-96-5084-2_1]
14. Höök K. Affective loop experiences – what are they? Presented at: PERSUASIVE Technology: Third International Conference, PERSUASIVE 2008; Jun 4-6, 2008. [doi: 10.1007/978-3-540-68504-3_1]
15. Maj M, Kowalski M, Gnat M, Łukasiewicz J, Pliszczuk D. An intelligent support system and emotional state tests for people who are sick or recovering. J Mod Sci 2024;57(3):371-387. [doi: 10.13166/jms/191148]
16. Striegl J, Richter JW, Grossmann L, et al. Deep learning-based dimensional emotion recognition for conversational agent-based cognitive behavioral therapy. PeerJ Comput Sci 2024;10:e2104. [doi: 10.7717/peerj-cs.2104] [Medline: 38983201]
17. Rashidy NE, Allogmani E, Hassan E, Alnowaiser K, Elmannai H, Ali ZH. Toward real-time emotion recognition in fog computing-based systems: leveraging interpretable PCA_CNN, YOLO with self-attention mechanism. Front Comput Sci 2026 Jan 14;7:1714394. [doi: 10.3389/fcomp.2025.1714394]
18. Chu M, Qu J, Zou T, Li Q, Bu L, Shen Y. Design and evaluation of AR-based adaptive human-computer interaction cognitive training. Int J Hum Comput Stud 2025 May;199:103504. [doi: 10.1016/j.ijhcs.2025.103504]
19. Mamodiya U, Kishor I, Ahmed Syed A, Sankalkar P, Naik N. An adaptive human–robot interaction framework using real-time emotion recognition and context-aware task planning. IEEE Access 2025;13:152219-152240. [doi: 10.1109/ACCESS.2025.3603738]
20. Wang CW, Zhang Y, Osawa K, Nakagawa K, Tanaka E. Development of emotion recognition for rehabilitation feedback system using wavelet transform and LSTM. Robomech J 2025;12(1):35. [doi: 10.1186/s40648-025-00322-5]
21. Hurst M, Jackson TW, Glencross M. Emotion recognition—theory or practicality. Presented at: 18th International Conference on Automation and Computing (ICAC); Sep 7-8, 2012 URL: https://ieeexplore.ieee.org/abstract/document/6330551 [accessed 2026-06-13]
22. Harrigan J, Rosenthal R, Scherer K, editors. The New Handbook of Methods in Nonverbal Behavior Research: Oxford University Press; 2008.
23. Khan FA, Graf S, Weippl ER, Tjoa AM. An approach for identifying affective states through behavioral patterns in web-based learning management systems. Presented at: 11th International Conference on Information Integration and Web Based Applications and Services (iiWAS2009); Dec 14-16, 2009. [doi: 10.1145/1806338.1806418]
24. Paiva A, Prada R, Chaves R. Towards tangibility in gameplay: building a tangible affective interface for a computer game. Presented at: 5th International Conference on Multimodal Interfaces, ICMI 2003; Nov 5-7, 2003. [doi: 10.1145/958444.958446]
25. Gonçalves VP, Giancristofaro GT, Filho GPR, et al. Assessing users' emotion at interaction time: a multimodal approach with multiple sensors. Soft Comput 2017 Sep;21(18):5309-5323. [doi: 10.1007/s00500-016-2115-0]
26. Harmon-Jones E, Harmon-Jones C, Summerell E. On the importance of both dimensional and discrete models of emotion. Behav Sci (Basel) 2017 Sep 29;7(4):66. [doi: 10.3390/bs7040066] [Medline: 28961185]
27. Ekman P. Facial expressions. In: Dalgleish T, Power MJ, editors. Handbook of Cognition and Emotion: John Wiley & Sons, Ltd; 1999:301-320. [doi: 10.1002/0470013494.ch16]
28. Plutchik R. Emotions and Life: Perspectives from Psychology, Biology, and Evolution: American Psychological Association; 2003.
29. Mehrabian A. Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. Current Psychology 1996 Dec;14(4):261-292. [doi: 10.1007/BF02686918]
30. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980 Dec;39(6):1161-1178. [doi: 10.1037/h0077714]
31. Russell JA. Core affect and the psychological construction of emotion. Psychol Rev 2003 Jan;110(1):145-172. [doi: 10.1037/0033-295x.110.1.145] [Medline: 12529060]
32. Russell JA, Mehrabian A. Evidence for a three-factor theory of emotions. J Res Pers 1977 Sep;11(3):273-294. [doi: 10.1016/0092-6566(77)90037-X]
33. Schrader C, Brich J, Frommel J, Riemer V, Rogers K. Rising to the challenge: an emotion-driven approach toward adaptive serious games. In: Ma M, Oikonomou A, editors. Serious Games and Edutainment Applications: Springer; 2017, Vol. II:3-28. [doi: 10.1007/978-3-319-51645-5_1]
34. Gual-Montolio P, Jaén I, Martínez-Borba V, Castilla D, Suso-Ribera C. Using artificial intelligence to enhance ongoing psychological interventions for emotional problems in real- or close to real-time: a systematic review. Int J Environ Res Public Health 2022 Jun 24;19(13):7737. [doi: 10.3390/ijerph19137737] [Medline: 35805395]
35. Leong SC, Tang YM, Lai CH, Lee CKM. Facial expression and body gesture emotion recognition: a systematic review on the use of visual data in affective computing. Comput Sci Rev 2023 May;48:100545. [doi: 10.1016/j.cosrev.2023.100545]
36. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]

XSL•FO
RenderX

37. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
38. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011 Oct 18;155(8):529-536. [doi: 10.7326/0003-4819-155-8-201110180-00009] [Medline: 22007046]
39. Tironi A, Mainetti R, Pezzera M, Borghese NA. An empathic virtual caregiver for assistance in exer-game-based rehabilitation therapies. Presented at: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); Aug 5-7, 2019. [doi: 10.1109/SeGAH.2019.8882477]
40. Rodriguez J, Del-Valle-Soto C, Gonzalez-Sanchez J. Affective states and virtual reality to improve gait rehabilitation: a preliminary study. Int J Environ Res Public Health 2022 Aug 3;19(15):9523. [doi: 10.3390/ijerph19159523] [Medline: 35954882]
41. Hou X, Sourina O. Emotion-enabled haptic-based serious game for post stroke rehabilitation. Presented at: VRST '13: The 19th ACM Symposium on Virtual Reality Software and Technology; Oct 6-9, 2013. [doi: 10.1145/2503713.2503738]
42. Rivas JJ, del Carmen Lara M, Castrejon L, et al. Multi-label and multimodal classifier for affective states recognition in virtual rehabilitation. IEEE Trans Affective Comput 2022;13(3):1183-1194. [doi: 10.1109/TAFFC.2021.3055790]
43. Rivas JJ, Orihuela-Espina F, Sucar LE, Williams A, Bianchi-Berthouze N. Automatic recognition of multiple affective states in virtual rehabilitation by exploiting the dependency relationships. Presented at: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII); Sep 3-6, 2019. [doi: 10.1109/ACII.2019.8925508]
44. Rivas J, Orihuela-Espina F, Sucar L, Palafox L, Hernández-Franco J, Bianchi-Berthouze N. Detecting affective states in virtual rehabilitation. Presented at: 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth); May 20-23, 2015. [doi: 10.4108/icst.pervasivehealth.2015.259250]
45. Rivas JJ, Orihuela-Espina F, Sucar LE. Recognition of affective states in virtual rehabilitation using late fusion with semi-naive bayesian classifier. Presented at: PervasiveHealth'19: The 13th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 20-23, 2019. [doi: 10.1145/3329189.3329222]
46. Rivas JJ, Orihuela-Espina F, Palafox L, et al. Unobtrusive inference of affective states in virtual rehabilitation from upper limb motions: a feasibility study. IEEE Trans Affective Comput 2020;11(3):470-481. [doi: 10.1109/TAFFC.2018.2808295]
47. Funabashi AMM, Aranha RV, Silva TD, Monteiro CBM, Silva WS, Nunes FLS. AGaR: a VR serious game to support the recovery of post-stroke patients. Presented at: 2017 19th Symposium on Virtual and Augmented Reality (SVR); Nov 1-4, 2017. [doi: 10.1109/SVR.2017.15]
48. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience: Harper Perennial Modern Classics; 2008.
49. Zhao S, Jia G, Yang J, Ding G, Keutzer K. Emotion recognition from multiple modalities: fundamentals and methodologies. IEEE Signal Process Mag ;38(6):59-73. [doi: 10.1109/MSP.2021.3106895]
50. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012 Feb;141(1):2-18. [doi: 10.1037/a0024338] [Medline: 21823805]
51. Ethical principles of psychologists and code of conduct. American Psychological Association (APA). URL: https://www.apa.org/ethics/code [accessed 2025-02-05]
52. Lumley MA, Cohen JL, Borszcz GS, et al. Pain and emotion: a biopsychosocial review of recent research. J Clin Psychol 2011 Sep;67(9):942-968. [doi: 10.1002/jclp.20816] [Medline: 21647882]
53. Keefe FJ, Lumley M, Anderson T, Lynch T, Studts JL, Carson KL. Pain and emotion: new research directions. J Clin Psychol 2001 Apr;57(4):587-607. [doi: 10.1002/jclp.1030] [Medline: 11255208]

Abbreviations

AI: artificial intelligence
EEG: electroencephalogram
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analysis–Extension for Scoping Review
QUADAS: Quality Assessment of Diagnostic Accuracy Studies
RGB-D: red, green, blue, and depth

XSL•FO
RenderX

Edited by S Munce; submitted 26.Jul.2025; peer-reviewed by A Serrano, T Plavoukou; revised version received 03.May.2026; accepted 01.Jun.2026; published 02.Jul.2026.

Please cite as:
Beristain-Colorado MDP, Batres-Mendoza P, Guerra-Hernández EI, Cano-Pérez JL, Perezcampos-Mayoral C, Vargas-Treviño M, Gutiérrez-Gutiérrez J, Ambros-Antemate JF
Affective Computing in Serious Games for Physical Rehabilitation: Scoping Review
JMIR Rehabil Assist Technol 2026;13:e81344
URL: https://rehab.jmir.org/2026/1/e81344
doi:10.2196/81344

© María del Pilar Beristain-Colorado, Patricia Batres-Mendoza, Erick Israel Guerra-Hernández, José Luis Cano-Pérez, Christian Perezcampos-Mayoral, Marciano Vargas-Treviño, Jaime Gutiérrez-Gutiérrez, Jorge Fernando Ambros-Antemate. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 2.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Technologies for Children With Hearing Impairments to Support Language Learning: Scoping Review

Amarria Dila Sari[1,2], MSc; Abdullah Al Mahmud[1], PhD; Johanna Renny Octavia[3], PhD; Shivani Tyagi[1], PhD

[1]Centre for Design Innovation, Department of Architectural and Industrial Design, Swinburne University of Technology, John St, Hawthorn, Melbourne, Australia

[2]Department of Industrial Engineering, Universitas Islam Indonesia, Yogyakarta, Indonesia

[3]Centre for Ergonomics, Parahyangan Catholic University, Bandung, Indonesia

Corresponding Author:
Abdullah Al Mahmud, PhD
Centre for Design Innovation, Department of Architectural and Industrial Design, Swinburne University of Technology, John St, Hawthorn, Melbourne, Australia

Abstract

Background: Children with hearing impairments (HIs) have traditionally faced difficulties with language acquisition due to various factors, such as difficulties in accessing early intervention and therapy, among others. There is an opportunity for digital technology to address this problem; however, how different technologies can facilitate language acquisition, as well as the level of evidence, is not well understood.

Objective: This study aimed to explore the types, roles, and outcomes of different digital technologies in supporting language development in children with HIs.

Methods: Following Arksey and O'Malley's framework for conducting a scoping review and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews), we systematically searched 6 online databases from 2014 to September 2024 for relevant studies on the role of digital technologies in supporting language learning among children with HIs (aged <18 years).

Results: A total of 45 studies met the inclusion criteria. Mobile apps were the most frequently reported technologies, followed by digital books, extended reality, computer-based programs, and tangible or robotic tools. The most commonly used signed-language apps focus on vocabulary development and visual communication. Spoken language tools were less common but included augmentative and alternative communication (AAC) systems and auditory training apps. The efficacy outcomes of 12 studies included in this review showed that AAC-based apps had stronger effects on expressive language, pragmatic communication skills, and vocabulary development. These outcomes occurred when AAC-based apps were used in structured contexts with a therapist or parent. In contrast, signed language apps often reported narrower, word-level outcomes. Notably, relatively few interventions targeted early spoken language development, despite early childhood being a critical period for language acquisition. In addition, limited attention was given to visual design considerations related to user comfort, cultural relevance, and contextual adaptation.

Conclusions: The evidence base reveals 3 structural patterns: the dominance of accessibility-driven mobile technologies; an imbalance between signed and spoken-language interventions, characterized by a greater number of signed-language tools but stronger and more consistent evidence for spoken-language outcomes, particularly in AAC-based interventions; and limited integration of digital technologies across home, school, and clinical contexts. Future research should prioritize the co-design and evaluation of culturally responsive, integrated, developmentally appropriate digital systems that support early spoken-language development and sustained, family-mediated language learning across contexts.

(JMIR Rehabil Assist Technol 2026;13:e85066) doi:10.2196/85066

KEYWORDS

assistive technologies; digital technologies; augmentative and alternative communication; information communication technology; scoping review; children; digital health intervention; youth; hearing impairment; hearing loss; deaf; hard of hearing

XSL•FO
RenderX

Introduction

Background

Globally, more than 5% of people worldwide, about 430 million, including 34 million children, are living with disabling hearing loss, and this number is expected to rise to more than 700 million by 2050 [1]. Children with hearing impairments (HIs) encounter considerable challenges in acquiring spoken-language skills. HI refers to a reduced ability to hear sounds at normal levels, with varying degrees of severity. In children, hearing loss is typically diagnosed when the pure tone average exceeds 15 dB across key frequencies [2]. The scope of HI ranges from mild to profound. Children with profound hearing loss, who are often referred to as deaf, are unable to hear conversational speech without hearing technology, such as hearing aids or cochlear implants [2,3]. Those with mild to moderately severe hearing loss, commonly referred to as being hard of hearing, can often benefit from hearing aids [4]. In the context of this review, the term "children with HI" is used to refer to children who are deaf, hard of hearing, or have varying levels of hearing loss. Where relevant, cultural terms such as "deaf" have been used as per their application in the literature.

Communication preferences of children with HIs vary based on their access to hearing devices and language training. Some children rely on signed languages, while others use auditory input to develop spoken language through devices such as cochlear implants or hearing aids [5]. However, even with assistive devices, many children experience persistent challenges in developing spoken language due to limited access to early intervention, variability in device effectiveness, and insufficient support at home or school, which can lead to academic, behavioral, and social difficulties [6]. These challenges are compounded by barriers such as limited access to therapy, high costs, and a lack of parental knowledge regarding effective language-learning methods [7-9]. Many parents lack access to appropriate materials and training, leaving them unprepared to support their child's language development [10]. Language learning can be defined as the process by which individuals develop the capacity to understand, create, and use language for communication purposes. This involves the acquisition of listening, speaking, reading, and writing skills, as well as the understanding of grammar, vocabulary, and social use of language [11]. Language learning progresses from basic abilities, which include auditory or visual awareness, to the acquisition of language skills, followed by the use of language for pragmatic purposes [12].

Assistive and digital tools offer a promising solution to address these challenges. According to ISO 9999:2016, assistive products are defined as tools, whether custom-made or commercially available, designed to support people with disabilities. These technologies, including mobile apps and tangible devices, are innovative ways to enhance spoken-language learning for children with HIs. These technologies provide accessible, cost-effective resources that can support both parents and therapists in speech therapy [8,13].

Recent studies have included augmentative and alternative communication (AAC) devices [14], virtual reality (VR) apps such as BEARS (Both Ears) developed by Vickers et al [15], and digital books designed for children with HIs [16]. Despite the potential of these tools, there remains a gap in the development of cost-effective, interactive technologies that children can use independently, with parental support [8,9,13,17]. In particular, mobile technology has tremendous potential for providing home-based auditory training and speech therapy [18].

Recent reviews have explored various digital tools for supporting children with HIs. Herrera et al [13] highlighted that using hearing technology in mobile learning for children was affordable and accessible, although efficacy analyses and attention to visual design were lacking. Frisby et al [19] reviewed 146 studies of mobile hearing health care services across all age groups, with most studies originating in high- and upper-middle-income countries. While the review highlighted a focus on screening, diagnosis, and treatment, it placed limited emphasis on language learning for children with HIs. Hajesmaeel-Gohari et al [20] investigated digital games and found that they could increase motivation; however, their therapeutic impact remains unclear. Similarly, Nanjundaswamy et al [21] reviewed computer-based auditory training programs and acknowledged a range of tools for different age groups, including children, but highlighted a critical lack of evidence proving their efficacy. These findings point to an important gap: despite the increasing presence of digital tools, there is a limited understanding of how they influence the language-learning outcomes of children with HIs, particularly those who rely on spoken language. This scoping review addresses this gap by mapping the landscape of assistive and digital technologies for children with HIs, examining their features, contexts of use, and reported outcomes. In particular, the extent to which these tools support spoken-language development, for which evidence is scarce but is urgently needed, is explored.

Aim

This scoping review seeks to establish the landscape of assistive and digital technologies used to facilitate language learning among children with HIs. The review covers the following areas: (1) types of digital technologies applied, (2) the fundamental design elements and purposes of these technologies in facilitating language learning among children with HIs, and (3) the scope and nature of the reported efficacy outcomes. Although the review focuses on language learning as the core area of analysis, some studies that integrate language technologies with other domains of learning, such as literacy, mathematics, science, and general classroom interaction, are also included if language development remains a core component. The review covers both spoken language and sign language technologies to accommodate the varied communication needs of children with HIs. In this review, we have defined technology-based language learning interventions as digital tools intentionally designed to aid individuals in learning or improving language skills such as vocabulary acquisition, speech production, phonological awareness, listening comprehension, or signed language proficiency. Technologies whose primary purpose was unrelated to language learning, such as general communication aids or educational

XSL•FO
RenderX

games, were included in this review only when the intervention explicitly targeted measurable language-learning outcomes.

Methods

Overview

This scoping review was conducted following the framework proposed by Arksey and O'Malley [22], which posits that scoping studies are useful for mapping key concepts, types of evidence, and gaps in the research within a given field, particularly when the area is complex or has not yet been comprehensively reviewed. In light of the variety of digital technologies as well as the research designs that have been implemented so far, this study does not intend to critically discuss the efficacy of specific interventions. Instead, it maps and summarizes the types of efficacy outcomes across studies. This design enabled the identification of patterns across the literature, highlighted areas with limited evidence, and informed recommendations for future research.

We used the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews; refer to Multimedia Appendix 1 for the complete Checklist 1) to guide the reporting in this scoping review [23] and developed a population, concept, and context framework [24] that is provided in Table 1 below. As this was a scoping review, no formal critical appraisal or statistical synthesis was performed.

Table . Population, concept, and context (PCC) framework.

PCC[a] aspect	Description
Population	Children aged <18 years with HIs[b], including children who are deaf, Deaf, or hard of hearing, as defined by the original studies.
Concept	Assistive and digital technology used to support language learning.
Context	Settings in which digital technology was used in controlled interventions to support the spoken-language learning of children with HIs. In addition to the contexts of intervention (home, school, or clinic), the geographical contexts (high-income countries or low- and middle-income countries) were also explored in descriptive terms under the component of "Context" in the PCC framework. However, this is only to look at the differences in access, scale, or learning environments and not to pose another research question in analytical terms.

[a]PCC: population, concept, and context.

[b]HI: hearing impairment.

Step 1: Identifying the Research Questions (RQs)

This review is based on the following research questions that seek to explore how digital technologies are used in facilitating language learning and academic engagement for children with HIs:

RQ1. What types of digital technologies support language learning for children with HIs?

RQ2. What is the role of digital technology in supporting language learning and language-mediated academic participation among children with HIs?

RQ3. What are the reported efficacy outcomes of these digital technologies in supporting language development for children with HIs?

Step 2: Search Strategy

The authors conducted a comprehensive search strategy to include all relevant literature published between January 2014 and September 2024 across the following databases: Ovid MEDLINE, PsycINFO, Web of Science, Scopus, CINAHL, and Computers and Applied Sciences Complete. An experienced librarian assisted in refining the subject headings and keywords to ensure broad coverage of technology-based interventions and HIs among children. The final search strategies are provided in Multimedia Appendix 1.

Step 3: Study Selection

Studies needed to focus on the identification, implementation, and evaluation of assistive technology and digital technology for children with HIs to be included in this review. Refer to Textbox 1 for details regarding the inclusion and exclusion criteria. Eligible studies were identified using a 3-step process. In step 1, one reviewer (ADS) collected the search results and then removed duplicates. In step 2, two reviewers (ADS and AAM) reviewed the titles and abstracts of the identified studies that appeared to meet the inclusion criteria. In step 3, the same reviewers evaluated the full texts of the remaining publications that were identified in our searches to include potentially relevant publications. Covidence software (Veritas Health Innovation Ltd) was used to facilitate the screening, selection, and data management processes during the scoping review [25].

Eligibility Criteria

Studies needed to focus on the identification, implementation, and evaluation of assistive technology and digital technology for children with HIs to be included in this review. Refer to Textbox 1 for details regarding the inclusion and exclusion criteria. Studies were included if they focused on identifying, implementing, or evaluating assistive or digital technologies for children with HIs. Studies were limited to English-language, peer-reviewed journal studies published between January 2014 and September 2024. Studies of adults, nondigital technologies, or those focusing exclusively on signed language were excluded.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Studies focused on hearing technology for children with hearing impairments (HIs).
- Studies involving digital technology, mobile apps, or computer-based interventions.
- Studies written in English.
- A study was only included in our review if the technology was specifically intended for language development outcomes such as vocabulary learning, grammar acquisition, speech production, or signed language comprehension.
- Digital tools used by children with HI may indirectly impact language development if they serve as communication aids or educational games. However, this review only included technologies mainly aimed at language learning or those directly measuring language-related outcomes.
- Studies published between 2014 and 2024 were included in this review. This period was chosen because mobile and app-based digital interventions have evolved rapidly since smartphones and tablets became widely used [26]. Earlier studies were largely based on technologies (eg, desktop-based or clinic-bound technologies such as computer-based training programs, CD-ROM–based tools, or fixed web platforms typically used in controlled settings) that are now quite different from the digital tools commonly used today (eg, app-based learning environments that support multimodal interaction and home-based language practice). From around 2014 onward, the growth of mobile learning technologies and app-based learning tools has had a substantial impact on digital environments for language learning, making this timeframe particularly relevant to the aims of our study.
- Journal studies.

Exclusion criteria:

- Studies were excluded from our review if the technology served primarily as a communication aid, accessibility tool, or general educational application without explicit language-learning objectives or language outcome measures.
- Abstract only, conferences and posters, study protocols without the published full-text study, and editorial letters.
- Studies exclusively focused on traditional sign-language instruction without a digital assistive learning component targeting broader communication, literacy, or educational participation outcomes were excluded.
- Studies focused on HIs in adults aged >18 years.
- Studies not written in English.
- Studies published before 2014.
- Nondigital technologies.

Screening and Selection Process

Eligible studies were identified using a 3-step process. In step 1, one reviewer (ADS) collected the search results and then removed duplicates. In step 2, two reviewers (ADS and AAM) reviewed the titles and abstracts of the identified studies that appeared to meet the inclusion criteria. In step 3, the same reviewers evaluated the full texts of the remaining publications that were identified in our searches to include potentially relevant publications. Covidence software (Veritas Health Innovation Ltd) was used to facilitate the screening, selection, and data management processes during the scoping review [25].

Step 4: Data Extraction and Charting the Data

Data from the eligible studies were extracted using standardized data extraction forms developed by the research team (refer to Multimedia Appendix 1). Two separate data extraction tools were used to capture key details about digital technologies, including the type of technology, the communication preference (signed language or spoken language), app settings, app purpose (research-based or commercial), and app functionality (eg, language learning, math learning, literacy, or educational assistance). In addition, the extracted data included study characteristics such as the year of publication, the country in which the study was conducted, and the purpose of the study, as well as participants' demographics, such as age, sex, and degree of HI. Outcome measures were also recorded to assess the impact of the technology.

Step 5: Collating, Summarizing, and Reporting the Results

We grouped the studies based on their technological focus and summarized the key characteristics, including the study design, target populations, settings, research purposes, and reported limitations. We identified and outlined the features of the digital tools used for each group, together with the intended outcomes. This process was intended to highlight existing gaps and to inform future research or intervention designs.

Results

Overview of the Studies

Initially, 2541 studies were imported into Covidence; after duplicate removal, 1783 citations were identified from electronic database searches and study references. After screening the titles and abstracts of these citations, 330 studies were included in the full-text review process. Of these 330 studies, 45 studies met the inclusion criteria, and 285 were excluded (refer to Figure 1). Data extraction was completed for the 45 studies that were considered eligible for this review. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart is shown in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart for the selection of the included studies.

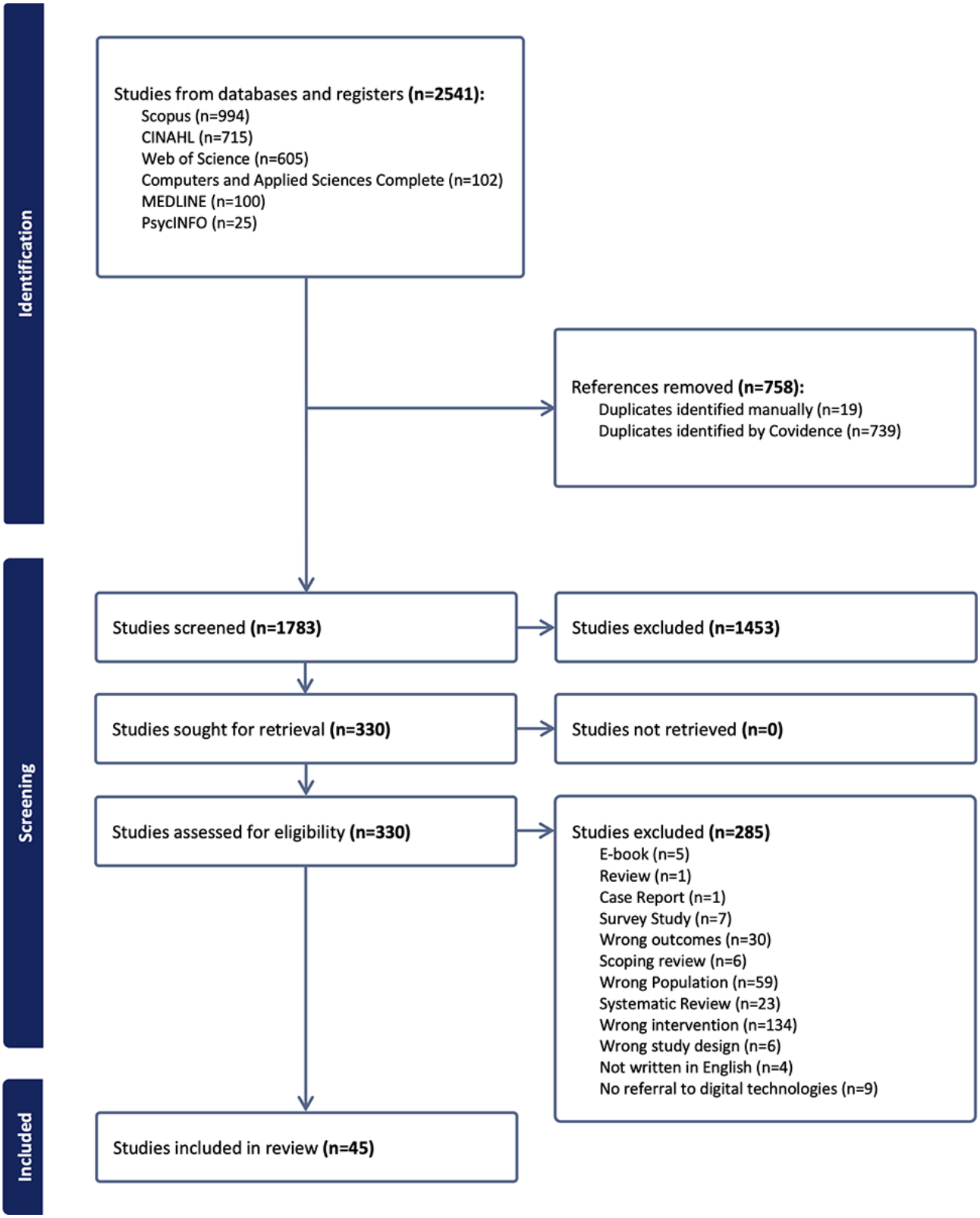


Figure 2 shows the geographical distribution of the included studies. Most studies were conducted in Asia (40%, 18/45), followed by Europe (22%, 10/45), and North America (20%, 9/45). All North American studies originated from the United States, which represented the largest national contribution. South America contributed 13% (6/45), primarily from Colombia (4/45), while Africa accounted for 4% (2/45), represented by Mauritius and Tunisia. Consistent with the population, concept, and context framework, geographical context (high-income vs low- and middle-income countries) is reported descriptively to contextualize access, scalability, and learning environments, rather than as a separate analytical focus.

Among the studies conducted in low- and middle-income countries (n=24), Indonesia and Colombia contributed the

XSL•FO
RenderX

highest number of studies (n=4 each), followed by Pakistan (n=3) (Table 2).

Of the 45 studies, only 18 studies involved parents and experts. Parents mainly supported progress monitoring [27-29] and were most engaged during the design stage [16,30-36]. A few studies also created dedicated interfaces for parents and experts [6,34], highlighting the role of technology in connecting home and professional contexts. In addition, testing was not limited to children, as parents and experts were also part of the evaluation process. Previous studies by Mood et al [30], Meinzen-Derr et al [37], and Shoaib et al [38] highlighted that technology requires the continuous input of parents, therapists, and teachers to maximize its impact.

Figure 2. Distribution of included studies by country, showing the number of studies conducted across different geographical regions.

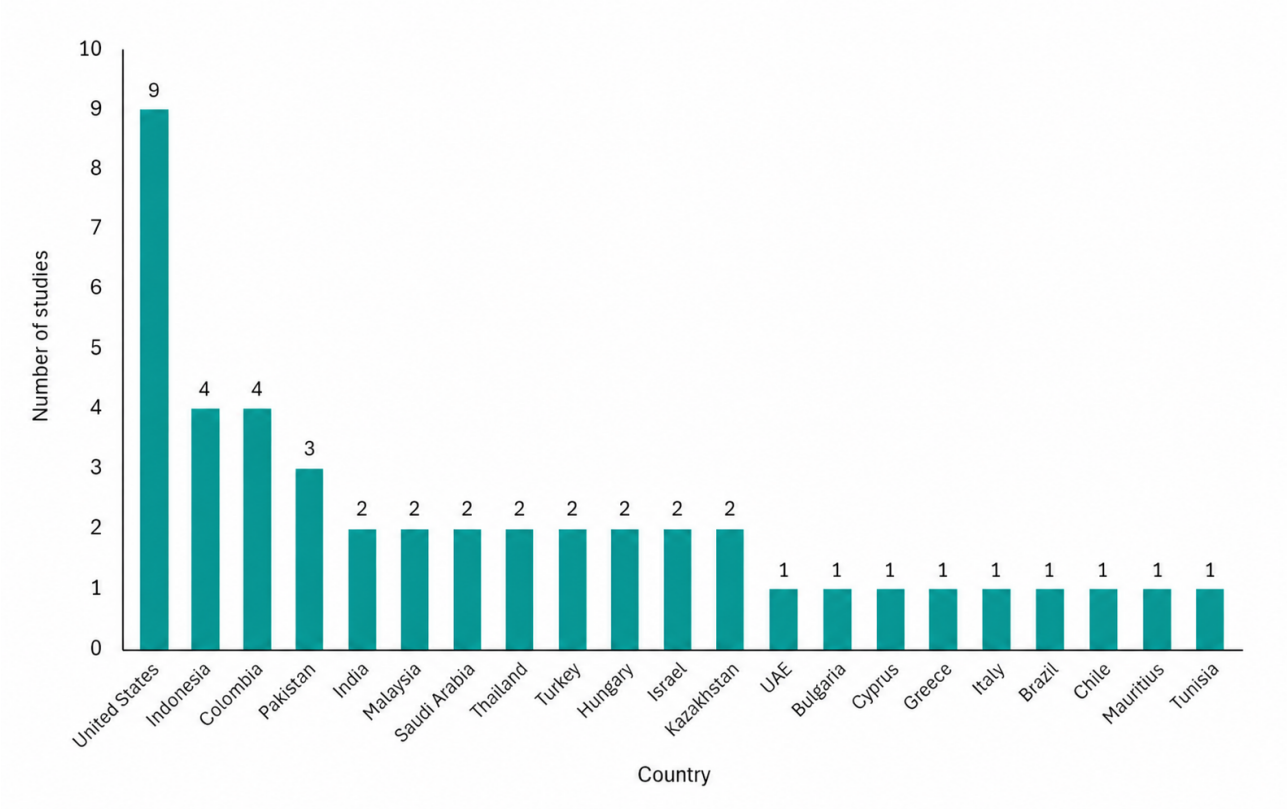


Table . Distribution of studies across low- and middle-income countries (N=24).

Country	Value, n (%)
Colombia	4 (16.7)
Indonesia	4 (16.7)
Pakistan	3 (12.5)
India	2 (8.3)
Kazakhstan	2 (8.3)
Malaysia	2 (8.3)
Thailand	2 (8.3)
Turkey	2 (8.3)
Brazil	1 (4.2)
Mauritius	1 (4.2)
Tunisia	1 (4.2)
Total	24 (100)

Participant Characteristics

The mean sample size across the included studies was relatively small, with an average of 23 participants. Participant ages ranged from 2 to 17 years, with a mean age of 10.6 years. The mean age of children was 11 years, and the mean sample size was 20 participants. Of the 45 included studies in this review, 27 reported participant gender, comprising 778 individuals. Of these individuals, 397 (51%) were male, and 381 (49%) were

XSL•FO
RenderX

female. The remaining 18 studies (221 participants) did not specify gender information. All studies involved participants with HIs; however, only 21 studies provided details on the severity of hearing loss. In total, 8 studies reported profound hearing loss, 1 severe, 5 moderate, and 7 mild. The remaining 24 studies did not provide specific information regarding the extent of hearing loss. There was an inconsistency in the reporting of gender and demographic data; therefore, the aggregate data only reflect studies that reported this information.

Types of Digital Technologies That Support Language Learning for Children With HIs

Five categories of digital technology were identified, namely mobile apps (n=26), digital books (n=6), extended reality, augmented reality, and virtual reality (n=5), computer-based programs (n=3), and other tools such as robots, tangible devices, or cross-platform systems (n=5; refer to Multimedia Appendix 1 and Table 3). Across the range of technologies, mobile apps consistently highlighted the importance of customization, interactivity, and visual reinforcement, whereas XR and tangible systems highlighted the importance of embodied and immersive approaches to learning.

XSL•FO
RenderX

Table . Digital technologies and features supporting spoken, signed, and bimodal language learning.

Purpose, type of technology, and key features	Frequency
Language learning, literacy, and math learning	
Mobile app	
Game-based approach: reward, level, and challenges.	4
Multimedia content: audio, video, images, text, animations, and colorful design.	8
Learning monitoring: activity log and progress report.	5
Interactive: real-time feedback, task manipulation, touch, speech, and gesture.	2
AI[a] feature: optical character recognition, AI image classification, and object detection.	2
Online and offline functionality.	3
Real-time image processing	2
Parental support	2
Professional support	1
Vocabulary learning	2
Fingerspelling	2
Pronunciation, speech, and assistant	
Computer	
Game-based approach	3
Real-time feedback and visual feedback.	3
Learning progression	2
Education support and language learning	
AR[b], MR[c], and VR[d]	
Interactive object manipulation	4
3D avatar-based interaction	5
Visual learning	3
Spoken-language learning	
AAC[e]	
Symbol-based communication, speech output, and custom interface.	5
Language learning and education support	
Other (tangible or robot)	
Gestures, poses, speech synthesis, and signed-language tutoring; tablet-based display for videos, animations, and interactive exercises; multimodal engagement (auditory-verbal, visual, and cued speech approaches); and autonomous navigation and interaction with children in an adaptive learning environment, with real-time feedback and rewards.	3
Physical board games and tactile features.	3

[a]AI: artificial intelligence.

[b]AR: augmented reality.

[c]MR: mixed reality.

[d]VR: virtual reality.

[e]AAC: augmentative and alternative communication.

Various digital technologies have been developed to support language learning for children with HIs, each offering distinct features and learning approaches. Mobile-based apps, which were the most frequently represented, tended to incorporate diverse interactive elements, such as color, animations, and accessible font styles that were designed to engage young users

XSL•FO
RenderX

[39,40]. Customization features appeared most frequently across all technology types, reflecting the need to accommodate variations in hearing levels and individual progress. Leveling systems were also suggested as essential additions, enabling content to be adapted according to the child's stage of development [41,42]. XR and tangible technologies, such as robots, provided immersive and hands-on experiences, while computer-based and cross-platform software supported more structured or instructional content.

The learning goals that were supported by these technologies typically aligned with either spoken-language or signed-language development. Technologies that were designed to support spoken language, which were frequently delivered through mobile apps or AAC-based platforms, often focused on auditory memory and visual reinforcement [30,37]. These tools included speech-to-text, audio playback, activity logging, and real-time feedback to support receptive and expressive language skills. By contrast, technologies that supported signed-language learning relied heavily on visual modalities and incorporated elements such as video demonstrations of signs, fingerspelling, and articulation animations [43-45].

While mobile-based technologies dominated in frequency and feature integration, each type of technology had unique value depending on the targeted communication mode. Tools that supported spoken languages tended to integrate multimodal input (auditory and visual), whereas those supporting signed language were predominantly visual. These differences highlight the importance of aligning technological features with the sensory needs of children with HIs, as well as with their language preferences [31]. Customization, interactivity, and adaptability remain central to the effectiveness of these digital tools in diverse learning contexts.

The Role of Digital Technology in Supporting Language Learning and Language-Mediated Academic Participation in Children With HIs

While this review focuses on language learning, many of the technologies described incorporate language development into more general learning activities, such as literacy, mathematics, science, and participation. As a result, these areas were also included whenever language comprehension, expression, or communication formed a core component of the learning process. Digital technologies served the following two main functions: (1) assistive and learning support (literacy, math, science, and classroom participation) and (2) language learning (spoken and signed languages).

Assistive and Learning Support

Across the reviewed studies, assistive and learning support technologies for children with HIs consistently featured multimodal, interactive tools that adapted traditional curricula and presented them in more accessible and engaging formats. These technologies span various domains, including literacy, mathematics, science, and classroom participation, with an emphasis on personalized and inclusive learning experiences.

In terms of literacy, the most common tools used are digital storytelling and gamified apps to enhance reading comprehension, vocabulary acquisition, and phonemic awareness. Standard features included animated narratives, interactive feedback, and adaptive content, which not only maintained learner engagement, but also supported comprehension in both signed- and spoken-language contexts [46-52]. Several studies highlighted the role of e-books and storytelling apps in facilitating bilingual or bimodal literacy, reflecting a shift toward culturally and linguistically responsive learning materials.

In terms of math, digital tools often incorporate multimedia elements, such as visual aids, auditory prompts, and gesture-based input, to improve the understanding of mathematical concepts. A shared feature across these tools was their ability to break down abstract mathematical ideas into concrete, interactive tasks. Many platforms also integrated signed language or visual modeling to accommodate different learning styles [32,53-56]. These interventions indicated a common goal of making math education more inclusive and accessible, particularly by addressing language-related barriers that may impede conceptual understanding.

In terms of science and classroom learning, a variety of assistive technologies, such as AR apps and humanoid robots, were used to simulate real-world experiences to foster inquiry-based learning and promote social interaction [43,57-59]. Of note, some of these tools, such as humanoid robots, were designed for classroom use and to support parents and caregivers at home, thereby enabling them to reinforce language and academic skills in informal settings. This highlighted the increasing recognition of the home environment as a crucial context for learning, particularly when access to professional therapy was limited.

Differences among these categories reflected their learning objectives and sensory modalities. Literacy tools often focus on phonemic and vocabulary development through fingerspelling for signed languages and audiovisual storytelling for spoken languages [6,49]. Math tools prioritized logical reasoning and problem-solving through interactive formats [53,54,60], while science-oriented technologies were geared toward experiential learning and stimulating curiosity [61]. Across the domains, most of the tools incorporated multimodal input (visual, auditory, or gestural) and adaptive content features.

Language Learning

A strong trend across spoken-language interventions for children with HIs was the increasing use of technology-enhanced tools that combined visual, auditory, and interactive elements to support diverse aspects of spoken-language development. These tools addressed various communication skills, ranging from auditory-verbal training and lip reading to speech articulation, phonemic awareness, and expressive language production.

One commonality among these interventions was the emphasis on multisensory input. Most apps integrated real-time visual and auditory feedback, allowing users to self-correct and reinforce their learning. For example, tools targeting articulation and auditory training often use game-based interfaces and visual prompts to help learners practice specific speech sounds and patterns in an engaging environment [5,27,33]. Similarly, lip-reading tools focused on visual speech cues, providing

XSL•FO
RenderX

guided repetition and clear visual models to support speech recognition without relying on signed language [28,38].

AAC-based systems were another common strand, particularly for supporting expressive spoken language and pragmatic communication. These tools typically provided audiovisual pairings, customizable vocabularies, and structured sentence construction to support communication across different developmental stages [14,30,37,62]. Longitudinal research has also shown that AAC tools can be adapted to a child's developing language abilities and can remain effective across contexts [63].

Despite these shared features, differences in focus and depth were noted. Some tools were narrowly designed for phoneme-level training or early auditory skills [27], while others provided comprehensive language systems with vocabulary expansion, grammar structuring, and social communication support [14,30]. The level of interaction also varied, ranging from passive repetition tasks to immersive, gamified environments with dynamic feedback loops [64].

Overall, these studies reflected a shift toward personalized, feedback-driven spoken-language learning, with tools tailored to different linguistic profiles, therapy goals, and developmental needs.

34

Efficacy of the Tools for Language Learning

Table 4 summarizes the characteristics, interventions, and key findings of the 12 studies that reported efficacy outcomes. These studies covered spoken-language learning (AAC, phonemic awareness, and prosody) and signed-language learning, with 1 study also reporting academic outcomes. Where available, statistical results are reported descriptively to reflect the outcomes presented in the original studies, rather than to infer comparative effectiveness.

XSL•FO
RenderX

Table . Efficacy of digital technologies.

Study	Aim of the study	Name of the platform	Age of the children (years)	Key results
Mood et al [30]	To evaluate AAC[a] on iPads for language development	AAC on iPad TALI[b]	3 - 12	Over 24 weeks, children aged ≥5 years in the TALI[b] group showed greater gains in pragmatic language (12.7-point increase in CELF-5[c]) vs a 6-point decrease in the TAU[d] group (P=.04). For children under 5, the TALI group improved significantly in VABS[e] Communication (86.7 to 99.1) and Social domains (91.8 to 97.4, P=.01), while the TAU group showed no significant change.
Meinzen-Derr et al [62]	To assess the TALI program's impact on communication skills	TouchChat HD: AAC with WordPower (used in TALI)	3 - 12	Over 24 weeks, children in the TALI group, compared to those in the TAU group, had significantly greater increases in the length of phrases they used to express themselves (β=.91 vs .15, respectively; P<.001). Similar findings were seen with conversational turn-taking and the number of different words spoken.
Meinzen-Derr et al [37]	To implement a language intervention that used AAC technology on iPad VR devices to enhance language development in children who are DHH[g] with language underperformance	AAC software on iPad TouchChat HD	5 - 10	[f]More than 24 weeks, median MLU[f] increased from 2.41 (range 1.09-6.63; mean 2.88) to 3.68 (range 1.97-6.81; mean 3.62; P=.002). Median total words were 251 (range 101-458) and median different words were 100 (range 36-100) at baseline, both increasing significantly over time. Total words: β=26.8 (7.1); P=.001 and different words: β=8.0 (2.7); P=.008. Mean (SD) values increased from 3.15 (2.09) to 3.86 (1.81) for MLU, from 265.4 (130.1) to 357.8 (159.2) for total words, and from 96.6 (44.2) to 125.2 (42.8) for different words; mean turn length showed minimal change.
Meinzen-Derr et al [14]	To determine whether integrating AAC core-word language strategies improves speech-language outcomes in young children who are deaf or hard of hearing	AAC strategies using TouchChat HD with WordPower by Silver Kite	4-11	The study reported high satisfaction with AAC strategies and notable language gains: MLU increased by 38%, NDW by 39%, and the Pragmatic Checklist Score rose from 68.8 to 95.5 (P=.01).

XSL•FO
RenderX

Study	Aim of the study	Name of the platform	Age of the children (years)	Key results
Mohammad et al [34]	To test a game-based mobile app for a signed language	My Sign	6 - 12	The study showed that children using the mobile app scored significantly higher in the post-quiz, with a 60% improvement in Group A. The *t* test (*P*=.01) confirmed the app's positive impact, and the Eta squared (0.585) indicated a strong effect size. The game-based learning approach led to higher engagement and better learning outcomes compared to the traditional method.
Ploylji w and Michel [35]	To use the AR[h] book to teach vocabulary and signs	AR book	5 - 9	The results showed that the efficiency of the lesson in the pretest and the exercise during the study (E1)[i] averaged 85.33%, while the posttest (E2) averaged 87%.
Sztahó et al [65]	To improve prosody using CAPT[j]	CAPT	6 - 10	The CAPT app significantly improved prosody in children with HIs (t_{14}=4.07; *P*<.001), with strong agreement between teachers' and automatic scores (*r*=0.87). Gains were observed across all hearing levels, particularly among children with mild hearing loss.
Vasel and Ragonis [32]	To assess multimedia impact on academic outcomes	Multimedia instruction	7 - 11	The program significantly improved students' language and math skills, with gains in semantics, reading (49.5→60.5), grammar (63.3→66.9; *P*=.001), and math mapping. Morphology showed slight improvement; syntax remained unchanged.
Lawal et al [39]	To evaluate English learning via a digital app	A mobile app named ''Hausar Kurma''	5-7	The study showed that the mobile app significantly improved English learning among Hausa-speaking students with hearing impairment (pretest median = 30; posttest median=90; *z*=2.82; *P*<.05; *r*=0.633; binomial *P*=.01).
Techaraungrong et al [55]	Math learning (counting, addition, and subtraction) with language components through a signed language	Multimedia app for arithmetic learning	7	The study revealed that multimedia-based learning significantly outperformed traditional instruction (*z*=4.545; *P*<.001), improving counting, addition, and subtraction. Students reported better motivation, understanding, and independence, supporting multimedia's effectiveness and the need for inclusive design.

XSL•FO
RenderX

Study	Aim of the study	Name of the platform	Age of the children (years)	Key results
Alsalem and Alzahrani [46]	Language learning (phonemic awareness, vocabulary, writing, and reading comprehension)	Custom-designed e-books	10 - 12	There were statistically significant between-group differences across all outcome measures, with the experimental group demonstrating superior performance in phonemic awareness (t_{39}=11.25; *P*<.001; *d*=1.77), writing (t_{39}=12.18; *P*<.001; *d*=2.03), vocabulary (t_{39}=7.25; *P*<.001; *d*=1.15), and reading comprehension (t_{39}=6.29; *P*<.001; *d*=0.99). All observed effects were large.
Joy et al [31]	To develop and evaluate a mobile-based tool for learning sign vocabulary to improve deaf and Deaf children's language acquisition.	SiLearn	School-age children	The SiLearn program significantly improved vocabulary, with higher gains in the experimental group (mean 21.43, SD 7.45) vs the control group (mean 11.07; $t_{22.91}$=2.926; *P*=.003), confirming its effectiveness.

[a]AAC: augmentative and alternative communication.

[b]TALI: Technology-Assisted Language Intervention.

[c]CELF-5: Clinical Evaluation of Language Fundamentals, Fifth Edition.

[d]TAU: treatment as usual.

[e]VABS: Vineland Adaptive Behavior Scale.

[f]MLU: mean length of utterance.

[g]DHH: deaf and hard of hearing.

[h]AR: augmented reality.

[i]E: efficiency index.

[j]CAPT: computer-aided pronunciation training.

Across the 5 AAC-focused studies, 4 studies [14,30,37,62] reported positive efficacy outcomes in specific language domains, including vocabulary growth, increased utterance length, and improvements in pragmatic communication, while 1 study did not include relevant efficacy data. For instance, in a 24-week study using AAC on iPads, the mean length of utterance increased from 2.41 to 3.68 (*P*=.002), the total words increased by 26.8 (*P*=.001), and different words increased by 8 (*P*=.008) [30]. In the randomized controlled trial of the Technology-Assisted Language Intervention (TALI), children aged ≥5 years showed a 12.7-point increase in Clinical Evaluation of Language Fundamentals, Fifth Edition Pragmatics Profile scores compared to a 6-point decrease in the treatment-as-usual group (*P*=.04) [62]. For children aged <5 years, significant improvements were observed in the Vineland Adaptive Behavior Scale (VABS) Communication (86.7-99.1) and social domain scores (91.8-97.4; *P*=.01). In addition, children in the TALI group demonstrated greater gains in the number of different words spoken (82-104) compared to the treatment-as-usual group, which increased from 82 to 87 [14]. This difference was statistically significant (β=11.04 vs β=2.65; *P*=.007), further emphasizing the more substantial impact of the TALI approach on language development [62]. Preliminary results suggest promise; however, most research designs had small sample sizes, short intervention durations, and heterogeneous outcome measures.

Only 12 studies reported efficacy outcomes; however, they had small sample sizes, short intervention periods, and methodological variability. Thus, the findings should be interpreted as exploratory rather than confirmatory. One study reported on the effectiveness of computer-assisted pronunciation training (CAPT) in supporting speech therapy for children with HIs [65]. Sztahó et al [65] found that, after a 3-month intervention, children using the CAPT app showed significantly improved prosody compared to the control group, with strong agreement between automatic assessments and teachers' ratings. Improvements were most notable in children with mild hearing loss, but were present across all levels [46]. Students in the treatment group showed large improvements in phonemic awareness, writing, vocabulary, and reading (all *P*<.05) with large effect sizes (*d*=0.99 - 2.03) [46].

Several studies (n=5) demonstrated the effectiveness of digital tools in enhancing the acquisition and development of signed languages [31,34,35,39,55]. A mobile app significantly improved signed-language learning compared to traditional methods (*P*=.01; η^2=0.585) because of its engaging, game-based features [34]. Ployjiw et al [35] reported vocabulary gains using

XSL•FO
RenderX

AR books, with scores increasing from 7.20 to 8.70 (P=.01) and instructional efficiency exceeding 85%. High satisfaction was noted, particularly with signed-language videos and parental involvement. The SiLearn app (developed by Joy et al [31]) also showed significant vocabulary improvements, confirming its positive impact. Another study [39] reported a sharp increase in English scores, while Techaraungrong et al [55] found that multimedia instruction outperformed traditional methods in basic math by enhancing motivation and independent learning. One study [32] reported on the efficacy of digital tools in supporting academic outcomes such as language, math, and classroom learning for children with HIs. Significant gains were observed in reading, grammar, and math mapping. Refer to Table 4 for the full results. Although several studies showed statistical significance, the limited number of studies on the efficacy of digital technology makes it difficult to draw conclusions on the comparative effectiveness of the intervention.

Discussion

Principal Findings

In this scoping review, we synthesized evidence regarding how digital technologies were used to support language learning for children with HIs across home, school, and clinical contexts. Of these technologies, mobile apps emerged as the most frequently used tools due to their accessibility, ease of use, and adaptability to different contexts [34,66]. This suggests that the ease of deployment is prioritized over the pedagogical depth and the mechanisms of language learning.

By contrast, other technologies, such as computer-based programs, AR, VR, tangible user interfaces, and social robotics, were implemented less frequently, probably due to higher costs, greater complexity, or limited availability [57,59,67]. The scoping review findings reveal 3 structural patterns in the existing evidence base, which affect the way in which children with HIs acquire languages through the use of digital technologies. These patterns relate to the dominance of mobile technologies, the relative focus on the use of signed-language vs spoken-language interventions, and the use of digital technologies in the classroom, the clinic, and the home.

One of the major findings of this review is that the majority of the mobile apps were created for signed-language learning, with a strong emphasis on enabling children to develop their communication skills through visual language [57,68]. This is also understandable since the relative ease of designing signed-language technologies may also be a factor, since it is much more easily implemented in mobile apps without the need for any complex auditory-processing capabilities. Besides the purpose of signed-language learning, other apps included speech-to-text functionality for general purposes of communication, as well as a variety of assistive learning apps such as literacy, reading, mathematics, and classroom apps [31,35,49].

By contrast, apps focusing on the development of spoken language were relatively limited, particularly for younger children. Existing spoken-language tools typically supported auditory training for children who used hearing aids or cochlear implants by using phoneme recognition, AAC, and interactive exercises to enhance speech, literacy, and pragmatic communication [27,30,48]. While the number of spoken-language apps was relatively small, both types of apps, whether focused on signed or spoken language, shared the same goal of helping children with HIs to communicate more effectively and strengthen their communication skills [30,32]. Despite this imbalance, both sign- and speech-oriented tools shared the overarching goal of improving children's ability to communicate effectively, which is essential for academic achievement, social inclusion, and overall quality of life [69].

There were also a number of studies (n=10) related to the design and development of spoken-language apps that were targeted at young children, which may indicate that there is a relative lack of digital interventions in this area. This has significant analytical connotations, as it has been demonstrated that early childhood is an important period in the development of spoken language, where input is maximally effective [70,71]. There may be several reasons why there is a relative lack of early childhood spoken language tools, including the developmental sensitivity, complexity, and therapy requirements to support early spoken-language development in children with HIs, rather than later ages or signed-language development. Nevertheless, digital interventions for early spoken-language development have the potential to serve as effective and low-cost tools to support therapy in contexts where access to professional support is limited. It is therefore essential to prioritize the development of the body of research on tools to support spoken-language development in association with family-based interventions and therapy in early childhood, as this is the period where interventions are maximally likely to have long-term effects.

An important finding of this study highlighted the value of designing mobile apps that were both interactive and engaging, particularly for children with HIs. Features such as games, modular activities, and visual feedback helped to maintain attention and to support learning [34]. This aligned with theories of multimodal learning, which suggest that combining auditory, visual, and kinesthetic inputs enhances comprehension and retention [72]. Evidence has shown that multimedia-based approaches demonstrated a stronger and more consistent impact on improving math skills compared to conventional teaching methods, as multimedia-based approaches enhanced students' motivation, improved comprehension, and promoted greater independence in learning [55]. Nevertheless, visual design must also take users' comfort and cultural relevance into account. Prolonged screen time may make children with HIs sensitive to harsh colors, such as bright red, which can cause eye strain [35]. In addition, culturally relevant content using familiar visuals, language, and contexts can improve engagement and comprehension [73]. Therefore, while rich interactive features enhance learning, they should be balanced with thoughtful visual and cultural design to ensure accessibility, comfort, and meaningful engagement.

We observed that most of the digital apps remained classroom-oriented and were rarely extended to a therapeutic setting or to home-based practice. With regard to spoken languages, multisensory approaches combining auditory and visual feedback in phoneme training, AAC platforms, or

XSL•FO
RenderX

gamified exercises were common. This approach aligned with evidence showing that pairing auditory training with visual reinforcement strengthened outcomes [74]. Signed-language apps typically emphasized vocabulary building and visual communication, but often relied on drill-based practice, with limited support for family engagement or integration into daily social and classroom interactions. Overall, many of the tools lacked progression, therapist dashboards, and parent-engagement features. Reviews of parent-mediated interventions have raised similar concerns, noting that family involvement and asynchronous support were often underused [75]. This highlights a gap in tools that actively involve families and provide opportunities for sustained, adaptive practice at home.

Most of the available evidence comes from high-income countries. Only a handful of studies examined interventions in low- and middle-income countries, despite these being the settings in which access to therapists and structured educational services is most limited. In these contexts, home-based practice and community-led learning frequently represent the primary pathways for children's communicative development. Addressing this challenge requires adapting and evaluating digital tools in resource-limited settings to ensure that they are relevant, sustainable, and responsive to the needs of both spoken- and signed-language users. Future work should therefore prioritize large-scale trials, long-term follow-up, and the cultural adaptation of digital interventions to ensure sustainability and broader impacts.

In terms of efficacy, AAC interventions demonstrated the strongest and most consistent evidence by improving utterance length, vocabulary, and pragmatic communication [14,37,62]. Signed-language apps, while effective for vocabulary and visual communication, often produced narrower outcomes that were limited to word-level learning [31,76]. Academic-oriented digital tools showed broader benefits in improving reading, writing, and mathematics, although their focus was less specifically on communication outcomes. Furthermore, CAPT interventions provided promising evidence for speech training and prosody, but the findings remain preliminary and were based on small-scale studies. Despite the positive outcomes, the evidence base remains limited. Many studies were short-term, small-scale, or lacked rigorous statistical testing, leaving questions about long-term efficacy and generalizability unanswered [13]. Few tools integrated therapist dashboards, adaptive progression, or structured parental involvement; these features could extend the benefits beyond drill-based practice. This review contributes to the field by providing an understanding of the role of current digital technologies in the facilitation of language learning across spoken-, signed-, and bimodal-language modalities, including the gaps in early spoken-language interventions, parent-mediated interventions, and culturally scalable interventions.

Limitations

This scoping review has several limitations. First, the search was limited to studies published in English, which may have led to the exclusion of relevant studies published in other languages. As a result, the findings may not fully represent global perspectives, particularly in non-English-speaking regions in which important work in this area may take place. Second, while efforts were made to include studies from reputable sources, some relevant interventions were reported in lower-impact journals, which may have affected the overall quality and reliability of the included evidence. Third, restricting the search to studies published after 2014 may have excluded earlier work on computer-assisted language-learning technologies. Nevertheless, the aim was to synthesize evidence relevant to contemporary digital learning environments that are more relevant to current educational and clinical practice. As this study is a scoping review, the purpose was not to evaluate the quality of the included studies; therefore, the results should not be considered indicative of the effectiveness of the interventions. Additionally, the results may have been influenced by publication bias, selective results reporting, and the exclusion of non-English publications, which may have contributed to the lack of representation of certain regions. Moreover, the development of technology may affect the longevity of the tools identified in the study.

Conclusion

This scoping review was intended to synthesize the existing literature concerning the use of digital technology to promote the development of languages among children with HIs in different settings such as the home, school, and clinic. The existing literature concerning the use of digital technology to promote language development among children with HIs indicates that there are 3 patterns that can be used to explain the use of digital technology to promote language development among children with HIs.

Mobile technology was the most widely used digital approach to promote language development among children with HIs due to its accessibility and adaptability across contexts. Among the mobile technologies, the use of signed-language interventions was the most commonly used to promote language development among children with HIs. The use of signed languages was effective in promoting vocabulary development among children with HIs. However, there was no existing literature concerning language development other than vocabulary among children with HIs. On the other hand, spoken-language interventions were less commonly used to promote language development among children with HIs. However, the existing literature concerning the use of AAC was effective in promoting vocabulary and language development among children with HIs in structured contexts and with the assistance of therapists and parents.

Another major gap that we identified is the early development of spoken language. Although it is acknowledged that the period of early childhood is a critical period for the development of language, there are still very few digital interventions for the development of spoken-language skills for young children with HIs. In addition, the interventions that were implemented did not have the essential characteristics of being adaptable, having therapist dashboards, and involving parents.

In conclusion, although digital technology is acknowledged as a potential facilitator of language development for children with HIs, this potential is limited by the imbalance between the 2 modalities and fragmented interventions. Further studies should

XSL•FO
RenderX

aim at the co-design and coevaluation of digital technology for the development of both signed and spoken languages for children with HIs, as well as the facilitation of the continuation of language learning for children both at home and school.

Acknowledgments

The author would like to thank the Swinburne University librarian for assistance with the search strategy. No generative artificial intelligence tools were used in the preparation of this paper.

Funding

No external financial support or grants were received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of the included studies.
[PDF File, 199 KB - rehab_v13i1e85066_app1.pdf]

Checklist 1
PRISMA ScR checklist.
[PDF File, 91 KB - rehab_v13i1e85066_app2.pdf]

References

1. Deafness and hearing loss. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss [accessed 2024-07-04]
2. Lieu JEC, Kenna M, Anne S, Davidson L. Hearing loss in children: a review. JAMA 2020 Dec 1;324(21):2195-2205. [doi: 10.1001/jama.2020.17647] [Medline: 33258894]
3. Smith RJ, Bale JF, White KR. Sensorineural hearing loss in children. The Lancet 2005 Mar;365(9462):879-890. [doi: 10.1016/S0140-6736(05)71047-3]
4. Tye-Murray N. Foundations of Aural Rehabilitation: Children, Adults, and Their Family Members: Plural Pub Inc; 2022.
5. Cano S, Collazos CA, Flórez Aristizábal L, Gonzalez CS, Moreira F. Towards a methodology for user experience assessment of serious games with children with cochlear implants. Telematics and Informatics 2018 Jul;35(4):993-1004. [doi: 10.1016/j.tele.2017.09.011]
6. DeForte S, Sezgin E, Huefner J, et al. Usability of a mobile app for improving literacy in children with hearing impairment: focus group study. JMIR Hum Factors 2020 May 28;7(2):e16310. [doi: 10.2196/16310] [Medline: 32205305]
7. Hermawati S, Pieri K. Assistive technologies for severe and profound hearing loss: beyond hearing aids and implants. Assist Technol 2020 Jul 3;32(4):182-193. [doi: 10.1080/10400435.2018.1522524] [Medline: 30652936]
8. Pratomo HTA, Adriani RB, Akhyar M. Association between parental education, occupation, income, language activity, and language proficiency in children. Indones J Med 2016;01(3):152-159. [doi: 10.26911/theijmed.2016.01.03.02]
9. Wahyuni T, Wahono S, Muryati S, Sari N. Speech therapy media modeling through contextual approach for Indonesian language and literature department students in Surakarta, Central Java. In: Sari N, editor. Presented at: 2nd Workshop on Language, Literature and Society for Education; Dec 21-22, 2018. [doi: 10.4108/eai.21-12-2018.2282769]
10. Walker EA, Holte L, McCreery RW, Spratford M, Page T, Moeller MP. The influence of hearing aid use on outcomes of children with mild hearing loss. J Speech Lang Hear Res 2015 Oct;58(5):1611-1625. [doi: 10.1044/2015_JSLHR-H-15-0043] [Medline: 26151927]
11. Lightbown PM, Spada N. How Languages Are Learned, 5th edition: Oxford University Press (OUP); 2021.
12. Owens Jr RE. Language Development: An Introduction: Pearson; 2000.
13. Herrera SI, Manresa-Yee C, Sanz CV. Mobile learning for hearing-impaired children: review and analysis. Univ Access Inf Soc 2023 Jun;22(2):635-653. [doi: 10.1007/s10209-021-00841-z]
14. Meinzen-Derr J, Sheldon RM, Henry S, et al. Enhancing language in children who are deaf/hard-of-hearing using augmentative and alternative communication technology strategies. Int J Pediatr Otorhinolaryngol 2019 Oct;125:23-31. [doi: 10.1016/j.ijporl.2019.06.015] [Medline: 31238158]
15. Vickers D, Salorio-Corbetto M, Driver S, et al. Involving children and teenagers with bilateral cochlear implants in the design of the BEARS (Both EARS) virtual reality training suite improves personalization. Front Digit Health 2021;3:759723. [doi: 10.3389/fdgth.2021.759723] [Medline: 34870270]

XSL•FO
RenderX

16. Véliz S, Espinoza V, Sauvalle I, Arroyo R, Pizarro M, Garolera M. Towards a participative approach for adapting multimodal digital books for deaf and hard of hearing people. Int J Child Comput Interact 2017 Jan;11:90-98. [doi: 10.1016/j.ijcci.2016.10.003]
17. Estabrooks W, MacIver-Lux K, Rhoades E. Auditory-Verbal Therapy: For Young Children with Hearing Loss and Their Families, and the Practitioners Who Guide Them: Plural Publishing; 2016. URL: https://books.google.com.au/books/about/Auditory_Verbal_Therapy.html?id=g89xDgAAQBAJ&redir_esc=y [accessed 2026-05-26]
18. Anusha MN, Sathveeka KS, Maria J, Sangeetha G, Saranyaa G. Communication challenges and implementation of telepractice for children with hearing impairment during lockdown- a parental perspective. JCDR 2021;15(4):MC06-MC11. [doi: 10.7860/JCDR/2021/47439.14812]
19. Frisby C, Eikelboom RH, Mahomed-Asmail F, et al. Community-based adult hearing care provided by community healthcare workers using mHealth technologies. Glob Health Action 2022 Dec 31;15(1):2095784. [doi: 10.1080/16549716.2022.2095784] [Medline: 35960191]
20. Hajesmaeel-Gohari S, Goharinejad S, Shafiei E, Bahaadinbeigy K. Digital games for rehabilitation of speech disorders: a scoping review. Health Sci Rep 2023 Jun;6(6):e1308. [doi: 10.1002/hsr2.1308] [Medline: 37283880]
21. Nanjundaswamy M, Prabhu P, Rajanna RK, Ningegowda RG, Sharma M. Computer-based auditory training programs for children with hearing impairment - a scoping review. Int Arch Otorhinolaryngol 2018 Jan;22(1):88-93. [doi: 10.1055/s-0037-1602797] [Medline: 29371904]
22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
23. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
24. Pollock D, Peters MDJ, Khalil H, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth 2023 Mar 1;21(3):520-532. [doi: 10.11124/JBIES-22-00123] [Medline: 36081365]
25. Homepage. Covidence. 2023. URL: https://www.covidence.org [accessed 2026-05-26]
26. Smith A. Smartphone ownership 2013. Pew Research Center. 2013. URL: https://www.pewresearch.org/internet/2013/06/05/smartphone-ownership-2013/ [accessed 2026-05-26]
27. Firdous S, Wahid M, Ud A, et al. Android based receptive language tracking tool for toddlers. Int J Adv Comput Sci Appl 2019;10(3):589-595. [doi: 10.14569/IJACSA.2019.0100375]
28. Novi Dayawati R, Dwi Sulistiyo M, Mustika Rani M, Maulidina Nistia R, Noor Linda D, Widyastuti Suwarsono L. The implementation of A-TooLips, a learning mobile application for deaf children to produce words. Jurnal Teknologi 2016;78(5):405-411. [doi: 10.11113/jt.v78.8345]
29. Alnasif H, Alyahya LA, Alromaih H, Alhelal G, Barnawi N, Altamim A. Mushir: an Arabic edutainment application for deaf and hard of hearing children using real-time image processing. IJCSNS 2022;22(9) [FREE Full text]
30. Mood D, Sheldon R, Tabangin M, Wiley S, Meinzen-Derr J. Technology assisted language intervention (TALI) for children who are deaf/hard of hearing: promising impact on pragmatic skills. Deafness Educ Int 2022;24(4):334-355. [doi: 10.1080/14643154.2022.2135731] [Medline: 37304207]
31. Joy J, Balakrishnan K, M. S. SiLearn: an intelligent sign vocabulary learning tool. JET 2019 Aug 14;13(3):173-187. [doi: 10.1108/JET-03-2019-0014]
32. Vasel H, Ragonis N. Empowering hearing‐impaired students: a mobile learning intervention in Israeli Arab elementary education. Research in Spec Educ Needs 2024 Oct;24(4):1093-1104. [doi: 10.1111/1471-3802.12697]
33. Navarro-Newball AA, Loaiza D, Oviedo C, et al. Talking to Teo: video game supported speech therapy. Entertain Comput 2014 Dec;5(4):401-412. [doi: 10.1016/j.entcom.2014.10.005]
34. Mohammad H, Tamimi H, Abu-Amara F. An educational Arabic sign language mobile application for children with hearing impairment. Int J Interact Mob Technol 2022;16(20):114-129. [doi: 10.3991/ijim.v16i20.32427]
35. Ployjiw U, Michel PC. Development of augmented reality learning materials for the hearing impaired students in primary I. Int J Inf Educ Technol 2023;13(11):1696-1703. [doi: 10.18178/ijiet.2023.13.11.1978]
36. Chan GL, Santally MI, Whitehead J. Gamification as technology enabler in SEN and DHH education. Educ Inf Technol (Dordr) 2022;27(7):9031-9064. [doi: 10.1007/s10639-022-10984-y] [Medline: 35345601]
37. Meinzen-Derr J, Wiley S, McAuley R, Smith L, Grether S. Technology-assisted language intervention for children who are deaf or hard-of-hearing; a pilot study of augmentative and alternative communication for enhancing language development. Disabil Rehabil: Assist Technol 2017 Nov;12(8):808-815. [doi: 10.1080/17483107.2016.1269210] [Medline: 27982714]
38. Shoaib L, Khan S, Abbas MA, Salman A. Enabling profound hearing impaired children to articulate words using lip-reading through software application. J Pak Med Assoc 2018 Mar;68(3):432-436. [Medline: 29540880]
39. Lawal A, Cavus N, Lawan AA, Sani I. Hausar Kurma: development and evaluation of interactive mobile app for the English-Hausa sign language alphabet. IEEE Access 2024;12:46012-46023. [doi: 10.1109/ACCESS.2024.3381538]
40. Yousaf K, Mehmood Z, Saba T, et al. A novel technique for speech recognition and visualization based mobile application to support two‐way communication between deaf‐mute and normal peoples. Wirel Commun Mob Comput 2018 Jan;2018(1):129762331. [doi: 10.1155/2018/1013234]

XSL•FO
RenderX

41. Levina E, Levin S, Koroleva I, Kuzovkov V, Shashukova E, Shaporova A. Rehabilitation of patients with mondini malformation and auditory neuropathy after cochlear implantation. Presented at: 15th International Conference on Cochlear Implants and Other Implantable Auditory Technologies; Jun 27-30, 2018.
42. Goker H, Ozaydin L, Tekedere H. The Effectiveness and Usability of the Educational Software on Concept Education for Young Children with Impaired Hearing. EURASIA Journal of Mathematics, Science and Technology Education 2016;12(1). [doi: 10.12973/eurasia.2016.1207a]
43. Imashev A, Kydyrbekova A, Oralbayeva N, Kenzhekhan A, Sandygulova A. Learning sign language with mixed reality applications - the exploratory case study with deaf students. Educ Inf Technol 2024 Sep;29(13):17261-17292. [doi: 10.1007/s10639-024-12525-1]
44. Bouzid Y, Khenissi MA, Essalmi F, Jemni M. Using Educational Games for Sign Language Learning -A SignWriting Learning Game: Case Study. Educational Technology & Society 2016 Jan:129-141 [FREE Full text]
45. Beisov N, Madyarova G, Kerimbayev N. Gesture recognition technology: a new dimension in human-computer interaction interface. IJEECS 2024;35(2):1311. [doi: 10.11591/ijeecs.v35.i2.pp1311-1324]
46. Alsalem MA, Alzahrani HA. In search of language development for students who are hard of hearing: measuring the effectiveness of assistive technologies through teaching practices. Assist Technol 2024 Jan 2;36(1):3-10. [doi: 10.1080/10400435.2023.2193761] [Medline: 36996032]
47. Flórez-Aristizábal L, Cano S, Collazos CA, Solano AF. DesignABILITY: framework for the design of accessible interactive tools to support teaching to children with disabilities. In: Brewster S, editor. 2019 Presented at: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019. [doi: 10.1145/3290605.3300240]
48. Brouwer K, Downing H, Westhoff S, et al. Effects of clinician-guided emergent literacy intervention using interactive tablet technology for preschool children with cochlear implants. Commun Disord Q 2017 Aug;38(4):195-205. [doi: 10.1177/1525740116666040]
49. Herzig M, Allen TE. Deaf children's engagement with American sign language-english bilingual storybook apps. J Deaf Stud Deaf Educ 2022 Dec 23;28(1):53-67. [doi: 10.1093/deafed/enac032] [Medline: 36223888]
50. Majorano M, Santangelo M, Redondi I, et al. The use of a computer-based program focused on the syllabic method to support early literacy in children with cochlear implants. Int J Pediatr Otorhinolaryngol 2024 Aug;183:112048. [doi: 10.1016/j.ijporl.2024.112048] [Medline: 39068706]
51. Mohamad N, Hashim NL. UX Testing for Mobile Learning Applications of Deaf Children. International Journal of Advanced Computer Science and Applications 2021;12(11). [doi: 10.14569/IJACSA.2021.0121134]
52. Yaman F, Dönmez O, Avcı E, Kabakçı Yurdakul I. Integrating Mobile Applications into Hearing Impaired Children's Literacy Instuction. Education and Science 2016;41(188):153-174. [doi: 10.15390/EB.2016.6687]
53. Parvez K, Khan M, Iqbal J, et al. Measuring effectiveness of mobile application in learning basic mathematical concepts using sign language. Sustainability 2019;11(11):3064. [doi: 10.3390/su11113064]
54. Suarsana IM. Developing interactive digital mathematics book with multi representation approach for deaf students. Int J Emerg Technol Learn 2021;16(13):128. [doi: 10.3991/ijet.v16i13.22459]
55. Techaraungrong P, Suksakulchai S, Kaewprapan W, Murphy E. The design and testing of multimedia for teaching arithmetic to deaf learners. Educ Inf Technol 2017 Jan;22(1):215-237. [doi: 10.1007/s10639-015-9441-1]
56. Shelton BE, Parlin MA. Teaching math to deaf/hard-of-hearing (DHH) children using mobile games: outcomes with student and teacher perspectives. Int J Mob Blended Learn 2016;8(1):1-17. [doi: 10.4018/IJMBL.2016010101]
57. Lekova AK, Tsvetkova PT, Tanev TK. Robot-assisted psychosocial techniques for language learning by hearing-impaired children. Int J Inf Technol Secur 2021;13:63-76 [FREE Full text]
58. Mohd Suhaimi NR, Megat Mohd Zainuddin N, Ahmad H, Maarop N, Ibrahim R, Wan Hassan WA. Mobile application dictionary for hearing impaired students. Int J Innov Comp 2019;9(2). [doi: 10.11113/ijic.v9n2.241]
59. Czap L, Pintér JM, Baksa-Varga E. Features and results of a speech improvement experiment on hard of hearing children. Speech Commun 2019 Jan;106:7-20. [doi: 10.1016/j.specom.2018.11.003]
60. Harisman Y, Dwina F, Longgom MN, Amiruddin MH, Syaputra H. The development of proton-electron math e-comic to improve special needs students' mathematical concepts understanding. IJ 2023;12(2):359-376 [FREE Full text] [doi: 10.22460/infinity.v12i2.p359-376]
61. Zainuddin NMM, Maarop N, Hassan WAW. Measuring satisfaction on augmented reality courseware for hearing-impaired students: adjustment formula form SUS. Asian J Univ Educ 2022 Apr 30;18(2):348-360. [doi: 10.24191/ajue.v18i2.17990]
62. Meinzen-Derr J, Sheldon R, Altaye M, Lane L, Mays L, Wiley S. A technology-assisted language intervention for children who are deaf or hard of hearing: a randomized clinical trial. Pediatrics 2021 Feb;147(2):e2020025734. [doi: 10.1542/peds.2020-025734] [Medline: 33452063]
63. King M, Romski M, Sevcik RA. Growing up with AAC in the digital age: a longitudinal profile of communication across contexts from toddler to teen. Augment Altern Commun 2020 Apr 2;36(2):128-141. [doi: 10.1080/07434618.2020.1782988]
64. Cano S, Flórez-Aristizábal L, Collazos CA, Fardaoun HM, Alghazzawi DM. Designing interactive experiences for children with cochlear implant. Sensors (Basel) 2018 Jul 4;18(7):2154. [doi: 10.3390/s18072154] [Medline: 29973549]
65. Sztahó D, Kiss G, Vicsi K. Computer based speech prosody teaching system. Comput Speech Lang 2018 Jul;50:126-140. [doi: 10.1016/j.csl.2017.12.010]

XSL•FO
RenderX

66. Hanafi Y, Hendrawan HJ, Hakim IN. Accelerating QURán reading fluency through learning using QURÁNI application for students with hearing impairments. Int J Emerg Technol Learn 2019;14(6):110. [doi: 10.3991/ijet.v14i06.9863]
67. Uluer P, Akalin N, Gurpinar C, Kose H. Interactive games with an assistive robotic system for hearing-impaired children. Stud Health Technol Inform 2017;14(6):298-305. [doi: 10.3233/978-1-61499-798-6-298] [Medline: 28873814]
68. Rocha DFS, Bittencourt II, de Amorim Silva R, Ospina PLE. An assistive technology based on Peirce's semiotics for the inclusive education of deaf and hearing children. Univ Access Inf Soc 2023 Nov;22(4):1097-1116. [doi: 10.1007/s10209-022-00919-2]
69. Roland L, Fischer C, Tran K, et al. Quality of life in children with hearing impairment: systematic review and meta-analysis. Otolaryngol Head Neck Surg 2016 Aug;155(2):208-219. [doi: 10.1177/0194599816640485] [Medline: 27118820]
70. Kuhl PK. Early language acquisition: cracking the speech code. Nat Rev Neurosci 2004 Nov;5(11):831-843. [doi: 10.1038/nrn1533] [Medline: 15496861]
71. Porcar-Gozalbo N, López-Zamora M, Valles-González B, Cano-Villagrasa A. Impact of hearing loss type on linguistic development in children: a cross-sectional study. Audiol Res 2024 Nov 27;14(6):1014-1027. [doi: 10.3390/audiolres14060084] [Medline: 39727607]
72. Mayer RE. The Cambridge Handbook of Multimedia Learning: Cambridge University Press; 2014.
73. Rau PLP. Cross-Cultural Design: 8th International Conference, CCD 2016, Held as Part of HCI International 2016, 1st edition: Springer International Publishing; 2016.
74. Stropahl M, Besser J, Launer S. Auditory training supports auditory rehabilitation: a state-of-the-art review. Ear Hear 2020;41(4):697-704. [doi: 10.1097/AUD.0000000000000806] [Medline: 31613823]
75. Zhao J, Neto I, Pires AC, Tomé-Pires C, Nicolau H. Digital technologies for deaf and hard of hearing children: a systematic review, critical reflections, and future research directions. 2025 Apr 26 Presented at: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; Apr 26 to May 1, 2025. [doi: 10.1145/3706598.3714241]
76. Joy J, Balakrishnan K, Madhavankutty S. Developing a bilingual mobile dictionary for Indian Sign Language and gathering users experience with SignDict. Assist Technol 2020 May 3;32(3):153-160. [doi: 10.1080/10400435.2018.1508093] [Medline: 30193075]

Abbreviations

AAC: augmentative and alternative communication
AR: augmented reality
CAPT: computer-assisted pronunciation training
HI: hearing impairment
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analysis Extension for Scoping Reviews
RQ: research question
TALI: Technology-Assisted Language Intervention
VR: virtual reality
XR: extended reality

Edited by S Munce; submitted 06.Oct.2025; peer-reviewed by H Gurteen, P Sooful; revised version received 12.Mar.2026; accepted 09.Apr.2026; published 16.Jun.2026.

Please cite as:
Sari AD, Al Mahmud A, Octavia JR, Tyagi S
Digital Technologies for Children With Hearing Impairments to Support Language Learning: Scoping Review
JMIR Rehabil Assist Technol 2026;13:e85066
URL: https://rehab.jmir.org/2026/1/e85066
doi:10.2196/85066

© Amarria Dila Sari, Abdullah Al Mahmud, Johanna Renny Octavia, Shivani Tyagi. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Ecological Momentary Assessment of Daily Activities in Individuals With Neurological Disorders: Scoping Review

Eva Delooz[1], MSc; Barbara Piškur[1,2], PhD; Aki Rintala[1,3], PhD; Bruno Bonnechère[1,4], PhD; Annemie Spooren[1,4], PhD

[1]Faculty of Rehabilitation Sciences, REVAL Rehabilitation Research Center, Hasselt University, Diepenbeek, Flanders, Belgium

[2]Research Centre for Autonomy and Participation, Zuyd University of Applied Sciences, Heerlen, Limburg, The Netherlands

[3]Physical Activity & Functional Capacity Research Group, LAB University of Applied Sciences, Lahti, Paijat-Hame, Finland

[4]Centre of Expertise in Care Innovation, Department of PXL-Healthcare, PXL University of Applied Sciences and Arts, Hasselt, Flanders, Belgium

Corresponding Author:
Eva Delooz, MSc
Faculty of Rehabilitation Sciences
REVAL Rehabilitation Research Center
Hasselt University
Wetenschapspark 7
Diepenbeek, Flanders, 3590
Belgium
Phone: 32 476377464
Email: eva.delooz@uhasselt.be

Abstract

Background: Daily activities shape individuals' health and well-being, reflecting functioning and lived health. For people with neurological conditions, these activities are often disrupted, affecting autonomy and quality of life. Traditional assessments miss subtle, real-time fluctuations, whereas ecological momentary assessment (EMA) captures moment-to-moment activity within natural contexts, offering insight into person-environment-occupation interactions. Despite its growing use, it remains unclear how EMA protocols conceptualize daily activities and integrate person-environment-occupation dimensions in applications for neurological populations.

Objective: This scoping review aims to map the existing literature on the use of EMA to capture daily activities, ranging from basic self-care to more complex activities, in individuals with neurological disorders.

Methods: A scoping review was conducted to map studies using EMA to capture daily activities in adults with neurological conditions, with a specific focus on content and practical application. A total of 341 articles were identified.

Results: A total of 20 studies using EMA to assess daily activities in neurological populations were included; most were observational, with 2 longitudinal studies and 2 randomized controlled trials. Daily activity questions and response formats varied, often using multiple-choice lists; only 1 study allowed open-ended responses. In addition to daily activity questions, EMA captured person (physical, affective, and cognitive), environment (physical and social), and occupation domains, as well as motivation and EMA disturbance. Protocols differed in setting, schedule, technology, and adherence, with most reporting completion rates above 70%.

Conclusions: EMA captures daily activities in neurological populations and demonstrates generally high adherence despite variability in study designs, questions, and technologies. The findings suggest that the phrasing of EMA items, the predominance of closed-response formats, and the narrow focus on the verb "doing" limit the depth and nuance of the data collected, often overlooking important aspects of performance, engagement, or both in daily activities.

(JMIR Rehabil Assist Technol 2026;13:e92797) doi:10.2196/92797

KEYWORDS

ecological momentary assessment; EMA; daily activities; neurological disorders; scoping review

XSL•FO
RenderX

Introduction

Daily life is made up of countless activities that shape how individuals experience their health and well-being. Understanding how people engage in these daily activities offers important insights into the broader concept of functioning, which has gained recognition as a key link between medical conditions and the lived experience of health [1]. The World Health Organization's International Classification of Functioning, Disability and Health (ICF) emphasizes functioning not merely as the absence of disease, but as lived health: the capacity to carry out activities and participate in life situations within real-world contexts [2]. This perspective shifts attention toward what individuals are actually able to do in their environments and how they engage with the world around them. However, while the ICF offers a valuable conceptual framework, it provides limited guidance on how personal, environmental, and activity-related factors interact dynamically in daily life. To address this gap, the Person-Environment-Occupation (PEO) model has been widely applied in occupational science and rehabilitation research [3-5]. This model highlights the complex and evolving relationship among the person, environment, and occupation. Meaningful engagement in daily activities arises at the intersection of these 3 domains [4-8]. Daily activities are context- and role-specific, ranging from basic self-care to more complex activities such as work, leisure, and caregiving. They are central to human identity, autonomy, and connectedness [7,9]. Through these activities, individuals express preferences, fulfill social roles, and structure their everyday lives, making engagement in daily activities a vital expression of functioning and lived health. In this review, the term "daily activities" refers to observable actions and tasks performed in everyday life [10]. Although related, this construct is distinct from broader concepts such as participation (involvement in life situations, as defined in the ICF) and occupation (as conceptualized within the PEO model) [2,4]. To ensure conceptual clarity, these terms are used consistently and are not treated as interchangeable.

However, for individuals living with neurological disorders, engagement in daily activities is frequently challenged or disrupted. Neurological disorders encompass a wide range of conditions, including multiple sclerosis (MS), spinal cord injury, stroke, traumatic brain injury, and neurodegenerative disorders such as Parkinson and Alzheimer disease [11]. Recent global data underscore the magnitude of this issue. A major study published in The Lancet Neurology in collaboration with the World Health Organization, drawing on the Global Burden of Disease 2021 dataset, revealed that more than 3 billion people worldwide were living with a neurological disorder in 2021 [12]. Neurological disorders have become the leading cause of disability worldwide. Since 1990, the burden associated with these disorders, measured in disability-adjusted life years—which combine years lost due to ill health, disability, or premature death—has increased by 18%, highlighting their profound impact on functional independence and quality of life [12]. While clinical symptoms such as motor deficits, cognitive decline, and sensory impairments are well documented, the real-world consequences of these disorders, particularly their effect on engagement in daily activities, remain a critical yet underexplored area in both research and clinical practice. For example, motor impairments, such as paralysis and loss of mobility, can disrupt basic self-care activities [13]. A similar pattern is observed in MS, where 42% of individuals reported a decline in their ability to manage daily life over a 2-year period [14]. These conditions differ substantially in symptom presentation, disease progression, and associated functional limitations, influencing the nature of engagement in daily activities.

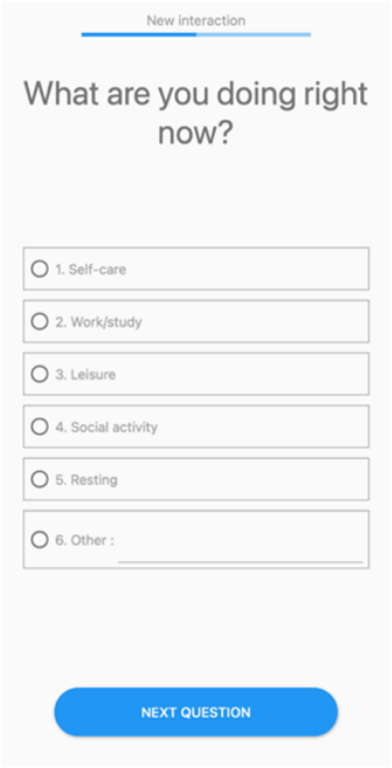
To effectively support individuals living with neurological disorders, it is essential to gain a nuanced understanding of how daily activities are experienced and performed. To this end, a range of established assessments and observational tools are used to examine the factors that influence engagement in daily activities, namely, the dynamic interaction among the person, the environment, and the activity itself [15]. Within the ICF, a distinction is made between capacity—what a person is able to do in a standardized or controlled environment—and performance—what a person actually does in their real-life context [2]. Assessments at the activity level, such as the Canadian Occupational Performance Measure [16], the Functional Independence Measure, and the Barthel Index [17], primarily address this domain, often with a focus on performance in everyday tasks. Assessments at the participation level of the ICF—referring to a person's involvement in everyday situations and society—are also limited for individuals with neurological disorders. In addition, diagnosis-specific instruments, such as the Parkinson's Disease Questionnaire, the Barthel Index, and the Quadriplegia Index of Functioning, are frequently used to capture functional status [2]. However, most of these tools rely on retrospective self-reports and structured observations, which have limited capacity to capture the subtle, moment-to-moment fluctuations that shape everyday functioning [18,19]. As they reflect past experiences rather than real-time activity, they are vulnerable to memory bias, social desirability effects, and inaccurate recall [20,21]. As a result, these methods often fail to detect microprocesses—small but meaningful behavioral and cognitive patterns—that play a crucial role in shaping engagement in daily activities [8,20]. One of the main challenges is the lack of assessments that directly measure how people perform their daily activities. In recent years, advances in digital health technologies have enabled more objective approaches to capturing aspects of daily behavior, such as mobility tracking, sensor-based monitoring, and smartphone interaction data [22]. While these methods provide valuable, continuous measurements of observable behavior, they offer limited insight into the subjective experience, meaning, and contextual factors that shape engagement in daily activities [9,20,22]. As a result, they may not fully capture how individuals perceive, interpret, and engage in their daily lives.

To overcome the challenge of accurately capturing daily activities, ecological momentary assessment (EMA) has emerged as a promising approach. EMA typically involves repeated assessments administered multiple times per day over several days or weeks, often delivered via smartphone-based apps or other digital devices. Participants receive prompts at fixed or random intervals and are asked to report on their current experiences, behaviors, or context in real time. This approach

XSL•FO
RenderX

minimizes recall bias and enhances ecological validity by capturing data in natural environments as events unfold. By collecting real-time data on behavior, emotions, and contextual factors in people's natural environments through regular, random prompts covering various topics, EMA offers a dynamic way to understand daily life as it unfolds [20,23,24]. Also referred to as ambulatory assessment or experience sampling methodology, EMA has been widely adopted across the health sciences, psychology, and clinical research. Its applications range from monitoring symptoms and evaluating treatments to designing personalized interventions, such as Just-In-Time Adaptive Interventions and Ecological Momentary Interventions, for individuals with and without disabilities [25-30]. To provide a concrete illustration of how EMA is experienced by participants, an example of an EMA prompt is presented in Figure 1. The illustration was created using the m-Path platform and serves as an example only; it was not derived from an included study. The potential of EMA to provide insight into daily activities was recently highlighted in a bibliometric review, which demonstrated that EMA has been applied across a wide range of domains, including neurological disorders. Notably, individuals with neurological disorders represented the third largest population in which daily activities were assessed using EMA [31]. However, while this review provided an overview of the fields in which EMA has been applied, it did not examine how daily activities are specifically measured or how daily activity–related questions, arising from the dynamic interaction among the person, occupation, and environment, are incorporated into EMA protocols for individuals with neurological disorders [31].

Figure 1. Example of an ecological momentary assessment question.



Note: The illustration was created using the m-Path platform and serves as an example only.

In addition to understanding what is measured, it is equally important to consider how EMA protocols that include questions about daily activities are conducted. Key design characteristics of EMA protocols, such as the sampling schedule, number of assessments, delivery tools (eg, smartphone apps), and participant adherence, play a crucial role in determining feasibility, participant burden, and data quality [20,21,24,32]. For example, a higher sampling frequency may increase ecological validity but also participant burden, potentially reducing adherence [32]. Despite the growing use of EMA in health research, there is currently no clear consensus on how these design characteristics should be optimized when assessing daily activities, particularly in populations with neurological disorders. As a result, substantial variability exists in how EMA protocols are applied, which may influence both the feasibility of implementation and the validity of the data collected.

Therefore, an analysis is needed to understand which aspects of daily activities are captured in EMA studies, particularly in relation to the PEO model. For example, it remains unclear whether current EMA methodologies adequately capture person-, environment-, and occupation-related aspects of daily activities. In addition, it is important to understand how EMA is implemented, including its key design characteristics. Therefore, this scoping review aims to map the existing literature on the use of EMA to capture daily activities, ranging from

XSL•FO
RenderX

basic self-care to more complex activities, in individuals with neurological disorders. Specifically, this review addresses 2 key dimensions:

- Content-related aspects (WHAT): identifying EMA items that directly assess daily activities and analyzing how additional EMA questions relate to daily activities through the lens of the PEO model; and
- Practical application (HOW): examining the key design characteristics of EMA protocols and their potential implications for feasibility and data quality.

Methods

Design

A scoping review was conducted in accordance with the framework of Arksey and O'Malley [33] and reported following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews; see Multimedia Appendix 1) guidelines to map the existing literature on how EMA is used to capture daily activities in individuals with neurological disorders [34].

Search Strategy

The search focused on neurological disorders and included the most prevalent conditions identified in Feigin et al [11]. This restriction was applied to ensure methodological coherence and feasibility. To ensure a comprehensive search, we combined the MeSH term "Nervous System Diseases" with a selection of specific neurological conditions. The MeSH term was used in its exploded form, thereby capturing a wide range of related disorders within the hierarchical MeSH structure (eg, cerebrovascular diseases and spinal cord diseases), including conditions not explicitly listed as separate search terms. Additional specific conditions (eg, tetanus, migraine, and brain neoplasms) were included to increase sensitivity, account for potential variability in indexing and terminology across databases, and capture studies not yet indexed with MeSH terms. This set of disorders was combined with terms related to daily activities and EMA. Daily activities were operationalized broadly to capture activities across different contexts and roles, ranging from basic self-care to more complex activities such as work, leisure, productivity, and caregiving. This approach aligns with an occupational perspective in which daily activities are understood as emerging from the dynamic interaction among the person, environment, and occupations performed, as described in the PEO model [3,4]. EMA was defined as a real-time assessment method, also referred to as ambulatory assessment or experience sampling methodology, designed to capture experiences, behaviors, and activities as they occur in everyday life [20,21,23]. The complete search string is presented in Multimedia Appendix 2.

The search was performed on February 27, 2025, across the following electronic databases: MEDLINE, Scopus, Embase, and Web of Science. These databases were selected to ensure broad coverage of biomedical and interdisciplinary research. The search strategies presented in Multimedia Appendix 2 were adapted to the indexing systems and technical requirements of each database. Although minor adjustments in syntax were made, the underlying search concepts and structure remained consistent across databases. The initial search yielded 341 records. After duplicates were removed, 207 unique articles remained. Following the screening of these 207 articles, an additional search for gray literature was performed using DANS Data Station Social Sciences and Humanities, DANS Data Station Life Sciences, and ProQuest Dissertations & Theses. However, this gray literature search did not identify any additional relevant records.

Data Screening

Following the search strategy, the screening process commenced. An initial screening of titles and abstracts was performed using the following inclusion criteria: (1) studies that used EMA, experience sampling method, or ambulatory assessment, or that met the core characteristics of EMA methodology. These characteristics included repeated data collection in real-world environments using prompts delivered at fixed or random intervals, with participants reporting on their current or very recent experiences, behaviors, or context. Given the variability in terminology, studies were not excluded solely based on the terminology used if their methodological characteristics aligned with EMA. (2) Neurological populations defined according to the classification of neurological disorders described in the Global Burden of Disease framework. Studies were included if the sample consisted predominantly of individuals with a neurological condition. Studies including mixed samples were considered only if data from participants with neurological conditions could be clearly identified. (3) At least one item related to daily activities in the EMA protocol, defined as questions explicitly referring to participants' engagement in everyday activities or tasks (eg, what the participant was doing at the moment or during a recent period). Only studies including at least one such item within the EMA protocol were retained. (4) Full-text research reports, regardless of publication type (eg, peer-reviewed articles or gray literature such as theses). The exclusion criteria were (1) poster or conference abstracts and (2) mHealth apps that did not implement EMA methodology.

The screening was performed independently by 2 reviewers. Two members of the research group each served as the first reviewer, while the lead researcher (ED) served as the second reviewer. Discrepancies were discussed during regular meetings until consensus was reached. After the initial title and abstract screening, 159 articles were excluded. Subsequently, a full-text screening was conducted, again using independent duplicate review. After this stage, an additional 28 articles were excluded. The final selection comprised 20 articles included for data analysis. The entire screening process was managed using Rayyan software.

Data Extraction and Analysis

Following screening, the included articles were systematically analyzed based on their full texts. Data analysis was conducted using structured extraction tables developed for this review. Data extraction and coding were performed independently by 2 reviewers, following the same approach used during screening. Any uncertainties or discrepancies regarding extracted variables (eg, PEO categorization, setting, or study type) were discussed

XSL•FO
RenderX

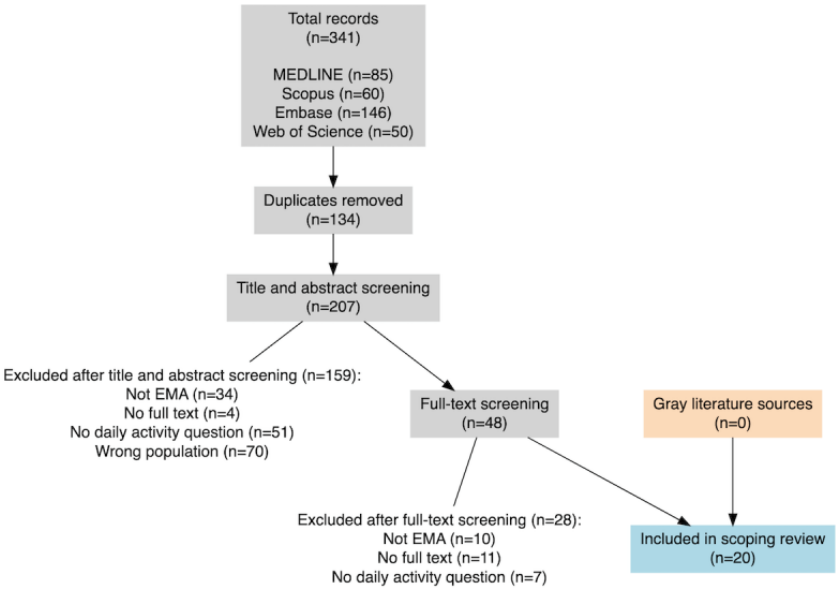
during regular meetings until consensus was reached. Three main domains were extracted to address the research question. First, general study characteristics were recorded for each included study, including article information, setting, study design, and population. Second, content-related aspects focused on how daily activities were assessed using EMA. All EMA items directly assessing daily activities were extracted, together with their response formats. In addition, supplementary EMA items were coded according to the components of the PEO model, distinguishing among person-, environment-, and occupation-related constructs. The categorization of EMA variables into the PEO domains was guided by the theoretical definitions of the PEO model [4]. Each item was assigned to the domain that best reflected its primary focus, based on predefined subdimensions (eg, physical, affective, and cognitive aspects for the person domain). Third, practical application aspects addressed key design characteristics of EMA protocols and their potential implications for feasibility and data quality. These data were summarized in a separate extraction table and included the EMA schedule, number and timing of assessments, and reported adherence rates. To explore broader trends in EMA design characteristics, exploratory simple linear regression analyses were conducted using RStudio (version 2024.12.0+467; Posit PBC). Three simple linear regression analyses were performed: (1) regressing the number of assessments per day on study duration (days), (2) regressing adherence on the total number of EMA assessments, and (3) regressing adherence on EMA schedule. Given the limited number of included studies and incomplete reporting of adherence data, the regression analyses were conducted in an exploratory manner to identify potential trends rather than establish confirmatory associations. The analyses were not weighted by sample size. For visualization purposes, study sample size was represented by point size. Model assumptions (eg, linearity, normality of residuals, and homoscedasticity) were assessed by visual inspection. Studies reporting nonfixed ranges for predictor variables were excluded from the relevant regression analyses to avoid arbitrary estimates.

Results

General Information of Included Articles

In total, 20 articles on the use of EMA to assess daily activities in populations with neurological disorders were included, representing 1128 participants (Figure 2). However, this number should be interpreted with caution, as several publications appeared to originate from overlapping datasets or closely related participant cohorts. Consequently, the reported participant total likely overestimates the number of unique individuals represented across studies. This overlap should also be considered when interpreting patterns in EMA protocol characteristics and adherence. The populations studied included individuals with stroke (n=11), acquired brain injury (n=3), traumatic brain injury (n=2), Parkinson disease (n=1), tinnitus (n=1), hereditary spastic paraplegia (n=1), and MS (n=1). Sample sizes varied considerably across studies, ranging from 7 to 212 participants. The majority of the included studies were observational, alongside 2 longitudinal studies and 2 randomized controlled trials. The studies were published between 2012 and 2024. Most studies (15/20) were conducted primarily in a home setting, often combined with brief laboratory visits. Three studies were conducted entirely in the home environment, 3 exclusively in community settings, 1 exclusively in a hospital, and 1 across both home and hospital settings. An overview of the included studies is presented in Table 1.

Figure 2. Study flowchart.



XSL•FO
RenderX

Table 1. General information about the included studies.

Study	Setting[a]	Study type	Population
Blome et al [35]	Home	Observational (qualitative and exploratory)	People with multiple sclerosis (n=17)
Broen et al [36]	Home	Observational (feasibility study)	Parkinson disease with motor fluctuations (n not stated)
Bui et al [37]	Home + laboratory visits	Observational	Mild-to-moderate stroke (n=202)
Bui et al [38]	Community + laboratory visits	Longitudinal observational	Mild-to-moderate stroke (n=212)
Ezekiel et al [39]	Community (not explicitly stated)	Observational (qualitative usability)	Acquired brain injury (n=7)
Forster et al [40]	Hospital	Observational (pilot study)	Acquired brain injury with cognitive or motor impairments (n=15)
Hart et al [41]	Community	Randomized controlled trial (protocol)	Moderate-to-severe traumatic brain injury (n=60)
Henry et al [42]	Home	Observational (pilot)	Tinnitus (n=24)
Jean et al [43]	Community	Observational (preliminary)	Stroke (first-ever ischemic/hemorrhagic, n=36)
Lau et al [44]	Home + laboratory visits	Observational/prospective cohort	Stroke survivors (n=40, aged 18-65)
Lau et al [45]	Home + laboratory visits	Observational	Stroke survivors (n=40, aged 18-65)
Lau et al [46]	Home + laboratory visits	Observational	Stroke survivors (n=40, aged 18-65)
Lau et al [47]	Home + laboratory visits	Observational	Stroke survivors (n=40, aged 18-65)
Lenaert et al [48]	Home + laboratory visits	Feasibility study/observational	Acquired brain injury (stroke or traumatic brain injury, n=17)
Lenaert et al [49]	Home + laboratory visits	Longitudinal observational	Stroke (n=30)
Sartori et al [50]	Home + laboratory visits	Cross-sectional observational	Hereditary spastic paraplegia (n=50)
Shi et al [51]	Home + laboratory visits	Prospective observational	Stroke (n=202)
Rabinowitz et al [52]	Home + laboratory visits	Randomized controlled trial (exploratory)	Traumatic brain injury (minimum 6 months after injury, n=23)
Villain et al [53]	Hospital (baseline) and home (follow-up)	Prospective observational	Mild ischemic stroke (baseline n=44; follow-up n=34)
Villain et al [54]	Home + laboratory visits	Observational	Minor ischemic stroke (n=34)

[a]Setting refers to the context in which ecological momentary assessments were conducted, as described in the included studies. “Home” indicates assessments primarily conducted in participants’ home environments; “community” refers to assessments conducted in everyday environments beyond the home (eg, during daily life in the community); “hospital” refers to inpatient or clinical settings; and “home + laboratory visits” indicates ecological momentary assessment conducted in daily life combined with scheduled research or clinical visits.

Content Related

Across the included studies, questions about daily activities differed in both temporal framing and linguistic phrasing (Table 2), yet all were designed to capture participants’ engagement in daily activities. Some studies used explicitly time-bound prompts with multiple-choice response options, such as “What are you doing right now?” [40,42], whereas others referred to a retrospective period, for example, “What have you been doing over the past hour?” [41,52]. The degree of temporal specificity ranged from minutes [39] to hours [41]. By contrast, Lau et al [45] used a general, non-time-specific question (What are you doing?). In terms of phrasing, the verb “doing” was consistently used to denote the performance of or engagement in daily activities. Six studies [39-42,45,52] addressed participants in the second person (you), whereas 3 studies [36,48,49] used first-person phrasing (eg, “What am I doing?”). Notably, in 11 [35,37,38,43,44,46,47,50,51,53,54] of the 20 included articles, the exact wording of the daily activity question was not reported, either in the main manuscript or in the available supplementary materials.

XSL•FO
RenderX

Table 2. Content-related daily activities questions and answer possibilities.

Study	Questions on daily activities	Answer possibilities	
		Multiple choice answers	Open box
Blome et al [35]	Usual activities (question not reported)	A lot of daily activity options (options not reported)	Not reported
Broen et al [36]	Currently I am:	Resting, working, housework, hygiene, eating, drinking, relaxing, conversating, or other	Not reported
Bui et al [37]	Daily activities (question not reported)	Physical, cognitive, social, activities of daily living, instrumental activities of daily living, vocational, or passive leisure activities. Further grouped into "total productive" versus "nonproductive, sedentary" activities.	Not reported
Bui et al [38]	Daily activities (question not reported)	Physical, cognitive, social, activities of daily living, instrumental activities of daily living, vocational, or passive leisure activities. Further grouped into "total productive" versus "nonproductive, sedentary" activities.	Not reported
Ezekiel et al [39]	What have you been doing for the last 10 minutes?	Looking after myself, sport exercise, working or studying, social activities, using a computer/tablet/phone, household or family tasks, hobby or quiet pastime, or traveling	Not reported
Forster et al [40]	What are you doing right now?	Exercising, therapy, drinking coffee, taking a break, watching TV, eating, listening to radio/music, reading, surfing the internet, talking to others, or being outside	Not reported
Hart et al [41]	What have you been doing over the past hour or so?	Learning/school activity, working, household chores/pet care, leisure activity, pray/worship/meditate, resting/sleeping, watching TV/video, walking, shopping/errands, talking or socializing, reading, listening to music/radio/podcast, playing videogames/e-games, active play/sports/exercise, intimate relations, social media/internet/email/texting, preparing food/cooking, care of children, or other	Not reported
Henry et al [42]	What are you doing right now?	Working at paid job, watching TV/computer, etc, eating a meal, physical activity, or other	Not reported
Jean et al [43]	Current activity (question not reported)	Doing nothing or resting, work, sports, watching television, listening to music, having an in-person conversation, having a phone conversation, nonphysical leisure, personal hygiene, eating, or other daily life activities	Not reported
Lau et al [44]	Current daily activity (question not reported)	Multiple choice options were not reported.	Not reported
Lau et al [45]	What are you doing?	List of 44 daily activities, developed based on validated ecological momentary assessment surveys and the Activity Card Sort. According to the Activity Card Sort, activities were categorized into 4 domains (instrumental activities of daily living [eg, shopping], high-physical-demand leisure activities [eg, gardening], low-physical-demand leisure activities [eg, watching TV], and social activities [eg, visiting family or friends]), or an additional domain of activities of daily living (eg, toileting).	Not reported
Lau et al [46]	Daily activities (question not reported)	Multiple choice options were not reported.	Not reported
Lau et al [47]	Current activity (question not reported)	List of 44 daily activities (eg, watching TV, eating out, gardening), developed based on validated ecological momentary assessment surveys and the Activity Card Sort.	Not reported
Lenaert et al [48]	What am I doing (just before the beep)? And also? And?	Doing nothing, resting, work, household, self-care, taking care of partner, relaxation, or something else	Not reported
Lenaert et al [49]	What am I doing?	Nothing, resting, working, household, self-care, relaxing, traveling, or other	Not reported
Sartori et al [50]	Ongoing activity (question not reported)	Not reported	Categorized into productive, leisure, interactions, personal care, and others (open-ended)

XSL•FO
RenderX

Study	Questions on daily activities	Answer possibilities	
		Multiple choice answers	Open box
Shi et al [51]	What activities they were engaging in (question not reported)	Productive (eg, eating, exercising, visiting health care providers) or nonproductive (eg, resting, smoking, and doing nothing). Options were not reported.	Not reported
Rabinowitz et al [52]	What have you been doing over the past hour or so?	Learning/school activity working, household chores/pet care, leisure activity, pray/worship/meditate, resting/sleeping, watching TV/video, walking, shopping/errands, talking or socializing, reading, listening to music/radio/podcast, playing videogames/e-games, active play/sports/exercise, intimate relations, social media/internet/email/texting, preparing food/cooking, care of children, or other (to be specified in textbox)	Not reported
Villain et al [53]	Activities of daily living (question not reported)	Leisure activities (sport, nonphysical leisure, watching television, or listening to music) and activities of daily living (working, shopping, cooking, or housework). Options were not reported.	Not reported
Villain et al [54]	Activities (question not reported)	Multiple choice options were not reported.	Not reported

As presented in Table 2, various response formats were used to assess daily activities. As many as 19 [35-49,51-54] of the 20 included studies relied on multiple-choice response options with predefined activities, whereas 1 study [50] allowed open-ended responses. Several studies offered highly detailed lists of daily activities, often based on validated instruments. For instance, Lau et al [45,47] presented a list of 44 discrete activities derived from validated EMA surveys and the Activity Card Sort, which were subsequently grouped into 4 functional domains—instrumental activities of daily living, high-physical-demand leisure, low-physical-demand leisure, and social activities—with an additional category for basic activities of daily living [45,47]. Bui et al [37,38] also used a multidimensional classification, grouping activities into physical, cognitive, social, activities of daily living, instrumental activities of daily living, vocational, and passive leisure activities, and further categorizing them as "total productive" versus "nonproductive/sedentary" [37,38]. Forster et al [40], Henry et al [42], Ezekiel et al [39], Rabinowitz et al [52], and Hart et al [41] provided varied lists encompassing both work-related and leisure activities, including digital media use (eg, listening to music, social media, and gaming) and household responsibilities (eg, cooking and caring for children). Additionally, 3 studies [43,48,49] explicitly included the option "doing nothing." The option "other" was included in 6 studies; however, it was not always clear whether participants could provide a written specification for this response or whether it simply indicated that none of the listed options applied [36,41-43,49,52]. In 5 studies, the specific multiple-choice response options were not reported in either the main manuscript or the supplementary material [44,46,50,53,54]. The complete set of detailed data is presented in Table 2.

Alongside items related to daily activities, the EMA protocols included measures of various additional constructs, as daily activities were not the primary focus of all included studies. To support synthesis and interpretation, these constructs were classified using the PEO model. Within this review, daily activities are conceptually understood as emerging from the dynamic interaction among the person, environment, and occupation, consistent with occupational science perspectives. However, for the purpose of this supplementary analysis, the PEO model was applied analytically by examining its components separately. This approach was used to categorize EMA items that addressed person-, environment-, or occupation-related aspects without constituting direct assessments of daily activities. Although the PEO domains are inherently interrelated, considering them as distinct analytical categories enabled the identification of common patterns and areas of emphasis across studies (Table 3).

XSL•FO
RenderX

Table 3. Additional questions related to components of the PEO[a] model.

Study	Person			Environment		Occupation
	Physical	Cognitive	Affective	Physical	Social	
Blome et al [35]	Mobility and pain discomfort	Not reported	Anxiety and depression	Not reported	Not reported	Self-care
Broen et al [36]	Parkinson-related symptoms and somatic	Not reported	Mood	Contextual factors (location)	Contextual factors (social presence and comfort)	Event related: importance of the event, pleasure
Bui et al [37]	Not reported	Not reported	Depressed mood	Current location	Social company	Activity quality and satisfaction of the activity
Bui et al [38]	Poststroke symptoms	Not reported	Not reported	Current location	Social company	Not reported
Ezekiel et al [39]	Reaction time (objective indicator of fatigue) and energy levels	Not reported	Not reported	Not reported	Not reported	Not reported
Forster et al [40]	Not reported	Self-reflection frequency	Mood	Not reported	Social context	Judgment of performance
Hart et al [41]	Not reported	Not reported	PANAS[b] items (eg, scared, alert, nervous)	Location	Social context	Enjoyment of the activity and sense of accomplishment
Henry et al [42]	Tinnitus perception	Not reported	Happiness-sadness and anxiety-calmness	Location and loudness of the context	Social context	Not reported
Jean et al [43]	Not reported	Not reported	Not reported	Immediate location	Social company	Not reported
Lau et al [44]	Not reported	Not reported	Basic psychological needs (autonomy, competence, and relatedness)	Not reported	Not reported	Not reported
Lau et al [45]	Physical fatigue and pain	Cognitive complaints and mental fatigue	Depressed mood and cheerfulness	Not reported	Not reported	Not reported
Lau et al [46]	Symptoms of stroke	Cognitive complaints	Depressed and cheerful affect	Not reported	Not reported	Not reported
Lau et al [47]	Physical fatigue and pain	Cognitive complaints and mental fatigue	Depressed mood	Not reported	Not reported	Not reported
Lenaert et al [48]	Pain, fatigue, and not feeling well	Not reported	Mood and self-esteem	Location	Social context	Important event that happened, pleasant activity, control over activity, unexpected activity, and important activity
Lenaert et al [49]	Physical well-being, current fatigue, and physical activity	Not reported	Perceived effort, enjoyment, and mood	Location	Not reported	Enjoyment of the activity and perceived effort
Sartori et al [50]	Not reported	Cognitive	Affective	Location, environmental challenges, and skills	Social context	Environmental challenges and skills
Shi et al [51]	Pain, tiredness, and stress	Concentration	Anxiety, worthlessness, and cheerfulness	Location	Social isolation and company	Not reported

XSL•FO
RenderX

Study	Person			Environment		Occupation
	Physical	Cognitive	Affective	Physical	Social	
Rabinowitz et al [52]	Not reported	Not reported	PANAS items (eg, scared, alert, nervous)	Physical context	Social context	Enjoyment of the activity and accomplishment of activity
Villain et al [53]	Not reported	Not reported	Mood and depression symptoms	Not reported	Not reported	Not reported
Villain et al [54]	Not reported	Not reported	Anxiety and depression symptoms	Not reported	Not reported	Not reported

[a]PEO: Person-Environment-Occupation.

[b]PANAS: Positive and Negative Affect Schedule.

Within the person domain, the PEO model distinguishes among physical, affective, and cognitive dimensions. Physical aspects were frequently assessed across the included studies and included symptoms such as pain, fatigue, mobility limitations, stress, and somatic complaints [35,36,38,42,44-49,51]. One study [39] also incorporated objective performance indicators, such as reaction time, as a proxy for fatigue. Affective measures were particularly prevalent, capturing mood states ranging from positive (eg, cheerfulness) to negative (eg, anxiety, depression, and feelings of worthlessness). Emotional valence was assessed using either discrete affective descriptors (eg, Positive and Negative Affect Schedule [PANAS] items) or global mood ratings. Cognitive constructs included concentration, cognitive complaints, mental fatigue, self-reflection frequency, and perceived effort (eg, Lau et al [45-47], Forster et al [40]; Sartori et al [50]; Shi et al [51]), although these were assessed far less frequently than physical or affective dimensions [40,45-47,50,51].

Within the environment domain, additional EMA items were categorized into physical, institutional, cultural, and social environments according to the PEO model. Only the physical and social environments were represented in the included studies; no EMA items addressed institutional or cultural environments. The physical environment was typically assessed through questions about participants' immediate location [37,38,41-43,48-51] and environmental conditions such as noise or loudness [36,42]. Assessments of the social environment focused on the presence or absence of social company, perceived social support, and experiences of social isolation [36-38,40-43,48,50-52].

Within the occupation domain, considerably fewer EMA items addressed this aspect than the person and environment domains. In the PEO framework, occupation is typically divided into self-care, productivity, and leisure. Only 1 study [35] explicitly assessed self-care activities. The remaining studies addressed the occupation domain more broadly by focusing on constructs such as activity quality, perceived importance of the event, satisfaction with the activity, judgment of performance, sense of accomplishment, perceived control over the activity, and whether the activity was considered important or unexpected [36,37,40,41,48-50,52].

In addition to the constructs categorized according to the PEO model, some EMA measures could not be directly assigned to a single PEO component, including motivation and EMA disturbance. In 4 studies [44,46,47,50], motivation was assessed alongside PEO-related constructs, consistent with self-determination theory. One study [48] also assessed EMA disturbance, defined as the extent to which the EMA protocol itself was perceived as disruptive during daily life. The complete set of detailed data on these additional constructs is presented in Table 3.

Practical Application

Across the included studies, a variety of technological platforms were used, including the PsyMate app (n=3), the RealLife Exp app (n=2), the PIEL Survey App (n=4), EQ-5D-AA on a mobile phone, custom Android apps, movisensXS, Palm Pilot, and PMAT v2.1.2. EMA schedules also varied considerably across studies. The complete set of detailed data is presented in Table 4.

XSL•FO
RenderX

Table 4. Practical application.

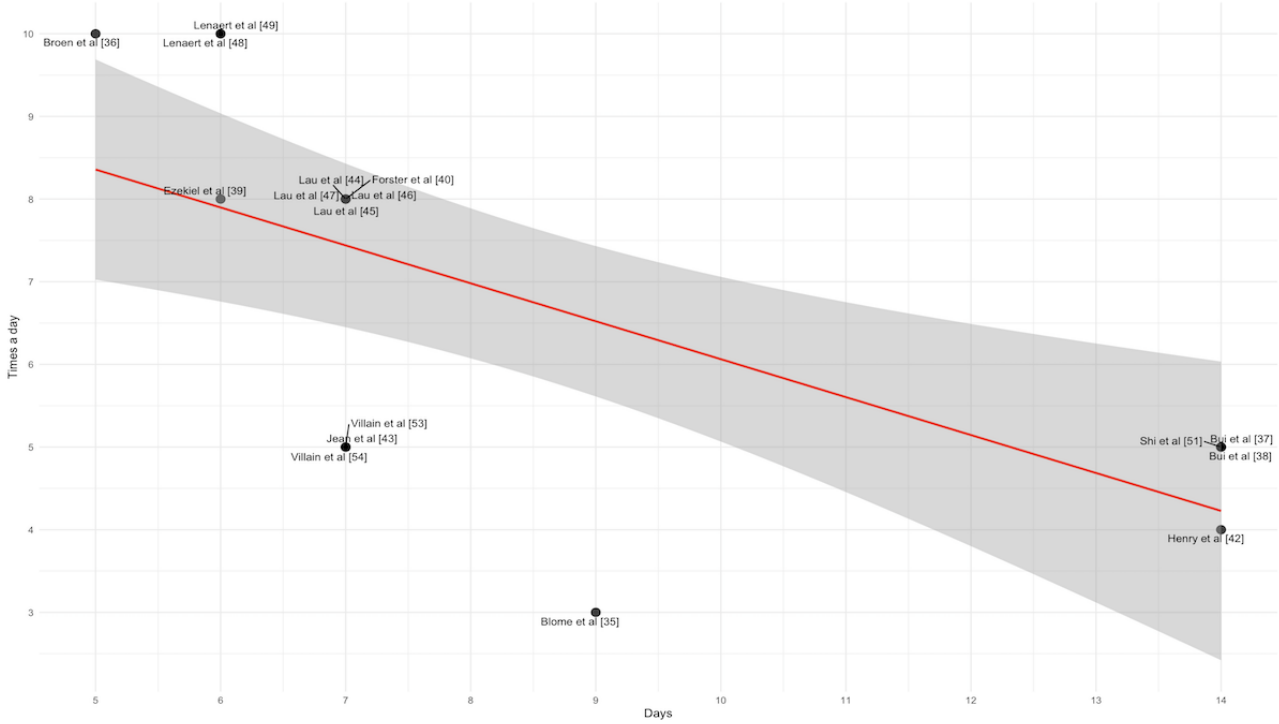
Study	EMA[a] schedule	EMA tool	Time points	Adherence, %
Blome et al [35]	3 times/day for 9 days (morning, midday, and evening)	EQ-5D-AA on mobile phone	27	92.6
Broen et al [36]	10 times/day for 5 days, semirandom within 90-minute blocks (7:30 AM-10:30 PM)	PsyMate app (iPod Touch or own smartphone)	50	84.0
Bui et al [37]	5 times/day for 14 days between fixed time blocks (8 AM-10 PM)	Not specified	70	86.3
Bui et al [38]	5 times/day for 14 days, random start within fixed 2-hour blocks	Status/Post app (iPod Touch or iPhone)	70	84.0
Ezekiel et al [39]	8 times/day for 6 days; 2 fixed + 6 stratified random (10 AM-8 PM)	Custom Android app	48	Not reported
Forster et al [40]	8 times/day for 7 days, semirandom (8 AM-8 PM)	movisensXS (Android)	56	71.6
Hart et al [41]	5 times/day, pseudo-randomized over a 14-hour window. Duration in days not reported.	RealLife Exp app	Not reported	Not reported
Henry et al [42]	4 times/day for 14 days, fixed intervals with reminders	Palm Pilot (CERTAS software)	56	90
Jean et al [43]	5 times/day for 7 days, fixed time blocks per participant	PMAT version 2.1.2 (PDA)	35	74.1
Lau et al [44]	8 times/day, 7 days, 2-hour intervals (8 AM-10 PM)	PIEL Survey App	56	93.6
Lau et al [45]	8 times/day, 7 days, 2-hour intervals (8 AM-10 PM)	PIEL Survey App	56	93.6
Lau et al [46]	8 times/day, 7 days, 2-hour intervals (8 AM-10 PM)	PIEL Survey App	56	93.6
Lau et al [47]	8 times/day, 7 days, 2-hour intervals (8 AM-10 PM)	PIEL Survey App	56	93.6
Lenaert et al [48]	10 signals/day for 6 days, semirandom (7:30 AM-10:30 PM)	PsyMate	60	71.2
Lenaert et al [49]	10 signals/day for 6 days, semirandom (7:30 AM-10:30 PM)	PsyMate	60	65
Sartori et al [50]	6-8 times/day for 7 days, randomized (8 AM-10 PM)	Not specified	42-56	54.9
Shi et al [51]	5 times/day for 14 days, about 2.5-hour intervals (8 AM-10 PM)	iPod Touch or iPhone app	70	84.3
Rabinowitz et al [52]	5 times/day, pseudo-random within a 14-hour window, 7-18 days	RealLife Exp app	35-90	65
Villain et al [53]	5 times/day, 7 days (fixed blocks: 9 AM-10 PM)	Palm Tungsten E2	35	74
Villain et al [54]	5 times/day for 7 days	Palm Tungsten E2	35	74

[a]EMA: ecological momentary assessment.

Given the variability in EMA schedules, Figure 3 (also see [35-40,42-49,51,53,54]) provides a visual representation of the included protocols. Regarding prompt timing within the day, 3 main scheduling patterns were identified: fixed intervals (eg, Henry et al [42]), semirandom or pseudorandom prompts within predefined time blocks (eg, Broen et al [36] and Lenaert et al [48,49]), and mixed fixed and random approaches (eg, Ezekiel et al [39]). The number of daily prompts also varied substantially across studies, ranging from fixed schedules (eg, Lau et al [46]) to variable schedules reported as ranges (eg, 6-8 prompts/day in Sartori et al [50]; 7-14 prompts/day in Rabinowitz et al [52]). Studies reporting a range rather than a fixed number of daily prompts were excluded from the exploratory regression analyses to avoid arbitrary estimates. In addition, Hart et al [41] was excluded because of insufficient reporting on EMA duration. Consequently, the exploratory regression analyses included 17 studies. The number of daily prompts ranged from 3 to 10, whereas study duration ranged from 5 to 14 days. Some studies combined shorter study periods with high sampling intensity. For instance, Broen et al [36] and Lenaert et al [48,49] collected data over only 5 days but prompted participants up to 10 times per day, representing the highest daily sampling intensity. Conversely, Henry et al [42], Shi et al [51], and Bui et al [37,38] collected data over 14 days but limited daily prompts to 4 or 5. A significant inverse association was observed (β=–.46, P=.004): studies with shorter durations generally used higher daily sampling frequencies, whereas those with longer study periods used fewer daily assessments. A slightly different pattern was observed in Blome et al [35], which used a 9-day protocol with 3 daily prompts. Approximately half of the included studies used a 7-day protocol with 5-8 prompts per day [40,43-47,50,53,54].

XSL•FO
RenderX

Figure 3. Regression analysis of the ecological momentary assessment schema of the included studies.



Additionally, adherence rates were examined in relation to the number of daily prompts. As shown in Figure 4 (see also [35-38,40,42-49,51,53,54]), adherence rates varied widely across studies, ranging from approximately 55% to more than 90%. No clear association was observed (β=–.919, *P*=.42). A slight positive trend was observed in Figure 5 ([see also 35,36,40-49,51-54]) between adherence and study duration, with longer studies tending to show somewhat higher adherence. However, this relationship did not reach statistical significance (β=.851, *P*=.28). Adherence rates were also examined in relation to the total number of EMA assessments (Figure 6; see also [35-38,40,42-49,51,53,54]). Overall, adherence was relatively high, with most studies reporting completion rates above 70%. The highest adherence (93.6%) was reported in studies with 56 assessments (eg, Lau et al [45-47]), whereas 1 study [35] reported similarly high adherence (92.6%) with 27 assessments. No significant linear association was observed between the total number of assessments and adherence (β=.087, *P*=.66). However, studies with a high number of assessments (>65) generally maintained adherence above 80% (eg, Bui et al [37] and Shi et al [51]), whereas studies with 50-60 assessments showed greater variability, with completion rates ranging from 55% to more than 90%. For consistency across the exploratory regression analyses, only studies reporting the complete set of required variables were included. Four studies were excluded because of incomplete reporting of adherence, EMA duration, or the total number of assessments. Specifically, Hart et al [41] and Ezekiel et al [39] did not report adherence data, whereas Rabinowitz et al [52] and Sartori et al [50] reported variable EMA schedules without a fixed study duration. To ensure comparability across analyses, these studies were excluded from the regression models and are therefore not represented in Figures 4-6. These findings should be interpreted as exploratory because of the limited number of included studies.

XSL•FO
RenderX

Figure 4. Adherence rates in relation to the total prompts a day.

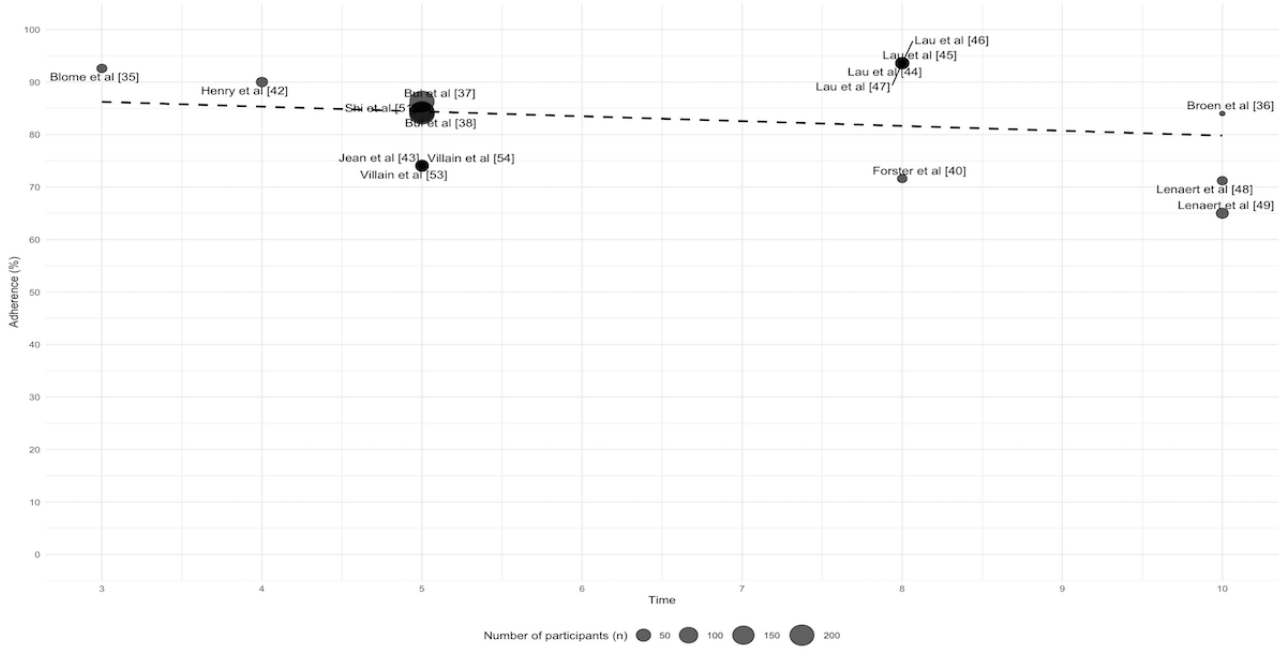
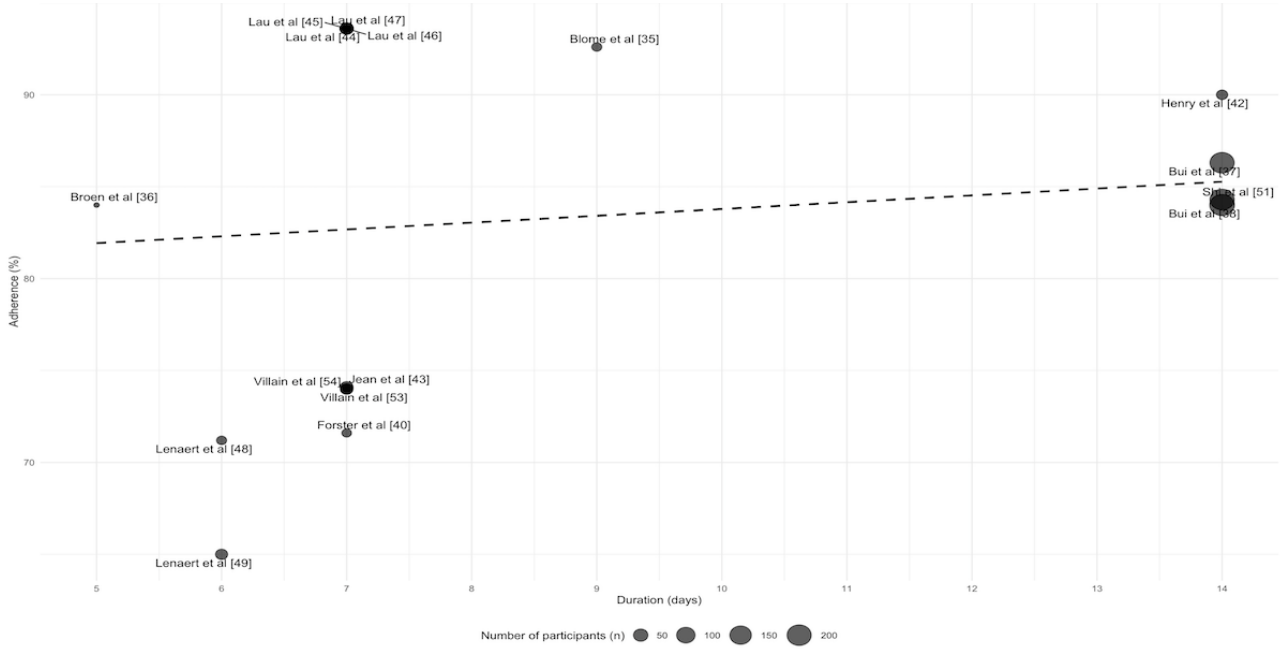
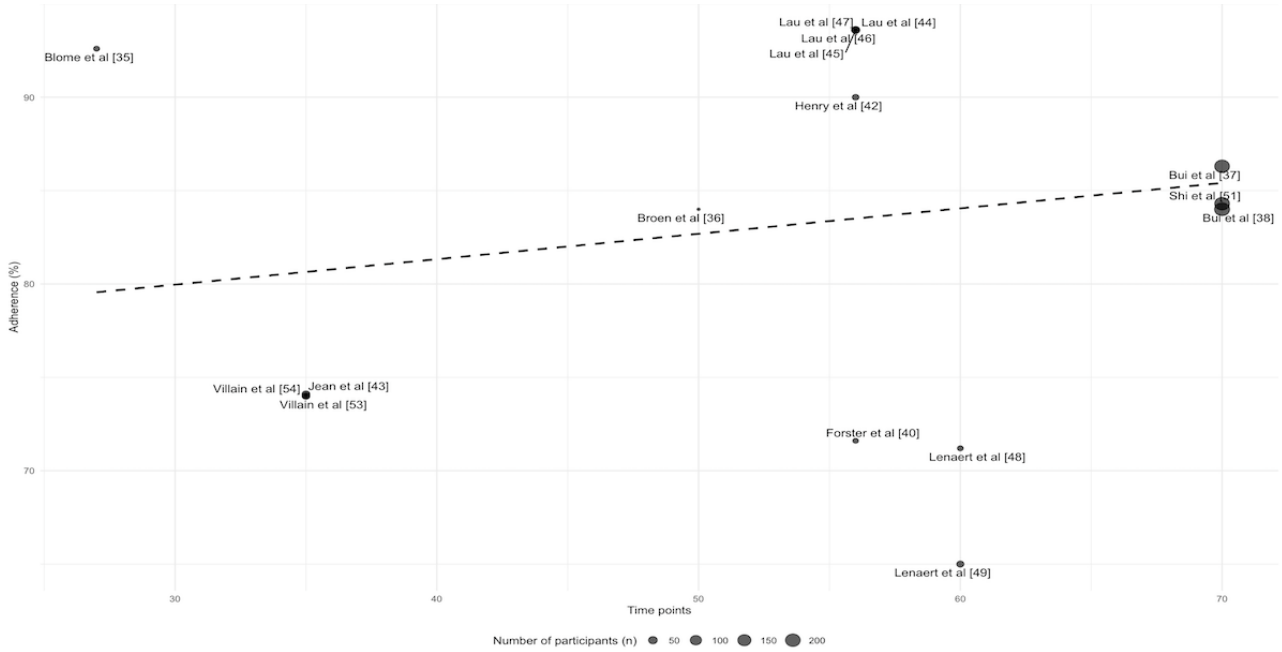


Figure 5. Adherence rates in relation to the total number of days.



XSL•FO
RenderX

Figure 6. Adherence rates in relation to the total number of time points.



Discussion

Principal Findings

This scoping review aimed to map the existing literature on the use of EMA to capture daily activities in individuals with neurological disorders, with a specific focus on content-related aspects and practical application. Twenty articles published between 2012 and 2024 were identified in which EMA was used to assess daily activities. Although EMA has been available for several decades since its introduction by Csikszentmihalyi and Larson in 1987 [21], only 20 studies [35-54] have focused on this population with the explicit aim of gaining insight into daily activities [55,56]. This relatively small body of literature is surprising because improving daily activities is a central goal of neurological rehabilitation and is essential for maximizing recovery, maintaining independence, and supporting long-term health, participation, and quality of life after formal rehabilitation has ended [2,9,57]. Given that EMA is specifically designed to capture experiences and behaviors in everyday life, its application to assessing daily activities in neurological rehabilitation remains surprisingly limited.

While EMA holds clear promise as a method for capturing daily activities, it is important to distinguish it from emerging passive sensing approaches. Although technologies such as wearable sensors and smartphone-based tracking can objectively capture aspects of behavior (eg, movement, location, or device use), they do not provide direct insight into the meaning, intention, or subjective experience of daily activities [22]. EMA uniquely captures these experiential and contextual dimensions, which are central to understanding engagement in daily activities from a PEO perspective.

The aim of this review was not only to consider the theoretical potential of EMA but also to examine how it is currently being applied to gather information about everyday life in individuals with neurological disorders, with the goal of identifying opportunities for broader and more effective use. EMA protocols appropriately vary according to the research aims and the daily-life phenomena being captured. Nevertheless, transparent reporting of the specific EMA items used is crucial. This review revealed that such transparency was often lacking, hindering interpretation and comparison across studies. As a result of the limited number of EMA studies focusing specifically on daily activities, we also included studies in which this construct was assessed as a secondary outcome to map its current operationalization. This may explain why daily activity items were often not described in detail, as they were not the primary focus of many included studies. In addition, a key principle of EMA is its ability to capture experiences and behaviors in the moment, thereby minimizing retrospective bias. Yet, several studies framed their questions retrospectively, for example, by asking participants, “What have you been doing in the past hour?” [52]. This approach undermines one of the central advantages of EMA, highlighting a gap between its theoretical potential and its practical implementation [20,21,24]. These findings also point to the need for more structured guidance on the design and reporting of EMA protocols. In particular, greater consistency in the timing, phrasing, and content of EMA items would improve comparability across studies and enhance interpretability. Developing a core set of reporting recommendations or a checklist for EMA studies assessing daily activities in neurological populations may therefore be a valuable next step.

This limitation in temporal framing also intersects with how daily activities are conceptualized in EMA studies. Questions frequently rely on the verb “doing” to capture engagement in daily activities. Although Wilcock [58] described doing as a core dimension of human occupation—alongside being, becoming, and, later, belonging—this perspective suggests that focusing primarily on “doing” may not fully capture the complexity of engagement in daily activities among individuals with neurological disorders. When physical performance is restricted, such a focus may overlook other meaningful forms

XSL•FO
RenderX

of engagement, such as planning, initiating, or contributing to daily activities. From a theoretical perspective, doing represents only 1 dimension of the broader concept of engagement. A narrow emphasis on doing may therefore overlook aspects such as occupational deprivation and occupational alienation. Although these concepts were not directly assessed in the included studies, the observed patterns in EMA item content can be interpreted in light of established occupational science frameworks. For example, occupational deprivation refers to a state in which individuals are prevented from participating in meaningful occupations because of external constraints, such as environmental barriers, social restrictions, or health-related limitations [10,59]. Similarly, occupational alienation occurs when engagement in daily activities lacks meaning, purpose, or personal satisfaction, even when the activity is physically possible [10]. By focusing primarily on "doing," EMA protocols may overlook the depth and richness of information that can be obtained. However, incorporating other PEO elements—personal factors, environmental contexts, and characteristics of the occupation itself—can help offset the limitations of focusing solely on "doing" and provide more meaningful insights into daily activities [4,10,60].

The way an EMA question is phrased directly shapes the type of responses it elicits. The multiple-choice response options used in the included studies resemble a classification system—such as self-care, leisure, or productivity—rather than providing insight into how a daily activity is performed or how the individual is engaged in that activity. While such categories may serve a useful descriptive purpose, they often remain at the level of occupational type rather than capturing the richer dimensions of daily activity. This limited framing can also obscure the so-called dark side of occupation—activities that may be harmful, socially undesirable, or associated with negative consequences [61]. However, some studies included an "other" response option, which could, in theory, capture such activities. Similarly, activities such as "doing nothing" are rarely explored in depth, even though they may represent diverse states such as rest, disengagement, or lack of opportunity. Recent developments in EMA methodology suggest that open-ended response formats can address some of these limitations. Open-ended questions are increasingly being used in other fields to capture richer data—for example, in autism research, where participants may respond not only through writing but also through drawing, providing insight into nuances that predefined categories may not capture [62,63]. Although open-ended EMA questions require more time and effort to complete, they can generate a broader range of responses and uncover insights that researchers may not have anticipated [63,64]. In this context, the value of open-ended questions lies in their ability to reveal the complexity, individuality, and sometimes unexpected aspects of daily activities that might otherwise remain hidden.

In addition to items targeting daily activities, many EMA protocols incorporated a range of supplementary questions. Classification of these questions using the PEO model revealed that a substantial proportion pertained to person- and environment-related factors. The strong representation of person-related factors is consistent with the findings of a recent bibliometric analysis examining the use of EMA to capture daily activities across health and disability populations [31]. With respect to environmental factors, a pronounced imbalance was observed. Most questions addressed the physical and social environment, whereas the cultural and institutional dimensions were largely absent. Institutional factors, such as organizational routines, policies, or the physical environment of institutions, received little attention despite evidence from neurological rehabilitation research demonstrating their relevance to participation and engagement [65]. Although the limited assessment of institutional factors may partly reflect the predominance of home-based EMA studies, in which these factors are less salient, this gap is more concerning in ambulatory and inpatient settings. Similarly, the cultural environment remains underrepresented in EMA research on daily activities among individuals with neurological disorders. For example, sociological research has highlighted the importance of incorporating cultural dimensions to achieve a more comprehensive understanding of human behavior and experience [66]. Expanding EMA methodologies to systematically address both cultural and institutional contexts could enhance the ecological validity of findings and provide a more comprehensive representation of the environments in which daily activities occur. However, one could argue that these contexts are relatively stable throughout the day and therefore may not require frequent assessment using EMA. At the same time, other environmental dimensions, such as the virtual environment, are becoming increasingly relevant in contemporary life, yet none of the included studies assessed this aspect [3].

Beyond the content-related aspects, this review also provides insight into the practical application of EMA for capturing daily activities. The results indicate consistently high adherence rates, supporting the feasibility of EMA in this context. Similar patterns have been reported in other populations. For example, a scoping review in perinatal mental health reported an average compliance rate of approximately 80%, ranging from 63% to 96%, indicating strong feasibility [67]. In populations with psychosis, a systematic review found a mean survey completion rate of 67%, with more than 86% of participants meeting minimum engagement criteria [68]. Likewise, a cross-diagnostic study reported median adherence rates above 80%, with early response behavior strongly predicting sustained participation over time [69]. Collectively, these findings suggest that well-designed EMA protocols—carefully balancing sampling frequency, participant burden, and interface usability—can achieve high adherence, supporting both comprehensive data collection and sustained participant engagement. However, the available data do not allow conclusions regarding the optimal number of assessments per day or the ideal duration of EMA protocols. Beyond adherence, this review also identified substantial variation in EMA design characteristics, including the number of daily prompts, study duration, sampling schedules, and response windows. Although these findings provide an overview of how EMA is currently implemented to assess daily activities in neurological populations and may inform the design of future studies, the available evidence remains insufficient to determine which protocol configurations are most effective for optimizing feasibility and data quality.

XSL•FO
RenderX

Limitations

The findings of this review should be interpreted in light of several methodological limitations. First, the search strategy may have been subject to selection bias. Although a broad search strategy was applied, including the use of the MeSH term "Nervous System Diseases," variability in indexing and the hierarchical structure of MeSH may have influenced the retrieval of specific conditions. As a result, some conditions were captured implicitly, whereas others were included explicitly, reflecting a balance between sensitivity and specificity. Searches were conducted using EMA-related terminology in the title and abstract, which may have resulted in the omission of relevant studies that used similar methodologies without explicitly using these terms. Although studies were screened for core EMA characteristics whenever possible, some eligible publications may not have been identified. Consequently, the size and scope of the available evidence base may have been underestimated. In addition, although this approach aligns with the exploratory nature of a scoping review, the absence of an extended manual search strategy (eg, reference checking or citation tracking), beyond the gray literature search, may have further limited the comprehensiveness of the evidence base. Incorporating such strategies in future reviews could yield additional insights. Second, this scoping review was conducted in accordance with the framework proposed by Arksey and O'Malley [33] and reported following the PRISMA-ScR guidelines, strengthening its methodological rigor and transparency. Nevertheless, the limitations related to the search strategy and the variability in reporting across the included studies should be considered when interpreting the findings. Third, the predefined inclusion criteria may have introduced conceptual ambiguity. The identification of EMA studies was based on the presence of terms such as ecological momentary assessment, experience sampling methodology, or ambulatory assessment in the title or abstract. However, the operationalization of EMA varies substantially across studies, particularly with respect to sampling frequency, duration, and intensity (eg, once daily vs multiple prompts per day). As a result, studies using markedly different methodological approaches were grouped under a single EMA label. Future research would benefit from a more explicit definition of EMA to improve conceptual clarity and comparability across studies. Furthermore, although 20 studies were included, several publications originated from the same research groups. This suggests that similar theoretical and methodological perspectives may have informed multiple articles. It also cannot be ruled out that overlapping datasets were used, as many studies relied on comparable samples and study parameters. Although definitive conclusions cannot be drawn, the body of evidence may therefore represent fewer than 20 fully independent perspectives. This observation highlights the importance of considering not only the number of studies but also the diversity of methodological approaches when interpreting the findings. Finally, substantial heterogeneity was observed across the included studies with respect to sample size, study duration, the number of EMA prompts, and adherence reporting. This variability complicates direct comparisons across studies and limits the extent to which the findings can be synthesized beyond a descriptive level, consistent with the aims of a scoping review. Furthermore, the regression analyses should be interpreted as exploratory because of the heterogeneity among studies, the limited number of included studies, and the incomplete reporting of adherence data.

Future Directions

Therefore, future studies should prioritize greater transparency in reporting EMA protocols and work toward developing more standardized guidelines or checklists to support consistency in EMA design and reporting. Researchers are encouraged to move beyond a narrow focus on "doing" and explore alternative question formats—including open-ended or mixed-format items—that can capture richer, more nuanced information about the context, meaning, and personal relevance of daily activities. In addition, when using predefined activity categories, it is important to include response options that allow participants to report less commonly addressed or potentially negative occupations, acknowledging that not all daily activities are inherently positive. Future research should also investigate the conditions under which adherence is highest to better understand the factors that support sustained engagement in EMA studies.

Methodologically, participatory approaches—such as co-designing EMA items with individuals who have lived experience of neurological disorders—may help ensure that questions reflect the diversity and complexity of real-world engagement. Combining EMA with complementary qualitative methods (eg, brief interviews or activity diaries) could further deepen insights and capture aspects of occupational engagement that might otherwise be overlooked. These strategies would enable EMA to more fully reflect the dynamic interplay among personal factors, environmental contexts, and occupations, ultimately broadening its value in both research and clinical practice.

Conclusions

This review systematically mapped both what EMA captures about daily activities in individuals with neurological disorders and how EMA is implemented in this context. Regarding the content of EMA protocols, the findings showed that daily activities were predominantly assessed using items focused on the verb "doing" and mainly closed-response formats. Additional EMA questions addressed several components of the PEO model, although physical and social environmental dimensions were represented more frequently than cultural and institutional dimensions. Together, these findings suggest that current EMA approaches may provide a relatively narrow representation of daily activity performance and engagement, potentially overlooking important nuances of individuals' lived experiences. Regarding practical application, considerable variability was observed in EMA protocol characteristics, including assessment frequency, study duration, prompting schedules, and adherence rates. Although adherence was generally high across studies, the heterogeneity of protocols and the exploratory nature of the available evidence make it difficult to draw firm conclusions about which design characteristics best support feasibility and data quality in neurological populations. Rather than identifying an optimal EMA approach, this review provides an overview of how EMA is currently being applied and highlights areas requiring further investigation. Future EMA research in

XSL•FO
RenderX

neurological populations would benefit from greater transparency in reporting EMA items, broader use of open-ended response formats, and more systematic consideration of underrepresented environmental dimensions. In addition, studies directly comparing EMA protocol characteristics are needed to better understand how protocol design influences feasibility, adherence, ecological validity, and the quality of data collected for individualized assessment of daily activities.

Acknowledgments

The authors acknowledge the use of ChatGPT (OpenAI) for language refinement and text editing. OpenAI did not contribute to the study design, data analysis, or interpretation, and all authors take full responsibility for the content of the manuscript. The authors also thank Ine Oosterbos and Ferre Maesen for their assistance with the data extraction process.

Data Availability

This scoping review does not include original data or materials, as it synthesizes and analyzes findings from previously published studies. Therefore, no datasets or study materials are available for sharing.

Funding

This study was supported by a doctoral research grant from Hasselt University awarded to the principal investigator (ED; grant BOF24DOC43). No other financial or material support was received for this study. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. AR was supported by the Research Council of Finland (grant 360087) as part of broader research activities; this support was not awarded specifically for this study. The other authors received no funding related to this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 177 KB - rehab_v13i1e92797_app1.pdf]

Multimedia Appendix 2
Search string.
[DOCX File , 14 KB - rehab_v13i1e92797_app2.docx]

References

1. Boggs D, Polack S, Kuper H, Foster A. Shifting the focus to functioning: essential for achieving Sustainable Development Goal 3, inclusive Universal Health Coverage and supporting COVID-19 survivors. Glob Health Action 2021 Jan 01;14(1):1903214 [FREE Full text] [doi: 10.1080/16549716.2021.1903214] [Medline: 33904370]
2. World Health Organization (WHO). International Classification of Functioning, Disability and Health. WHO. 2001. URL: https://iris.who.int/server/api/core/bitstreams/83e1c3e0-cf9e-4063-a424-71ccc9a84ded/content [accessed 2025-07-24]
3. Hinojosa P, Kramer C, Royeen CB. Perspectives on Human Occupations: Theories Underlying Practice. Philadelphia, PA: F. A. Davis; 2017.
4. Law M, Cooper B, Strong S, Stewart D, Rigby P, Letts L. The Person-Environment-Occupation Model: a transactive approach to occupational performance. Can J Occup Ther 2016 Apr 22;63(1):9-23. [doi: 10.1177/000841749606300103]
5. Strong S, Rigby P, Stewart D, Law M, Letts L, Cooper B. Application of the Person-Environment-Occupation Model: a practical tool. Can J Occup Ther 1999 Jun 01;66(3):122-133. [doi: 10.1177/000841749906600304] [Medline: 10462885]
6. Hagedorn R. Tools for Practice in Occupational Therapy: A Structured Approach to Core Skills and Processes. Edinburgh, Scotland, UK: Churchill Livingstone; 2000.
7. Hammell KRW. Belonging, occupation, and human well-being: an exploration. Can J Occup Ther 2014 Feb 14;81(1):39-50. [doi: 10.1177/0008417413520489] [Medline: 24783487]
8. Unsworth C. Research in occupational therapy: methods of inquiry for enhancing practice. Aus Occup Therapy J 2007 Jun;54(2):164-164. [doi: 10.1111/j.1440-1630.2007.00657.x]
9. Wilcock A. An Occupational Perspective of Health. London, UK: Routledge; 2006.
10. Townsend E, Polatajko H. Enabling Occupation II: Advancing an Occupational Therapy Vision for Health, Well-being, & Justice Through Occupation. Ottawa, ON: CAOT Publications; 2007.
11. Feigin VL, Vos T, Nichols E, Owolabi MO, Carroll WM, Dichgans M, et al. The global burden of neurological disorders: translating evidence into policy. The Lancet Neurology 2020 Mar;19(3):255-265. [doi: 10.1016/s1474-4422(19)30411-9]

XSL•FO
RenderX

12. GBD 2021 Nervous System Disorders Collaborators. Global, regional, and national burden of disorders affecting the nervous system, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Neurol 2024 Apr;23(4):344-381 [FREE Full text] [doi: 10.1016/S1474-4422(24)00038-3] [Medline: 38493795]
13. Yetzer E, Blake K, Goetsch N, Shook M, St Paul M. SAFE medication management for patients with physical impairments of stroke, part two. Rehabil Nurs 2017;42(5):282-289. [doi: 10.1002/rnj.286] [Medline: 27353573]
14. Bass AD, Van Wijmeersch B, Mayer L, Mäurer M, Boster A, Mandel M, et al. Effect of multiple sclerosis on daily activities, emotional well-being, and relationships: The Global vsMS Survey. Int J MS Care 2020;22(4):158-164 [FREE Full text] [doi: 10.7224/1537-2073.2018-087] [Medline: 32863783]
15. Law M, Baptiste S, Carswell A, McColl MA, Polatajko H, Pollock N. Canadian Occupational Performance Measure 5th ed. (COPM). Ottawa, ON: CAOT Publications; 2014.
16. Law M, Baptiste S, McColl M, Opzoomer A, Polatajko H, Pollock N. The Canadian occupational performance measure: an outcome measure for occupational therapy. Can J Occup Ther 1990 Apr 22;57(2):82-87. [doi: 10.1177/000841749005700207] [Medline: 10104738]
17. Cardol M, de Haan RJ, van den Bos GA, de Jong BA, de Groot IJ. The development of a handicap assessment questionnaire: the Impact on Participation and Autonomy (IPA). Clin Rehabil 1999 Oct;13(5):411-419. [doi: 10.1191/026921599668601325] [Medline: 10498348]
18. AOTA. Occupational Therapy Practice Framework: domain and process—fourth edition. American Journal of Occupational Therapy 2020 Aug 31;74(2_suppl):7412410010p1-7412410010p87. [doi: 10.5014/AJOT.2020.74S2001]
19. Nielsen KT, Wæhrens EE. Occupational therapy evaluation: use of self-report and/or observation? Scand J Occup Ther 2015 Jan 20;22(1):13-23 [FREE Full text] [doi: 10.3109/11038128.2014.961547] [Medline: 25327781]
20. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
21. Csikszentmihalyi M, Larson R. Validity and reliability of the Experience-Sampling Method. J Nerv Ment Dis 1987 Sep;175(9):526-536. [doi: 10.1097/00005053-198709000-00004] [Medline: 3655778]
22. Trifan A, Oliveira M, Oliveira JL. Passive sensing of health outcomes through smartphones: systematic review of current solutions and possible limitations. JMIR Mhealth Uhealth 2019 Aug 23;7(8):e12649 [FREE Full text] [doi: 10.2196/12649] [Medline: 31444874]
23. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006 Jan;31(1):13-20. [doi: 10.1139/jpn.0602]
24. Stone A, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Annals of Behavioral Medicine 1994;16(3):199-202. [doi: 10.1093/abm/16.3.199]
25. Hall M, Scherner PV, Kreidel Y, Rubel JA. A systematic review of momentary assessment designs for mood and anxiety symptoms. Front Psychol 2021 May 17;12:642044 [FREE Full text] [doi: 10.3389/fpsyg.2021.642044] [Medline: 34079492]
26. Kwasnicka D, Kale D, Schneider V, Keller J, Yeboah-Asiamah Asare B, Powell D, et al. Systematic review of ecological momentary assessment (EMA) studies of five public health-related behaviours: review protocol. BMJ Open 2021 Jul 16;11(7):e046435 [FREE Full text] [doi: 10.1136/bmjopen-2020-046435] [Medline: 34272218]
27. Mink F, Hehlmann M, Lutz W. Ecological momentary assessment in clinical research: a systematic review. ResearchGate. 2024. URL: https://www.researchgate.net/publication/383677144_Ecological_Momentary_Assessment_in_clinical_research_A_systematic_review [accessed 2026-08-09]
28. Perski O, Keller J, Kale D, Asare BY, Schneider V, Powell D, et al. Understanding health behaviours in context: a systematic review and meta-analysis of ecological momentary assessment studies of five key health behaviours. Health Psychol Rev 2022 Dec 15;16(4):576-601 [FREE Full text] [doi: 10.1080/17437199.2022.2112258] [Medline: 35975950]
29. Schneider S, Junghaenel DU, Smyth JM, Fred Wen CK, Stone AA. Just-in-time adaptive ecological momentary assessment (JITA-EMA). Behav Res Methods 2024 Feb 25;56(2):765-783 [FREE Full text] [doi: 10.3758/s13428-023-02083-8] [Medline: 36840916]
30. Mei X, Li Y, Yeung WF, Hu Y, Li J, Li M, et al. Effectiveness of ecological momentary interventions on pain, mental health, and quality of life in individuals with rheumatic diseases: a systematic review and meta-analysis of randomized controlled trials. J Nurs Manag 2025 Jul 30;2025(1):9923240. [doi: 10.1155/jonm/9923240] [Medline: 40766076]
31. Delooz E, Bonnechère B, Piškur B, Spooren A. Tracking daily activities with ecological momentary assessment: a bibliometric analysis of current use in health Mapping daily activities with ecological momentary assessment. PLOS Digit Health 2026 Mar 31;5(3):e0001331 [FREE Full text] [doi: 10.1371/journal.pdig.0001331] [Medline: 41915636]
32. Stinson L, Liu Y, Dallery J. Ecological momentary assessment: a systematic review of validity research. Perspect Behav Sci 2022 Jun 06;45(2):469-493 [FREE Full text] [doi: 10.1007/s40614-022-00339-w] [Medline: 35719870]
33. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
34. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64 [FREE Full text] [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]

XSL•FO
RenderX

35. Blome C, Carlton J, Heesen C, Janssen MF, Lloyd A, Otten M, et al. How to measure fluctuating impairments in people with MS: development of an ambulatory assessment version of the EQ-5D-5L in an exploratory study. Qual Life Res 2021 Jul 12;30(7):2081-2096 [FREE Full text] [doi: 10.1007/s11136-021-02802-8] [Medline: 33710593]
36. Broen MPG, Marsman VAM, Kuijf ML, Van Oostenbrugge RJ, van Os J, Leentjens AFG. Unraveling the relationship between motor symptoms, affective states and contextual factors in Parkinson's disease: a feasibility study of the experience sampling method. PLoS One 2016;11(3):e0151195 [FREE Full text] [doi: 10.1371/journal.pone.0151195] [Medline: 26962853]
37. Bui Q, Kaufman KJ, Munsell EG, Lenze EJ, Lee J, Mohr DC, et al. Smartphone assessment uncovers real-time relationships between depressed mood and daily functional behaviors after stroke. J Telemed Telecare 2024 Jun;30(5):871-884 [FREE Full text] [doi: 10.1177/1357633X221100061] [Medline: 35549589]
38. Bui Q, Kaufman KJ, Pham V, Lenze EJ, Lee J, Mohr DC, et al. Ecological momentary assessment of real-world functional behaviors in individuals with stroke: a longitudinal observational study. Arch Phys Med Rehabil 2022 Jul;103(7):1327-1337 [FREE Full text] [doi: 10.1016/j.apmr.2022.02.011] [Medline: 35278467]
39. Ezekiel L, Veiga JJD, Ward T, Dawes H, Collett J. Exploring the usability of a smartphone application to monitor fatigue and activity for people with acquired brain injury. Br J Occup Ther 2023 Nov 12;86(11):767-776. [doi: 10.1177/03080226231183293] [Medline: 40336795]
40. Forster SD, Gauggel S, Petershofer A, Völzke V, Mainz V. Ecological momentary assessment in patients with an acquired brain injury: a pilot study on compliance and fluctuations. Front Neurol 2020 Mar 5;11:115 [FREE Full text] [doi: 10.3389/fneur.2020.00115] [Medline: 32194494]
41. Hart T, Rabinowitz A, Vaccaro M, Chervoneva I, Wilson J. Behavioral activation augmented with mobile technology for depression and anxiety in chronic moderate-severe traumatic brain injury: protocol for a randomized controlled trial. Arch Rehabil Res Clin Transl 2019 Dec;1(3-4):100027 [FREE Full text] [doi: 10.1016/j.arrct.2019.100027] [Medline: 33543057]
42. Henry JA, Galvez G, Turbin MB, Thielman EJ, McMillan GP, Istvan JA. Pilot study to evaluate ecological momentary assessment of tinnitus. Ear Hear 2012;33(2):179-290 [FREE Full text] [doi: 10.1097/AUD.0b013e31822f6740] [Medline: 21960147]
43. Jean FAM, Swendsen JD, Sibon I, Fehér K, Husky M. Daily life behaviors and depression risk following stroke: a preliminary study using ecological momentary assessment. J Geriatr Psychiatry Neurol 2013 Sep;26(3):138-143. [doi: 10.1177/0891988713484193] [Medline: 23584854]
44. Lau SC, Connor LT, Baum CM. Associations between basic psychological need satisfaction and motivation underpinning daily activity participation among community-dwelling survivors of stroke: an ecological momentary assessment study. Arch Phys Med Rehabil 2023 Feb;104(2):229-236. [doi: 10.1016/j.apmr.2022.07.011] [Medline: 35934048]
45. Lau SC, Connor LT, King AA, Baum CM. Multimodal ambulatory monitoring of daily activity and health-related symptoms in community-dwelling survivors of stroke: feasibility, acceptability, and validity. Arch Phys Med Rehabil 2022 Oct;103(10):1992-2000 [FREE Full text] [doi: 10.1016/j.apmr.2022.06.002] [Medline: 35780826]
46. Lau SCL, Connor LT, Skidmore ER. Associations of circadian rest-activity rhythms with affect and cognition in community-dwelling stroke survivors: an ambulatory assessment study. Neurorehabil Neural Repair 2024 Mar 06;38(3):197-206. [doi: 10.1177/15459683241230027] [Medline: 38318642]
47. Lau SC, Connor LT, Skidmore ER, King AA, Lee J, Baum CM. The moderating role of motivation in the real-time associations of fatigue, cognitive complaints, and pain with depressed mood among stroke survivors: an ecological momentary assessment study. Arch Phys Med Rehabil 2023 May;104(5):761-768 [FREE Full text] [doi: 10.1016/j.apmr.2022.11.012] [Medline: 36535421]
48. Lenaert B, Colombi M, van Heugten C, Rasquin S, Kasanova Z, Ponds R. Exploring the feasibility and usability of the experience sampling method to examine the daily lives of patients with acquired brain injury. Neuropsychol Rehabil 2019 Jun;29(5):754-766 [FREE Full text] [doi: 10.1080/09602011.2017.1330214] [Medline: 28562164]
49. Lenaert B, Neijmeijer M, van Kampen N, van Heugten C, Ponds R. Poststroke fatigue and daily activity patterns during outpatient rehabilitation: an experience sampling method study. Arch Phys Med Rehabil 2020 Jun;101(6):1001-1008. [doi: 10.1016/j.apmr.2019.12.014] [Medline: 32001258]
50. Sartori RD, Marelli M, D'Angelo MG, Delle Fave A. Autonomy level and quality of everyday experience of people with hereditary spastic paraplegia. Health Soc Care Community 2019 Sep 17;27(5):e850-e860. [doi: 10.1111/hsc.12812] [Medline: 31313875]
51. Shi Y, Fong MW, Metts CL, LaVela SL, Bombardier C, Hu L, et al. Dynamics of perceived social isolation, secondary conditions, and daily activity patterns among individuals with stroke: a network analysis of ecological momentary assessment data. Arch Phys Med Rehabil 2024 Jul;105(7):1314-1321. [doi: 10.1016/j.apmr.2024.02.733] [Medline: 38458373]
52. Rabinowitz A, Hart T, Wilson J. Ecological momentary assessment of affect in context after traumatic brain injury. Rehabil Psychol 2021 Nov;66(4):442-449. [doi: 10.1037/rep0000403] [Medline: 34516166]
53. Villain M, Sibon I, Renou P, Poli M, Swendsen J. Very early social support following mild stroke is associated with emotional and behavioral outcomes three months later. Clin Rehabil 2017 Jan 10;31(1):135-141. [doi: 10.1177/0269215515623600] [Medline: 26851250]

XSL•FO
RenderX

54. Villain M, Sibon I, Renou P, Poli M, Swendsen J. Depression and routinization following stroke. Rev Neurol (Paris) 2021 Oct;177(8):964-968. [doi: 10.1016/j.neurol.2020.08.012] [Medline: 34215428]
55. Larson R, Csikszentmihalyi M. Flow and the foundations of positive psychology. In: The Experience Sampling Method. Dordrecht, The Netherlands: Springer; Jan 01, 2014:21-34.
56. Prescott S, Csikszentmihalyi M. Environmental effects on cognitive and affective states: the experiential time sampling approach. Soc Behav Pers 1981 Jan 01;9(1):23-32. [doi: 10.2224/sbp.1981.9.1.23]
57. Barker RN, Sealey CJ, Polley ML, Mervin MC, Comans T. Impact of a person-centred community rehabilitation service on outcomes for individuals with a neurological condition. Disabil Rehabil 2017 Jun 09;39(11):1136-1142. [doi: 10.1080/09638288.2016.1185803] [Medline: 27281692]
58. Ross MW. Typing, doing, and being: sexuality and the internet. J Sex Res 2005 Nov;42(4):342-352. [doi: 10.1080/00224490509552290] [Medline: 19827239]
59. Whiteford G. Occupational deprivation: global challenge in the new millennium. British Journal of Occupational Therapy 2000 May 01;63(5):200-204. [doi: 10.1177/030802260006300503]
60. Christiansen C, Townsend E. Introduction to Occupation: The Art of Science and Living. London, UK: Pearson Education; 2013.
61. Twinley R. The dark side of occupation: a concept for consideration. Aust Occup Ther J 2013 Aug;60(4):301-303 [FREE Full text] [doi: 10.1111/1440-1630.12026] [Medline: 23888980]
62. Chen Y, Xi Z, Greene T, Mandy W. A systematic review of ecological momentary assessment in autism research. Autism 2025 Jun 18;29(6):1374-1389 [FREE Full text] [doi: 10.1177/13623613241305722] [Medline: 39696839]
63. Stadel M, Langener A, Hoemann K, Bringmann L. Assessing daily life activities with experience sampling methodology (ESM): scoring predefined categories or qualitative analysis of open-ended responses? Methods in Psychology 2025 Jun;12:100177 [FREE Full text] [doi: 10.1016/j.metip.2025.100177]
64. van Roekel E, Keijsers L, Chung JM. A review of current ambulatory assessment studies in adolescent samples and practical recommendations. J Res Adolesc 2019 Sep;29(3):560-577 [FREE Full text] [doi: 10.1111/jora.12471] [Medline: 31573762]
65. Beck M, Engelke E, Birkelund R, Martinsen B. Becoming a nomad when hospitalized with a neurological disease: a phenomenological study. International Journal of Qualitative Studies on Health and Well-being 2020 Sep 15;15(1):1815487 [FREE Full text] [doi: 10.1080/17482631.2020.1815487] [Medline: 32930071]
66. Lohani M, Blodgett G. Innovative and ecological: integrating ecological momentary assessment into environmental science research. Front Psychol 2025 Apr 15;16:1557055 [FREE Full text] [doi: 10.3389/fpsyg.2025.1557055] [Medline: 40302913]
67. Zhang X, Cui W, Yao J, Zhang Y, Wang Y. Feasibility and utility of ecological momentary assessment to measure mental health issues in perinatal women: scoping review. Psychiatry Res 2025 Jun;348:116400. [doi: 10.1016/j.psychres.2025.116400] [Medline: 40220631]
68. Bell IH, Eisner E, Allan S, Cartner S, Torous J, Bucci S, et al. Methodological characteristics and feasibility of ecological momentary assessment studies in psychosis: a systematic review and meta-analysis. Schizophr Bull 2024 Mar 07;50(2):238-265 [FREE Full text] [doi: 10.1093/schbul/sbad127] [Medline: 37606276]
69. Jones SE, Moore RC, Pinkham AE, Depp CA, Granholm E, Harvey PD. A cross-diagnostic study of adherence to ecological momentary assessment: comparisons across study length and daily survey frequency find that early adherence is a potent predictor of study-long adherence. Pers Med Psychiatry 2021;29-30:100085 [FREE Full text] [doi: 10.1016/j.pmip.2021.100085] [Medline: 34541425]

Abbreviations

EMA: ecological momentary assessment
ICF: International Classification of Functioning, Disability and Health
MS: multiple sclerosis
PANAS: Positive and Negative Affect Schedule
PEO: Person-Environment-Occupation
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

XSL•FO
RenderX

Edited by S Munce; submitted 03.Feb.2026; peer-reviewed by L Peeters, A Hartanto, TAK Manne; comments to author 22.Mar.2026; revised version received 08.Jul.2026; accepted 13.Jul.2026; published 14.Aug.2026.

Please cite as:
Delooz E, Piškur B, Rintala A, Bonnechère B, Spooren A
Ecological Momentary Assessment of Daily Activities in Individuals With Neurological Disorders: Scoping Review
JMIR Rehabil Assist Technol 2026;13:e92797
URL: https://rehab.jmir.org/2026/1/e92797
doi:10.2196/92797
PMID:

©Eva Delooz, Barbara Piškur, Aki Rintala, Bruno Bonnechère, Annemie Spooren. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 14.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Leisure-Time Physical Activity Interventions for Health Promotion in Individuals With Rett Syndrome: Scoping Review

Christen John Mendonca[1], PhD; James H Rimmer[2], PhD; Ashley Wright[3], MS; Raven Young[3], MS; Amitha Ananth[4], MD; Byron Lai[3], PhD

[1]SHP Research Collaborative, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States

[2]Dean's Office, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States

[3]Division of Pediatric Rehabilitation Medicine, University of Alabama at Birmingham, Birmingham, AL, United States

[4]Heersink School of Medicine, Division of Pediatric Neurology, University of Alabama at Birmingham, Birmingham, AL, United States

Corresponding Author:
Christen John Mendonca, PhD
SHP Research Collaborative
School of Health Professions
University of Alabama at Birmingham
3810 Ridgeway Drive
Birmingham, AL, 35209
United States
Phone: 1 (205) 975 6415
Email: christenjm@uab.edu

Abstract

Background: While half of individuals with Rett syndrome (RTT) are older than 50 years, research shows that they have low levels of physical activity, especially for those of advanced age and poor ambulation. Despite this evidence, recent studies underscore the potential of leisure-time physical activity interventions to improve health outcomes and quality of life for individuals with RTT.

Objective: This scoping review aimed to summarize the state of the science regarding physical activity interventions for individuals with RTT and their families across the life span; map the extent, range, and nature of research activity in telehealth; and describe outcomes targeted by interventions with a focus on health-related fitness.

Methods: Systematic searches were performed in the MEDLINE, Scopus, Google Scholar, and CINAHL Plus databases. The data charted from eligible studies included specific details about the participants, study design, setting, intervention characteristics, outcomes, and relevant key findings. Inclusion criteria were (1) original, peer-reviewed research; (2) full-text articles in English; (3) studies reporting a leisure-time physical activity intervention; and (4) a sample of individuals with RTT and/or caregivers.

Results: A total of 23 studies enrolled 200 participants in total (mean age 13, SD 7, range 2-48 years). Participants were all female individuals, and studies included mostly those who were ambulatory. RTT severity was reported in less than half (9/23, 39%) of the studies; however, participants with mild to severe scores were represented. Most of the studies (20/23, 87%) used a case study design or single-group repeated measures, and only 13% (3/23) were randomized trials. The focus was primarily on gross motor function, walking ability, and physical activity outcomes. In total, 39% (9/23) of the interventions implemented telehealth into their design, primarily using remote video calls between a physical therapist and the participants' parents.

Conclusions: The findings of this review provide a foundation for designing evidence-based, scalable physical activity programs tailored to the unique needs of individuals with RTT. The integration of telehealth strategies offers a promising avenue for enhancing accessibility and caregiver engagement.

(JMIR Rehabil Assist Technol 2026;13:e91550) doi:10.2196/91550

KEYWORDS

Rett syndrome; physical activity; neurodevelopmental disorder; rehabilitation; telehealth

XSL•FO
RenderX

Introduction

Rett syndrome (RTT) is a rare progressive neurodevelopmental disorder that is prevalent among female individuals [1,2]. RTT is characterized by impairments in motor function, communication, and cognitive abilities [3-5]. Half of individuals with RTT are older than 50 years [6]. Considering that mortality in RTT has been linked with several preventable factors such as declines in physical function and cardiorespiratory issues [7], there is a need to identify interventions that can improve and manage physical health and wellness.

Over 40 years of research have found that leisure-time physical activity (LTPA) participation has been identified as a key behavior for improving and managing physical and mental health among people with mobility disabilities [8,9]. LTPA refers to physical activities performed outside of regular work, including exercise; sports; and physically active hobbies such as walking for exercise, running, and similar activities done during discretionary time [10]. While there is a growing volume of research related to LTPA in more prevalent disability groups (eg, stroke, Parkinson disease, cerebral palsy, and multiple sclerosis) [8,9,11,12], less is known about its potential impact in preserving or improving health in individuals with RTT.

Although there are systematic reviews that highlight the multifaceted benefits of specific rehabilitation interventions in people with RTT [13-18], there has not been a review of LTPA interventions for this population. Compiling a single resource of LTPA interventions for RTT can help health professionals and caregivers quickly identify programs that extend outside the rehabilitation context and promote physical health and wellness across the life span. To provide a foundation that can guide future research in LTPA and telehealth for individuals with RTT, the objectives of this scoping review were as follows:

1. To aggregate and summarize the state of the science regarding LTPA for people with RTT across the lifespan
2. To describe LTPA intervention characteristics and their potential outcomes on physical health, function, and fitness
3. To explore telehealth technologies and their use in LTPA interventions

Methods

Study Design

This scoping review was conducted using the methodological framework developed by Arksey and O'Malley [19], which consists of (1) identifying a research question; (2) identifying relevant studies; (3) selecting studies; (4) charting the data; (5) collating, summarizing, and reporting the results; and (6) an optional consultation exercise. Additionally, recommendations and considerations to improve scoping reviews in health research were implemented [20,21]. This research adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for reporting, and a completed PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist can be found in Multimedia Appendix 1.

Review Questions

The population, intervention, comparison, and outcome framework [22] was used to develop the guiding questions for this scoping review: (1) What is the state of the science regarding LTPA interventions for individuals with RTT? (2) What are the intervention delivery strategies and their outcomes, with a specific focus on health-related fitness [23] (cardiorespiratory fitness, muscular strength, muscular endurance, body composition, and flexibility)? (3) What telehealth strategies have been implemented?

Inclusion Criteria

Studies were deemed eligible for inclusion if they met the following criteria: (1) publications that were primary sources of original, peer-reviewed research; (2) LTPA interventions that incorporated the deliberate use of repetitive large body movements to improve, maintain, or change a component of health (exercise, recreation, and conditioning); (3) interventions that targeted people with RTT and/or caregivers; and (4) LTPA interventions that could be delivered in the community outside a rehabilitation setting and did not require the supervision of a licensed therapist. This study included all types of research designs, from case studies to randomized controlled trials (RCTs). Gray literature articles were excluded from review.

Search Process

A preliminary search of the PubMed database, Cochrane Database of Systematic Reviews, and Scopus database was conducted (CM), and no current or underway systematic or scoping reviews on the topic were identified. The text words contained in the titles and abstracts of relevant articles and the index terms used to describe the articles were used to develop a full search strategy.

The article search was conducted (CM) for all years up to September 30, 2025, in the following databases: MEDLINE, Scopus, and CINAHL Plus. Hand searches also included research found through Google Scholar and from the reference lists of existing RTT reviews.

Search Strategy

Article searches were generally conducted by integrating RTT terms (eg, "Rett syndrome" and "neurodevelopmental disorder") and LTPA key terms (eg, "physical activity," "exercise," and "sedentary activity"). Boolean operators were used to connect RTT terms with physical activity terms depending on the database. The MEDLINE database search string was as follows: "Rett syndrome AND physical activity."

Study and Source of Evidence Selection

Following the search, all identified citations were collated and uploaded into EndNote 2025 (Clarivate Analytics), and duplicates were removed. Titles and abstracts were screened by 2 independent reviewers (CM and BL) for assessment against the inclusion criteria for the review. Potentially relevant sources were retrieved in full text and saved as a PDF to a shared digital folder. Rationales for the exclusion of articles were recorded. Any disagreements that arose between the reviewers at each stage of the selection process were resolved through discussion with an additional reviewer (JR) during weekly meetings. The

XSL•FO
RenderX

results of the search and the study inclusion process are reported in full and presented in a PRISMA flow diagram.

Data Charting

Data were extracted from the full texts of the articles by 3 reviewers. Extraction was guided by a data template in Microsoft Excel. One reviewer (CM) extracted data from all articles. Two reviewers (RY and AW) each reviewed half of the total articles as a way to ensure that all articles were reviewed by 2 reviewers. After reviewers charted data from the first 2 studies, the research team met to resolve any disputes and aggregate the data extraction results. Extracted data included specific details about the participants, concept, context, study methods, interventions, outcomes, and key findings relevant to the review questions. The final data extraction chart was cross-checked for accuracy by 2 analysts (CM and BL).

Data Analysis and Presentation

The data were organized graphically and in tabular form. These data included study parameters (design and country of origin), participant demographics and clinical characteristics (age, RTT severity, and ambulatory ability), intervention characteristics (duration, type, frequency, setting, and mode of delivery), outcomes and measures, and telehealth characteristics where applicable (mode of delivery and contact frequency). In instances in which there were multiple publications that were associated with a single trial, the extracted data were considered as a single article for the purposes of summation calculation or frequency counts.

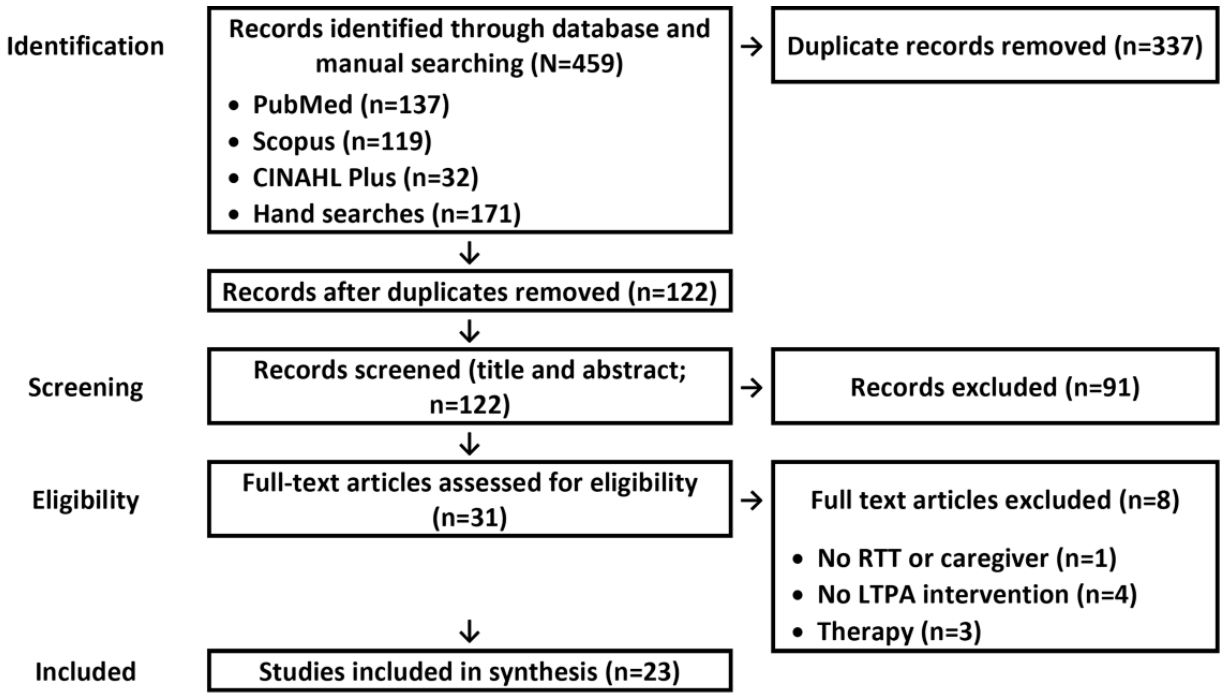
A draft of the preliminary results and findings of this work was reviewed by the mother of an individual with RTT. Consistent with expert recommendations [20], she provided feedback to confirm whether the findings aligned with her experiences, identify any misinterpretations, suggest alternative conceptualizations, and add context to this work. Her feedback was documented in writing through email and integrated into this final work.

Results

Search Results

The search strategy yielded 459 records from 122 original research studies. After screening, of the 459 records, 31 (7%) were assessed for eligibility. Of these 31 publications, 8 (26%) were excluded due to not targeting individuals with RTT or caregivers (n=1) [24], not incorporating an LTPA intervention (n=4) [17,25-27], or administering an intervention that required a licensed therapist (n=3) [28-30]. After assessment, 23 articles were included in the final synthesis (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the scoping review. LTPA: leisure-time physical activity; RTT: Rett syndrome.



Study Parameters

General study parameters are shown in Table 1. Most studies (21/23, 91%) were conducted in the last 20 years. The most common research designs were case studies (11/23, 48%) and within-group repeated measures (9/23, 39%). Among the 23 studies, 3 (13%) randomized designs were identified, with only 2 (9%) being RCTs. Almost every study reported at least one favorable result of their interventions, with the most common benefits being improved gross motor function (eg, sitting independently and standing transitions), improved walking ability, or an increase in physical activity or decrease in sedentary time.

XSL•FO
RenderX

Table 1. General study parameters.

Study	Year	Design	Country	Notable results
Bumin et al [31]	2003	Case report	Turkey	Hand stereotypies reduced, feeding and hand skills improved, walking balance improved, environmental interaction increased, hyperactive behavior decreased, and anxiety decreased
Schaefer-Campion and Johnson [32]	2015	Case study	United States	Walked furthest using a toy shopping cart and anterior walker compared to other walking aids
Lerma-Castaño et al [33]	2024	Case study	Colombia	Improved gross motor function, balance, obstacle avoidance, and behavior
Downs et al [34]	2018	Randomized controlled trial	China	Improved gross motor function and BMI slightly decreased
Downs et al [35]	2023	Randomized controlled trial	Australia and Denmark	Sedentary time decreased
Imamura et al [36]	2020	Case study	Japan	Improved 10-m walk test performance
Kapel et al [18]	2022	Case study	Slovenia	Improved gross motor function
Lotan et al [37]	2004	Repeated measures	Israel	Improved gross motor function and cardiorespiratory fitness
Lotan and Barmatz [38]	2009	Case study	Israel	Improved communication and bodily control
Lotan et al [39]	2015	Case study	Israel	N/A[a]
Lotan et al [40]	2021	Repeated measures	Italy	Improved motor function, rehabilitation goal achievement, and good parental satisfaction
Lotan et al [41]	2021	Repeated measures	Ireland	Favorable telehealth rating by parents and positive rehabilitation goal attainment
Panzeri et al [42]	2024	Repeated measures	Italy	Improved endurance, participant preference for autonomous activities, and good parental satisfaction
Rocco et al [43]	2023	Case study	United States	Improved participation
Romano et al [44]	2022	Repeated measures	Italy	Most rehabilitation goals attained, improved gross motor function, increased physical activity, and higher motor functioning correlated with higher PA[b] level
Romano et al [45]	2022	Repeated measures	Italy	Scoliosis progression prevented and improved gross motor function
Romano et al [46]	2022	Repeated measures	Italy	Positive rehabilitation goal achievement; improved gross motor function; and high usefulness, adherence, and satisfaction
Romano et al [47]	2025	Repeated measures	Italy	Improved gross motor function
Stahlhut et al [48]	2020	Repeated measures	Denmark	Intervention rated as feasible; reduced sedentary time; and increased step count, walking capacity, quality of life, and rehabilitation goal attainment
Stasolla and Caffo [49]	2013	Case study	Italy	Increased ambulation and happiness and reduced stereotypies
Stasolla et al [50]	2018	Case study	Italy	Improved functional ability, mood, and social relationships
Escobar Torres et al [51]	2019	Case study	Spain	Improved functional ability, mood, social relationships, mobility, strength, and endurance
Zwilling et al [52]	2022	Randomized between-group intervention	Italy	COVID-19 lockdown did not negatively affect mothers' well-being

[a]N/A: not applicable.

[b]PA: physical activity.

Participants

Participant information is shown in Table 2. A total of 200 individuals with RTT were enrolled across the 23 included studies. In total, 13% (3/23) of the publications overlapped and reported separate results from the same sample of participants [40,44,52]. Sample sizes ranged from 1 to 42. Adults were underrepresented, with the mean participant age being 13 (SD 7) years and ages ranging from as young as 2 up to 48 years. A total of 65% (130/200) of the participants were ambulatory (independent or with assistance); 25% (50/200) were nonambulatory; and for 10% (20/200), ambulation was not reported or unknown. In total, 39% (9/23) of the studies reported a measure of RTT severity, which included at least one

XSL•FO
RenderX

participant with severe RTT. Only 22% (2/9) of these studies enrolled participants with mild RTT.

Table 2. Sample information (N=23).

Study	Year	Sample size, n	Mean age (y)	Age range (y)	Ambulatory	RTT[a] severity	Waitlist control
Bumin et al [31]	2003	1	11	N/A[b]	1 yes	—[c]	—
Schaefer-Campion and Johnson [32]	2015	1	5	N/A	1 yes	—	—
Lerma-Castaño et al [33]	2024	1	10	N/A	1 yes	—	—
Downs et al [34]	2018	12	3	2-6	9 yes; 3 no	—	Crossover design
Downs et al [35]	2023	39	20	6-41	39 yes; 0 no	—	Control group
Imamura et al [36]	2020	1	17	N/A	1 yes	—	N/A
Kapel et al [18]	2022	5	31	17-44	3 yes; 2 no	—	N/A
Lotan et al [37]	2004	4	10	8-11	4 yes	—	Multiple baseline
Lotan and Barmatz [38]	2009	1	5	N/A	1 no	—	—
Lotan et al [39]	2015	1	—	N/A	1 yes	—	—
Lotan et al [40][d]	2021	42	15.7	2-41	21 yes; 21 no	Moderate to severe	Multiple baseline
Lotan et al [41]	2021	5	9	5-18	4 yes; 1 no	—	—
Panzeri et al [42]	2024	9	8.6	6-44	9 yes	Moderate to severe	—
Rocco et al [43]	2023	1	8	N/A	1 yes	—	—
Romano et al [44][d]	2022	42	15.7	2-41	21 yes; 21 no	Moderate to severe	Multiple baseline
Romano et al [45]	2022	20	15.6	3-39	Unknown	Moderate to severe	
Romano et al [46]	2022	13	17.9	3-35	8 yes; 5 no	Mild to severe	Multiple baseline
Romano et al [47]	2025	20	11.5	6-16	9 yes; 11 no	Mild to severe	Multiple baseline
Stahlhut et al [48]	2020	14	18.7	5-48	9 yes; 5 no	Moderate to severe	Multiple baseline
Stasolla and Caffo [49]	2013	2	9.5	2-17	2 yes	—	Multiple baseline
Stasolla et al [50]	2018	5	15	13-17	5 yes	Severe	Multiple baseline
Escobar Torres et al [51]	2019	3	Unknown	4-7	2 yes; 1 no	—	—
Zwilling et al [52][d]	2022	42	15.7	2-41	21 yes; 21 no	Moderate to severe	Multiple baseline

[a]RTT: Rett syndrome.

[b]N/A: not applicable.

[c]Indicates staggered introduction of the intervention to establish experimental control.

[d]Indicates overlap in publications stemming from a single sample.

Intervention Characteristics

The intervention characteristics are shown in Table 3. The mean intervention duration was 30 (SD 35) weeks, ranging from a single week to 3 years. The most common setting was in the home and community, followed by physical therapy clinics. Most interventions were delivered by health professionals directly or indirectly as an activity program designed for a caregiver or parent to deliver. Interventions were delivered from twice per week to daily, with the typical frequency being 5 times per week. Mean session duration was 55 (SD 36) minutes per week, with some durations as short as 5 to 10 minutes. The types of interventions varied. Some studies (20/23, 87%) included therapy-driven programs, hydrotherapy, walking, exercise training, and motor activities. While most studies (19/23, 83%) integrated a single type of intervention, some (4/23, 17%) were multimodal. For example, one study included 5 different intervention types (motor activities, hippotherapy, hydrotherapy, exercise training, and walking) [18]. While intervention characteristics were typically reported in detail (duration, setting, and delivery), protocols were often superficially described and lacked sufficient details to allow for replication. For example, frequency was often reported as recommended times per week without providing actual adherence. Additionally, time per session was typically reported as a range that could vary based on progression or unknown factors. Finally, most interventions relied on tailored, individualized exercise programs that listed the prescribed activities but did not provide specific dose information or movement adaptations.

XSL•FO
RenderX

Table 3. Intervention characteristics.

Study	Year	Intervention length	Setting	Delivery	Frequency (d per wk)	Duration (min per session)	Modality (activity)	Assessment
Bumin et al [31]	2003	8 wk	Swimming pool	Therapist	2	Unknown	Hydrotherapy (Halliwick concept)	Baseline and after the intervention
Schaefer-Campion and Johnson [32]	2015	24 wk	Elementary school	Special education physiotherapist	4	10 to 15	Walking trials with various assistive devices	—[a]
Lerma-Castaño et al [33]	2024	12 mo	Clinic	Physical therapist and behavioral therapist	3	45	PNF[b], exercise activities, behavioral therapy, and physiotherapy	Baseline and after the intervention
Downs et al [34]	2018	6 mo	Clinic	Physical therapist	6	120 to 180	Motor activity with environmental stimulation	Baseline, after the intervention, 2-mo follow-up, and 4-mo follow-up
Downs et al [35]	2023	12 wk	Home, community, and school based	Caregiver	4	Variable	Standing and walking activities	Baseline and after the intervention
Imamura et al [36]	2020	4 mo	Clinic (weekdays) and home and community (weekends)	Physical therapist (weekdays) and caregiver (weekends)	7	40 at the clinic and 80 at home and in the community	Physical therapy, occupational therapy (clinic), and walking (community)	Each weekend day (MVPA[c] and steps) and completion of first and final week of each intervention month (10-m walk test)
Kapel et al [18]	2022	12 mo	Various	Physical therapists	7	30 to 60	Neurodevelopmental, hippotherapy, hydrotherapy, exercise training, and walking	Baseline and after the intervention
Lotan et al [37]	2004	8 wk	Educational facility	National service person	7	5 to 30	Treadmill walking	Baseline, before the intervention, and after the intervention
Lotan and Barmatz [38]	2009	3 y	Clinic pool	Physical therapist	Unknown	Unknown	Hydrotherapy	Summation of reports for years 1, 2, and 3
Lotan et al [39]	2015	3 mo	Residential care facility	Caregiver	7	Unknown	Walking program enhanced with ABA[d]	Baseline and every session
Lotan et al [40]	2021	12 wk	Home based	Caregiver	5	60	Active and passive exercise and functional motoring activity	Baseline, before the intervention, after the intervention, and 3-mo follow-up
Lotan et al [41]	2021	6 mo	Home based	Caregiver	Unknown	Unknown	Physical therapy	Baseline and after the intervention
Panzeri et al [42]	2024	1 wk	Laboratory	Physical therapist	4	Unknown	Treadmill walking with semi-immersive virtual reality (GRAIL)	Baseline and every visit
Rocco et al [43]	2023	12 wk	Clinic	Physical therapist	3	60	Motor play activity with Ayres Sensory Integration	Every session
Romano et al [44]	2022	3 mo	Home based	Caregiver	5	60	Active and passive exercise and functional motoring activity	Baseline, before the intervention, after the intervention, and 3-mo follow-up
Romano et al [45]	2022	6 mo	Home based	Caregiver	5	60	Motor activities and postural program	Baseline and after the intervention

XSL•FO
RenderX

Study	Year	Intervention length	Setting	Delivery	Frequency (d per wk)	Duration (min per session)	Modality (activity)	Assessment
Romano et al [46]	2022	3 mo	Home based	Caregiver	5	60	Motor activities	Baseline, before the intervention, after the intervention, and 3-mo follow-up
Romano et al [47]	2025	10 mo	Home based	Caregiver	5	60	Postural and rehabilitation program	Baseline, before the intervention, and after the intervention
Stahlhut et al [48]	2020	12 wk	Home and community based	Caregiver	7	Unknown	Standing and walking activities	Baseline, before the intervention, after the intervention, and 3-mo follow-up
Stasolla and Caffo [49]	2013	6 mo	Home based	Research assistants	4	60	Microswitch-based program	Every session
Stasolla et al [50]	2018	3 mo	Rehabilitation facility	Research assistants	5	60 to 90	Microswitch-based program	Every session
Escobar Torres et al [51]	2019	20 mo	Community pool	Trainer	3	30	Hydrotherapy (WaterFit MITAF[e] program)	Beginning of years 1 and 2 (functional performance) and end of year 1 (physical)
Zwilling et al [52]	2022	12 wk	Home based	Caregiver	5	60	Motor activity program	Baseline, before the intervention, after the intervention, and 3-mo follow-up

[a]Not applicable.

[b]PNF: Proprioceptive Neuromuscular Facilitation.

[c]MVPA: moderate to vigorous physical activity.

[d]ABA: Applied Behavior Analysis.

[e]MITAF: Integral Method of Functional Aquatic Work.

Measured Outcomes

The primary and secondary outcomes and measures reported in the included studies are shown in Table 4, and frequencies are shown in Table 5. The most frequent primary or secondary outcome was gross motor function (11/23, 48%), followed by LTPA (6/23, 26%) or walking ability (6/23, 26%). Several of the studies (7/23, 30%) assessed behavior, goal attainment, caregiver satisfaction and well-being, fine motor function, or communication as a primary or secondary outcome. Health-related fitness outcomes were not measured thoroughly. Some studies (9/23, 39%) reported body composition, balance, quality of life, flexibility, cardiorespiratory health, intervention feasibility, or mood; however, these variables were typically secondary or tertiary measures. Additionally, very few objective measures were used to support the primary or secondary objectives, with most of the studies (21/23, 91%) relying on clinical observational performance or self-report (parent or caregiver) measures. There were no studies that reported direct measures of strength or cardiorespiratory health.

XSL•FO
RenderX

Table 4. Primary and secondary outcomes and measures.

Study	Year	Primary outcomes	Secondary outcomes	Measures
Bumin et al [31]	2003	Stereotypical hand movement, hand function, and feeding skills	Gait, balance, hyperactive behavior, communication, and social interaction	Observation
Schaefer-Campion and Johnson [32]	2015	Walking ability	Reason for discontinuing use	Initiation of walking, distance walked, and observation
Lerma-Castaño et al [33]	2024	Gross motor function	Behavior	EAD-3[a] and behavioral observation scale
Downs et al [34]	2018	Gross motor function	BDNF[b]	RSGMS[c], BMI, sleep disturbance scale for children, RTT[d] behavior questionnaire (mood subscale), height, and blood level of BDNF protein
Downs et al [35]	2023	Sedentary time and number of steps	QOL[e] and behavior	Accelerometer, step count, Quality of Life Inventory–Disability, sleep disturbance scale, and Mood and fear/anxiety subscales of the Rett Syndrome Behavior Questionnaire
Imamura et al [36]	2020	Time in MVPA[f]	Steps	Step and activity counts and 10-m walk test
Kapel et al [18]	2022	Gross motor function	N/A[g]	RSGMS and Gross Motor Function Measure–88
Lotan et al [37]	2004	Cardiorespiratory fitness	Gross motor function	RHR[h] to peak HR[i] and motor function scale (researcher made)
Lotan and Barmatz [38]	2009	Gross motor function	Communication and behavior	Observation
Lotan et al [39]	2015	Daily steps	—[j]	Accelerometer
Lotan et al [40]	2021	Parental satisfaction	Goal attainment and gross motor function	GAS[k], questionnaire, and semistructured interview
Lotan et al [41]	2021	Feasibility	Goal attainment and parental satisfaction	GAS and satisfaction survey
Panzeri et al [42]	2024	Feasibility	Suitability, happiness, endurance, speed, and attention focus	Happiness index, time played until discomfort, peak treadmill speed, observation, and SEQ[l]
Rocco et al [43]	2023	Barriers to and facilitators of participation	N/A	Observation
Romano et al [44]	2022	Goal attainment and gross motor function	PA[m] level	GAS, RESMES[n], and mBAR[o]
Romano et al [45]	2022	Scoliosis severity	PA level and gross motor function	Clinical scoliosis score, mBAR, and RESMES
Romano et al [46]	2022	Gross motor function and goal attainment	Caregiver satisfaction, scoliosis severity, and ROM[p]	GAS, RESMES, joint angle measurement, clinical scoliosis score, and caregiver survey
Romano et al [47]	2025	Scoliosis severity	Gross motor function	Clinical scoliosis score and RESMES
Stahlhut et al [48]	2020	Feasibility, sedentary time, and daily PA	Gross motor skills, walking capacity, and QOL	Step and activity counts, caregiver activity diary, RSGMS, 2-min walk test, and QOL—disability
Stasolla and Caffo [49]	2013	Walking ability	Happiness and stereotyped behavior	Total microswitch activations, indexes of happiness (observed), and observation
Stasolla et al [50]	2018	Walking ability	Happiness, stereotyped behavior, and social validation	Total microswitch activations, indexes of happiness (observed), and observation
Escobar Torres et al [51]	2019	Resting HR and exercise HR	Body composition, joint mobility, ADLs[q], and functional independence	HR, skinfolds, Barthel ADL index, functional independence measure, and clinical and psychological evaluation

XSL•FO
RenderX

Study	Year	Primary outcomes	Secondary outcomes	Measures
Zwilling et al [52]	2022	Mothers' well-being	N/A	Caregiver well-being scale

[a]EAD-3: Evaluation of Aquatic Abilities–Third Edition.
[b]BDNF: brain-derived neurotropic factor.
[c]RSGMS: Rett Syndrome Gross Motor Scale.
[d]RTT: Rett syndrome.
[e]QOL: quality of life.
[f]MVPA: moderate to vigorous physical activity.
[g]N/A: not applicable.
[h]RHR: resting heart rate.
[i]HR: heart rate.
[j]Not available.
[k]GAS: Goal Attainment Scaling.
[l]SEQ: Suitability Evaluation Questionnaire.
[m]PA: physical activity.
[n]RESMES: Rett Syndrome Motor Evaluation Scale.
[o]mBAR: modified Bouchard activity record.
[p]ROM: range of motion.
[q]ADL: activity of daily living.

Table 5. Frequency of outcome measures across included studies (n=23).

Outcome measure	Frequency, n (%)
Gross motor function	11 (47.8)
Physical activity	6 (26.1)
Walking ability	6 (26.1)
Behavior	5 (21.7)
Caregiver satisfaction and well-being	5 (21.7)
Fine motor function	4 (17.4)
Goal attainment	4 (17.4)
Communication	3 (13)
Mood	2 (8.7)
Feasibility	2 (8.7)
Cardiorespiratory health	2 (8.7)
Flexibility	2 (8.7)
Quality of life	2 (8.7)
Balance	1 (4.3)
Body composition	1 (4.3)

The studies that reported gross motor function were split between the use of RTT-specific motor assessment tools, including the Rett Syndrome Gross Motor Scale [53] and the Rett Syndrome Motor Evaluation Scale [54]. Physical activity was typically measured through step or activity counts using an accelerometer; however, 9% (2/23) of the studies estimated LTPA using a modified Bouchard activity record [55].

Telehealth Implementation

Among the 23 studies included in this review, 9 (39%) interventions were delivered through telehealth. These studies accounted for 77% (153/200) of the total participant population. Ambulatory participants made up approximately 59% of the total telehealth enrollment. Every telehealth study used a remote video call delivered by a health professional to a caregiver. Only one study augmented the video calls with an online monitoring platform (smartphone app) [47]. The most common frequency of video calls was once every other week (67%). Several studies only used video calls as needed or monthly. Very few telehealth procedural details (eg, communication, behavior change, or technical implementation strategies) were reported across these studies to support replication.

XSL•FO
RenderX

Discussion

State of Research

This scoping review summarized 23 studies involving LTPA for individuals with RTT, with a secondary aim of charting telehealth implementation. The findings reveal a growing body of evidence supporting the feasibility and potential benefits of LTPA interventions across diverse settings, delivery modes, and outcome domains. We discuss the current state of research on LTPA interventions, telehealth implementation, expected benefits, and recommendations for future research that could enhance LTPA interventions for individuals with RTT.

These study findings, similar to those of a recent scoping review of rehabilitation therapies, suggest a lack of studies that could be classified as high-quality evidence [56]. Regarding quality, there appeared to be 2 key limitations: sample size and study design. While many global regions were represented in our review (Table 1), study samples were predominantly small. Over half (12/23, 52%) of the studies included in this review did not incorporate a control or comparison group, and many studies (9/23, 39%) instead used a multiple-baseline design in which the intervention was introduced at different times across participants, settings, or behaviors to demonstrate experimental control (Table 2). Single-group within-subject designs may be appropriate for evaluating the effects of LTPA interventions in neuromuscular disorders, particularly in the context of small sample sizes (eg, 10-40 participants) [57]. On a positive note, our findings suggest that several research design decisions across these studies were consistent with recommendations from the International Rare Diseases Research Consortium for small-population clinical trials, including making use of longitudinal data to show how treatment effects evolve over time, using multiple end points, adjusting follow-up outcomes for baseline values rather than relying on a single change-from-baseline measure, and using registry data to aid in sample size estimates [58]. The only RCTs (2/23, 9%) were implemented by one researcher [34,35]. These studies were able to demonstrate an increase in gross motor skills with environmental enrichment and a small reduction in sedentary time through a telehealth-supported program.

Disappointingly, low rates of physical activity are common in RTT, particularly among people with RTT who have a mobility disability and a higher age [59,60]. A notable finding was that several types of home- and community-based LTPA interventions were successful in achieving their intended outcomes. An encouraging finding was that these interventions were adapted to the needs of individuals with RTT who were ambulatory and nonambulatory. Nonambulatory participants represented 65% of the sample, which was representative of the population [61]. Such adaptations and programs were particularly noteworthy for their inclusion of individuals with more severe levels of disability. However, the lack of depth in reporting intervention and implementation characteristics limits replication of current research efforts in this area. We encourage future LTPA research in RTT to provide more transparent details regarding intervention tailoring strategies, along with successful and unsuccessful lessons that have been learned from intervention delivery.

We found that caregivers were often the primary target for LTPA interventions. This approach likely reflects the severity of the children's disability, which limits the feasibility of direct participation or communication. Although caregivers play a vital role in fostering LTPA, particularly when communication difficulty in the child is a concern, caregivers experience substantial burden with LTPA interventions, which can manifest in various barriers, including limited time and resources, difficulty obtaining mobility aids, few trained professionals, and a lack of knowledge [62]. Caregivers of children or adult patients with RTT report barriers to LTPA that are consistent with those of other disability groups [62,63], including limited resources in the community, lack of transportation, lack of trained professionals, lack of suitable LTPA modalities and resources, and lack of knowledge [64,65]. Therefore, LTPA programs that are accessible and adaptable to the unique needs of individuals with RTT may promote LTPA and reduce caregiver burden [4].

Many of the targeted outcomes reported across the studies in this review were not objectively measured. There was a lack of health and fitness outcomes that are typical in LTPA interventions. For instance, there was not a single example of an objective muscle strength measure even though arm and leg muscle area measurements are significantly lower in female patients with RTT than in their male counterparts [3]. Gross motor function was a consistent positive finding [18,33,34]. There was little evidence to demonstrate an improvement in cardiopulmonary health, which is concerning because RTT is associated with cardiac arrhythmias that are a potential cause of sudden death [3]. Therefore, benefits as a result of LTPA for people with RTT should be interpreted with caution.

A notable gap identified in this study was the limited information on telehealth implementation. The use of telehealth as an intervention delivery platform may expand intervention reach and address enrollment barriers. Delivery of interventions through telehealth has been shown to be effective and important in neurodevelopmental rehabilitation and disability management [66,67]. Remote interventions in this review demonstrated positive goal achievement and caregiver satisfaction [35,40,41,47,48]. Consistent with findings across all included studies, we observed little to no direct communication with participants with RTT via remote telecommunication. While impaired communication is common in individuals with RTT [68], a major concern of their caregivers [69], and a barrier to LTPA participation [62], it is unclear whether direct participant communication through telehealth is feasible. For example, one study demonstrated significant potential in supplementing telehealth coaching using alternative augmentative communication [70]. While direct interaction with participants is typical in telehealth exercise interventions for individuals with disabilities [71], achieving meaningful communication with participants with RTT may require additional support from speech-language therapists or a family member. Although the intervention characteristics were often superficially reported across studies, the description of telehealth delivery and engagement of participants with RTT was even more limited.

XSL•FO
RenderX

A deeper understanding of the factors contributing to telehealth implementation success and failure could help guide future research in this area.

Consultation Exercise

The consultation exercise reinforced and contextualized several key gaps identified in the literature. The consulted caregiver emphasized that effective LTPA interventions for individuals with RTT must be tailored to age and ambulation status, echoing the heterogeneity observed across the included studies. Their insights highlighted the potential value of telehealth in supporting participation and accountability for families, particularly given the heavy caregiving demands that limit opportunities for structured physical activity. Importantly, the caregiver identified a critical service gap wherein adolescents and adults with RTT often lose access to outpatient therapy services, underscoring the need for accessible, home-based, or remotely delivered programs. The consultation also drew attention to the lack of standardized severity classifications in existing research and the absence of key outcome domains—such as muscle strength, cardiovascular health, balance, and quality of life—that families consider essential for daily functioning. These perspectives align with the broad variability in intervention characteristics and outcome measures found in our review and further support the need for consistent reporting and comprehensive assessment frameworks. Collectively, the consultation emphasized the urgency of developing adaptable, evidence-informed LTPA resources and best practice guidelines across the life span, particularly for nonambulatory individuals and those aging out of pediatric services.

Future Recommendations

The most promising outcomes observed in this scoping review were improvements in physical activity and walking ability. To advance the field and address existing gaps, future research should consider the following. First, studies should include more adults with RTT to better understand age-related changes in LTPA participation and intervention responsiveness across the life span. Second, while the feasibility of interventions was often reported as good, future research should add greater levels of detail in their protocols and modalities to increase the reproducibility of their design. For example, the walking trials administered by Schaefer-Campion and Johnson [32] illustrated nuanced differences in walking aids for a child with RTT that can inform decisions for future studies to foster walking ability. Conducting larger and more frequent multisite RCTs with appropriate control conditions may strengthen causal inferences. Adopting objective measures for aspects of health-related fitness such as strength across studies would facilitate cross-study comparisons and meta-analyses. Because low physical fitness, gait abnormalities, and comorbidities such as scoliosis and osteoporosis can be potential barriers to participation in LTPA [4,72,73], it is important to confirm the short-term efficacy of interventions that may mitigate barriers and increase LTPA in RTT.

Moreover, the addition of telehealth communications and remotely delivered interventions at home may be ideal for overcoming community barriers to LTPA among people with RTT. Telehealth LTPA clinical trials bypass the need for on-site visitation and have reached the largest sample sizes among people with physical disabilities, which would be ideal for rare diseases and conditions such as RTT [74,75].

Limitations

Several limitations should be acknowledged. Many of the included studies partially reported or did not report at all critical variables such as RTT severity, exercise protocols, or outcome frequency, thereby making it difficult to generalize the findings of this scoping review. Due to the lack of randomized trials, this review focused on non-RCTs, and the heterogeneity in study designs, sample sizes, and intervention protocols limits the generalizability of the findings. Classifying interventions as LTPA was challenging because the studies often involved a mixture of community-based exercise and physical therapy. This limitation reflects the field itself rather than being a limitation of this review specifically. Different operational definitions of LTPA could yield different conclusions. The heterogeneity of RTT (eg, variability in severity, age, and functional ability), combined with small sample sizes and diverse study designs, limits comparability across studies and complicates the synthesis of findings and interpretation of overall effects. Reliance on observational measures and caregiver-reported outcomes introduces potential bias. Finally, with a strong focus on pediatric populations, this review highlights the lack of evidence related to adults with RTT despite over half of individuals with RTT living beyond 50 years of age [6].

Conclusions

This review underscores the potential of LTPA interventions to enhance wellness and functional outcomes in individuals with RTT. Telehealth strategies show promise for improving accessibility and caregiver engagement. Nevertheless, further research is needed to establish best practices, optimize delivery models, and ensure inclusivity across the life span. Future studies should incorporate objectively measured health outcomes to verify both short- and long-term benefits while also providing detailed reporting on intervention tailoring and implementation strategies, including factors contributing to both successes and challenges.

Acknowledgments

The authors wish to thank the members of the Adaptive Human Performance Lab within the Center for Engagement in Disability Health and Rehabilitation Sciences, the Suki Foundation, and Children's of Alabama for their help and assistance.

XSL•FO
RenderX

Funding

The research reported in this publication was supported by the Suki Foundation and the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under award P50HD118626. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

Data are available upon request to the corresponding author.

Authors' Contributions

CJM contributed to developing the research aims, designing the analysis, collecting data, performing the analysis, and manuscript writing. JHR contributed to developing the research aims, designing the analysis, manuscript writing, and mentorship. AW contributed to collecting data, performing the analysis, and manuscript writing. RY contributed to collecting data, performing the analysis, and manuscript writing. AA contributed to manuscript writing and mentorship. BL contributed to developing the research aims, designing the analysis, collecting data, performing the analysis, manuscript writing, and mentorship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 173 KB - rehab_v13i1e91550_app1.pdf]

References

1. Rett A. On a unusual brain atrophy syndrome in hyperammonemia in childhood [Article in German]. Wien Med Wochenschr 1966 Sep 10;116(37):723-726. [Medline: 5300597]
2. Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Ann Neurol 1983 Oct;14(4):471-479. [doi: 10.1002/ana.410140412] [Medline: 6638958]
3. Percy AK, Benke TA, Marsh ED, Neul JL. Rett syndrome: the natural history study journey. Ann Child Neurol Soc 2024 Aug 11;2(3):189-205. [doi: 10.1002/cns3.20086]
4. Lotan M, Ben-Zeev B. Rett syndrome. A review with emphasis on clinical characteristics and intervention. ScientificWorldJournal 2006 Dec 06;6:1517-1541 [FREE Full text] [doi: 10.1100/tsw.2006.249] [Medline: 17160339]
5. Neul JL, Kaufmann WE, Glaze DG, Christodoulou J, Clarke AJ, Bahi-Buisson N, et al. Rett syndrome: revised diagnostic criteria and nomenclature. Ann Neurol 2010 Dec;68(6):944-950 [FREE Full text] [doi: 10.1002/ana.22124] [Medline: 21154482]
6. Kirby RS, Lane JB, Childers J, Skinner SA, Annese F, Barrish JO, et al. Longevity in Rett syndrome: analysis of the North American Database. J Pediatr 2010 Jan;156(1):135-8.e1 [FREE Full text] [doi: 10.1016/j.jpeds.2009.07.015] [Medline: 19772971]
7. Tarquinio DC, Hou W, Neul JL, Kaufmann WE, Glaze DG, Motil KJ, et al. The changing face of survival in Rett syndrome and MECP2-related disorders. Pediatr Neurol 2015 Nov;53(5):402-411 [FREE Full text] [doi: 10.1016/j.pediatrneurol.2015.06.003] [Medline: 26278631]
8. Martin Ginis KA, van der Ploeg HP, Foster C, Lai B, McBride CB, Ng K, et al. Participation of people living with disabilities in physical activity: a global perspective. Lancet 2021 Jul 31;398(10298):443-455 [FREE Full text] [doi: 10.1016/S0140-6736(21)01164-8] [Medline: 34302764]
9. Lai B, Lee E, Kim Y, Matthews C, Swanson-Kimani E, Davis D, et al. Leisure-time physical activity interventions for children and adults with cerebral palsy: a scoping review. Dev Med Child Neurol 2021 Feb;63(2):162-171 [FREE Full text] [doi: 10.1111/dmcn.14751] [Medline: 33241561]
10. Steinbach D, Graf C. Leisure time physical activity and sedentariness. In: Kirch W, editor. Encyclopedia of Public Health. Dordrecht, The Netherlands: Springer; 2008.
11. Jeng B, DuBose NG, Martin TB, Šilić P, Flores VA, Zheng P, et al. Updated systematic review and quantitative synthesis of physical activity levels in multiple sclerosis. Am J Phys Med Rehabil 2024 Apr 01;103(4):284-292. [doi: 10.1097/PHM.0000000000002312] [Medline: 37408136]
12. Ernst M, Folkerts AK, Gollan R, Lieker E, Caro-Valenzuela J, Adams A, et al. Physical exercise for people with Parkinson's disease: a systematic review and network meta-analysis. Cochrane Database Syst Rev 2023 Jan 05;1(1):CD013856 [FREE Full text] [doi: 10.1002/14651858.CD013856.pub2] [Medline: 36602886]
13. Fonzo M, Sirico F, Corrado B. Evidence-based physical therapy for individuals with Rett syndrome: a systematic review. Brain Sci 2020 Jun 30;10(7):410 [FREE Full text] [doi: 10.3390/brainsci10070410] [Medline: 32630125]

XSL•FO
RenderX

14. Lim J, Greenspoon D, Hunt A, McAdam L. Rehabilitation interventions in Rett syndrome: a scoping review. Dev Med Child Neurol 2020 Aug;62(8):906-916 [FREE Full text] [doi: 10.1111/dmcn.14565] [Medline: 32472972]
15. Romano A, Caprì T, Semino M, Bizzego I, Di Rosa G, Fabio RA. Gross motor, physical activity and musculoskeletal disorder evaluation tools for Rett syndrome: a systematic review. Dev Neurorehabil 2020 Nov;23(8):485-501. [doi: 10.1080/17518423.2019.1680761] [Medline: 31668104]
16. Amoako AN, Hare DJ. Non-medical interventions for individuals with Rett syndrome: a systematic review. J Appl Res Intellect Disabil 2020 Sep;33(5):808-827. [doi: 10.1111/jar.12694] [Medline: 31833197]
17. Larsson G, Julu PO, Witt Engerström I, Sandlund M, Lindström B. Walking on treadmill with Rett syndrome-effects on the autonomic nervous system. Res Dev Disabil 2018 Dec;83:99-107 [FREE Full text] [doi: 10.1016/j.ridd.2018.08.010] [Medline: 30193160]
18. Kapel A, Kovacic T, Kos N, Velnar T. Impact of a 12-month multifaceted neurological physiotherapy intervention on gross motor function in women with Rett syndrome. J Integr Neurosci 2022 Mar 22;21(2):59 [FREE Full text] [doi: 10.31083/j.jin2102059] [Medline: 35364647]
19. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
20. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
21. Thomas A, Lubarsky S, Varpio L, Durning SJ, Young ME. Scoping reviews in health professions education: challenges, considerations and lessons learned about epistemology and methodology. Adv Health Sci Educ Theory Pract 2020 Oct;25(4):989-1002. [doi: 10.1007/s10459-019-09932-2] [Medline: 31768787]
22. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club 1995;123(3):A12-A13. [Medline: 7582737]
23. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]
24. Borst HE, Townend GS, van Eck M, Smeets E, van den Berg M, Laan A, et al. Abnormal foot position and standing and walking ability in Rett syndrome: an exploratory study. J Dev Phys Disabil 2018;30(2):281-295 [FREE Full text] [doi: 10.1007/s10882-017-9585-6] [Medline: 29563764]
25. Layne CS, Lee BC, Young DR, Glaze DG, Schwabe A, Suter B. Temporal gait measures associated with overground and treadmill walking in Rett syndrome. J Child Neurol 2018 Jan 01;33(10):667-674. [doi: 10.1177/0883073818780471] [Medline: 29926771]
26. Layne CS, Young DR, Lee BC, Glaze DG, Schwabe A, Suter B. Kinematics associated with treadmill walking in Rett syndrome. Disabil Rehabil 2021 Jun;43(11):1585-1593. [doi: 10.1080/09638288.2019.1674389] [Medline: 31613656]
27. Fabio RA, Giannatiempo S, Caprì T, Semino M. Repeated motor training on attention reaching skills and stereotypies in Rett syndrome. Mov Disord Clin Pract 2022 May 19;9(5):637-646 [FREE Full text] [doi: 10.1002/mdc3.13465] [Medline: 35844285]
28. Larsson G, Engerström IW. Gross motor ability in Rett syndrome--the power of expectation, motivation and planning. Brain Dev 2001 Dec;23 Suppl 1:S77-S81. [doi: 10.1016/s0387-7604(01)00334-5] [Medline: 11738846]
29. Jacobsen K, Viken A, von Tetzchner S. Rett syndrome and ageing: a case study. Disabil Rehabil 2001;23(3-4):160-166. [doi: 10.1080/09638280150504234] [Medline: 11247011]
30. Lotan M, Gootman A. Regaining walking ability in individuals with Rett syndrome: a case study. Int J Disabil Hum Dev 2012 May;11(2). [doi: 10.1515/ijdhd-2012-0020]
31. Bumin G, Uyanik M, Yilmaz I, Kayihan H, Topçu M. Hydrotherapy for Rett syndrome. J Rehabil Med 2003 Jan;35(1):44-45 [FREE Full text] [doi: 10.1080/16501970306107] [Medline: 12610848]
32. Schaefer-Campion C, Johnson NL. Fostering ambulation for a preschool child with Rett syndrome: a case report. Phys Occup Ther Pediatr 2015;35(4):354-364. [doi: 10.3109/01942638.2014.899287] [Medline: 24649976]
33. Lerma-Castaño PR, Roldán-González E, Mantilla-Toloza SC, Hernández-Cachaya JP, Romaña-Cabrera LF. Rett syndrome: an approach from physiotherapy and behavioral therapy [Article in Spanish]. Rev Ecuat Neurol 2024;33(2):57-63. [doi: 10.46997/revecuatneurol33200057]
34. Downs J, Rodger J, Li C, Tan X, Hu N, Wong K, et al. Environmental enrichment intervention for Rett syndrome: an individually randomised stepped wedge trial. Orphanet J Rare Dis 2018 Jan 10;13(1):3 [FREE Full text] [doi: 10.1186/s13023-017-0752-8] [Medline: 29321033]
35. Downs J, Blackmore AM, Wong K, Buckley N, Lotan M, Elefant C, et al. Can telehealth increase physical activity in individuals with Rett syndrome? A multicentre randomized controlled trial. Dev Med Child Neurol 2023 Apr;65(4):489-497 [FREE Full text] [doi: 10.1111/dmcn.15436] [Medline: 36284370]
36. Imamura T, Nakayama T, Nakayama J, Iwasaki N. A patient with Rett syndrome maintained motor function by periodic rehabilitation therapy and proactive daily activities. Prog Rehabil Med 2020;5:20200014. [doi: 10.2490/prm.20200014] [Medline: 32844127]
37. Lotan M, Isakov E, Merrick J. Improving functional skills and physical fitness in children with Rett syndrome. J Intellect Disabil Res 2004 Nov;48(Pt 8):730-735. [doi: 10.1111/j.1365-2788.2003.00589.x] [Medline: 15494062]

XSL•FO
RenderX

38. Lotan M, Barmatz C. Hydrotherapy for a young child with Rett syndrome. Review of the literature and a case study. Int J Disabil Hum Dev 2009 Dec 01;8(4):349-358. [doi: 10.1515/IJDHD.2009.8.4.349]
39. Lotan M, Shavit E, Merrick J. Enhancing walking ability in individuals with Rett syndrome through the use of applied behavioral analysis (ABA): review and a case study. Open Rehabil J 2015;8:1-8 [FREE Full text]
40. Lotan M, Ippolito E, Favetta M, Romano A. Skype supervised, individualized, home-based rehabilitation programs for individuals with Rett syndrome and their families - parental satisfaction and point of view. Front Psychol 2021 Sep 16;12:720927 [FREE Full text] [doi: 10.3389/fpsyg.2021.720927] [Medline: 34603144]
41. Lotan M, Downs J, Elefant C. A pilot study delivering physiotherapy support for Rett syndrome using a telehealth framework suitable for COVID-19 lockdown. Dev Neurorehabil 2021 Aug;24(6):429-434. [doi: 10.1080/17518423.2021.1914762] [Medline: 33853477]
42. Panzeri D, Perina M, Biffi E, Semino M, Diella E, Caprì T. Effects of immersive virtual reality with treadmill in subjects with Rett syndrome: a pilot study. Children (Basel) 2024 Sep 11;11(9):1110 [FREE Full text] [doi: 10.3390/children11091110] [Medline: 39334642]
43. Rocco K, Drobnyk W, Bruce S, Soumerai SB. Ayres Sensory Integration therapy for a child with Rett syndrome: a case report. Clin Med Insights Pediatr 2023 Jul 29;17:11795565231188939 [FREE Full text] [doi: 10.1177/11795565231188939] [Medline: 37529622]
44. Romano A, Ippolito E, Favetta M, Lotan M, Moran DS. Individualized remotely supervised motor activity programs promote rehabilitation goal achievement, motor functioning, and physical activity of people with Rett syndrome-a single-cohort study. Int J Environ Res Public Health 2022 Dec 30;20(1):659 [FREE Full text] [doi: 10.3390/ijerph20010659] [Medline: 36612980]
45. Romano A, Ippolito E, Risoli C, Malerba E, Favetta M, Sancesario A, et al. Intensive postural and motor activity program reduces scoliosis progression in people with Rett syndrome. J Clin Med 2022 Jan 22;11(3):559 [FREE Full text] [doi: 10.3390/jcm11030559] [Medline: 35160011]
46. Romano A, Di Rosa G, Tisano A, Fabio RA, Lotan M. Effects of a remotely supervised motor rehabilitation program for individuals with Rett syndrome at home. Disabil Rehabil 2022 Oct;44(20):5898-5908. [doi: 10.1080/09638288.2021.1949398] [Medline: 34282992]
47. Romano A, Rodocanachi Roidi ML, Savini MN, Viganò I, Dziubak M, Pietrogrande L, et al. Effects of a supervised-as-needed home exercise program on scoliosis and motor function in Rett syndrome: a multiple-baseline study. J Clin Med 2025 Mar 11;14(6):1873 [FREE Full text] [doi: 10.3390/jcm14061873] [Medline: 40142681]
48. Stahlhut M, Downs J, Wong K, Bisgaard AM, Nordmark E. Feasibility and effectiveness of an individualized 12-week "uptime" participation (U-PART) intervention in girls and women with Rett syndrome. Phys Ther 2020 Jan 23;100(1):168-179 [FREE Full text] [doi: 10.1093/ptj/pzz138] [Medline: 31584667]
49. Stasolla F, Caffò AO. Promoting adaptive behaviors by two girls with Rett syndrome through a microswitch-based program. Res Autism Spectr Disord 2013 Oct;7(10):1265-1272. [doi: 10.1016/j.rasd.2013.07.010]
50. Stasolla F, Caffò AO, Perilli V, Boccasini A, Damiani R, D'Amico F. Fostering locomotion fluency of five adolescents with Rett syndrome through a microswitch-based program: contingency awareness and social rating. J Dev Phys Disabil 2017 Dec 5;30:239-258. [doi: 10.1007/s10882-017-9582-9]
51. Escobar Torres L, Sanders M, Belenguer Benitez C, Melendez Ortega A. Efficacy of an aquatic exercise program for 3 cases of Rett syndrome. Pediatr Phys Ther 2019 Oct;31(4):E6-E13. [doi: 10.1097/PEP.0000000000000655] [Medline: 31568388]
52. Zwilling M, Romano A, Favetta M, Ippolito E, Lotan M. Impact of a remotely supervised motor rehabilitation program on maternal well-being during the COVID-19 Italian lockdown. Front Psychol 2022 Mar 7;13:834419 [FREE Full text] [doi: 10.3389/fpsyg.2022.834419] [Medline: 35345633]
53. Downs J, Stahlhut M, Wong K, Syhler B, Bisgaard AM, Jacoby P, et al. Validating the Rett Syndrome Gross Motor Scale. PLoS One 2016 Jan 22;11(1):e0147555 [FREE Full text] [doi: 10.1371/journal.pone.0147555] [Medline: 26800272]
54. Rodocanachi Roidi ML, Isaias IU, Cozzi F, Grange F, Scotti FM, Gestra VF, et al. A new scale to evaluate motor function in Rett syndrome: validation and psychometric properties. Pediatr Neurol 2019 Nov;100:80-86. [doi: 10.1016/j.pediatrneurol.2019.03.005] [Medline: 31047758]
55. Bouchard C, Tremblay A, Leblanc C, Lortie G, Savard R, Thériault G. A method to assess energy expenditure in children and adults. Am J Clin Nutr 1983 Mar;37(3):461-467. [doi: 10.1093/ajcn/37.3.461] [Medline: 6829488]
56. Yang D, Robertson HL, Condliffe EG, Carter MT, Dewan T, Gnanakumar V. Rehabilitation therapies in Rett syndrome across the lifespan: a scoping review of human and animal studies. J Pediatr Rehabil Med 2021;14(1):69-96. [doi: 10.3233/prm-200683]
57. Stefanetti RJ, Blain A, Jimenez-Moreno C, Errington L, Ng YS, McFarland R, et al. Measuring the effects of exercise in neuromuscular disorders: a systematic review and meta-analyses. Wellcome Open Res 2020 May 4;5:84 [FREE Full text] [doi: 10.12688/wellcomeopenres.15825.1] [Medline: 32671231]
58. Day S, Jonker AH, Lau LP, Hilgers RD, Irony I, Larsson K, et al. Recommendations for the design of small population clinical trials. Orphanet J Rare Dis 2018 Nov 06;13(1):195 [FREE Full text] [doi: 10.1186/s13023-018-0931-2] [Medline: 30400970]

XSL•FO
RenderX

59. Downs J, Leonard H, Wong K, Newton N, Hill K. Quantification of walking-based physical activity and sedentary time in individuals with Rett syndrome. Dev Med Child Neurol 2017 Jun;59(6):605-611 [FREE Full text] [doi: 10.1111/dmcn.13398] [Medline: 28164278]
60. Stahlhut M, Hill K, Bisgaard A, Jensen AK, Andersen M, Leonard H, et al. Measurement of sedentary behaviors or "downtime" in Rett syndrome. J Child Neurol 2017 Oct;32(12):1009-1013. [doi: 10.1177/0883073817728861] [Medline: 28911278]
61. Downs J, Bebbington A, Jacoby P, Msall ME, McIlroy O, Fyfe S, et al. Mobility skills in Rett syndrome as determined by video analysis. J Paediatr Child Health 2007;43(8):A4-A5. [doi: 10.1111/j.1440-1754.2007.01182.x]
62. Stahlhut M, Esbensen BA, Larsen JL, Bisgaard AM, Downs J, Nordmark E. Facilitators and barriers of participation in "uptime" activities in girls and women with Rett syndrome: perspectives from parents and professionals. Qual Health Res 2019 Mar;29(4):609-619. [doi: 10.1177/1049732318803358] [Medline: 30304998]
63. Schönewolf-Greulich B, Stahlhut M, Larsen JL, Syhler B, Bisgaard AM. Functional abilities in aging women with Rett syndrome - the Danish cohort. Disabil Rehabil 2017 May;39(9):911-918. [doi: 10.3109/09638288.2016.1170896] [Medline: 27206693]
64. Rimmer JH, Riley B, Wang E, Rauworth A, Jurkowski J. Physical activity participation among persons with disabilities: barriers and facilitators. Am J Prev Med 2004 Jun;26(5):419-425. [doi: 10.1016/j.amepre.2004.02.002] [Medline: 15165658]
65. Martin Ginis KA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. Health Psychol Rev 2016 Dec;10(4):478-494. [doi: 10.1080/17437199.2016.1198240] [Medline: 27265062]
66. Hsu N, Monasterio E, Rolin O. Telehealth in pediatric rehabilitation. Phys Med Rehabil Clin N Am 2021 May;32(2):307-317. [doi: 10.1016/j.pmr.2020.12.010] [Medline: 33814060]
67. Mayston M. Telehealth for disability management: what really matters? Dev Med Child Neurol 2021 Feb;63(2):124 [FREE Full text] [doi: 10.1111/dmcn.14734] [Medline: 33393700]
68. Neul JL, Lane JB, Lee HS, Geerts S, Barrish JO, Annese F, et al. Developmental delay in Rett syndrome: data from the natural history study. J Neurodev Disord 2014;6(1):20 [FREE Full text] [doi: 10.1186/1866-1955-6-20] [Medline: 25071871]
69. Neul JL, Benke TA, Marsh ED, Suter B, Silveira L, Fu C, et al. Top caregiver concerns in Rett syndrome and related disorders: data from the US Natural History Study. Research Square Preprint posted online on March 20, 2023 [FREE Full text] [doi: 10.21203/rs.3.rs-2566253/v1] [Medline: 36993737]
70. Kolb RL, McComas JJ, Girtler SN, Simacek J, Dimian AF, Unholz-Bowden EK, et al. Teaching requesting to individuals with Rett syndrome using alternative augmentative communication (AAC) through caregiver coaching via telehealth. J Dev Phys Disabil 2023 Dec;35(6):1063-1090 [FREE Full text] [doi: 10.1007/s10882-023-09894-9] [Medline: 38053943]
71. Dawson R, Oliveira JS, Kwok WS, Bratland M, Rajendran IM, Srinivasan A, et al. Exercise interventions delivered through telehealth to improve physical functioning for older adults with frailty, cognitive, or mobility disability: a systematic review and meta-analysis. Telemed J E Health 2024 Apr 08;30(4):940-950 [FREE Full text] [doi: 10.1089/tmj.2023.0177] [Medline: 37975811]
72. Motil KJ, Geerts S, Annese F, Neul JL, Benke T, Marsh E, et al. Anthropometric measures correspond with functional motor outcomes in females with Rett syndrome. J Pediatr 2022 May;244:169-77.e3 [FREE Full text] [doi: 10.1016/j.jpeds.2022.01.009] [Medline: 35063470]
73. Fu C, Armstrong D, Marsh E, Lieberman D, Motil K, Witt R, et al. Multisystem comorbidities in classic Rett syndrome: a scoping review. BMJ Paediatr Open 2020 Sep 22;4(1):e000731 [FREE Full text] [doi: 10.1136/bmjpo-2020-000731] [Medline: 33024833]
74. Lai B, Lee E, Wagatsuma M, Frey G, Stanish H, Jung T, et al. Research trends and recommendations for physical activity interventions among children and youth with disabilities: a review of reviews. Adapt Phys Activ Q 2020 Apr 01;37(2):211-234. [doi: 10.1123/apaq.2019-0081] [Medline: 32131052]
75. Lai B, Young HJ, Bickel CS, Motl RW, Rimmer JH. Current trends in exercise intervention research, technology, and behavioral change strategies for people with disabilities: a scoping review. Am J Phys Med Rehabil 2017 Oct;96(10):748-761. [doi: 10.1097/PHM.0000000000000743] [Medline: 28398967]

Abbreviations

LTPA: leisure-time physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RCT: randomized controlled trial
RTT: Rett syndrome

XSL•FO
RenderX

Edited by S Munce; submitted 15.Jan.2026; peer-reviewed by S Jiang, Å Elden, M Stahlhut; comments to author 09.Mar.2026; accepted 01.Jun.2026; published 21.Sep.2026.

Please cite as:
Mendonca CJ, Rimmer JH, Wright A, Young R, Ananth A, Lai B
Leisure-Time Physical Activity Interventions for Health Promotion in Individuals With Rett Syndrome: Scoping Review
JMIR Rehabil Assist Technol 2026;13:e91550
URL: https://rehab.jmir.org/2026/1/e91550
doi:10.2196/91550
PMID:42767625

©Christen John Mendonca, James H Rimmer, Ashley Wright, Raven Young, Amitha Ananth, Byron Lai. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 21.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Ankle-Foot Orthosis and Footwear Interactions in Adult Healthy and Patient Populations: Systematic Review of Biomechanical Outcomes

Alexandra Millar[1]; Christian Greve[1,2], PhD; Juha M Hijmans[1], PhD

[1]University of Groningen, University Medical Center Groningen, Department of Rehabilitation Medicine, Groningen, The Netherlands

[2]University of Groningen, University Medical Center Groningen, Department of Human Movement Sciences, Groningen, The Netherlands

Corresponding Author:
Alexandra Millar
University of Groningen, University Medical Center Groningen, Department of Rehabilitation Medicine
Hanzeplein 1
Groningen, 9713 GZ
The Netherlands
Phone: 31 686558487
Email: a.h.millar@umcg.nl

Abstract

Background: Both ankle-foot orthoses (AFOs) and footwear adaptations are used to improve gait across a broad range of neuromuscular disorders. AFOs and footwear can each independently influence gait through the manipulation of certain design parameters, such as AFO footplate stiffness or the use of a forefoot rocker. While the effects of design parameters from AFOs and footwear on gait have been investigated independently, how these parameters interact to influence walking biomechanics remains poorly understood.

Objective: This systematic literature review aims to provide an overview of the interacting effects between AFOs and footwear on gait biomechanics from both a theoretical standpoint and through the synthesis of existing literature.

Methods: A search of existing literature was conducted across 4 major databases. Studies were eligible if they included adult participants from healthy or patient populations. Studies included should have investigated biomechanical gait outcomes recorded during walking with an AFO and at least two different footwear conditions. Quality assessment was performed using the Downs and Black checklist and AFO reporting guidelines.

Results: A total of 14 studies were included in this review. Across these 14 studies, a total of 204 participants (108 healthy, 84 patients after stroke, and 12 unilateral ankle weakness or pain) were included. The majority of these studies were within-subjects designs. Footwear adaptations were primarily changes to the heel-sole differential (HSD; 7/14), followed by forefoot rockers (5/14), with some studies on changes to heel shape and hardness (3/14), and a study on mediolateral wedging. These adaptations were used in combination with primarily rigid AFOs (10/14), with studies also using dorsiflexion-free and frontal plane restrictive AFOs.

Conclusions: Increasing the HSD can be used to increase knee flexion across the stance phase in rigid AFO users, including in the third rocker when HSD is manipulated between the shoe and AFO. Beyond changing the HSD, there is a lack of well-designed experimental studies exploring footwear adaptations. Detailed reporting of both AFO and footwear design choices is necessary for future work in order to make informed statements on biomechanics.

(JMIR Rehabil Assist Technol 2026;13:e93423) doi:10.2196/93423

KEYWORDS

gait; adult; orthotic devices; shoes; biomechanical phenomena; ankle-foot orthoses; AFO-FC; ankle-foot orthosis and footwear combination

XSL•FO
RenderX

Introduction

Background

The ankle-foot complex plays a crucial role in human walking due to its contributions to both balance and forward propulsion [1]. Plantarflexor muscle weakness, due to neuromuscular disorders such as stroke, can disrupt the biomechanics of gait, leading to slower walking speeds and a less efficient gait pattern [2,3]. To facilitate mobility and an efficient walking pattern, ankle-foot orthoses (AFOs) may be prescribed.

There exists a wide range of AFOs in a variety of shapes, materials, and stiffnesses, with varying effects on patient gait biomechanics [4,5]. These design parameters are often used to optimize biomechanical outcomes for individual patients. For example, increasing the bending stiffness of an AFO can reduce knee flexion during the second rocker in patients with crouch gait [6]. Similarly, increasing footplate stiffness can maximize energy storage and return, reducing the energetic cost of walking [7].

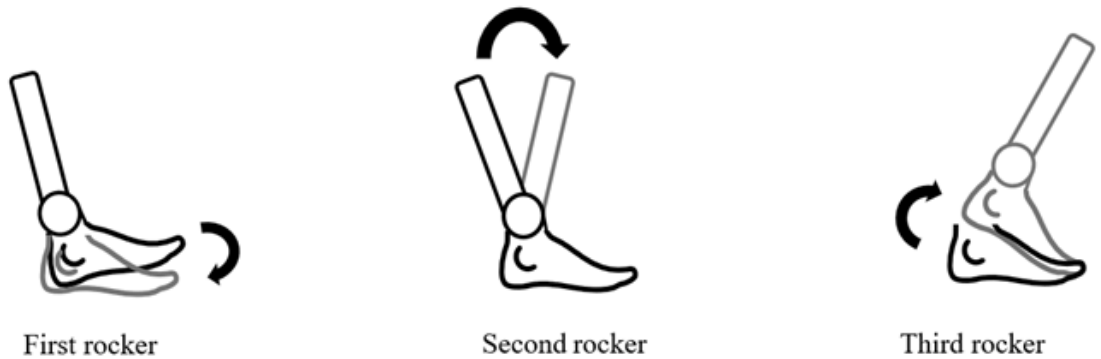
Adapted footwear such as rocker shoes have been investigated independent of AFO use and shown to affect gait kinematics and kinetics [8,9]. As AFOs are typically worn with footwear, it is essential to consider how footwear design interacts with AFOs to influence gait. When combined, footwear and AFO form an integrated orthotic system where the shape and bending stiffness of both devices play a combined role in influencing how forces are transmitted through the lower limb during gait. Consequently, footwear adaptations can alter the biomechanical effects of AFOs on gait. Despite these interactions, the specific impact of different footwear parameters on AFO performance remains unclear [10]. The footwear design can either complement or counteract the intended mechanical function of the AFO. For instance, when the ankle is immobilized with a rigid AFO, increasing the heel-sole differential (HSD) can increase knee flexion during midstance [11]. Conversely, a rocker shoe with a proximal apex position may shorten stance phase duration [12] and reduce the energy storage and return of a carbon fiber AFO, potentially negating its benefits.

Although the importance of the interaction between AFOs and footwear is recognized, most existing research has focused on sagittal plane alignment, particularly the shank-to-vertical angle (SVA), and often within specific patient populations [4,10,11]. While systematic reviews have examined the biomechanical effects of footwear or AFOs independently [13,14], no systematic review to date has examined the biomechanical effects resulting from the interactions between ankle-foot orthosis and footwear combinations (AFO-FCs) in adults. The aim of this review is to provide a systematic overview of the current evidence on AFO-FCs. Specifically, it evaluates how footwear modifications, when paired with AFOs, influence gait biomechanics in both healthy adults, who provide insight into gait adaptations as controlled models, and patients with gait pathology. These findings are intended to inform clinical recommendations and identify directions for future research. A background on ankle-foot biomechanics and the theoretical effects of AFO-FCs in cases of dorsiflexor and plantarflexor weakness is first introduced to provide context for the findings of this review. Existing literature on AFO-footwear interactions in both patient populations and healthy adults is then synthesized to establish an evidence-based overview of the biomechanical outcomes resulting from AFO-footwear interactions.

Locomotion

Human locomotion can be broken down into 2 main phases, stance and swing. The stance phase is typically defined as the moment the heel of the foot strikes the ground until the foot leaves the ground again [1]. During stance, the ankle-foot complex can be likened to the rolling of a wheel as the center of pressure (COP) moves from heel to toe [15]. Within stance, 3 subphases, known as rockers, can be distinguished as the COP progresses from heel to ankle to forefoot in a curved path relative to a fixed lower-leg coordinate system [16]. The period between initial contact and the forefoot touching the ground makes up the first rocker, also known as the heel rocker. Following this is the second, or ankle, rocker, where dorsiflexion takes place, and the foot remains flat on the ground. The third, forefoot rocker, begins as the heel leaves the ground and concludes with toe-off. The 3 rockers are illustrated in Figure 1.

Figure 1. The 3 rockers.



XSL•FO
RenderX

Together, these 3 rockers make up part of a phenomenon known as the "roll-over shape" (ROS). This term refers to the ellipsoidal trajectory of the COP relative to a fixed frame of reference at the ankle [17]. Figure 2 illustrates the transformation from the laboratory reference frame (A) to the ankle-foot reference frame (B). The ROS provides a useful framework for evaluating gait adaptations to AFOs, as it can demonstrate changes in timing and kinematics, providing insight into gait adaptations with a single measure. Figure 3 provides examples of the ROS for healthy individuals (A), two theoretical ROS for patients with calf muscle weakness (B), and visualizations of the ROS when walking with a rigid AFO, both without (C) and with an increased HSD (D). In healthy walking, the ROS is an arc that maintains a relatively constant radius even when walking speed or body weight is altered, due to physiological adaptations of the ankle-foot complex [17,18]. This consistent shape provides a predictable rollover, smoothing step-to-step transitions, allowing humans to walk with minimal energy expenditure [15,19]. Individuals with calf muscle weakness may have a shortened or flattened ROS of the ankle-foot complex compared to the typical arc [20]. Dorsiflexor weakness primarily affects the first rocker, while plantarflexor weakness impairs forward progression of the COP during the second rocker and reduces push-off power during the third rocker [1].

Figure 2. ROS across the 3 rockers. (A) The anteroposterior trajectory of the COP viewed from the global reference frame. (B) The circular trajectory of the COP when the origin is set at the ankle and aligned with the shank, creating the ankle-foot reference frame and forming the roll-over shape. COP: center of pressure; ROS: roll-over shape.

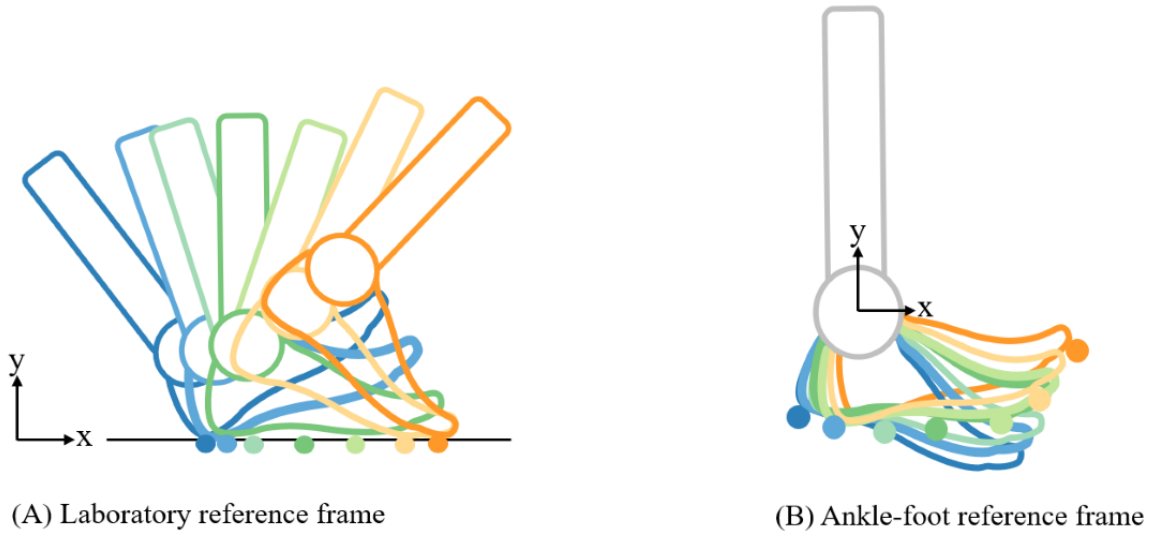
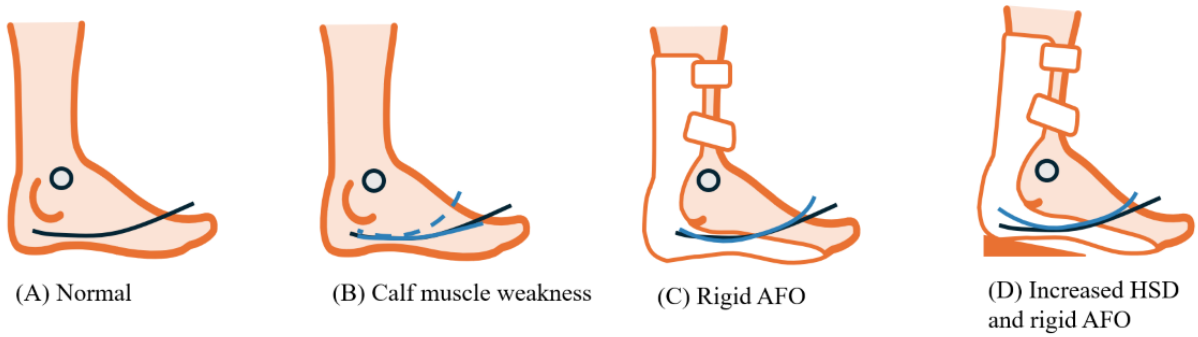


Figure 3. Schematic representation of ROS across clinical and AFO conditions. The gray circle indicates the origin of the ankle-foot reference frame. The black line indicates the normal ROS, and blue lines indicate the ROS under different conditions. (A) ROS for normal healthy walking, based on data from Hansen et al (2004) [18]. (B) Two theoretical ROS for patients with calf muscle weakness. The solid line represents patients with slow COP progression, while the dashed line represents patients with rapid tibial inclination and slow COP progression. (C) ROS when using a rigid AFO, based on data from Fatone et al (2007) [20] and Esposito et al (2021) [16]. (D) ROS when using a rigid AFO with increased HSD, based on data from Esposito et al (2021) [16]. AFO: ankle-foot orthosis; COP: center of pressure; ROS: roll-over shape.



AFOs are commonly prescribed to support patients with calf muscle weakness during gait [3,10,22]. While both passive and active AFOs (exoskeletons) exist, this review focuses on passive AFOs, which are more typically used by patients in daily life. Passive AFOs range from solid polypropylene designs to energy-returning carbon fiber leaf springs and can be tailored to control ankle joint motion to restore a more optimal gait pattern. However, while AFOs can improve stability and help re-establish a smooth ROS [20], their increased stiffness and

XSL•FO
RenderX

limited range of plantarflexion motion may reduce push-off power and limit gains in energy efficiency [23].

Footwear plays an important role in gait biomechanics, with design features such as the HSD, rocker soles, and heel hardness influencing how the COP progresses through each rocker relative to the ankle coordinate system [16,24,25]. When combined with AFOs, footwear characteristics can influence orthotic effects. Together, AFO-footwear adaptations offer a powerful means to restore the ROS and improve gait biomechanics in patients with plantar- or dorsiflexor- muscle weakness. However, their effects may differ across the 3 rockers, leading to complex interaction effects between AFO-FCs and gait biomechanics. The following sections summarize the biomechanical effects of the most commonly used AFO-footwear adaptations in the sagittal plane.

First Rocker

During the first rocker, the dorsiflexors generate a resistive moment that controls plantarflexion, gradually lowering the forefoot to the ground [1]. This negative work from the dorsiflexors leads to smooth knee and hip flexion and supports the controlled transition of the center of mass (COM) from one stance limb to the next [19].

When plantarflexion-restrictive AFOs are used to compensate for dorsiflexor weakness, this gradual transition of the COM is disrupted [26]. The reduced dorsiflexor contribution to the energy transfer between limbs and absence of controlled plantarflexion motion increases angular momentum at the knee and hip. This can result in rapid COP progression toward the midfoot, accompanied by increased knee and hip flexion during early stance [27]. Certain footwear adaptations may mitigate the adverse biomechanical effects of absent plantarflexion during the first rocker. For instance, a shoe with a rocker distal to the heel positions the COP more anteriorly at initial contact [28]. This reduces the ground reaction force (GRF) moment arm at the ankle and knee [28], reducing angular momentum and potentially promoting a smoother COM transition in patients using rigid AFOs. In contrast, a rocker proximal to the heel with a small radius combined with a rigid AFO could theoretically increase the GRF moment arm, thereby increasing angular momentum at the knee and hip.

Second Rocker

In healthy gait, once the foot is fully in contact with the ground, the gradually increasing plantarflexor activity generates a plantarflexor moment about the ankle joint, decelerating shank inclination and knee flexion [29]. During this phase, near-isometric activation of the plantarflexors combined with ankle dorsiflexion stretches the Achilles tendon, allowing elastic potential energy to be stored and later released as push-off power in terminal stance [29].

Throughout the second rocker, gradually increasing plantarflexor activity shifts the COP toward the forefoot. This progression lengthens the GRF moment arm relative to the ankle joint, contributing to knee and hip extension during terminal stance.

In the case of plantarflexor muscle weakness, insufficient force generation at the ankle may lead to rapid and excessive shank inclination, ankle dorsiflexion, and knee flexion, as well as increased knee and hip extensor moments [1,3]. Excessive knee flexion impairs elastic energy storage in the Achilles tendon and reduces the moment-producing capacities of the ankle plantarflexors, resulting in a flatter ROS and reduced push-off power [20]. A dorsiflexion-resistant AFO can help modulate the rate of forward shank inclination by providing an external plantarflexion moment. This reduces excessive knee and hip flexion, leading to a smoother ROS [3]. Positioning the forefoot apex more distally can prolong the second rocker and extend COP progression [30], potentially increasing external plantarflexion moments and hip and knee extension in AFO users. Conversely, a more proximal apex shortens the second rocker, leading to restricted COP progression and reduced knee extension and plantarflexor moments [12,31].

Third Rocker

During the third rocker, the heel lifts off the ground and COP progression and ankle plantarflexor moments approach their peak. At this stage, body weight rapidly transitions from single-limb to double-limb support and elastic energy stored in the Achilles tendon is rapidly released, generating positive ankle power and facilitating a smooth redirection of the COM into the subsequent stance phase, as well as rapid swing initiation of the trailing limb [1]. Adequate timing and sufficient magnitude of peak ankle power ensure rapid rotation of the shank-foot system over the forefoot [19].

The increase in plantarflexion moment during the second rocker contributes to forward progression of the COP and a gradual increase in ankle angular velocities [1]. As the COP progresses anteriorly toward the forefoot, it approaches the shoe's forefoot apex, which effectively acts as a rotational axis for the foot-shoe system. When the COP reaches this apex position, the external dorsiflexion moment generated by the GRF can no longer balance the internally generated plantarflexion moment, resulting in rotation of the foot about the forefoot apex and heel lift-off, marking the onset of the third rocker.

AFOs with energy-storing elements (such as springs or carbon fiber elements) can store mechanical energy during the second rocker and release it at push-off (third rocker) to aid propulsion. A more distal apex position could delay the onset of the third rocker, increasing COP progression, peak GRF moment arms, and ankle dorsiflexion, leading to greater energy storage in spring-like AFOs and enhanced push-off power when the energy is released in the third rocker. Additionally, changes in rocker radius can modulate timing and duration of foot-to-ground contact and shank kinematics during the third rocker, where a smaller radius accelerates shank inclination and energy release (smaller ROS arc), while a larger radius decelerates it, prolonging the third rocker phase (larger ROS arc). Table 1 summarizes the theoretical biomechanical effects across all 3 rockers for 3 types of footwear adaptations when used with an AFO that restricts both plantar- and dorsiflexion.

XSL•FO
RenderX

Table 1. Overview of theoretical kinematic and kinetic outcomes for different footwear adaptations with the use of an AFO[a] restricting plantarflexion and dorsiflexion.

	Heel-sole differential	Heel rocker apex position, radius/angle	Forefoot rocker apex position, radius/angle
Kinematics			
First rocker	Increasing HSD[b] increases SVA[c] and knee flexion	Larger radius/smaller rocker angle decreases rate of forefoot lowering	N/A[d]
Second rocker	Increasing HSD increases SVA and knee flexion	N/A	N/A
Third rocker	N/A	N/A	Smaller radius/larger rocker angle increases SVA and knee flexion and increases angular velocity of the foot-ankle-shank complex
Kinetics			
First rocker	N/A	• More distal heel AP[e,f] reduces ankle DF[g] moment • Smaller radius increases ankle DF moment	N/A
Second rocker	Increasing HSD increases knee extensor moment	N/A	N/A
Third rocker	N/A	N/A	• Smaller radius/larger rocker angle decreases peak knee extensor moment • More distal forefoot position increases PF[h] moment

[a]AFO: ankle-foot orthosis.

[b]HSD: heel-sole differential.

[c]SVA: shank-to-vertical angle.

[d]N/A: not applicable.

[e]AP: apex position

[f]For the purpose of this review, a more proximal heel radius refers to an apex position proximal to the heel.

[g]DF: dorsiflexion.

[h]PF: plantarflexion.

Methods

Study Design

This systematic review was registered with PROSPERO (International Prospective Register of Systematic Reviews; CRD42024625819) and is reported in line with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guideline [32].

Search Strategy

A systematic database search was performed by the primary reviewer using PubMed, Embase, Scopus, and Web of Knowledge. The search was performed using the following phrases: ankle foot ortho*, AFO, AFO-FC, orthotic*, orthosis, orthoses, splint, combined with biomech*, gait, locomotion, walk*, kinemat*, and adaptation, shoe, footwear, rocker, roll-over, insert, wedge, and tuning. Papers were retrieved from the beginning of the respective database history until December 2024. A follow-up search was conducted in May 2026. The complete search string for Scopus is:

TITLE-ABS-KEY((“ankle foot ortho*” OR “orthotic device” OR AFO OR “AFO-FC” OR (ankle AND (orthotic* OR orthosis OR orthoses OR splint))) AND (biomech* OR gait OR “gait analys*” OR locomotion OR walk* OR kinemat* OR tuning OR manipulation) AND (shoe* OR footwear OR rocker* OR “roll-over” OR insert* OR wedge* OR tuning))

The remaining search strings, including database-specific keywords, can be found in Multimedia Appendix 1.

Eligibility Criteria

Titles were excluded if they mentioned nonadult participants, animal participants, or cadavers. Review papers, protocols, or papers that investigated only movements other than walking, surgical techniques or surgery-related outcomes, or devices that were not AFOs, lower leg casts, or footwear were also excluded in the title exclusion process. Rigid lower leg (plaster) casts were considered to behave in a similar manner to rigid AFOs and were eligible for inclusion.

Abstracts were excluded if the abstract did not mention both footwear and AFOs and did not describe effects of footwear adaptation on gait while using an AFO. The abstract must have described walking with an AFO in at least two footwear conditions. Abstracts were excluded when footwear was manipulated in only the leg contralateral to the AFO.

XSL•FO
RenderX

Papers written in English, French, German, or Dutch were eligible for full-text selection. Inclusion criteria required that the study described or visualized the AFOs, footwear, and footwear modifications. To qualify, the footwear adaptation must have been applied externally to the AFO, either between the AFO and the shoe, the shoe itself, or between the shoe and the ground. Additionally, at least one of the following outcome measures must have been examined: sagittal plane lower-limb joint kinetics and kinematics, sagittal plane whole-body COM kinetics and kinematics, spatiotemporal walking parameters, muscle-tendon dynamics, lower-limb EMG (electromyography), COP, GRF, plantar pressure, or metabolic cost.

Selection Process

Screening was performed with the use of Catchii, a systematic review management tool [33]. Two reviewers independently screened papers at each stage of the inclusion process. Reviewers AM and CG performed the title and full-text screening, and the abstract inclusion process was performed by AM and JMH. Disagreements were resolved through discussion between each pair of reviewers. For each stage of the selection process, papers were excluded based on the first exclusion criterion that was met, and only that criterion was recorded.

Quality Assessment and Data Extraction

The quality of the included research was assessed using a modified Downs and Black Checklist [34]. This version of the checklist scored question 27 binarily, rather than out of the original 5 points, and is more suited to observational studies. Questions regarding follow-up were omitted from the checklist as gait studies are typically within-participant designs without follow-up. The final checklist was scored out of a potential 25 points. Adjusting for the omitted questions, a study could be considered “excellent” quality if it scored 23-25 points, “good” between 17 and 22 points, “fair” between 12 and 16 points, and “poor” if marked less than 12 points [35]. Only studies scoring “fair” or higher would be considered to have sufficient quality for data extraction.

A modified version of the AFO reporting checklist developed by Ridgewell et al [36] was used to evaluate the quality of AFO reporting. The checklist evaluates 3 categories: participant details, AFO details, and the experimental protocol. This checklist was developed for studies on children with cerebral palsy, and as such the section on participant details was excluded for the purpose of this review. Scoring followed the guideline provided by Ridgewell et al [36]. Reviewers AM and CG independently scored both checklists. Any disagreements were resolved through discussion.

Data were extracted by AM. A custom checklist was used to extract footwear design parameters. This checklist includes design descriptors such as outsole hardness and apex position but does not include manufacturing items such as last type or sole material. Only outcome measures listed in the inclusion criteria were included in this review. Data were systematically synthesized using a structured biomechanically informed framework. Data were synthesized first by footwear adaptation (eg, HSD), then by the phase of gait in which outcomes were measured, and finally by the specific outcome measure. For clarity purposes, all moments are reported as internal moments, with any values reported as external moments transformed by AM.

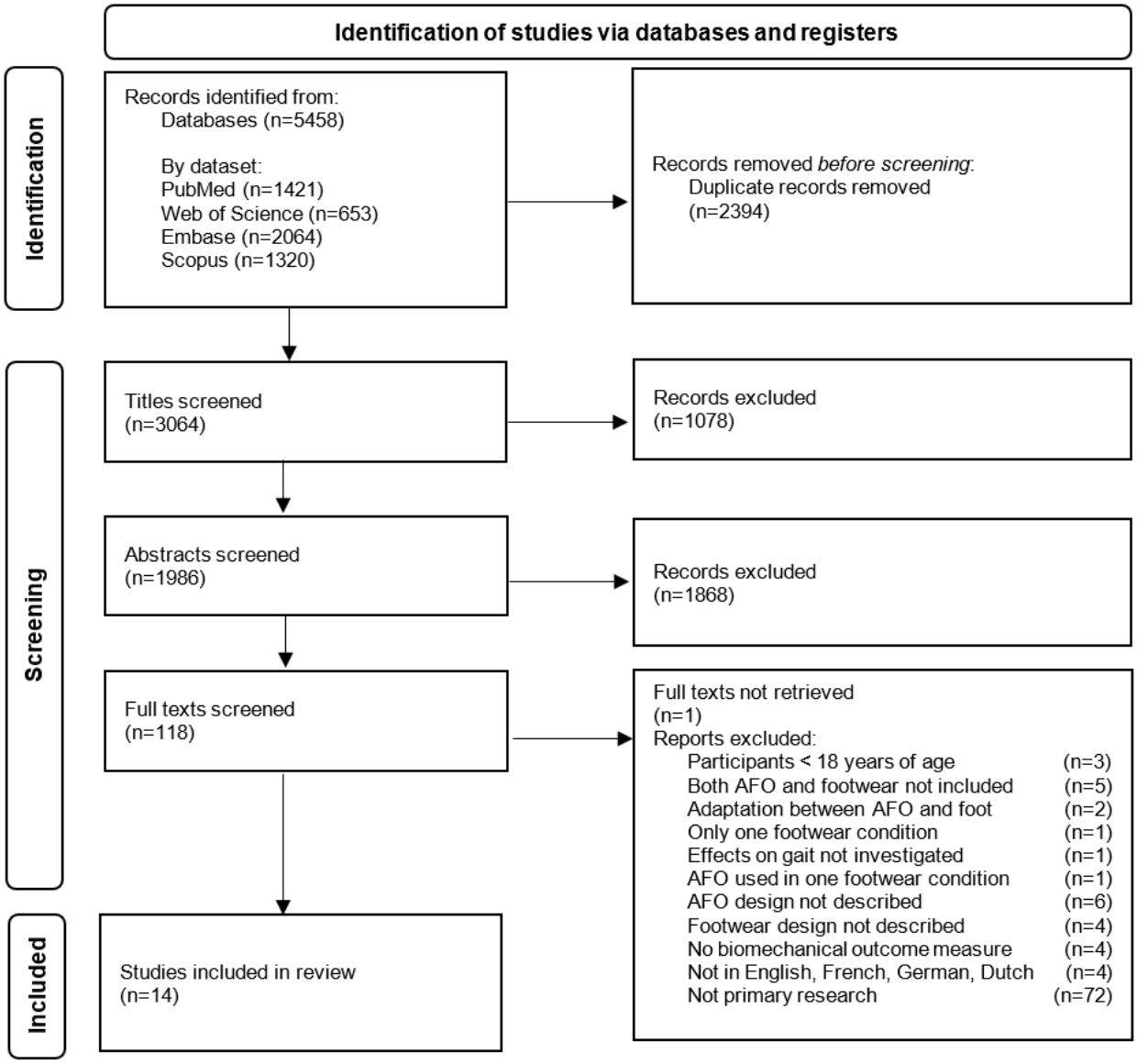
Results

Search Results

The systematic database search, including the updated search, resulted in 3064 papers after duplicates were removed. A total of 118 full texts were evaluated for inclusion in the review. Of these, 9 papers were written in German, and 1 paper in French, all of which were excluded due to not being primary research. A total of 14 studies were included in the review, following the inclusion procedure outlined in Figure 4.

XSL•FO
RenderX

Figure 4. Study selection flowchart including updated search. Exclusion criteria shown here were those used to exclude full texts. AFO: ankle-foot orthosis.



Study Characteristics

An overview of the participants, interventions, and study designs can be found in Table 2. The majority (10/14) of studies were within-subjects designs evaluating hemiparetic patients after stroke using solid dorsal AFOs.

XSL•FO
RenderX

Table 2. Summary of included studies.

Author, year	Participant characteristics	Pathology	Study design	Gait speed[a]	Type of AFO[b]	Footwear modification	Quality score (0-25)
Wiest et al [28], 1979	9 adults, aged 19-46 years	None	Prospective within-subjects	Outcome	Bichannel adjustable AFO	Heel hardness and heel shape	17
Farmani et al [37], 2016	18 adults aged 57.86 ± 10.44 years	Hemiplegia secondary to stroke	Prospective within-subjects	Outcome	Custom rigid AFO	Rocker outsole attached directly to AFO footplate	19
Sutdet et al [38], 2019	4 adults aged 51 ± 15.2 years	Right-sided hemiplegia and genu recurvatum secondary to stroke	within-subjects pilot	Outcome	Custom rigid AFO	Insole wedge height, location unspecified	18
Dhalla et al [39], 2003	28 adults, age not given	None	Prospective within-subjects	NS[c]	Custom fabricated total contact cast	Conventional cast shoe, rigid rocker cast shoe, flat rubber heel attached directly to cast, rubber rocker heel attached directly to cast	16
Farmani et al [40], 2016	30 adults, aged 49-68 years, mean weight 64.7 kg	Hemiplegia secondary to stroke	Prospective block randomized	Outcome	Prefabricated solid polypropylene AFO	High-top rocker shoes, high-top flat-soled shoes	19
Kearney et al [41], 2011	15 adults aged 31.3 ± 4.7 years, height 172 ± 9 cm, and weight 70 ± 15 kg	None	Prospective within-subjects	Self-selected per condition	Rocker bottom walking cast and 2 different carbon-fiber AFOs	Insole heel wedge height	18
Arvin et al [42], 2013	24 adults aged 24.7 ± 4.84 years	None	Prospective within-subjects	Outcome	Modified solid AFO with capability of locking at 5°-7° dorsiflexion	Standard shoe and AFO compared to AFO with inclined shank without a shoe	18
Jagadamma et al [43], 2010	Female person aged 61 years	Left side hemiplegia secondary to stroke	Case study	Outcome	Custom solid polypropylene AFO	External heel wedge height with and without a stiff rocker outsole	17
Masoumi et al [44], 2020	10 adults aged 53.6 ± 7.5 years, height 167.1 ± 5.93 cm, weight 71.3 ± 8.53 kg	Hemiplegia secondary to stroke	Prospective within-subjects	Outcome	Custom rigid AFO	Rocker shoe	19
Kerkum et al [45], 2015	10 adults, aged 24 ± 3 years	None	Prospective within-subjects	Self-selected[d]	Prefabricated carbon AFO	Insole heel wedge height, AFO footplate stiffness	19
Ikeda et al [46], 2018	24 adults aged 18-55 years	Unilateral foot or ankle weakness or pain; controls with no pathologies	Prospective repeated-measures crossover design	Nonself-selected[d]	Carbon fiber AFO	Insole heel wedge height and hardness	16
Choi et al [47], 2016	Male person aged 56 years, height 190 cm, weight 88.5 kg	Left-sided hemiplegia secondary to stroke	Case study and model	Outcome	Prefabricated posterior leaf spring AFO; solid polypropylene AFO	Shoe heel height	14

XSL•FO
RenderX

Author, year	Participant characteristics	Pathology	Study design	Gait speed[a]	Type of AFO[b]	Footwear modification	Quality score (0-25)
Daryabor et al [48], 2021	20 adults aged 33-65 years, height 1.62 ± 0.1 / 1.6 ± 0.1 m, weight 67.8 ± 8.7 / 67.1 ± 10.5 kg	Hemiplegia secondary to stroke	Randomized controlled trial	Outcome	Custom PF[e]-stop AFO; custom PF-resistant AFO	Rocker shoe	16
Schmalz et al [49], 2006	10 adults aged 34 ± 9 years, height 178 ± 4 cm, weight 73 ± 9 kg	None	Prospective within-subjects	Self-selected per condition	Rigid AFO; semirigid supramalleolar orthosis	Medial and lateral wedges external to shoe	17

[a]The gait speed column indicates whether gait speed was an outcome variable or controlled across conditions.

[b]AFO: ankle-foot orthosis.

[c]NS: not stated.

[d]Denotes the gait speed was controlled across conditions.

[e]PF: plantarflexion.

Participant Characteristics

A total of 204 participants (108 healthy, 84 patients after stroke, and 12 with unilateral ankle weakness or pain) were included across all 14 studies. Seven papers examined AFO-footwear combinations in patients with hemiparesis secondary to stroke [37,38,40,43,44,47,48]. Hemiparetic participants ranged across studies from 11 months to over 10 years post stroke onset. Six studies recruited only healthy participants [28,39,41,42,45,49], and one study enrolled both healthy controls and participants with unilateral foot or ankle weakness or pain [46]. Of the 8 studies that included patients, 5 studies recruited participants who used AFOs in their daily lives [37,38,43,46,47], 1 study included only participants who had not used AFOs in their daily lives [48], and 2 studies did not specify whether AFOs were used in daily life [40,44]. Nine studies included gait speed as an outcome measure [28,37,38,40,42-44,47,48]. Three studies let participants walk at a self-selected walking speed and did not report it as an outcome. Of these 3 studies, one study controlled the speed between the trials [45], and 2 studies did not specify whether speed was controlled between conditions [41,49]. One study set a target gait speed based on the participants' leg length and a Froude number of 0.16 and used this for all conditions [46]. One study did not report gait speed as an outcome or specify if gait speed was controlled [39].

AFO Properties

A summary of the AFOs included in this review and their properties can be found in Table 3. A total of 19 different AFOs were used across the 14 included studies. Eight of the AFOs used a solid dorsal shell design constructed out of polypropylene [37,38,40,42-44,47]. One study used 2 AFOs with a "plastic shell," which are assumed to be polypropylene dorsal shell AFOs, one with plantarflexor resistance and one with a plantarflexion stop [48]. Four AFOs were carbon-fiber, of which 2 AFOs had ventral shells [41], 1 AFO a proximal cuff [46], and 1 AFO a dorsal shell [45]. While a walking boot was also investigated by Kearney et al [41], results from the cast were not included, as it did not meet the inclusion criteria of having the footwear adaptation be between the AFO and the shoe or ground. One AFO had a hinged ankle joint with adjustable medial and lateral struts which attached directly to an experimental shoe [28]. One study used 2 AFOs with the same design with either a medial or lateral strut [49]. These 2 AFOs were counted as 1 AFO in Table 3 as they were otherwise identical. The AFO had free dorsiflexion and plantarflexion but did not permit frontal plane motion. No information about any parameters other than strut position was given or possible to infer for this AFO. This study also examined a stirrup brace which restricted frontal plane movement and was considered an AFO. A plaster cast [39] was reported and categorized as an AFO for the purpose of this review. Including the cast, 11 AFOs had full-length footplates [37,39-42,44,46,47]. The length of the footplate was not stated or possible to infer from images for 7 AFOs [38,43,45,47-49]. Eight of the footplates were rigid [37-39,42,44,45,47], one was flexible throughout the footplate [45], and one was flexible at the metatarsophalangeal joint [43]. Two AFOs had indeterminable footplate stiffnesses, with the only description given as one AFO being "more flexible," and the other AFO being "less flexible" [41]. The footplate stiffness was not given or impossible to infer for 7 AFOs [28,40,47-49].

XSL•FO
RenderX

Table 3. Overview of AFOsa.

Author, year	Design	Materials	Movement restrictions	Ankle joint (rigid, leaf spring, stop)	Ankle angle[b] (°)	Footplate length	Footplate stiffness
Wiest et al [28], 1979	Custom[c] instrumented bichannel with adjustable height and locking ankle joint	NS[d]	DF[e] and PF[f]	Locked hinge	90	Directly attached to shoe	NS
Farmani et al [37], 2016	Custom solid dorsal	3-mm thick polypropylene	DF and PF	Rigid	90	Full length (heel to tip of toes)	Rigid (inferred)
Sutdet et al [38], 2019	Custom solid dorsal	5-mm thick polypropylene	All directions	Rigid, reinforced with 3-mm polypropylene	90	NS	Rigid (inferred)
Dhalla et al [39], 2003	Custom total contact cast	Orthoflex plaster with fiberglass shell	All directions	Rigid	90	Full length (heel to slightly beyond tip of toes)	Rigid
Farmani et al [40], 2016	Prefab[g] solid dorsal	Polypropylene	NS	Rigid	90	Full length	NS
Kearney et al [41], 2011	Prefab ventral	Carbon-fiber	NS	Leaf spring	90	Full length	“Less flexible”
Kearney et al [41], 2011	Prefab ventral	Carbon-fiber	NS	Leaf spring	90	Full length	“More flexible”
Arvin et al [42], 2013	Custom solid dorsal	4-mm thick polypropylene	All directions	Rigid	90	Full length	Rigid (inferred)
Arvin et al [42], 2013	Custom solid dorsal	4-mm thick polypropylene	All directions	Rigid	83-84	Full length	Rigid (inferred)
Jagadamma et al [43], 2010	Custom solid dorsal	4-mm thick polypropylene	NS	Rigid	90	NS	Flexible at MTP[h] joints
Masoumi et al [44], 2020	Custom solid dorsal	4-mm thick polypropylene	NS	Rigid	90	Full length	Rigid (inferred)
Kerkum et al [45], 2015	Prefab solid dorsal	Prepreg carbon	All directions	Rigid (7.9 Nm/deg)	90	NS	0.06 Nm/deg and 0.89 Nm/deg + 0.06 Nm/deg
Ikeda et al [46], 2018	Custom proximal cuff	Carbon-fiber	PF	Flexible strut	NS	Full length	Rigid
Choi et al [47], 2016	Prefab posterior leaf spring	NS	NS	Leaf spring	NS	NS	NS
Choi et al [47], 2016	Custom solid dorsal	Polypropylene	NS	Rigid	82	Full length	NS
Daryabor et al [48], 2021	Custom, dorsal	Plastic	PF	PF stop	90	NS	NS
Daryabor et al [48], 2021	Custom, dorsal	Plastic	PF	PF resistive spring	90	NS	NS
Schmalz et al [49], 2006	Prefab stirrup ankle brace	Textile and plastic	Mediolateral	Flexible	N/A[i]	N/A	N/A
Schmalz et al [49], 2006	Customization unknown, medial/lateral strut	NS	Mediolateral	NS	NS	NS	NS

[a]AFO: ankle-foot orthosis.

[b]Ankle angle is taken as the angle between the footplate and the shank of the ankle-foot orthosis.

[c]All participants used the same ankle-foot orthosis, which was a custom-made experimental ankle-foot orthosis that could be adjusted to each participant.

[d]NS: not stated.

[e]DF: dorsiflexion.

[f]PF: plantarflexion.

XSL•FO
RenderX

[g]Prefab: prefabricated.

[h]MTP: metatarsophalangeal.

[i]N/A: not applicable.

Footwear Adaptations

The footwear adaptations used for each study can be found in Table 4. Footwear adaptations are reported in the same manner as the respective study. Seven papers investigated adaptations at the heel of the shoe. The majority of these focused on increasing the HSD, either with insole wedges [41,45,46] or through altering the heel height of the shoe itself [47]. An overview of these adaptations can be examined in Table 5. One study used insoles to manipulate the SVA, but did not specify the shape, height, or location of these insoles [38]. One paper examined the effects of changing the shape and hardness of the heel of an experimental shoe [28]. One paper examined the effects of increasing the HSD using wedges, which were then translated to the design of a custom shoe and combined with a stiff rocker outsole [43]. Design parameters for that rocker outsole were not given; however, from the image provided, it can be inferred that an apex position proximal to the metatarsophalangeal joint was likely used.

XSL•FO
RenderX

Table 4. Overview of footwear adaptations with respect to the shoe profile. Comparison condition refers to the condition with the AFOa without any footwear adaptations and may not be the control condition. Footwear adaptations follow the naming conventions used per respective study.

Author, year	Comparison condition	Footwear adaptation	Forefoot/heel rocker	Rocker angle/radius	Apex position	Heel sole hardness	Outsole hardness	Midsole hardness and distribution
Wiest et al [28], 1979	Dress shoe with hard (Shore-A 74) rubber heel and AFO	SACH[b]	None	N/A[c]	N/A	Medium density polyurethane center	NS[d]	NS
Wiest et al [28], 1979	Dress shoe with hard (Shore-A 74) rubber heel and AFO	Heel rocker	Heel	20°	3.5 cm from distal edge of heel	Shore-A 74	NS	NS
Wiest et al [28], 1979	Dress shoe with hard (Shore-A 74) rubber heel and AFO	Compressible heel	None	N/A	N/A	Shore-A 40	NS	NS
Farmani et al [37], 2016	AFO with thin nonslip rubber outsole attached directly to AFO footplate	Rocker sole applied directly to AFO (no shoe)	FF[e]	15°	65%	N/A	Shore-A 50-60	Shore-A 30-40, length of footplate
Sutdet et al [38], 2019	Sport shoe with AFO without wedges	Increasing HSD[f]	None	N/A	N/A	NS	NS	NS
Dhalla et al [39], 2003	Cast with no shoe	Conventional cast shoe	FF and heel	NS	NS	N/A	NS	NS
Dhalla et al [39], 2003	Cast with no shoe	Rigid rocker bottom shoe	FF and heel	NS	NS	N/A	“Rigid”	N/A
Dhalla et al [39], 2003	Cast with no shoe	Rubber rocker bottom applied directly to cast, centered on longitudinal arch of the foot	Heel	NS	NS	NS	NS	N/A
Dhalla et al [39], 2003	Cast with no shoe	Flat rubber heel applied directly to cast in line with the tibia	N/A	N/A	N/A	N/A	NS	N/A
Farmani et al [40], 2016	High-top standard shoe with AFO	Rocker sole applied to standard shoe	FF	15°	65%	NS	NS	Shore-A 30-40, length of shoe
Kearney et al [41], 2011	Participant’s own shoe with AFO; without wedges	Increasing HSD	None	NS	NS	NS	NS	NS
Arvin et al [42], 2013	Sport shoe with 1 cm HSD and AFO	AFO with inclined shank (5°-7° anterior inclination)	None	N/A	N/A	N/A	N/A	N/A
Jagadamma et al [43], 2010	Standard shoe with AFO	Rocker sole applied to standard shoe and changes to HSD	FF	NS	NS	NS	Rigid	NS
Masoumi et al [44], 2020	Standard shoe with AFO	Rocker sole applied to standard shoe	FF	15°	65%	NS	Shore-A 50-60	Shore-A 30-40, length of shoe
Kerkum et al [45], 2015	Flexible flat-soled shoe (Converse-style sneakers)	Increasing HSD	None	N/A	N/A	NS	NS	NS

XSL•FO
RenderX

Author, year	Comparison condition	Footwear adaptation	Forefoot/heel rocker	Rocker angle/radius	Apex position	Heel sole hardness	Outsole hardness	Midsole hardness and distribution
Ikeda et al [46], 2018	Sports shoe and AFO without wedges	Increasing HSD	None	N/A	N/A	NS	NS	NS
Choi et al [47], 2016	Sports shoe and AFO without wedge	Material removed from heel of shoe	None	N/A	N/A	NS	NS	NS
Daryabor et al [48], 2021	Standard shoe with AFO	Rocker sole applied to standard shoe	FF	15°	65%	NS	Shore-A 50-60	Shore-A 30-40, length of shoe
Schmalz et al [49], 2006	Dress shoe and AFO without wedge	Mediolateral differential	None	N/A	N/A	NS	NS	NS

[a]AFO: ankle-foot orthosis.

[b]SACH: solid ankle cushion heel

[c]N/A: not applicable.

[d]NS: not stated.

[e]FF: forefoot.

[f]HSD: heel-sole differential.

Table 5. Overview of heel-sole differential adaptations.

Author, year	Wedge used[a] (Y[b]/N[c])	Insole or external	Location	Shape	Thickness (cm)	Hardness
Sutdet et al [38], 2019	Y	Insole	NS[d]	NS	Inferred from SVA[e]	NS
Kearney et al [41], 2011	Y	NS	Heel	NS	1.5, 3.0, 4.5	NS
Jagadamma et al [43], 2010	Y	NS	Heel	NS	Inferred from SVA	NS
Kerkum et al [45], 2015	Y	Insole	Heel	NS	0.6/1.3; 2.8/2.8; 4.9/5.3 cm	NS
Ikeda et al [46], 2018	Y	Insole	Heel	NS	1.0, 2.0, 3.0	Per each thickness: 50-57, 75-78 on OO[f] scale
Choi et al [47], 2016	N	External	Heel	NS	NS	N/A[g]
Schmalz et al [49], 2006	Y	External	Medial/lateral; highest point at heel of shoe	Tapered 14° angle of horizontal inclination	10 mm (at highest point)	Firm

[a]"Wedge used" refers to whether a separate piece of material was used to increase the heel-sole differential.

[b]Y: yes.

[c]N: no.

[d]NS: not stated.

[e]SVA: shank-to-vertical angle.

[f]OO: shore OO hardness.

[g]N/A: not applicable.

Three studies compared standard shoes to rocker shoes [40,44,48]. All 3 studies used the same forefoot rocker design of an apex position at approximately 65% of the length of the shoe from the heel, and a rocker angle of 15° with a curved forefoot and stiff outsole. Farmani et al (2016) [37] studied the effects of adding a rocker sole directly to the footplate of a rigid polypropylene AFO compared to the same AFO with only a layer of nonslip rubber added to the footplate [37]. This rocker sole also had an apex position at approximately 65% of the shoe length with a beveled forefoot with a rocker angle of 15°, and as such is considered functionally a rocker shoe for the purpose of this review.

Dhalla et al [39] examined the effects of different types of cast shoes when used with a rigid walking cast. Schmalz et al [49] examined the effects of mediolateral wedging, in this instance

XSL•FO
RenderX

applying the wedges externally to the shoe. It is unclear whether the footplate of the AFO and wedge were the same length. A study from Arvin et al [42] compared a conventional solid dorsal AFO with a shoe to an AFO of the same manufacturer but with a shank inclined in dorsiflexion without a shoe. Both the traditional AFO with a shoe and the modified AFO worn without a shoe had the same SVA of 5°-7° inclination [42].

Quality Assessment

Study Quality

The scores for each included paper can be found in Table 2. The score each paper received per question can be examined in Multimedia Appendix 2 [28,37-49]. The majority of papers (10/14) were considered "good" quality, with the remaining papers (4/14) considered "fair." The lowest score given was 14. No paper received a "yes" for stating whether an attempt was made to blind either the participants or those measuring the outcomes.

AFO Reporting

Multimedia Appendix 2 provides an overview of the results of the AFO Reporting Checklist. A total of 19 different AFOs were used across the 14 included studies, with 2 studies using 2 different AFO designs, and 1 study comparing a walking boot to 2 different AFOs. The majority (10/14) of studies applied their modifications in a randomized order, 2 studies did not state whether randomization occurred, and 2 studies were case studies where randomization was not applicable. Over half (12/19) of the AFOs were custom manufactured, with only 1 AFO's manufacturing status being unclear. A clear motivation for why a specific AFO design was used was given for all AFOs. No consistent method of reporting AFO details was present across the included studies. AFO design details often had to be inferred from photos, with 8 of the AFOs having completely clear trimlines. The stiffness and ankle position were described in detail for 1 AFO. A clear description of the movements assisted or prevented was given for 13 of the AFOs.

Biomechanical Effects

Table 6 provides an overview of the outcome measures extracted for this review. Only outcome measures fitting the inclusion criteria outlined above were included for the purpose of this review. These outcome measures can be divided into 5 main categories: kinematics, kinetics, plantar pressure, spatiotemporal parameters, and GRF. Table 7 provides the effects of each type of footwear adaptation per outcome measure across the 3 rockers. All results are given for rigid AFOs unless otherwise specified.

Table 6. Outcome measures of interest.

Outcome measure	Included from reference number	Total patients (n)	Total healthy (n)
Kinematics			
Ankle	[44[a],46-48]	1 S[b], 12 U[c]	20
Knee	[27,36[a],42,44,46-48]	40 S	29
Hip	[36[a],44,46,47]	39 S	10
Kinetics			
Ankle	[45-49]	45 S	20
Knee	[28,38,43,45-49]	26 S, 12 U	29
Hip	[45,47,48]	21 S	10
Plantar pressure	[39,41]	0	43
Spatio-temporal	[27,36[a],37,39-43[a],45-48]	84 S, 12 U	58
GRF[d]	[28,48,49]	20 S	19
Other	[45,46]	12 U	10

[a]Denotes a significant difference between the standard shoe with AFO and at least one adapted shoe with AFO condition. Statistically significant results comparing a non-AFO condition to adapted shoe conditions are not indicated.

[b]S: patients after stroke.

[c]U: patients with unilateral weakness or pain.

[d]GRF: ground reaction force.

XSL•FO
RenderX

Table 7. Overview of outcomes across the gait cycle for different footwear adaptations when used with a rigid AFOa.b

Heel-sole differential	Heel rocker	Heel hardness	Forefoot rocker
Kinematics			
First rocker			
Increasing HSD[c] increases knee flexion [43] (n=18S)[d]	Use of heel rocker decreases knee flexion [28] (n=9H)	More compressible heel decreases knee flexion [28] (n=9H)	N/A[e]
Second rocker			
Increasing HSD increases knee flexion [38] (n=4S), [43] (n=1S), [45] (n=10H) and hip flexion [45][d]	N/A	N/A	N/A
Third rocker			
Increasing HSD increases knee flexion [43] (n=18S), [47] (n=1S)	N/A	N/A	• Increases knee flexion at toe-off [37] (n=18S)d • Earlier knee flexion during preswing [43] (n=1S)
Kinetics			
First rocker			
Increasing HSD increases knee extensor moment [46] (n=12U)[d]	Decreases knee extensor moment [28] (n=9H)[d]	More compressible heel decreases ankle DF[f] [46] (n=12U) and knee extensor moments [28] (n=9H)	N/A
Second rocker			
Increasing HSD increases knee extensor moment [47] (n=1S)[d]	N/A	N/A	N/A
Third rocker			
N/A	N/A	N/A	GRF[g] moved posterior to the knee joint during terminal stance and preswing [43] (n=1S)
Spatio-temporal			
No effect on walking speed [38] (n=4S)	No effect on cadence or walking speed [28] (n=9H)	—[h]	Increases walking speed with rigid AFO [37] (n=18S), [44] (n=10S)
Plantar pressure			
—	Decreases peak hindfoot plantar pressure [39] (n=28H)	—	Decreases peak forefoot plantar pressure [39] (n=28H)

[a]AFO: ankle-foot orthosis.

[b]The number of participants is given in brackets beside the respective reference. Whether the participants are healthy (H), patients after stroke (S), or have unilateral weakness or pain (U) is indicated.

[c]HSD: heel-sole differential.

[d]Effects are in line with the theory presented in Table 1.

[e]N/A: not applicable.

[f]DF: dorsiflexion.

[g]GRF: ground reaction force.

[h]Not available.

HSD

Increasing the HSD showed effects on gait across all 3 rocker phases, with insole wedges being used to influence biomechanics in the third rocker [43]. Increasing the HSD from 1 to 3 cm increases the peak dorsiflexor moment at the beginning of the second rocker, shifting the second rocker later in stance [46]. This change in HSD also increases the time required for the COP to reach peak anteroposterior velocity. Peak COP velocity occurred approximately 10% later in the stance phase when using a soft 3 cm wedge compared to a soft 1 cm wedge, and just over 30% of the stance phase later for a firm 3 cm wedge compared to a firm 1 cm wedge [46]. The peak knee extensor

XSL•FO
RenderX

moment during loading response increases and occurs later in the gait cycle as the HSD increases [46].

An increase in HSD increases the SVA in both healthy individuals and patients after stroke, resulting in greater knee flexion [38,43,45]. Similarly, a larger HSD increases the ankle dorsiflexion angle, plantarflexor moment, and hip flexion angle at midstance [45].

In the third rocker, the use of an increased HSD and rigid AFO elicited increases in knee flexion at toe-off when compared to a posterior leaf spring–AFO and standard shoes in a patient who had a stroke [47]. Larger plantarflexor and knee extensor moments in a patient who had a stroke were also observed at toe-off when compared to walking in only unmodified shoes [47].

Increasing the HSD linearly increased the radius of curvature of the ROS normalized to participant body height, with 3 cm wedges producing a radius of curvature in 4% of body height greater than 1 cm wedges [46].

A linear increase in forefoot pressure and simultaneous decrease in hindfoot pressure was found with increasing HSD [41]. There was no difference found between AFO stiffnesses when the same changes in HSD were applied [41].

Increasing the HSD did not affect walking speed but did decrease cadence in patients after stroke when compared to walking with shoes only, with a significant decrease in cadence when the shank was inclined by 10° [38]. There was also no difference in the duration of terminal stance regardless of AFO stiffness or HSD [41]. Similarly, there were no significant differences for speed, step length, or step width in healthy participants when the SVA is held constant between an AFO with an inclined shank and no shoe, and an AFO with a neutral SVA and a shoe with a 1 cm HSD [42].

Heel Rocker

Three different heel rockers were investigated across 2 studies. A heel with a 20-degree bevel reduced the knee extensor moment during loading response by 24% compared to a conventional men's dress shoe hard rubber heel when both were used with a rigid AFO [28]. The addition of a heel rocker to a walking cast allowed for a reduction in peak hindfoot pressure [39]. Conversely, a rigid whole-foot rocker applied to the same cast did not decrease hindfoot pressure; however, it did decrease peak forefoot plantar pressure [39].

Heel Hardness

Increasing the hardness of the heel external to the shoe increased the knee extensor moment during the loading response, with a hard rubber heel having a moment 31% higher than a SACH (solid ankle cushion heel), and 12% higher than a compressible crepe heel [28]. Harder heel wedges increased the peak ankle dorsiflexor and knee extensor moments during second rocker [46]. Increasing both the height and hardness of a heel wedge significantly delayed the timing during the stance phase of peak COP velocity, with the peak occurring approximately 15% later in stance for a firm 3 cm wedge compared to a compressible 1 cm wedge [46].

Forefoot Rocker

The use of a proximal forefoot rocker increased knee flexion at toe-off when used with a rigid AFO [37,43]. This increase in knee flexion also resulted in the GRF being more posterior to the knee joint during third rocker compared to a shoe with a standard forefoot [43]. In both plantarflexion stop and plantarflexion resistant AFOs, the addition of a proximal forefoot rocker significantly reduced the peak ankle plantarflexor moment when compared to the same AFOs and standard shoes [48].

In rigid AFOs, the addition of a proximal forefoot rocker increased walking speed when compared to the same AFO with a standard shoe [40,44]. Conversely, the use of a proximal forefoot rocker with dorsiflexion-free AFOs resulted in slower walking speeds when compared to those same AFOs and a standard shoe [48]. The stride length was also reduced in patients walking with dorsiflexion-free AFOs when paired with a proximal forefoot rocker compared to when paired with a standard shoe.

Mediolateral Wedging

Mediolateral wedging had limited effects on sagittal plane biomechanics. The peak knee extensor moment during third rocker was reduced with the use of either a medial or lateral wedge, although only significantly with the medial wedge, when paired with an AFO that restricted frontal plane motion at the ankle joint [49]. Medial or lateral wedging with either an ankle support or rigid AFO decreased the vertical component of the GRF when compared to a shoe without any wedging or AFO [49].

Discussion

Principal Findings

This review used a biomechanical framework to synthesize evidence on the interacting effects of footwear parameters and AFOs on gait outcomes in patients and healthy young adults. While it is evident that footwear adaptations can optimize gait biomechanics beyond what can be achieved by an AFO and standard shoe alone, large knowledge gaps remain, preventing empirical validation of many theoretical and biomechanical assumptions and adoption in clinical practice. The effects of plantar pressure [39,41] will be omitted from the discussion to focus on kinetics and kinematics.

Participants both with and without pathologies were included in this review, as healthy walkers can provide a valuable model for the biomechanical effects of AFO-FCs. Indeed, in the few studies that examined the same AFO type, footwear adaptation, and outcome measure, it was found that both patients after stroke and healthy walkers have increased knee flexion during the second rocker when the HSD is increased [38,43,45]. However, this heterogeneity in participants does require discretion when interpreting the results, as adaptations may not be transferable to all patient populations.

A wide range of AFO designs were included in this study; however, the majority (11/19) could be generalized as rigid AFOs which restrict both plantarflexion and dorsiflexion and

XSL•FO
RenderX

can be assumed to have negligible differences in their mechanical properties when interpreting sagittal plane gait variables.

Biomechanical Effects

Heel-Sole Differential

Adaptations increasing the HSD were comparatively well represented in this review, likely due to the SVA being considered a core component of the AFO tuning process [11,50]. Indeed, increasing the SVA, such as with a larger HSD, causes a cascade effect on lower limb kinematics and kinetics when the ankle joint is fixed. Notably, a larger HSD increased knee and hip flexion during the second rocker [45]. This increase in knee flexion was seen across all 3 papers that evaluated knee angle as a result of changing HSD in both healthy participants and patients after stroke [38,43,45]. Although a larger HSD was reported to increase ankle dorsiflexion despite the use of a rigid AFO, it should be noted that this increase may be due to marker placement on the shoe while using insoles to increase the HSD, causing an increase in SVA without a change in the participants' ankle angle [45].

With a constrained ankle joint, changes in HSD affect both the knee and hip joints, indicating that shank kinematics are also influenced. This is supported by the study of Kerkum et al (2021) [7], which examined the dynamic SVA at midstance during walking with an AFO and found that, with a rigid AFO, the SVA increases as HSD increases [45]. Intuitively, the HSD can be used to influence lower limb joint kinematics in the first and second rockers; however, when the HSD is altered within the shoe, kinematics and kinetics at the ankle, knee, and hip joints are also influenced in the third rocker for rigid AFO users [47]. This is likely due to the fixation of the ankle joint, causing the shank to become more inclined as the foot-to-floor angle increases with a larger HSD, even as the heel leaves the ground in the third rocker. When the ankle joint is not constrained, it is able to adapt to changes in the HSD, leaving shank kinematics unaffected [21,51]. These findings suggest that patients who wear rigid AFOs and experience excessive knee extension may benefit from an increased HSD.

Ikeda et al (2018) [46] demonstrated that the ROS radius of curvature could be linearly increased with a constrained ankle and increasing HSD. Conversely, past research has shown that ROS are not affected by an increased HSD for unconstrained ankles [21]. As ROS radii between 20% and 30% leg length have been shown to optimize metabolic rate [15], the use of HSD to tune the ROS radius in rigid AFO users may be something to consider when optimizing energy consumption. Walking speed and cadence remained unaffected across different HSD settings, suggesting that while HSD can be used to manipulate joint kinetics and kinematics, it may have a limited impact on spatio-temporal parameters.

Heel Rocker

It was shown that a heel with a 20-degree bevel could reduce the knee extensor moment in the first rocker [28]. This follows biomechanical theory, which suggests that a heel rocker will reduce the moment arm of the GRF, thereby reducing the peak moments and angular velocities at the knee and ankle joints, facilitating smooth COM transfer. In other words, the use of a heel rocker could elicit a steeper ROS arc during first rocker, increasing the rate of forward progression of the COP. This suggests that such a heel rocker could be used to improve the walking efficiency of patients using rigid AFOs who struggle with step-to-step transitions or walk with excessive knee flexion at the start of the second rocker.

Changes in walking speed were not observed when a heel rocker was used with a rigid AFO [28]. As the other study that examined heel rockers did not look at temporal effects, no conclusion can be drawn based on the evidence presented in this review.

Heel Hardness

A more compliant heel was associated with reduced ankle dorsiflexor and knee extensor moments at initial contact, likely due to a shortened moment arm of the GRF [28]. Following biomechanical theory, it was seen that a more compliant heel had a similar effect to a rocker heel in terms of decreasing the knee extensor moment during the loading response [28,46]. The study by Wiest et al (1979) [28] demonstrated that a SACH could reduce the knee extensor moment during initial contact more than both a crepe heel and 20-degree beveled heel. As the SACH was constructed out of medium-density polyurethane, it may be more compressible and have greater force distribution properties than the crepe heel and incompressible heel, resulting in a smaller GRF moment arm at the knee and ankle joints at initial contact. The use of a compressible heel may therefore help reduce joint loading and smooth step-to-step transitions at initial contact for individuals with a heavy heel strike.

Increasing heel hardness was associated with increased peak dorsiflexor and knee extensor moments during the first rocker, as well as delayed peak moment timing [46]. These effects were most pronounced with wedges thicker than 1 cm, suggesting a threshold beyond which mechanical influence plays a greater role due to the increased vertical component of the GRF.

Ikeda et al [46] found that a more compliant heel placed the ROS center of curvature more posteriorly when compared to a stiffer heel. This is in line with the work of Hansen et al [52], which demonstrated that more compliant feet have a more curved ROS compared to less compressible prosthetic feet. This suggests that the use of a compressible heel wedge could be used to decrease the rate of COP forward progression during first rocker when using a rigid AFO, thereby increasing the steepness of the ROS arc. Using a more compliant heel may be a method of increasing the steepness of the ROS arc without introducing greater moments at the joint, such as with a harder heel. While harder heels were associated with greater gait speeds, they may also increase joint loading, which could be detrimental in populations with joint degeneration or pain [28]. When joint loading is not a concern, harder heels may be used to slightly increase gait speed due to the increased moments and angular momentum at the knee and ankle, leading to a faster COM transition between steps.

Forefoot Rocker

In patients after stroke, the use of a proximal forefoot rocker demonstrated differing effects depending on the type of AFO

XSL•FO
RenderX

used. When paired with rigid AFOs, it increased knee flexion at toe-off as well as walking speed when compared to the same AFO with a standard shoe. This increase in walking speed may be due to improved rollover through the restoration of a more typical ROS radius caused by an increase in the COP velocity during third rocker. In line with this finding, it has been shown that a proximal forefoot rocker used without an AFO can facilitate improved weight transfer and a steeper ROS arc in healthy participants [25]. Increasing knee flexion at toe-off may also be beneficial in patients with limited foot clearance using rigid AFOs.

When combined with dorsiflexion-free AFOs, the peak ankle plantarflexor moment decreased when compared to the same AFO used with a standard shoe. This is likely due to the moment arm of the GRF being shortened by the proximal position of the rocker apex. These dorsiflexion-free AFOs were also shown to decrease stride length and walking speed when used with a proximal forefoot rocker instead of a standard shoe, possibly due to the reduced plantarflexor moment during third rocker.

Future Steps

A key finding of this review is the lack of research into varying rocker parameters such as both heel and forefoot apex positions and rocker radii when AFOs are used. The evidence from the limited studies included in this review suggests that rocker outsole parameters can have a significant impact on both healthy and pathological gait, indicating that further research into these areas could benefit AFO users. Of the 5 studies that used forefoot rockers, 4 studies used the same design [37,40,44,48], with limited reasoning behind their choice, and the fifth did not provide any descriptor of the rocker design [43]. It would be beneficial when using different AFOs with the same footwear adaptations to compare the same outcome measures, as this furthers the understanding of the complex interactions between AFOs and footwear combinations.

Future studies would benefit from studying the effects of differing forefoot rocker design choices, as these have been shown to have a significant impact on gait biomechanics [8,9,25]. In particular, understanding the biomechanical effects of differing apex positions and radii across a variety of AFO designs beyond rigid polypropylene AFOs would be beneficial. These parameters directly affect the GRF lever arms and shank kinematics, in addition to the timing of key gait events, and as such may have significant contributions to improving gait efficiency. A core understanding of the mechanisms driving biomechanical adaptations could lead to improvements in tuning AFOs to improve step-to-step transitions and minimize energy expenditure for patients with a range of pathologies.

An aim of this review was to consolidate evidence across diverse pathologies; however, the majority of included studies focused on patients after stroke and healthy participants. Broader representation would improve the generalizability of findings, as gait impairments in stroke differ from those in conditions such as cerebral palsy, multiple sclerosis, or peripheral neuropathy. Furthermore, focusing on a single pathology overlooks the considerable variability that exists within that group, potentially masking important individual differences. This narrow scope may hinder clinical decision-making, as practitioners often treat diverse patient populations and require evidence that supports personalized, condition-appropriate prescriptions. Future work should consider broadening the scope to include a wider range of pathologies in order to improve the clinical relevance and applicability of findings.

Limitations

A very specific set of exclusion criteria was used for this review. While the number of included studies reflects similar reviews and the existing knowledge gap, a broader inclusion protocol could have been beneficial. For instance, the inclusion of papers that did not use an AFO for all conditions but had multiple footwear adaptations could still have provided insight into the effects different footwear adaptations may have on biomechanics. Furthermore, data were extracted by a single author, and while the included results were examined by the other 2 authors of this review, it is possible that some data were overlooked or inconsistently collected during the inclusion process, introducing a risk of subjective bias and potentially reducing the reproducibility of this review's findings. As all studies included at least one significant finding, there may be a risk of publication bias. However, due to the design of the included studies, a formal assessment of publication bias was not performed.

While the papers included in this study were mostly of "good" quality, none could be considered "excellent." This is largely due to no studies being blinded; however, due to the physical nature of the interventions, it would be difficult to truly blind participants and researchers. It is important to note that 4 of the 5 studies using forefoot rockers belong to the same institution. While each paper from that group reported different outcome measures, the same rocker parameters were used for all 4 papers.

A major limitation of this review is the lack of empirical descriptors given for interventions. Without empirical descriptors, it is difficult to draw scientifically sound conclusions from the presented outcomes. Many AFO details were inferred, or the details provided were too vague to derive conclusions from. For instance, 2 AFO designs were described as "more stiff" or "less stiff" than the other [41]. As AFO stiffness has a significant impact on knee and ankle kinematics, whether that difference in stiffness is 1 or 5 Nm/deg can have a substantial influence on gait outcomes [53]. Compounding this issue, very few studies gave sufficient detail about their adaptations, and there was often minimal description of the footwear being used. This lack of transparency makes it difficult to replicate findings or understand the specific mechanical contributions of the footwear-AFO combinations being tested.

Conclusions

The results demonstrate that footwear adaptations can be used to optimize AFO effects for individual patients. However, while the biomechanical effects of increasing the HSD with a rigid AFO are well explored, there exists limited information regarding the effects of other outsole design parameters and compliant AFOs. Further understanding of the interactions between AFOs and rocker outsole parameters is required to maximize AFO effectiveness across different patient populations. Future research should investigate the effects of

XSL•FO
RenderX

varying forefoot parameters, including rocker radius, apex position, and footplate stiffness, in combination with a wider range of AFO designs including energy-storing AFOs. Explicit and consistent reporting of AFO and footwear design properties is necessary to make meaningful conclusions on biomechanical outcomes.

Acknowledgments

The authors used the generative AI tool Copilot from Microsoft to do minor grammar edits, which were further reviewed and revised by the authors.

Funding

The authors have no funding information to disclose.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings.
[DOCX File , 21 KB - rehab_v13i1e93423_app1.docx]

Multimedia Appendix 2
Results per question for the Downs and Black Checklist and AFO reporting checklist. AFO: ankle-foot orthosis.
[DOCX File , 39 KB - rehab_v13i1e93423_app2.docx]

Multimedia Appendix 3
PRISMA checklist. Items 24b and 24c can be added during the revision process.
[PDF File (Adobe PDF File), 554 KB - rehab_v13i1e93423_app3.pdf]

References

1. Perry J, Burnfield J. Gait Analysis: Normal and Pathological Function. Boca Raton, FL: CRC Press; Jun 1, 2024.
2. Demir YP. Neuromuscular diseases and rehabilitation. In: Suzuki T, editor. Neurological Physical Therapy. London, United Kingdom: Intechopen; May 10, 2017.
3. Ploeger HE, Bus SA, Brehm M, Nollet F. Ankle-foot orthoses that restrict dorsiflexion improve walking in polio survivors with calf muscle weakness. Gait Posture 2014 Jul;40(3):391-398. [doi: 10.1016/j.gaitpost.2014.05.016] [Medline: 24947072]
4. Malas BS. What variables influence the ability of an AFO to improve function and when are they indicated? Clin Orthop Relat Res 2011 May;469(5):1308-1314 [FREE Full text] [doi: 10.1007/s11999-010-1684-y] [Medline: 21116758]
5. Daryabor A, Arazpour M, Aminian G. Effect of different designs of ankle-foot orthoses on gait in patients with stroke: a systematic review. Gait Posture 2018 May;62:268-279. [doi: 10.1016/j.gaitpost.2018.03.026] [Medline: 29587246]
6. Ploeger HE, Waterval NFJ, Nollet F, Bus SA, Brehm M. Stiffness modification of two ankle-foot orthosis types to optimize gait in individuals with non-spastic calf muscle weakness - a proof-of-concept study. J Foot Ankle Res 2019;12:41 [FREE Full text] [doi: 10.1186/s13047-019-0348-8] [Medline: 31406508]
7. Kerkum Y, Philippart W, Houdijk H. The effects of footplate stiffness on push-off power when walking with posterior leaf spring ankle-foot orthoses. Clin Biomech (Bristol) 2021;88:105422. [doi: 10.1016/j.clinbiomech.2021.105422] [Medline: 34271367]
8. van Kouwenhove L, Verkerke G, Postema K, Dekker R, Hijmans J. Effect of different forefoot rocker radii on lower-limb joint biomechanics in healthy individuals. Gait Posture 2021;86:150-156 [FREE Full text] [doi: 10.1016/j.gaitpost.2021.03.008] [Medline: 33725583]
9. Hutchins S, Bowker P, Geary N, Richards J. The biomechanics and clinical efficacy of footwear adapted with rocker profiles—evidence in the literature. Foot (Edinb) 2009;19(3):165-170. [doi: 10.1016/j.foot.2009.01.001] [Medline: 20307470]
10. Eddison N, Chockalingam N. The effect of tuning ankle foot orthoses-footwear combination on the gait parameters of children with cerebral palsy. Prosthet Orthotics Int 2013;37(2):95-107. [doi: 10.1177/0309364612450706] [Medline: 22833518]
11. Owen E. The importance of being earnest about shank and thigh kinematics especially when using ankle-foot orthoses. Prosthet Orthotics Int 2010;34(3):254-269. [doi: 10.3109/03093646.2010.485597] [Medline: 20738230]
12. Sobhani S, Hijmans J, van den Heuvel E, Zwerver J, Dekker R, Postema K. Biomechanics of slow running and walking with a rocker shoe. Gait Posture 2013;38(4):998-1004. [doi: 10.1016/j.gaitpost.2013.05.008] [Medline: 23770233]

XSL•FO
RenderX

13. Paton J, Hatton AL, Rome K, Kent B. Effects of foot and ankle devices on balance, gait and falls in adults with sensory perception loss: a systematic review. JBI Database System Rev Implementations Rep 2016;14(12):127-162 [FREE Full text] [doi: 10.11124/JBISRIR-2016-003229] [Medline: 28009675]
14. Wilkins RA, Chapman LS, Emmel JC, Flannery T, Chapman GJ, Walwyn REA, et al. A systematic review and narrative synthesis of footwear and orthotic devices used in the management of ankle haemarthrosis and haemarthropathy in haemophilia. Haemophilia 2022;28(3):422-436 [FREE Full text] [doi: 10.1111/hae.14521] [Medline: 35245413]
15. Adamczyk PG, Collins SH, Kuo AD. The advantages of a rolling foot in human walking. J Exp Biol 2006;209(Pt 20):3953-3963. [doi: 10.1242/jeb.02455] [Medline: 17023589]
16. Esposito ER, Ruble MD, Ikeda AJ, Wilken JM. The effect of custom carbon ankle-foot orthosis alignment on roll-over shape and center of pressure velocity. Prosthet Orthotics Int 2021;45(2):147-152. [doi: 10.1177/0309364620971407] [Medline: 33225810]
17. Hansen AH, Childress DS, Knox EH. Roll-over shapes of human locomotor systems: effects of walking speed. Clin Biomech (Bristol) 2004;19(4):407-414. [doi: 10.1016/j.clinbiomech.2003.12.001] [Medline: 15109762]
18. Hansen AH, Childress DS. Effects of adding weight to the torso on roll-over characteristics of walking. J Rehabil Res Dev 2005;42(3):381-390. [doi: 10.1682/jrrd.2004.04.0048] [Medline: 16187250]
19. Kuo AD, Donelan JM, Ruina A. Energetic consequences of walking like an inverted pendulum: step-to-step transitions. Exerc Sport Sci Rev 2005;33(2):88-97. [doi: 10.1097/00003677-200504000-00006] [Medline: 15821430]
20. Fatone S, Hansen AH. Effect of ankle-foot orthosis on roll-over shape in adults with hemiplegia. J Rehabil Res Dev 2007;44(1):11-20. [doi: 10.1682/jrrd.2006.08.0090] [Medline: 17551854]
21. Hansen AH, Childress DS. Effects of shoe heel height on biologic rollover characteristics during walking. J Rehabil Res Dev 2004;41(4):547-554. [doi: 10.1682/jrrd.2003.06.0098] [Medline: 15558383]
22. Waterval NFJ, Brehm MA, Harlaar J, Nollet F. Description of orthotic properties and effect evaluation of ankle-foot orthoses in non-spastic calf muscle weakness. J Rehabil Med 2020;52(3):jrm00026 [FREE Full text] [doi: 10.2340/16501977-2642] [Medline: 31934727]
23. Fatone S, Sorci E, Hansen A. Effects of clinically prescribed ankle foot orthoses on ankle-foot roll-over shapes: a case series. JPO J Prosthetics Orthotics 2009;21(4):196-203. [doi: 10.1097/JPO.0b013e3181ba89c0]
24. Shang J, Geng X, Wang C, Chen L, Zhang C, Huang J, et al. Influences of high-heeled shoe parameters on gait cycle, center of pressure trajectory, and plantar pressure in young females during treadmill walking. J Orthop Surg (Hong Kong) 2020;28(2):2309499020921978 [FREE Full text] [doi: 10.1177/2309499020921978] [Medline: 32390534]
25. Forghany S, Nester CJ, Richards B, Hatton AL, Liu A. Rollover footwear affects lower limb biomechanics during walking. Gait Posture 2014;39(1):205-212. [doi: 10.1016/j.gaitpost.2013.07.009] [Medline: 23910726]
26. Singer ML, Kobayashi T, Lincoln LS, Orendurff MS, Foreman KB. The effect of ankle-foot orthosis plantarflexion stiffness on ankle and knee joint kinematics and kinetics during first and second rockers of gait in individuals with stroke. Clin Biomech (Bristol) 2014;29(9):1077-1080 [FREE Full text] [doi: 10.1016/j.clinbiomech.2014.09.001] [Medline: 25241248]
27. Waterval NFJ, Nollet F, Harlaar J, Brehm M. Modifying ankle foot orthosis stiffness in patients with calf muscle weakness: gait responses on group and individual level. J Neuroeng Rehabil 2019;16(1):120 [FREE Full text] [doi: 10.1186/s12984-019-0600-2] [Medline: 31623670]
28. Wiest DR, Waters RL, Bontrager EL, Quigley MJ. The influence of heel design on a rigid ankle-foot orthosis. Orthotics Prosthetics 1979;33(4):3-10 [FREE Full text]
29. Torricelli D, Gonzalez J, Weckx M, Jiménez-Fabián R, Vanderborght B, Sartori M, et al. Human-like compliant locomotion: state of the art of robotic implementations. Bioinspiration Biomimetics 2016;11(5):051002 [FREE Full text] [doi: 10.1088/1748-3190/11/5/051002] [Medline: 27545108]
30. Brown D, Wertsch JJ, Harris GF, Klein J, Janisse D. Effect of rocker soles on plantar pressures. Arch Phys Med Rehabil 2004;85(1):81-86. [doi: 10.1016/s0003-9993(03)00374-5] [Medline: 14970973]
31. Greve C, Schuitema D, Otten B, van Kouwenhove L, Verhaar E, Postema K, et al. Biomechanical effects of rocker shoes on plantar aponeurosis strain in patients with plantar fasciitis and healthy controls. PLoS One 2019;14(10):e0222388 [FREE Full text] [doi: 10.1371/journal.pone.0222388] [Medline: 31600227]
32. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
33. Halman A, Oshlack A. Catchii: empowering literature review screening in healthcare. Res Synth Methods 2024;15(1):157-165. [doi: 10.1002/jrsm.1675] [Medline: 37771210]
34. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998;52(6):377-384 [FREE Full text] [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
35. Hooper P, Jutai JW, Strong G, Russell-Minda E. Age-related macular degeneration and low-vision rehabilitation: a systematic review. Can J Ophthalmol 2008;43(2):180-187. [doi: 10.3129/i08-001] [Medline: 18347620]
36. Ridgewell E, Dobson F, Bach T, Baker R. A systematic review to determine best practice reporting guidelines for AFO interventions in studies involving children with cerebral palsy. Prosthet Orthotics Int 2010;34(2):129-145. [doi: 10.3109/03093641003674288] [Medline: 20384548]

XSL•FO
RenderX

37. Farmani F, Mohseni-Bandpei M, Bahramizadeh M, Aminian G, Abdoli A, Sadeghi-Goghari M. The influence of rocker bar ankle foot orthosis on gait in patients with chronic hemiplegia. J Stroke Cerebrovasc Dis 2016;25(8):2078-2082. [doi: 10.1016/j.jstrokecerebrovasdis.2016.02.021] [Medline: 27265095]
38. Sutdet P, Polhan N, Prakotmongkol V, Poomulna J, Trifani F, Iswana T, et al. The effect of tibia inclination on gait pattern of stroke patients with genu recurvatum who wear an ankle foot orthosis: a pilot study. Technol Disabil 2019;31(3):77-82. [doi: 10.3233/tad-180211]
39. Dhalla R, Johnson JE, Engsberg J. Can the use of a terminal device augment plantar pressure reduction with a total contact cast? Foot Ankle Int 2003;24(6):500-505. [doi: 10.1177/107110070302400609] [Medline: 12854672]
40. Farmani F, Bandpei MAM, Bahramizadeh M, Aminian G, Nikoo MR, Sadeghi-Goghari M. The effect of different shoes on functional mobility and energy expenditure in post-stroke hemiplegic patients using ankle-foot orthosis. Prosthet Orthotics Int 2016;40(5):591-597. [doi: 10.1177/0309364615592704] [Medline: 26184035]
41. Kearney RS, Lamb SE, Achten J, Parsons NR, Costa ML. In-shoe plantar pressures within ankle-foot orthoses: implications for the management of achilles tendon ruptures. Am J Sports Med 2011;39(12):2679-2685. [doi: 10.1177/0363546511420809] [Medline: 21908718]
42. Arvin M, Kamyab M, Moradi V, Hajiaghaei B, Maroufi N. Influence of modified solid ankle-foot orthosis to be used with and without shoe on dynamic balance and gait characteristic in asymptomatic people. Prosthet Orthotics Int 2013;37(2):145-151. [doi: 10.1177/0309364612454159] [Medline: 22907949]
43. Jagadamma KC, Owen E, Coutts FJ, Herman J, Yirrell J, Mercer TH, et al. The effects of tuning an ankle-foot orthosis footwear combination on kinematics and kinetics of the knee joint of an adult with hemiplegia. Prosthet Orthotics Int 2010;34(3):270-276. [doi: 10.3109/03093646.2010.503225] [Medline: 20738231]
44. Masoumi S, Aminian G, Daryabor A. The effects on functional balance in hemiplegic stroke individuals wearing an ankle foot orthosis with rocker bottom shoes. J Rehabil Sci Res (JRSR) 2020;7(3):141-145. [doi: 10.30476/jrsr.2020.86778.1094]
45. Kerkum YL, Houdijk H, Brehm M, Buizer AI, Kessels ML, Sterk A, et al. The shank-to-vertical-angle as a parameter to evaluate tuning of ankle-foot orthoses. Gait Posture 2015;42(3):269-274. [doi: 10.1016/j.gaitpost.2015.05.016] [Medline: 26050873]
46. Ikeda AJ, Fergason JR, Wilken JM. Effects of altering heel wedge properties on gait with the intrepid dynamic exoskeletal orthosis. Prosthet Orthotics Int 2018;42(3):265-274. [doi: 10.1177/0309364617728116] [Medline: 28870146]
47. Choi H, Bjornson K, Fatone S, Steele KM. Using musculoskeletal modeling to evaluate the effect of ankle foot orthosis tuning on musculotendon dynamics: a case study. Disabil Rehabil Assist Technol 2016;11(7):613-618 [FREE Full text] [doi: 10.3109/17483107.2015.1005030] [Medline: 25640240]
48. Daryabor A, Aminian G, Arazpour M, Baniasad M, Yamamoto S. The effects of ankle-foot orthoses with plantar flexion stop and plantar flexion resistance using rocker-sole shoes on stroke gait: a randomized-controlled trial. Turk J Phys Med Rehabil 2021;67(4):449-461 [FREE Full text] [doi: 10.5606/tftrd.2021.6448] [Medline: 35141485]
49. Schmalz T, Blumentritt S, Drewitz H, Freslier M. The influence of sole wedges on frontal plane knee kinetics, in isolation and in combination with representative rigid and semi-rigid ankle-foot-orthoses. Clin Biomech (Bristol) 2006;21(6):631-639. [doi: 10.1016/j.clinbiomech.2006.02.004] [Medline: 16567026]
50. Eddison N, Chockalingam N, Osborne S. Ankle foot orthosis-footwear combination tuning: an investigation into common clinical practice in the United Kingdom. Prosthet Orthotics Int 2015;39(2):126-133. [doi: 10.1177/0309364613516486] [Medline: 24567349]
51. Owen E, Fatone S, Hansen A. Effect of walking in footwear with varying heel sole differentials on shank and foot segment kinematics. Prosthet Orthotics Int 2018;42(4):394-401. [doi: 10.1177/0309364617728119] [Medline: 28884616]
52. Hansen AH, Childress DS, Knox EH. Prosthetic foot roll-over shapes with implications for alignment of trans-tibial prostheses. Prosthet Orthotics Int 2000;24(3):205-215. [doi: 10.1080/03093640008726549] [Medline: 11195355]
53. Totah D, Menon M, Jones-Hershinow C, Barton K, Gates DH. The impact of ankle-foot orthosis stiffness on gait: a systematic literature review. Gait Posture 2019;69:101-111. [doi: 10.1016/j.gaitpost.2019.01.020] [Medline: 30708092]

Abbreviations

AFO: ankle-foot orthosis
AFO-FC: ankle-foot orthosis and footwear combination
COM: center of mass
COP: center of pressure
EMG: electromyography
GRF: ground reaction force
HSD: heel-sole differential
PROSPERO: International Prospective Register of Systematic Reviews
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
ROS: roll-over shape
SACH: solid ankle cushion heel

XSL•FO
RenderX

SVA: shank-to-vertical angle

Edited by B Schmitz; submitted 13.Feb.2026; peer-reviewed by J Fei, N Waterval, KM Anderson; comments to author 28.Apr.2026; accepted 10.Jun.2026; published 13.Oct.2026.

Please cite as:
Millar A, Greve C, Hijmans JM
Ankle-Foot Orthosis and Footwear Interactions in Adult Healthy and Patient Populations: Systematic Review of Biomechanical Outcomes
JMIR Rehabil Assist Technol 2026;13:e93423
URL: https://rehab.jmir.org/2026/1/e93423
doi:10.2196/93423
PMID:

©Alexandra Millar, Christian Greve, Juha M Hijmans. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 13.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Feasibility of Technology-Based Interventions in Disability and Rehabilitation: Definitions, Considerations, and Dimensions

Josh Shore[1], MSc; Amy S Hwang[2,3], PhD; Rosalie H Wang[1,2], PhD

[1]Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[2]Department of Occupational Science and Occupational Therapy, Temerty Faculty of Medicine, University of Toronto, 500 University Avenue, Toronto, ON, Canada

[3]Centre de recherche de l'Institut universitaire de gériatrie de Montréal (CRIUGM), CIUSSS Centre-Sud-de-l'île-de-Montréal, Montréal, QC, Canada

Corresponding Author:
Rosalie H Wang, PhD
Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Abstract

Technology-based interventions in the field of disability and rehabilitation, which serve assistive, therapeutic, and/or service delivery functions, are considered complex due to the skills required of providers and recipients, degree of individual tailoring, and diversity of use settings. Feasibility studies are an important step in the evolution of complex interventions that can help refine the intervention, inform implementation, and prevent wasted resources. However, guidance is lacking regarding specific considerations for feasibility studies of technology-based interventions in disability and rehabilitation, which leaves researchers and developers reliant on resources from other fields that do not address important technology properties. To advance the field, context-specific definitions, considerations, and evaluation dimensions must be explicitly outlined to ensure that feasibility studies are constructively designed to meet the unique needs of these interventions. In this viewpoint article, we (1) propose a definition and framework for feasibility studies within the specific context of technology-based disability and rehabilitation interventions, (2) highlight important and unique imperatives for feasibility studies of these interventions, and (3) articulate relevant feasibility dimensions and associated evaluation criteria for these interventions. Building on previous work, we distinguish between feasibility studies, wherein we focus on iterative intervention refinement by addressing key development questions (eg, usability), and pilot studies, which are small-scale versions of a larger study that will evaluate intervention outcomes. Integrating previous typologies, we present 13 feasibility dimensions relevant to technology-based interventions and provide sample evaluation criteria, focusing on the intervention itself rather than study design considerations (eg, trial management). This information may be useful for research and development communities (academic, clinical, or industry) to inform comprehensive feasibility studies that examine unique aspects of technology-based interventions to promote real-world impact. This contribution encourages greater harmonization of terminology and evaluation methods to streamline interpretation and comparison across studies.

(JMIR Rehabil Assist Technol 2026;13:e79026) doi:10.2196/79026

KEYWORDS

rehabilitation; disabled persons; self-help devices; biomedical technology; feasibility studies; research design

Introduction

Technology-Based Interventions in Disability and Rehabilitation

Innovative technology-based solutions are increasingly sought to meet the evolving needs of health and social service systems, particularly since the onset of the COVID-19 pandemic [1-3]. In the field of disability and rehabilitation, technology-based interventions involve structured use of technological devices or systems to achieve targeted goals related to health, function, or participation. Technology-based interventions can be broadly classified as serving assistive, therapeutic, and/or service delivery purposes [4]. Assistive products, such as wheelchairs, hearing aids, and communication devices, help optimize a person's functioning to reduce disability [5]. Therapeutic interventions restore or improve function through the use of technologies including robotics, alternative reality systems, and exercise-based video games. Technology-mediated service delivery models replace or augment existing services to improve care access or maximize efficiency through approaches like telerehabilitation, web or mobile apps, and smart home monitoring systems. Technology-based interventions encompass the core technology products, such as the devices, software, or platform, as well as services (eg, training) and systems (eg, smart home system) that enable use of the intervention. Target users of such interventions may include rehabilitation patients, people with disabilities, clinicians, or care partners.

XSL•FO
RenderX

Regardless of intended function and target population, technology-based interventions are considered complex for several reasons. They often require providers and recipients to learn specific skills and necessitate customization or tailoring to individuals and/or specific contexts or settings of use [6]. Varying these factors may also engender technical complexity (eg, privacy, data security, reliability, usability, and scalability) that needs to be managed. Rapid testing may help uncover design requirements that are not identified during initial design activities and guide iterative intervention refinement.

Feasibility and Pilot Studies

Feasibility is a vital facet in the Medical Research Council (MRC) framework for complex intervention development and evaluation [7]. Dictionaries define feasibility as the possibility, capability, likelihood, or reasonableness that something can be made, done, or achieved [8,9]. In the context of health research, feasibility studies refer to preliminary research before a main study to determine whether future work can and should be done, and to inform the design of such work [10]. According to the MRC framework, feasibility studies should address both the intervention itself (ie, content and delivery), as well as study design and process considerations for future evaluation (eg, recruitment, data collection, outcome measurement) [7]. Common aspects of feasibility studies include a focus on process rather than outcomes, exploration of key areas of uncertainty, and evaluation of multiple dimensions based on predetermined success or progression criteria [11,12]. Thorough feasibility studies prior to larger-scale evaluation contribute to better interventions, stronger evaluation designs, greater likelihood of successful implementation, and fewer wasted resources [7,11].

Historically, the terms feasibility and pilot have been used interchangeably and inconsistently in the literature, causing confusion about their scope and objectives [13,14]. In response to this lack of terminological clarity and growing interest in the field, Eldridge et al [10] developed an international consensus–based conceptual framework for feasibility and pilot studies that has been widely cited. According to their framework, feasibility studies ask whether something can be done, if it should be done, and *how* it should be done. Pilot studies are considered a subset of feasibility studies that ask similar questions but do so by conducting a small-scale version of a future study (or parts of it) [10]. It is now generally accepted that feasibility studies represent an overarching type of research concerning whether something will work, with pilot studies as a subset that specifically aim to test the operation of a future larger study [10,14,15].

Rooted in phased clinical trial practices, the Eldridge et al [10] framework focuses on studies conducted in preparation for a randomized controlled trial (RCT). However, RCTs are often inappropriate for technology-based interventions in disability and rehabilitation due to practical challenges (eg, need for prolonged intervention use across diverse settings, difficulty matching the speed of technology advances), scientific limitations (eg, lack of masking, and limited generalizability to real-world contexts), and ethical concerns of withholding interventions with observable benefits (see Wang et al [4] for detailed description of these issues). Given the challenges associated with reliance on phased clinical trial practices, new approaches have been proposed to generate evidence for technology-based interventions in disability and rehabilitation [4,16]. An alternate field-specific conceptualization of feasibility studies is thus warranted.

Feasibility Evaluation of Technology-Based Interventions

To address the challenges above, Wang et al [4] proposed the Framework for Accelerated and Systematic Technology-Based Intervention Development and Evaluation Research (FASTER), a three-phased approach to evidence generation informed by the MRC guidance: (1) development, (2) progressive usability and feasibility evaluation, and (3) scaled evaluation and implementation [4]. The aim of the feasibility phase is to progressively refine the intervention prototype through small-scale iterative usability testing and assessment of other process concerns (eg, acceptability, adherence, and intervention safety) [4]. By uncovering insights on how to improve intervention features, the feasibility phase is essential to promote future uptake and sustained use [4]. Within the feasibility phase, FASTER promotes a variety of research designs and methodologies to gain a comprehensive understanding of the intervention, its users, and their contexts of use. Teams are encouraged to progressively address issues identified in previous work and tailor the feasibility evaluation to key research questions, uncertainties, and intervention maturity [4]. However, no guidance is offered on what aspects of feasibility may be relevant.

To our knowledge, no existing frameworks delineate aspects of feasibility relevant for technology-based interventions. While previous work has provided guidance on the design, key features, and areas of evaluation for feasibility studies across various health-related fields (eg, pharmaceuticals [17], public health [12], psychology [11], and rehabilitation [18,19]), these resources do not account for aspects of feasibility related to interactions between users, technology, and the environment (eg, usability and technical performance), or unique considerations for disability and rehabilitation populations (eg, duration and diversity of intervention use). Specification is important since feasibility studies are applied differently across fields. Technical feasibility studies within engineering or computer science disciplines generally assess whether a system can be developed to perform its intended functions within defined constraints and are therefore typically undertaken in highly controlled settings. Within business settings, feasibility studies assess the likelihood of success for a proposed commercial product or service considering factors like market demand and financial viability. A comprehensive harmonizing framework is needed to provide field-specific guidance for feasibility studies of technology-based interventions, insofar as their potential to perform as expected and deliver intended outcomes.

A harmonized framework could promote more efficient and theoretically driven intervention design, development, and evaluation. Currently, researchers and developers resort to borrowing and “translating” concepts from sources in other

XSL•FO
RenderX

fields—a time-consuming undertaking that requires deep familiarization with new literature. This translation leads to differences in interpretation and application of concepts, making it difficult to compare and consolidate across studies. Without unified and cohesive resources, research in this area remains conceptually and empirically fragmented, delaying the pace and quality of innovation. To advance the field amidst rapid digitization of tools, services, and programs, a comprehensive framework is needed that outlines feasibility considerations, dimensions, and evaluation criteria pertinent to technology-based interventions. Such a framework will help researchers and developers achieve consistency and comparability, providing a foundation to develop best practices for studying the feasibility of technology-based interventions for disability and rehabilitation.

Aims

This viewpoint article aims to (1) propose a definition and framework for feasibility studies within the context of technology-based disability and rehabilitation interventions, (2) highlight important and unique imperatives for feasibility studies of these interventions, and (3) articulate feasibility dimensions and evaluation criteria relevant to these interventions. We focus on the feasibility of the intervention itself rather than study design and process considerations (eg, trial management) for which suitable field-specific guidance is available [18-21]. It is hoped that this discussion encourages comprehensive feasibility studies that examine unique aspects of technology-based interventions to promote real-world research impact.

Authorship Identity Statement

For transparency, we highlight key identities of the authorship team informing this viewpoint. The authors have varied disciplinary backgrounds (kinesiology, occupational therapy, cognitive sciences, rehabilitation science, biomedical engineering, and computer science) and degrees of clinical, research, and lived experiences related to rehabilitation and assistive technologies. Our experiences span diverse populations (eg, traumatic brain injury, respiratory conditions, stroke, and dementia), technologies (eg, telerehabilitation, smart homes, mobility devices, artificial intelligence, robotic rehabilitation, and assistive devices), and user groups (eg, care recipients, partners, providers, industry, and policymakers), across the lifespan (children to older adults). These experiences shape our assumptions about the iterative multifaceted nature of technology-based intervention research and inform our interpretation of the literature.

Defining Feasibility Studies of Technology-Based Interventions

Building on conceptual definitions from Eldridge et al [10] and FASTER [4], we propose the following adapted definitions of feasibility and pilot studies for application to the context of technology-based intervention research in the field of disability and rehabilitation (Textbox 1).

Together, feasibility and pilot studies aim to gather insights about intervention features and their delivery [4,11,19]. This involves a focus on issues related to the intervention (eg, practicality, usability, and safety) and its implementation, as well as gaining a comprehensive understanding of user experience with the intervention across use contexts [4,19,22]. Once an intervention is sufficiently developed and data support its feasibility, pilot studies may be undertaken as part of the scaled evaluation and implementation phase [4,11]. Pilot studies have a greater focus on outcomes and may be more likely to consider additional scientific, logistical, or study process factors relevant to the design of future larger definitive studies, such as recruitment and retention rates, data completion and quality, and effect size estimates for efficacy measures [11,19]. At this stage, technology-based interventions should be evaluated with target users in real-world settings to enable realistic measurement of use and abandonment, effects on health or social outcomes, and economic considerations [4,7]. Pilot studies can be particularly valuable to inform such research and scale evaluation across time periods, user groups, care contexts, or geographic locations.

The next section highlights imperatives for evaluating the feasibility of technology-based interventions in disability and rehabilitation, followed by a summary of potential feasibility dimensions and associated evaluation criteria. To reiterate, we focus on the feasibility of the technology-based intervention itself, rather than study design and process considerations for future larger evaluation studies. Hence, factors related to study procedures (eg, recruitment, data collection, research infrastructure) are not included. Readers are directed to other works that discuss these aspects of study feasibility for intervention research, which are readily applicable to technology-based interventions [17-20]. We recognize there is overlap between intervention and study feasibility, and as such, information presented below may be useful for either purpose.

Textbox 1. Proposed definitions for technology-based interventions in disability and rehabilitation.

Feasibility study: preliminary investigation of a concept of, a part of, or a whole new technology-based intervention to assess its viability and inform its iterative refinement by addressing uncertainty about key development and design questions, such as how well an intervention prototype can be used (ie, usability), whether it should be used, and how it should be used.

Pilot study: specific type of feasibility study that is a small-scale version of a part of or whole, larger future study that evaluates outcomes, such as the use, effectiveness, or impacts of a new technology-based intervention. Pilot studies may be conducted in either controlled or naturalistic real-world contexts and require the intervention to be administered. Pilot studies involve target end-users, or participants who have some of the characteristics or experiences of the target end-users or knowledge of the experiences and contexts of target end-users. Decisions about the study context and participants depend on factors such as technology readiness or ethical concerns (eg, anticipated risk to target end users).

XSL•FO
RenderX

Imperatives for Evaluating the Feasibility of Technology-Based Interventions

Developing a new technology-based intervention is an iterative nonlinear process that often involves cycles of prototyping, evaluation, and refinement [4]. Several rounds of feasibility testing and design modification may be warranted before an intervention is considered for larger-scale evaluation. While most interventions require some form of feasibility study, the inclusion of technology and unique needs of disability and rehabilitation populations introduces additional complexities that merit consideration. Based on our collective research and clinical experiences, literature review, and application of the FASTER framework [4], we synthesize 10 imperatives that feasibility studies for technology-based disability and rehabilitation interventions should address. These imperatives relate to target user populations; technology and intervention design; and regulatory, policy, and funding constraints (Textbox 2). While some are not entirely unique to technology-based interventions, the inclusion of technology complicates how these imperatives are addressed.

XSL•FO
RenderX

Textbox 2. Imperatives for feasibility studies of technology-based interventions in disability and rehabilitation.

Target user population needs

1. Accounting for diversity in user groups and target aims.
   - Technology-based interventions in disability and rehabilitation may have several different user groups and various target health or functional outcomes. [7] The profile of individuals receiving these interventions may also be complicated by heterogeneity, comorbidities or concurrent needs, and involvement in multiple interventions[23].
   - Triangulation of multiple research methods and/or data sources provides a more complete understanding of feasibility and maximizes knowledge gained about the intervention from smaller sample sizes [24].
2. Identifying meaningful user outcomes for health and social function.
   - Technology-based interventions for disability and rehabilitation populations serve functions beyond the medical model (curing symptoms or impairments), addressing issues related to health and social outcomes, such as management of chronic conditions, participation, and social inclusion.
   - While the focus of feasibility studies is not to evaluate intervention efficacy, exploring mechanisms of action and identifying the potential impacts and meaningful outcomes for users can help refine intervention features and inform future studies [4].
3. Creating conditions for meaningful feedback from different user groups.
   - Multiple user groups may interact with or be impacted by rehabilitation and assistive technology interventions, such as persons with a disability or clinical condition, clinicians, caregivers, and family members.
   - Given that different user groups may have distinct needs, values, capacities, and perspectives, creative testing approaches may be needed to simulate the experiences of various groups using the intervention in real-world conditions for them to provide meaningful contextualized feedback.
   - There may be stigma associated with the intervention that requires care to ensure that users feel comfortable testing it and providing feedback.

Technology and intervention considerations

1. Navigating technological complexities.
   - Technology-based interventions in disability and rehabilitation may be complex due to the number of technology components encompassed and their flexibility (ie, customizability) [7], skills required by service providers and recipients,[7] degree of individual tailoring required [7], and diversity of use settings [7]. These factors are complicated by variable levels of digital literacy and technical expertise among service providers and recipients who must learn and adapt to the technology.
   - Diverse study designs and methods from technical and clinical disciplines are thus needed to generate evidence for the feasibility and benefits of technology-based interventions [4].
2. Evaluating technology properties in real-world settings.
   - Evaluating technical performance (eg, system performance and reliability) and usability is critical to the development of new technology-based interventions [25]. These factors should be considered in early feasibility testing and require unique approaches to evaluation.
   - At the feasibility stage, intervention performance and usability should ideally be explored beyond a controlled lab environment [4,25]. Real-world testing considers how technology properties (eg, components and functions) may be affected by environmental conditions such as lighting, flooring, or background noise.
3. Monitoring adaptations needed to accommodate user needs.
   - Modifications to a rehabilitation or assistive intervention and its application may be needed to accommodate individual user needs or specific use settings. Adaptability to enable such modifications is a key feature of implementation success that should be addressed in feasibility testing.
   - Monitoring the frequency, extent, and impact of intervention modifications at the feasibility stage can provide important information regarding adaptability and help identify overly rigid features as well as opportunities for increased flexibility.
4. Incorporating relevant support services.
   - Support services are often required to facilitate technology-based interventions, including eligibility and fit assessment, device customization, user training, product servicing or maintenance, and technical support [4].
   - Incorporating aspects of these support services during feasibility testing can provide a more realistic evaluation of how the intervention may function in real-world environments.
5. Capturing appropriate duration and diversity of use.
   -

XSL•FO
RenderX

Technology-based rehabilitation and assistive intervention use is often not discrete but rather may be used for prolonged or lifelong periods in various settings (eg, hospital or care center, home, workplace, community) that may each be associated with unique privacy, ethical, social, or security concerns.

- Evaluating feasibility at appropriate intervals in diverse real-world contexts is important to identify changes in user experience that may occur during activities with differing demands, in unique environments, or due to changing individual circumstances over time (eg, decline in health status or new social environment).

Regulatory, policy, and funding constraints

1. Navigating and obtaining regulatory approvals.
   - Some rehabilitation and assistive technologies may be classified as medical devices requiring regulatory approval (eg, Health Canada, Federal Drug Administration, and European Medicines Agency).
   - Depending on jurisdictional requirements, specific outcomes or data may be needed to obtain relevant approvals, documentation of which may be initiated at the feasibility stage (eg, history of risk management and safety plans).
   - Development teams should be aware of applicable device classifications and associated requirements so they can be addressed in subsequent studies.
2. Examining variability in funding and reimbursement structures.
   - Funding sources (eg, publicly funded health care, third-party insurance, and self-pay) and reimbursement policies for technology-based interventions vary across jurisdictions, meaning there may be uncertainty about optimal approaches to providing access and establishing commercial feasibility.
   - Data related to the cost of the intervention and/or potential cost savings may support market research and commercial feasibility studies to establish a clear pathway for providing access.

Evaluating the Feasibility of Technology-Based Interventions in Disability and Rehabilitation

Here we outline our approach to consolidate potential areas of focus for feasibility studies of technology-based interventions in disability and rehabilitation. Broadly, we "translated" existing concepts from other fields to better fit technology-based interventions by addressing important considerations in this field (see Imperatives for Evaluating the Feasibility of Technology-Based Interventions) and identified additional areas where needed. The first author identified sources by searching Google Scholar and PubMed using keywords including "technology," "intervention," and "feasibility" and reviewing forward and backward references. Commonalities across sources were clustered, synthesized, and organized into dimensions that were reviewed by other authors and refined through discussion. Based on our knowledge and experiences in the field, we developed guiding questions for each dimension by adapting existing descriptions from the literature where possible and selected sample evaluation criteria. Throughout this process, conflicting ideas were iteratively resolved through deliberation until consensus was reached.

Most feasibility aspects included here are not new, but rather a collection of those described in other fields [11,12,17-20], which we curated and adapted to this unique context. In particular, we drew upon guidance from Bowen et al [12] who proposed 8 areas of focus for feasibility studies in public health research that are relevant to technology-based interventions: (1) acceptability, (2) demand, (3) implementation, (4) practicality, (5) adaptation, (6) integration, (7) expansion, and (8) limited efficacy testing [12]. These areas were a foundation upon which we supplemented additional properties relevant to technology-based interventions.

Dimensions Overview

Table 1 summarizes feasibility dimensions for technology-based interventions with associated guiding questions and sample evaluation criteria. We use the term "dimension" to describe aspects of feasibility, while evaluation criteria refer to different types of information to address the guiding questions. These criteria include relevant constructs or properties for qualitative inquiry, specific measured parameters and outcomes, and potential analysis tools or frameworks. Each dimension is discussed within the subsections below, including literature examples to illustrate specific points. These dimensions encompass clinical, technical, and commercial properties, but are not exhaustive or prescriptive. Rather, it is a starting point for determining pertinent issues and research questions on an individual basis, which may be addressed in a single study or through the course of multiple studies. As some dimensions are highly interrelated and evaluation criteria may be relevant across multiple dimensions, readers are encouraged to select dimensions and criteria based on the unique context and maturity of their intervention, providing appropriate justification.

XSL•FO
RenderX

Table . Proposed feasibility dimensions for technology-based interventions with associated guiding questions and sample evaluation criteria.

Dimension	Guiding question	Sample evaluation criteria
Technical performance[a]	To what extent does the performance of a technology-based intervention satisfy its requirements and meet its intended goals over a specified period of time in a given context?	• Accuracy • Responsivity and responsiveness • Reliability • Sensitivity • Error resolution • Maintenance and repair
Safety	To what extent is a technology-based intervention associated with risk of harm?	• Frequency and severity of adverse events • Perceived safety • Risk management framework
Tolerability	To what extent can a technology-based intervention be used without experiencing unacceptable discomfort?	• Device comfort and fit • Serial symptom assessment • Ecological momentary assessment • Use sessions terminated due to discomfort
Acceptability[b]	To what extent is a technology-based intervention judged by users and other interest holders as suitable, appropriate, or satisfactory towards addressing its intended aim(s)?	• Satisfaction • Perceived appropriateness • Interest or intention to use • Recommendations to others
Demand[b]	To what extent is a technology-based intervention likely to be used by its intended users for its intended purpose?	• Engagement • Sustained use and abandonment • Perceived value • Interest or intention to use • Accessibility • Availability of alternatives
Usability[c]	To what extent does a technology-based intervention enable users to accomplish its stated goal(s) effectively, efficiently, and to their satisfaction?	• Usability attributes • Performance measures (eg, tasks completed, errors) • Behaviors or affective responses (eg, observed or verbalized) • Perceived usability • Degree of exertion, demand, or workload
Adaptability[b]	To what extent can a technology-based intervention be modified to meet the needs of diverse users and use contexts?	• Frequency, extent, and impact of device modifications • Changes in user experience or impact resulting from modifications • Protocol deviations • Device fit assessments • User settings and configuration • Accessibility features
Practicality[b]	To what extent can a technology-based intervention be used in its intended setting(s) given resource constraints such as space, equipment, personnel, and time?	• Space required for use and storage • Duration and ease of equipment set-up and take-down • Compatibility with existing equipment • Personnel and training requirements
Implementation[b]	To what extent can a technology-based intervention be deployed and used as intended in environments with varying levels of control?	• Adherence • Fidelity • Contexts of use • Implementation barriers and facilitators

XSL•FO
RenderX

Dimension	Guiding question	Sample evaluation criteria
Integration[b]	To what extent can a technology-based intervention be incorporated into existing systems?	• Perceived fit within organizational capacities and workflow • Fit within organizational culture • Perceived fit within daily routine • Sustainability
Impact[b]	To what extent does a technology-based intervention show promise in eliciting positive changes for its target population?	• Changes in clinical measures • Appropriateness of measures • Maintenance of changes • Perceived benefits • Caregiver burden • System efficiency
Ethicality	To what extent is a technology-based intervention aligned with ethical values?	• Bioethics (eg, Principlism) • Technology ethics • Prognosis Research Strategy framework • Model for the Ethical Evaluation of Socio-Technical Arrangements framework • Reflective questioning • Socratic approach
Commercial viability	To what extent does a technology-based intervention show promise towards being economically viable for commercialization?	• Production and operation costs • Cost-effectiveness • Economic modeling • Product-market fit

[a]Adapted from Garrett et al [26]

[b]Adapted from Bowen et al [12]

[c]Based on ISO 9241-11:2018 usability definition

Technical Performance

Users of technology-based interventions rely on systems that perform consistently to achieve their goals. Technical flaws and system errors are frequently cited as barriers to technology adoption [27,28]. Technical performance examines how well a technology-based intervention satisfies its functional requirements and meets its intended goals over a specified time period in a given context. This reflects aspects of technology performance evaluation from systems and design engineering [29]. Whereas usability focuses on interaction between the user and the technology, technical performance is more concerned with the function of the technology itself.

In systems engineering, technical performance measures refer to measurable qualities that determine whether a system satisfies key technical requirements [26]. Relevant technical performance measures for a project depend on the type of technology and its intended functions. Table 2 summarizes select technical performance measures with examples from interventions in disability and rehabilitation. While these properties may be initially addressed during intervention development, the feasibility stage extends performance evaluation to real-world contexts to identify features requiring refinement [4].

XSL•FO
RenderX

Table . Select technical performance measures relevant to technology-based interventions in disability and rehabilitation.

Measure	Definition[a]	Example
Accuracy	The closeness of agreement between an observed value and an accepted reference value (ground truth).	In a feasibility study of a sensor motion detection system for fall prevention in hospitalized older adults, Ferrari et al [30] assessed the accuracy of sensor output data by establishing its correlation with observer ratings of participant movement from video recordings.
Responsivity and Responsiveness	How well a system adjusts to commands or cues promptly.	Fang et al [31] evaluated the responsiveness of a Rotational Orthosis for Walking with Arm Swing system by assessing whether a measurable movement was produced that was close to the user's target joint trajectory.
Reliability	Likelihood of a system performing its intended function without error for a given period of time under a specified set of conditions.	Awad et al [32] assessed the reliability of the ReWalk ReStore soft robotic exosuit by measuring device malfunctions during usage.
Availability	Likelihood that a system is operational at a given point in time under a given set of environmental conditions.	Marschollek et al [33] assessed the technical feasibility of a smart home monitoring system by measuring the downtime of installed systems in relation to their up-times.

[a]According to American Society for Quality [34].

Safety

Several authors identify safety as a critical component of intervention feasibility, and it is among the most common measures reported by feasibility studies in rehabilitation research [20]. Safety refers to the risk of harm imposed on intervention recipients or others affected by the intervention [35]. In some cases, safety is closely related to technical performance. "Design for safety" is an approach whereby products and systems are purposefully engineered to prioritize the safety of people who interact with them, with specific features designed to proactively address safety risks [36]. For example, a robotic exoskeleton could be made "fail-safe" by defaulting into a safe system state upon detecting an operational error to prevent injury caused by excessive force applied to a limb [37]. Thus, intervention safety may depend on appropriate technical performance.

Safety is typically operationalized as the extent to which an intervention is associated with adverse events, such as injuries or disease exacerbations. Measuring the frequency and severity of adverse events provides information about the risks and potential harm of an intervention in a given setting. Defining adverse events prior to commencing evaluation helps ensure relevant incidents are captured and properly managed. For technology-based interventions, researchers may wish to evaluate users' ability to complete tasks safely while using the device. This can be accomplished through behavioral observation and coding safety incidents and near misses. Perceived safety from the perspective of users and members of their support system may also provide valuable insight, particularly for devices that involve loss of control (eg, robotic exoskeletons), or those that target vulnerable populations like people with dementia.

Wang et al [38] provide an example of multidimensional safety evaluation in a study evaluating contact sensors for powered wheelchairs among older adults with dementia [38]. A multiple single-subject mixed methods design was used, involving quantification of safety-related incidents observed by the investigator during testing with older adults in nursing homes, and qualitative exploration of perceived safety via debrief interviews with participants and focus groups with staff. Investigator observations showed that while the sensors successfully detected most obstacles, minor design faults (eg, gaps between panels) led to some performance errors, and more sensors were needed for complete environmental coverage to minimize injury risk. The qualitative analysis revealed that resident and staff perceptions of safety were mixed, leading the authors to identify a need for further safety improvements by increasing sensor reliability and coverage [38]. This study demonstrates the close relationship between safety and technical performance and underscores the importance of assessing and redesigning for safety when developing technology-based interventions for people with cognitive impairment.

Tolerability

Related to safety is the concept of tolerability. The US Food and Drug Administration defines tolerability as the degree to which adverse effects can be tolerated by participants [35]. For technology-based interventions, we conceptualize tolerability as the degree to which users can engage with the intervention without experiencing unacceptable discomfort. Tolerability is particularly important for interventions that are associated with discomfort, such as neuromodulation [39], brain-computer interfaces [40], and prosthetic devices [41], although it may also be relevant for other types of interventions that serve populations with sensitivities (eg, chronic pain, headaches, and autism spectrum disorder). It has important implications for other dimensions such as acceptability, usability, demand, implementation, and ultimately adoption. For example, device fit and comfort are critical factors in long-term use and abandonment of upper limb prosthetics [41]. Assessing tolerability can also help determine appropriate treatment dosage.

XSL•FO
RenderX

Self-report measures are useful to capture side effects or changes in participants' condition while interacting with the intervention. For example, symptoms or discomfort can be rated before, during, and after intervention use with condition-specific instruments, or generic subjective rating scales (eg, Visual Analog Scale, Visual Rating Scale, Numeric Rating Scale, and Borg Scale) [42]. Validated device-specific comfort questionnaires are also available for some interventions, such as wheelchairs [43] and wearable computers [44]. When testing interventions in real-world settings, ecological momentary assessment (EMA) can be a useful approach. EMA involves repeatedly sampling participants' current condition in real-time within their natural environment [45]. Technology-based interventions offer unique opportunities to embed EMA (eg, mobile app prompts), allowing teams to gather detailed and context-rich data on participant discomfort, stress, fatigue, and frustration. By assessing users' condition at various time points in everyday environments, EMA can provide valuable insight into factors that influence tolerability and help monitor cumulative stress from long-term use.

Some objective methods of assessing comfort have been described. Physiological measures (eg, heart rate, skin conductance, and muscle activity) can identify adverse stress responses associated with the intervention. Pressure sensors and movement analysis techniques have been used to capture compensatory movements that indicate discomfort [42]. Finally, observational analysis of participants' behavior during the intervention can help identify signs of discomfort, fatigue, or frustration [42].

Intervention usage (eg, frequency, duration, or maintenance) is sometimes reported as a measure of tolerability by indicating whether demands are appropriate, or if participants avoid the intervention due to discomfort. In such cases, it should be clear why usage metrics are framed as measures of tolerability rather than engagement or adherence, and how information about tolerability is inferred. For example, a study on a smart glasses social communication aid for people with autism assessed tolerability based on participants' ability to wear the device during coaching sessions [46]. The authors were interested in how well participants tolerated the smart glasses due to their sensory and cognitive challenges, rather than whether they chose to use the smart glasses (ie, engagement) or comply with a prescribed treatment plan (adherence).

A good example of tolerability evaluation can be found in a case study from Ryan et al [47] on the integration of transcranial direct current stimulation (tDCS) into a pediatric inpatient brain injury rehabilitation program [47]. Individual session tolerance was monitored using pre- and postratings of common tDCS side effects (eg, headache, fatigue, itching, and tingling). A postintervention Tolerability Questionnaire [48] was also used to assess tolerance across the study period, which involved the participant ordering a list of common activities from most to least enjoyable (eg, playing games, watching TV, going to the dentist, and getting a needle), and rating their tDCS experience relative to those activities. With minimal symptom exacerbation within sessions and a favorable rating of tDCS relative to other activities, the authors concluded that tDCS was tolerated well by the participant and could be incorporated into inpatient rehabilitation [47].

Acceptability

Acceptability is a foundational aspect of feasibility that describes how individuals involved in an intervention react to it [12]. While empathizing with users to understand their needs and preferences is a vital step in designing an intervention, perspectives may change once users actually experience it, underscoring the need for user involvement throughout development and evaluation [4]. According to Nielsen [49], technology is accepted when it satisfies the needs and requirements of its users and other interest holders. Thus, acceptability can be defined as the extent to which users of a technology-based intervention and other interest holders judge it to be suitable, appropriate, or satisfactory for addressing its intended aim. Acceptability is an important predictor of use behavior that precedes technology adoption [25]. Ultimately, successful implementation relies on procedures that are acceptable and meet the needs of individuals who deliver, receive, or otherwise engage with an intervention [11,50].

Efforts have been made to better understand and operationalize acceptability. Several theories have been proposed to explain technology acceptance, the most prominent of which include Technology Acceptance Model (TAM and TAM2), Combined Technology Acceptance and Theory of Planned Behavior Model, Motivational Model, and Innovation Diffusion Theory [51]. Combining principles from many models, the Unified Theory of Acceptance and Use of Technology (UTAUT and UTAUT2) posits that acceptance and intention to use technology are dependent on one's beliefs about performance expectancy (utility toward one's goals), effort expectancy (ease of use), and facilitating conditions (organizational and technical support), as well as subjective norms regarding their use of the technology, all of which are moderated by gender, age, experience, and voluntariness of use [52]. For health care interventions, Sekhon et al [50] developed the Theoretical Framework of Acceptability, which conceptualizes acceptability through 7 constructs: (1) affective attitude, (2) burden, (3) perceived effectiveness, (4) ethicality, (5) intervention coherence, (6) opportunity costs, and (7) self-efficacy [50]. It is clear that the acceptability of technology-based interventions in health care is complex and multidimensional, with several interacting factors related to beliefs and attitudes that overlap with other feasibility dimensions.

Satisfaction is considered an indicator of acceptability that can be measured quantitatively using questionnaires [50]. Sekhon et al [50] developed the Theoretical Framework of Acceptability Questionnaire, a general measure reflecting the Theoretical Framework of Acceptability constructs that can be adapted for various health care interventions [53]. Other generic satisfaction scales include the Client Satisfaction Questionnaire [54], Acceptability of Intervention Measure [55], and Intervention Appropriateness Measure [55]. For technology-based interventions, the TAM questionnaire evaluates acceptance based on perceived usefulness and ease of use [56]. Textbox 3 highlights acceptability and satisfaction questionnaires relevant to technology-based interventions [4]. Comprehensive reviews

XSL•FO
RenderX

of satisfaction and acceptability tools for assistive technologies [57] and mobile health programs [58] are also available.

Textbox 3. Sample acceptability and satisfaction questionnaires relevant to technology-based interventions in disability and rehabilitation.

Generic health care intervention measures

- Theoretical Framework of Acceptability Questionnaire [53]
- Client Satisfaction Questionnaire [54]
- Acceptability of Intervention Measure [55]
- Intervention Appropriateness Measure [55]

Generic technology measures

- Technology Acceptance Model Questionnaire [56]
- Unified Theory of Acceptance and Use of Technology Questionnaire [59]

Assistive technologies

- Quebec User Evaluation of Satisfaction with Assistive Technology 2.0 [60]
- Almere Technology Acceptance Questionnaire [61]
- Remote Monitoring Satisfaction Survey [62]
- Satisfaction with Prosthesis Questionnaire [63]
- Satisfaction with Amplification in Daily Life Questionnaire [64]

Therapeutic technologies

- Physiotherapy Mobile Acceptance Questionnaire [65]
- Exergame Enjoyment Questionnaire [66]
- Virtual Rehabilitation User Satisfaction Evaluation Questionnaire [67]
- Immersive Experience Questionnaire [68]
- Questionnaire for User Interaction Satisfaction [69]

Telehealth and mobile service delivery

- Telemedicine Satisfaction Questionnaire [70]
- Telemedicine Satisfaction and Usefulness Questionnaire [71]
- Telehealth Satisfaction Scale [1-3]
- Service User Technology Acceptability Questionnaire [72]
- Website Satisfaction Scale [73]
- Digital Health Acceptability Questionnaire [74]

Although satisfaction questionnaires are simple quantifiable measures, qualitative methods often elicit a deeper understanding of user experiences and acceptability at the feasibility stage [11,75]. Interviews and focus groups can be conducted after users have tested an intervention to gather perspectives regarding appropriateness of the intervention and delivery approach to inform iterative refinement. Qualitative inquiry is particularly valuable to probe the unique and nuanced aspects of technology acceptability in the field of disability and rehabilitation, such as social implications of observable assistive technologies across use contexts.

Importantly, intervention maturity and stage of development influence acceptability. For instance, early stages may involve low-fidelity intervention prototypes that users trial on a single occasion and must imagine the experience of living with in real-life circumstances. Conversely, high-fidelity prototypes in later stages allow users to trial the intervention in real-world environments, over longer periods, and in more diverse contexts. Acceptability should thus be viewed as a dynamic attribute that warrants consideration throughout evaluation.

Demand

Technology innovations in disability and rehabilitation are designed to address unmet needs by providing a new solution or service not available through conventional means. However, the likelihood that an intervention will be used to serve this function cannot be fully understood until in the hands of its target population. Based on interaction and experience with a prototype, demand evaluates the likelihood that a technology-based intervention will be used by intended users [12]. Demand is closely tied to acceptability, as users are more likely to engage with an intervention that is deemed suitable or

XSL•FO
RenderX

satisfactory. By estimating future use, demand also has implications for the commercial viability of technology-based interventions.

A relatively straightforward indicator of demand is user engagement. Cavanaugh [76] describes engagement as efforts by users to start and continue with an intervention. Engagement differs from adherence in that it implies an active and volitional choice to interact with an intervention in a meaningful way, rather than passive compliance with a prescribed plan. Therefore, it reflects the extent to which a technology-based intervention is likely to be used, particularly for those that involve free use in uncontrolled contexts, such as home-based exergames, virtual reality programs, and assistive devices. Evidence across various types of technology-based interventions shows that user engagement is an important predictor of future use and intervention success [77]. In digital health research, engagement has been conceptualized as a multidimensional construct involving observable behavior (ie, usage parameters) and subjective experience characterized by attention, interest, and affect [78]. Behavioral usage parameters relevant to gauging demand include the frequency, duration, and depth (ie, variety or intensity) of usage [78]. For interventions that use websites, mobile apps, gaming systems, or other digital devices, these metrics may be captured automatically through system logs. The User Engagement Scale measures engagement experience based on aesthetic appeal, focused attention, novelty, perceived usability, felt involvement, and endurability [79].

Engagement should ideally be monitored longitudinally, as diminished use and abandonment are common in digital health and assistive technology interventions [80,81]. While engagement may be high initially, it often tapers over time as novelty fades. Considering this "law of attrition," Eysenbach [81] recommends analyzing determinants of engagement with eHealth programs and characteristics of subgroups that maintain and discontinue engagement [81]. Feedback from participants regarding perceived value of the intervention and interest in continuing its use in their routine practice or daily activities can help understand demand. Qualitative inquiry is valuable to obtain a nuanced understanding of engagement by addressing questions such as *how*, *why*, and *when* users choose to engage with the intervention, and to appreciate psychological or social experiences of interacting with the intervention. This can help drive intervention refinement to promote more meaningful and sustained use, ultimately increasing demand.

Usability

Rehabilitation and assistive technology interventions aim to provide a useful and *usable* tool to accomplish the goal of optimizing function and reducing disability. According to the International Organization for Standardization (ISO) 9241-11:2018, usability refers to how well a system, product, or service can be used by its target users in a specific context to achieve their goals with effectiveness, efficiency, and satisfaction [82]. Usability is a common area of evaluation for rehabilitation and assistive technologies with important implications for their acceptance, adoption, and effectiveness [83,84]. It is a key component of feasibility studies that is particularly important for iteratively refining intervention features [4].

The ISO emphasizes the multidimensional nature of usability, highlighting effectiveness, efficiency, and satisfaction as key elements [82]. Similarly, Nielsen [85] identifies 5 usability heuristics in his textbook Usability Engineering: (1) learnability, (2) efficiency, (3) memorability, (4) errors, and (5) satisfaction. Building on these conceptualizations, Quesenbery [86] coined the 5E's of usability for user experience design: effectiveness, efficiency, engagement, error tolerance, and ease of learning. Table 3 summarizes these usability attributes. While not exhaustive, these attributes are a useful starting point for evaluating usability.

XSL•FO
RenderX

Table . Usability attributes and definitions.

	Usability attributes	Definitions
ISO Standard 9241-11:2018 [82]	Effectiveness	Accuracy and completeness with which users achieve intended goals or outcomes
	Efficiency	Resources used (eg, time, human effort, cost, and materials) relative to results achieved
	Satisfaction	Extent to which the user's physical, cognitive, and emotional responses that result from the use of a system meet the user's needs and expectations
Nielsen's Usability Engineering [85]	Learnability	How easily a technology can be learned such that the user can rapidly begin to use the technology to perform the necessary tasks
	Efficiency	How well the user can achieve a high level of productivity with minimum resource requirements (eg, time, money, effort)
	Memorability	How easily the user can remember how to use the technology such that they can return to the product easily with minimal relearning after a period of disuse
	Errors	The extent to which errors are minimized and catastrophic errors are avoided
	Satisfaction	How pleased the user is with their experience of the technology
Quesenbery's 5E's of Usability [86]	Efficiency	The speed with which the user can complete work accurately using the technology
	Engaging	How pleasant, satisfying, or interesting it is for the user to interact with technology
	Error tolerance	How well the technology prevents errors and helps the user recover from errors that occur
	Easy to learn	How well the product supports the user in initial orientation and deeper learning

Design engineering practices provide various approaches for evaluating usability. According to Nielsen [85], usability can be tested directly through behavioral measures (what people do) or indirectly assessed based on attitudinal beliefs of the user (what people think). Direct methods of usability testing consider how technology is actually used by measuring behavior or performance [25,85]. This can involve quantifiable performance metrics or qualitative methods such as participant observation and/or review of materials produced by the user [25]. Relevant performance measures depend on goals of the intervention and may include accuracy, task completion time, and errors (recoverable and unrecoverable) [85]. Think-aloud protocols, which involve users verbalizing their thoughts while interacting with technology, are a valuable strategy to identify usability problems with small samples. Finally, observational analysis techniques involve coding participant behavior while using the technology [85].

Indirect or attitudinal methods of usability evaluation explore user perspectives regarding their experience interacting with technology, which can be done through questionnaires, interviews, focus groups, or other methods [25,85]. Some argue these measures of user attitudes reflect technology acceptability rather than usability per se [25]. Nevertheless, subjective experience questionnaires are the most common method of usability evaluation across assistive and rehabilitation technology intervention studies [87,88]. Textbox 4 lists usability questionnaires for technology-based interventions. Qualitative feedback through posttesting interviews or focus groups is recommended to gain a deeper understanding of why usability problems exist and how they should be addressed [25,89].

The most appropriate usability evaluation methods depend on project context and intervention maturity. Available options also depend on the context in which the intervention is studied, for example, under natural settings in real-world target environments, scripted use within a controlled laboratory environment, or only perceptions and expectations about hypothetical use [89]. Combining direct and indirect measurement often provides a more comprehensive understanding of usability by not only identifying apparent usability issues but also exploring why they exist and how they can be resolved [25]. Early usability studies might benefit most from think-aloud, observation, and focus group methods within scripted use contexts to gather rich data targeting specific issues with small sample sizes [25]. At the feasibility stage, real-world use is important, and usability evaluation might involve a combination of performance measures, observation, questionnaires, and interviews [4]. For larger and longer-term outcome studies, performance measures are likely most

XSL•FO
RenderX

informative [25]. Readers are directed to Liu et al [25] for a thorough discussion of usability measurement, including examples from rehabilitation and assistive technology interventions.

Textbox 4. Example usability questionnaires for technology-based interventions.

- System Usability Scale [90]
- Usefulness, Satisfaction, and Ease of Use Questionnaire [91]
- NASA Task Load Index [92]
- Telehealth Usability Questionnaire [92]
- mHealth Application Usability Questionnaire [93]
- Usability Scale for Assistive Technology-Wheeled Mobility [94]
- Usability Scale for Assistive Technology-Computer Access [95]
- Accessible Usability Scale [96]

Adaptability

Technology-based interventions for disability and rehabilitation are characterized by diversity in user populations and use settings [4]. Device modifications or personalized intervention features may therefore be needed to accommodate individual users. Adaptability is the degree to which a technology-based intervention can be modified to meet the needs of diverse users and use contexts [12]. Evaluating adaptability can help define the scope of the target population, identify overly rigid features, and determine opportunities for increased flexibility to enable personalization.

Adaptability encompasses the frequency, extent, and reasons for intervention modifications, as well as any unique user needs that could not be accommodated through modifications. Adaptable interventions are expected to achieve similar outcomes despite changes in features or delivery format to accommodate individual needs [12]. Thus, it is important to assess the impact of modifications, including whether implementation processes, user experiences, and outcomes are consistent across modified versions [11]. Addressing adaptability also provides an opportunity to test the processes that enable personalized modifications, such as fit assessments, trial periods, technical support, and professional follow-up.

The importance of considering the adaptability of early prototypes is evident in a study from Pellichero et al [97], which evaluated the usability and usefulness of a blind spot sensor system for powered wheelchairs [97]. Although customization of sensor mounting, feedback configuration (ie, sounds, lights, or vibration), and sensor parameters (detection distances) were available through an associated smartphone app, the study was conducted with a standardized set-up preventing customization. The authors subsequently postulated that insufficient customization may have limited participants' use of the system during the 2-month home trial period and influenced perceived usability and usefulness. Customization of the sensor system according to individuals' abilities, preferences, and environment was thus recommended for future research and implementation [97]. This example illustrates the interrelatedness of feasibility dimensions, highlighting how adaptability has important implications for evaluating other aspects of intervention feasibility (eg, usability and demand). It also demonstrates how allowing opportunities for customization and assessing subsequent modifications during early-stage evaluations can help optimize interventions to accommodate diverse user needs.

Practicality

The best of interventions will not be successful unless they are practical within user and organizational capacities. Practicality refers to how well an intervention can be used in its intended setting given resource constraints like space, equipment, personnel, and time [11,12]. Literature across an array of technology-based interventions consistently highlights practical factors as important barriers to clinical implementation [98-100]. Disability and rehabilitation interventions may be deployed in hospitals or care centers, clinics, homes, or other community settings that have variable resource capacities, making practicality an important consideration.

Physical factors related to practicality include space and equipment needs for intervention use and storage [98,100]. System interoperability and compatibility with existing equipment and infrastructure is also important. For example, whether an exergame deployed for home use is suitable for conventional recreational gaming systems and whether it successfully integrates with commercial wearable heart rate monitors. Personnel required to facilitate intervention use should be assessed, including the level of training and expertise needed [98,101]. Time limitations are a well-documented barrier to implementation of rehabilitation technologies, which can encompass the duration of training, set-up, use, and take-down [98,101].

Implementation

To understand whether an intervention will work, it is important to explore the process of how it is delivered and received [102]. The implementation dimension evaluates how well a technology-based intervention can be deployed and used as intended in target settings with varying levels of control. This is important given the diversity of contexts in which technology-based disability and rehabilitation interventions are used [4]. Examining implementation helps understand how well the intervention can be delivered and what supports are needed to enable successful delivery [11]. It can also inform the training and support required for users to learn and apply the technology appropriately. The context of use, within a controlled lab

XSL•FO
RenderX

environment, home, or community setting, is an important consideration to determine the conditions under which the intervention is likely to succeed and identify features requiring adaptability [4,11,102].

Process evaluations assessing implementation describe the quality or integrity with which an intervention is delivered by monitoring adherence and fidelity [102]. Adherence is traditionally defined as the extent to which the behavior of an intervention recipient corresponds with the prescribed program [103]. Whereas we describe *engagement* above as an active choice to interact with the intervention in a meaningful way, we use the term *adherence* in reference to whether participants use the intervention as prescribed in its delivery [76]. Adherence is a common measure among feasibility studies of rehabilitation interventions that can be quantified via self-report or observation as the proportion of intervention components (eg, appointments, training sessions, modules, and tasks) completed by participants relative to those prescribed [20]. These measures, which translate well to technology-based interventions serving therapeutic or service delivery functions and may be captured through automated system logs, can help infer the burden of intervention demands and appropriateness of procedures [19]. For assistive technologies, adherence is poorly defined [104] and infrequently reported [105]. In reporting guidelines for assistive technology outcomes, Tuazon and Jutai [105] identify frequency, duration, and continuance (or discontinuance) of use as measures of adherence. We argue that these usage metrics reflect adherence when the intervention is used *as* recommended, but may be more appropriately viewed as measures of engagement in conditions of free voluntary use. Given the apparent similarity and overlap between adherence and engagement, we encourage authors to thoughtfully consider and rationalize their application.

Fidelity refers to whether the intervention is delivered as intended [102]. Self-report checklists and/or observational methods may be used to assess the degree to which intervention delivery and use align with key features of the intended design. This includes components or ingredients of the intervention and the level of competence demonstrated by intervention providers [11], which can help determine the training required to achieve proficiency and fidelity. A benefit of technology-based interventions is the possible integration of screen, audio, or video recordings that capture user behavior, enabling assessment of fidelity or adherence. For interventions that require complex technology operation, user perceptions regarding the ease of achieving fidelity may be informative and can be assessed using the Agency for Healthcare Research and Quality Feasibility Assessment for Wide-Scale Implementation (Criterion 2) [106]. These approaches to fidelity assessment work well for technology-based interventions with therapeutic or service delivery functions, for which treatment theory can be readily defined and interventions can be manualized, allowing active ingredients to be monitored [107]. In a provocative discussion, Lenker et al [108] highlight historical inadequacy of fidelity reporting in assistive technology research, which they attribute to poorly defined treatment protocols, temporal delays between device use and evaluation, and lack of available measurement tools to capture intervention attributes. To strengthen the quality of assistive technology research, they recommend greater inclusion of treatment theory, better specificity in intervention characterization, and routine fidelity monitoring [108].

The distinction between intervention and study feasibility is particularly blurred within the implementation dimension. For example, the fidelity by which clinicians deliver a tele-rehabilitation intervention and the extent to which recipients adhere to the program will likely have implications for both the success of the intervention itself (ie, whether benefits are achieved by recipients), as well as the scientific integrity and outcome of research studies evaluating the intervention. A clear rationale for evaluating implementation should be specified based on development stage (eg, refining intervention or study procedures).

Integration

Organizational support and alignment with system-level priorities are frequently cited as facilitators to technology implementation in disability and rehabilitation practices [98-101]. The integration dimension considers the “fit” of a technology-based intervention in terms of how well it can be integrated into existing systems within its target settings [12]. A focus on system-level integration is important as it likely affects implementation quality and eventual technology adoption. Importantly, there may be multiple systems into which a technology-based intervention must integrate. For example, therapeutic exergames may be initially introduced by therapists within a clinical setting and subsequently integrated into individual or family routines at home for continued practice. Assessing integration across all relevant systems can help resolve incongruencies and proactively enact strategies that support technology transfer.

Several system-level factors have been reported in the literature that influence integration of technology-based interventions in disability or rehabilitation practice. Reviews identify a lack of intervention fit within organizational capacities and workflows as important barriers to technology adoption, which may include referral processes [100], scheduling [98,100], service delivery models [100], and length of therapy sessions [98,100]. Organizational culture and policy, particularly regarding technology in practice, is another important factor that can facilitate or hinder integration [99,100].

The Agency for Healthcare Research and Quality Implementation Feasibility Assessment evaluates integration factors by having service providers rate intervention fit with organizational capabilities (Criterion 1) and alignment with policies and incentives (Criterion 2) [106]. The Program Sustainability Assessment Tool can be used to assess organizational structures and processes that promote sustained integration [109]. From the perspective of clients and families, the ease of integrating a technology-based intervention in their home environment and daily routine is a priority [110].

Impact

Performance expectancy, or perceived usefulness, is a leading determinant of acceptance and use of technology-based interventions among therapists, caregivers, and people with disabilities [101,111,112]. The impact dimension evaluates the

XSL•FO
RenderX

promise or potential that a technology-based intervention shows in eliciting positive change for its target population. This dimension has been renamed from *limited efficacy* described by Bowen et al [12] to recognize the diverse aims of technology-based interventions and methods through which their benefits may be evaluated. Although the focus of feasibility research is process rather than outcomes, this stage is an opportunity to gather preliminary evidence on participant responses to the intervention [11,12,18,19].

Various research designs involving quantitative, qualitative, or mixed methods approaches may be used at the feasibility stage to explore potential impacts of technology-based interventions. Preliminary effects can be measured using descriptive observational studies, case reports or series, single-subject designs, interrupted time series designs, or one-group pretest-posttest designs [4]. Outcomes of interest for care recipients include symptoms, functional performance, participation, mobility, quality of life, and other psychosocial constructs [57]. The purpose of evaluating impact during the feasibility stage is not to definitively establish effectiveness but rather to test assessment procedures that accompany the intervention, evaluate appropriateness and sensitivity of outcome measures, and inform features of intervention design including duration and dosage [11]. Qualitative feedback is essential to accomplish these goals by exploring user perceptions regarding perceived benefits and the meaningfulness of outcome measures [75].

For care providers, the usefulness of technology in supporting care is an important driver of adoption [101], indicating that perceived job performance or quality of care are relevant measures of impact. For technologies designed to reduce therapist workload, such as autonomous robots and exoskeletons, impact may also be measured in subjective task effort or stress and burnout, which are common among care providers [113]. Alleviating caregiver burden is a target of some interventions for which impact may be investigated using the Caregiver Assistive Technology Outcome Measure [114]. System-level impacts are also relevant to technology-based interventions, which may include measures of resource use, service efficiency, and access to care [115-117].

Ethicality

Ethical concerns are barriers to adoption and commercialization across a variety of technology-based approaches in disability and rehabilitation [118-120]. Common ethical concerns include challenges related to privacy, security, equity, stigma, autonomy, and power [118-120]. There are calls for more proactive consideration of ethics during development, evaluation, implementation, and policymaking processes for technology-based interventions [120]. The ethicality dimension considers the degree to which a technology-based intervention aligns with values of biomedical and social ethics. Ethical evaluation during feasibility studies can help anticipate and address concerns that may affect eventual adoption or policy [4].

Ethical concerns occur in the field of disability and rehabilitation when there is conflict between two or more biomedical or societal values. For example, protecting the safety of older adults using smart home monitoring technologies while maintaining personal privacy [121]. Applied ethics frameworks outline considerations for identifying and navigating these conflicts. Principles of bioethics and technology ethics may be useful for identifying value conflicts among technology-based interventions. Bioethics, concerned with ethical issues related to biomedicine and health care, includes 4 primary principles: (1) autonomy (personal choice), (2) beneficence (provision of benefit), (3) nonmaleficence (avoidance of harm), and (4) justice (fairness and equity) [122]. Technology ethics address issues related to technological advancements; relevant concepts overlap with bioethics and include human welfare, ownership and property, privacy, freedom from bias, universal usability, trust, autonomy, informed consent, accountability, identity, calmness, and environmental sustainability [123].

Based on these concepts, ethical domains relevant to the intervention can be compiled in consultation with users and other interest holders [121]. As the intervention is trialed during feasibility testing, conflicts across these domains can be monitored through discussion with users and reflective questioning about how the technology is used and any unexpected secondary effects [121]. A Socratic approach to questioning can be used, which poses morally relevant questions to highlight value issues related to health technologies [124]. The Model for Ethical Evaluation of Sociotechnical Arrangements (MEESTAR) is another tool to guide dialog on ethical analysis, designed specifically for assistive technologies [125]. It comprises seven ethical dimensions (care, autonomy, safety, justice, privacy, participation, and self-conception) across four levels (harmless, ethically sensitive with possibility of compensation, extreme ethical sensitivity requiring monitoring, opposition of use) and three perspectives (individual, organizational, society). Key questions are provided to support ethical discussion across the seven dimensions.

An example of ethical evaluation during feasibility testing can be found in a study from Kernebeck et al of an app- and sensor-based assistive technology intervention for caregivers of people with dementia [126]. In addition to exploratory outcome measures, the study involved a process evaluation of implementation, mechanisms of impact, and intervention context, including ethical implications. The ethical evaluation was a workshop based on MEESTAR conducted after the intervention with the study team, interest holders, and advisory board members. The authors highlight how this process supports the design of their intervention to reflect appropriate ethical orientation considering the values of independence, privacy, and safety from multiple perspectives [126].

Commercial Viability

Cost and affordability are priorities in decisions regarding rehabilitation and assistive technology interventions from the perspective of clinicians, care recipients, families, health care organizations, and policymakers [98,99,127,128]. Yet, economic factors are infrequently and inadequately captured in research on these interventions [129]. The commercial dimension evaluates the degree to which a technology-based intervention shows promise toward being economically viable for commercialization. While often reserved for later

XSL•FO
RenderX

implementation stages, gathering data on economic considerations at the feasibility stage can help refine intervention features to better fit the market, be more cost-efficient, and support commercialization.

Various economic evaluation methods are described in the complex intervention and health technology assessment literature. Cost-utility and cost-effectiveness analyses compare the costs and health outcomes of a new intervention against a comparator (eg, current standard of care), but use different measures of health outcome [130]. Cost-effectiveness analyses employ natural units of clinical effect, such as life-years or cases of disability, while cost-utility analyses, a subtype of cost-effectiveness analysis, use generic measures of health status reflecting both morbidity and mortality (eg, disability-adjusted life-years) [130]. Cost-benefit analyses express costs and benefits in financial terms and aim to capture a broader range of economic costs, including those not directly related to health outcomes (eg, cost of travel to care) [130]. Cost-utility analyses are favored in policy decision-making regarding health technologies because they allow comparison across conditions and intervention approaches [128].

The comprehensive economic evaluations described above may not be appropriate at the feasibility stage. During early development and testing of health technologies, Sculpher et al [131] recommend that economic evaluation should focus on defining and understanding the costs of existing practice that the new technology may help alleviate, and exploratory economic modeling of potential cost-effectiveness using data from existing studies or limited prospective data collection [131]. Relevant costs to consider for assistive and rehabilitation interventions include device production, assessment and training services, staffing, and servicing or maintenance. In terms of economic value, cost savings can be characterized through factors such as improved health outcomes, reduced costs of care, increased productivity, and decreased caregiver burden [129]. These exploratory economic feasibility approaches can inform intervention refinement by helping to redesign device features using more cost-efficient methods and improve efficiency of delivery services.

Discussion

Overview

Feasibility studies are of growing importance in intervention research and frameworks are available to guide this work across health disciplines, yet none specifically address technology-based interventions, which are inherently complex and necessitate iterative development and evaluation approaches [4]. This viewpoint helps define the scope of feasibility studies for technology-based interventions and highlights considerations for this work in the field of disability and rehabilitation. Key contributions and practical uses are discussed below.

Contributions and Recommendations

Chiefly, this article is intended to support researchers and developers (in academic, clinical, or industry settings) by providing clarity on terminology in feasibility studies of technology-based interventions. Reviews of feasibility research across health disciplines have identified inconsistent use of terminology and variability in methods of evaluation [13,20]. Synthesizing concepts from feasibility literature and research on rehabilitation and assistive technology interventions, we outline various dimensions of feasibility specifically relevant to technology-based interventions and highlight potential criteria for their evaluation (Table 1). This can inform the design of comprehensive feasibility studies that reflect unique technological considerations and the complex factors that influence rehabilitation intervention outcomes. Integrating engineering practices and digital health concepts, we collated conceptual definitions and practical examples (Tables 2 and 3) to facilitate dialog between technical developers and clinical research teams. The compiled measurement instruments (Textboxes 3 and 4) can help tailor data collection approaches and tools for specific types of technology-based interventions. Ultimately, our contributions serve to harmonize terminology and evaluation methods, enable critical selection of feasibility dimensions and criteria, and facilitate interpretation and comparison across studies.

To our knowledge, this is the first attempt to consolidate information on feasibility studies specifically for technology-based interventions. While further refinement may occur in time, this work provides essential background material to support study design decisions. Those seeking to apply information in this article should be aware of general best practices for feasibility studies [10,12,17]. We recommend that teams carefully select and justify the most appropriate dimensions and evaluation criteria according to their project needs, relevant uncertainties, and field-specific considerations (Textbox 2). These decisions are influenced by various contextual factors like intervention maturity, risk profile, and setting; Multimedia Appendix 1 provides sample reflection questions to support critical decision-making. Setting *a priori* success criteria for each feasibility dimension promotes rigor within and between studies for evidence building. Importantly, failure to achieve criteria may signal the need for intervention refinement (eg, software improvements to enhance usability). The process of feasibility evaluation is thus highly iterative, whereby multiple rounds of testing may be warranted to progressively refine intervention prototypes until comprehensive success criteria are met [4]. For instance, it may be appropriate to first establish user acceptability and demand with early-stage prototypes for interventions involving technology that is expensive or difficult to produce, while high-risk interventions with potential harms or discomforts necessitate early focus on safety and technical performance.

This work is also relevant for improving alignment between research and policy on technology-based interventions. The feasibility dimensions we delineate align with and extend current practices in health technology assessment—the comprehensive review process that informs policy recommendations on adoption of new technologies in health systems [127]. While health technology assessment approaches differ between jurisdictions, there is a common shared focus on considering factors beyond only clinical evidence on safety and effectiveness [127,132]. Health technology assessment processes around the world recognize the importance of organizational integration,

XSL•FO
RenderX

economic and resource efficiency, social implications, equity, ethical ramifications, and user and public perspectives in decision-making [127,132]. These multidimensional factors and their application to technological interventions are not specifically articulated in existing feasibility literature and thus may not be considered until their implementation, after significant resources have already been invested in development and outcome evaluation. By considering these factors in the early stages of evaluation, potential issues can be proactively identified and addressed through refinement of intervention features and delivery approach, which we emphasize as a key aim of feasibility studies for technology-based intervention. In doing so, innovations can be advanced with the greatest likelihood of attaining policy support for widespread adoption to the benefit of people with disabilities or rehabilitation needs and the communities that support them.

Future Directions

The ideas in this viewpoint represent the experiences and opinions of a small authorship team. While a useful guide for research and development, it is a starting point for future work. There is a need for broader discussion and engagement among interest holders in this space to establish consensus on feasibility considerations and dimensions in this field. It is our hope that this article serves as a catalyst for such dialog. Future directions building on this work may include the development of decision aids (eg, checklists, workflows, and case examples) and reporting guidelines to ensure adequate consideration, justification, and operationalization of relevant feasibility issues during the intervention development process. Decision tools may also help guide the process of iteratively refining and advancing intervention prototypes. Methods of weighting or prioritizing feasibility dimensions could also be useful in the context of limited resources or competing priorities.

Acknowledgments

This article was written in partial fulfillment of academic requirements for a graduate-level individual reading and research course (Foundations in Rehabilitation and Assistive Technology) within the Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto. The authors wish to thank the other students from this course, Erica Dove and Chaitali Desai, for their support and feedback related to this work. The authors also thank the anonymous reviewers for their feedback, which has helped to improve this article.

Funding

No direct funding was obtained to support this work. JS received graduate scholarship funding support from the Temerty Faculty of Medicine (University of Toronto), Ontario Graduate Scholarship, and Canadian Institutes of Health Research.

Authors' Contributions

JS researched and drafted the manuscript, which was subsequently reviewed and edited in collaboration with ASH and RHW. All authors approved the final version of the manuscript prior to submission for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample feasibility reflection questions.
[DOCX File, 19 KB - rehab_v13i1e79026_app1.docx]

References

1. Dahl T, Boulos M. Robots in health and social care: a complementary technology to home care and telehealthcare? Robotics 2013;3(1):1-21. [doi: 10.3390/robotics3010001]
2. Martin S, Kelly G, Kernohan WG, McCreight B, Nugent C. Smart home technologies for health and social care support. Cochrane Database Syst Rev 2008 Oct 8;2008(4):CD006412. [doi: 10.1002/14651858.CD006412.pub2] [Medline: 18843715]
3. Blandford A, Wesson J, Amalberti R, AlHazme R, Allwihan R. Opportunities and challenges for telehealth within, and beyond, a pandemic. Lancet Glob Health 2020 Nov;8(11):e1364-e1365. [doi: 10.1016/S2214-109X(20)30362-4]
4. Wang RH, Kenyon LK, McGilton KS, et al. The time is now: a FASTER approach to generate research evidence for technology-based interventions in the field of disability and rehabilitation. Arch Phys Med Rehabil 2021 Sep;102(9):1848-1859. [doi: 10.1016/j.apmr.2021.04.009] [Medline: 33992634]
5. Bauer SM, Elsaesser LJ, Arthanat S. Assistive technology device classification based upon the World Health Organization's, International Classification of Functioning, Disability and Health (ICF). Disabil Rehabil Assist Technol 2011;6(3):243-259. [doi: 10.3109/17483107.2010.529631] [Medline: 21446850]
6. Petticrew M. When are complex interventions "complex"? When are simple interventions "simple"? Eur J Public Health 2011 Aug;21(4):397-398. [doi: 10.1093/eurpub/ckr084] [Medline: 21771736]

XSL•FO
RenderX

7. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
8. Feasibility English meaning. Cambridge Dictionary. URL: https://dictionary.cambridge.org/dictionary/english/feasibility [accessed 2026-03-14]
9. Dictionary.com. URL: https://www.dictionary.com/browse/feasibility [accessed 2025-02-02]
10. Eldridge SM, Lancaster GA, Campbell MJ, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. PLoS ONE 2016;11(3):e0150205. [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]
11. Gadke DL, Kratochwill TR, Gettinger M. Incorporating feasibility protocols in intervention research. J Sch Psychol 2021 Feb;84:1-18. [doi: 10.1016/j.jsp.2020.11.004] [Medline: 33581765]
12. Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457. [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
13. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010 Jul 16;10(1):67. [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
14. Whitehead AL, Sully BGO, Campbell MJ. Pilot and feasibility studies: is there a difference from each other and from a randomised controlled trial? Contemp Clin Trials 2014 May;38(1):130-133. [doi: 10.1016/j.cct.2014.04.001] [Medline: 24735841]
15. Donald G. A brief summary of pilot and feasibility studies: Exploring terminology, aims, and methods. Eur J Integr Med 2018 Dec;24:65-70. [doi: 10.1016/j.eujim.2018.10.017]
16. Viswanathan P, Wang RH, Sutcliffe A. Smart wheelchairs in assessment and training (SWAT): state of the field. : AGE-WELL NCE; 2018 URL: https://agewell-nce.ca/wp-content/uploads/2019/01/SWAT-State-of-the-field_FINAL.pdf [accessed 2026-03-25]
17. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 6;10(1):1. [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
18. Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. Am J Occup Ther 2013;67(2):171-176. [doi: 10.5014/ajot.2013.006270] [Medline: 23433271]
19. Orsmond GI, Cohn ES. The distinctive features of a feasibility study: objectives and guiding questions. OTJR (Thorofare N J) 2015 Jul;35(3):169-177. [doi: 10.1177/1539449215578649] [Medline: 26594739]
20. Lawson DO, Mellor K, Eddy S, et al. Pilot and feasibility studies in rehabilitation research: a review and educational primer for the physiatrist researcher. Am J Phys Med Rehabil 2022 Apr 1;101(4):372-383. [doi: 10.1097/PHM.0000000000001797] [Medline: 34091466]
21. Kho ME, Thabane L. Pilot and Feasibility Studies in Rehabilitation: Moving into the Next Decade: University of Toronto Press; 2020. [doi: 10.3138/ptc-72-3-gee2]
22. Dobkin BH. Progressive staging of pilot studies to improve phase III trials for motor interventions. Neurorehabil Neural Repair 2009;23(3):197-206. [doi: 10.1177/1545968309331863] [Medline: 19240197]
23. Clayback D, Stanley R, Leahy J, et al. Standards for assistive technology funding: what are the right criteria? Knowledge Translation for Technology Transfer 2014(KT4TT) [FREE Full text]
24. Aschbrenner KA, Kruse G, Gallo JJ, Plano Clark VL. Applying mixed methods to pilot feasibility studies to inform intervention trials. Pilot Feasibility Stud 2022;8(1):1-13. [doi: 10.1186/s40814-022-01178-x]
25. Liu L, Cruz AM, Rincon AMR. Technology Acceptance, Adoption, and Usability: Arriving at Consistent Terminologies and Measurement Approaches: Everyday Technologies in Healthcare CRC Press; 2019:319-338. [doi: 10.1201/9781351032186-17]
26. Garrett C, Levack D, Rhodes R. Using technical performance measures. Presented at: 47th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit; Jul 31, 2011. [doi: 10.2514/6.2011-5500]
27. Alper S, Raharinirina S. Assistive technology for individuals with disabilities: a review and synthesis of the literature. J Spec Educ Technol 2006 Mar;21(2):47-64. [doi: 10.1177/016264340602100204]
28. Fager SK, Burnfield JM. Patients' experiences with technology during inpatient rehabilitation: opportunities to support independence and therapeutic engagement. Disability and Rehabilitation: Assistive Technology 2014 Mar;9(2):121-127. [doi: 10.3109/17483107.2013.787124]
29. Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design: Springer; 2009. URL: http://link.springer.com/10.1007/978-1-84800-175-6 [doi: 10.1007/978-1-84800-175-6]
30. Ferrari M, Harrison B, Rawashdeh O, et al. Clinical feasibility trial of a motion detection system for fall prevention in hospitalized older adult patients. Geriatr Nurs (Lond) 2012 May;33(3):177-183. [doi: 10.1016/j.gerinurse.2011.11.011]
31. Fang J, Xie Q, Yang GY, Xie L. Development and feasibility assessment of a rotational orthosis for walking with arm swing. Front Neurosci 2017;11:32. [doi: 10.3389/fnins.2017.00032] [Medline: 28203142]
32. Awad LN, Esquenazi A, Francisco GE, Nolan KJ, Jayaraman A. The ReWalk ReStore TM soft robotic exosuit: A multi-site clinical trial of the safety, reliability, and feasibility of exosuit-augmented post-stroke gait rehabilitation. J NeuroEngineering Rehabil 2020 Dec;17(1). [doi: 10.1186/s12984-020-00702-5]

XSL•FO
RenderX

33. Marschollek M, Becker M, Bauer JM, et al. Multimodal activity monitoring for home rehabilitation of geriatric fracture patients – feasibility and acceptance of sensor systems in the GAL-NATARS study. Informatics for Health and Social Care 2014 Sep;39(3-4):262-271. [doi: 10.3109/17538157.2014.931852]
34. Quality glossary of terms, acronyms & definitions. ASQ. URL: https://asq.org/quality-resources/quality-glossary [accessed 2023-04-11]
35. Guidance for Industry E9 Statistical Principles for Clinical Trials. 1998. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e9-statistical-principles-clinical-trials [accessed 2023-03-29]
36. About design for safety. URL: https://www.tal.sg/wshc/topics/design-for-safety/about-design-for-safety [accessed 2025-02-18]
37. Roderick SN, Carignan CR. An approach to designing software safety systems for rehabilitation robots. Presented at: 9th International Conference on Rehabilitation Robotics, 2005 ICORR 2005; Jun 28, 2005. [doi: 10.1109/ICORR.2005.1501096]
38. Wang RH, Gorski SM, Holliday PJ, Fernie GR. Evaluation of a contact sensor skirt for an anti-collision power wheelchair for older adult nursing home residents with dementia: safety and mobility. Assist Technol 2011 Sep;23(3):117-134. [doi: 10.1080/10400435.2010.541406]
39. Friel KM, Gordon AM, Carmel JB, Kirton A, Gillick BT. Pediatric issues in neuromodulation: safety, tolerability and ethical considerations. In: Mapping and Modulating The Developing Brain: Academic Press; 2016:131-149. [doi: 10.1016/B978-0-12-802001-2.00007-2]
40. Kögel J, Wolbring G. What it takes to be a pioneer: ability expectations from brain-computer interface users. Nanoethics 2020 Dec;14(3):227-239. [doi: 10.1007/s11569-020-00378-0]
41. Biddiss E, Chau T. Upper-limb prosthetics: critical factors in device abandonment. Am J Phys Med Rehabil 2007 Dec;86(12):977-987. [doi: 10.1097/PHM.0b013e3181587f6c] [Medline: 18090439]
42. Pearson EJM. Comfort and its measurement--a literature review. Disabil Rehabil Assist Technol 2009 Sep;4(5):301-310. [doi: 10.1080/17483100902980950] [Medline: 19565373]
43. Crane BA, Holm MB, Hobson D, Cooper RA, Reed MP. Responsiveness of the TAWC tool for assessing wheelchair discomfort. Disability and Rehabilitation: Assistive Technology 2007 Jan;2(2):97-103. [doi: 10.1080/17483100601178484]
44. Knight JF, Baber C. A tool to assess the comfort of wearable computers. Hum Factors 2005;47(1):77-91. [doi: 10.1518/0018720053653875] [Medline: 15960088]
45. McKeon A, McCue M, Skidmore E, Schein M, Kulzer J. Ecological momentary assessment for rehabilitation of chronic illness and disability. Disabil Rehabil 2018 Apr;40(8):974-987. [doi: 10.1080/09638288.2017.1280545] [Medline: 28637126]
46. Keshav NU, Salisbury JP, Vahabzadeh A, Sahin NT, Power B. Social communication coaching smartglasses: well tolerated in a diverse sample of children and adults with autism. JMIR Mhealth Uhealth 2017 Sep 21;5(9):e140. [doi: 10.2196/mhealth.8534] [Medline: 28935618]
47. Ryan JL, Beal DS, Levac DE, Fehlings DL, Wright FV. Integrating transcranial direct current stimulation into an existing inpatient physiotherapy program to enhance motor learning in an adolescent with traumatic brain injury: a case report. Phys Occup Ther Pediatr 2023;43(4):463-481. [doi: 10.1080/01942638.2022.2163214] [Medline: 36624962]
48. Garvey MA, Kaczynski KJ, Becker DA, Bartko JJ. Subjective reactions of children to single-pulse transcranial magnetic stimulation. J Child Neurol 2001 Dec;16(12):891-894. [doi: 10.1177/088307380101601205] [Medline: 11785502]
49. Nielsen J. Usability Engineering: Morgan Kaufmann. URL: https://books.google.ca/books?hl=en&lr=&id=fA2OF6fOC&oi=fnd&pg=PR8&dq=nielsen+usability+engineering&ots=3GxDeSY&sig=8B8ZRw8WZSWBXsBnHkv-opg&q=nielsen%20usability%20engineering&f=false [accessed 2026-03-14]
50. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
51. Momani AM, Jamous MM, Hilles SMS. Technology acceptance theories. International Journal of Cyber Behavior, Psychology and Learning 2017;7(2):1-14 [FREE Full text] [doi: 10.4018/IJCBPL.2017040101]
52. Venkatesh V, Thong JYL, Xu X. Unified theory of acceptance and use of technology: a synthesis and the road ahead. Journal of the Association for Information Systems 2016;17(5):328-376 [FREE Full text]
53. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022 Dec;22(1). [doi: 10.1186/s12913-022-07577-3]
54. Attkisson CC, Zwick R. The client satisfaction questionnaire. Eval Program Plann 1982 Jan;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-X]
55. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implementation Sci 2017 Dec;12(1). [doi: 10.1186/s13012-017-0635-3]
56. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
57. Tao G, Charm G, Kabacińska K, Miller WC, Robillard JM. Evaluation tools for assistive technologies: a scoping review. Arch Phys Med Rehabil 2020 Jun;101(6):1025-1040. [doi: 10.1016/j.apmr.2020.01.008] [Medline: 32059944]
58. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022 Jan 27;22(1):22. [doi: 10.1186/s12911-022-01764-2] [Medline: 35081953]

XSL•FO
RenderX

59. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
60. Demers L, Weiss-Lambrou R, Ska B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): an overview and recent progress. TAD 2002 Jan 1;14(3):101-105. [doi: 10.3233/TAD-2002-14304]
61. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int J of Soc Robotics 2010 Dec;2(4):361-375. [doi: 10.1007/s12369-010-0068-5]
62. Finkelstein SM, MacMahon K, Lindgren BR, et al. Development of a remote monitoring satisfaction survey and its use in a clinical trial with lung transplant recipients. J Telemed Telecare 2012 Jan;18(1):42-46. [doi: 10.1258/jtt.2011.110413]
63. Samitier CB, Guirao L, Costea M, Camós JM, Pleguezuelos E. The benefits of using a vacuum-assisted socket system to improve balance and gait in elderly transtibial amputees. Prosthetics & Orthotics International 2016;40(1):83-88. [doi: 10.1177/0309364614546927]
64. Cox RM, Alexander GC. Validation of the SADL Questionnaire. Ear Hear 2001 Apr;22(2):151-160. [doi: 10.1097/00003446-200104000-00008]
65. Blumenthal J, Wilkinson A, Chignell M. Physiotherapists' and physiotherapy students' perspectives on the use of mobile or wearable technology in their practice. Physiother Can 2018;70(3):251-261. [doi: 10.3138/ptc.2016-100.e] [Medline: 30275650]
66. Fitzgerald A, Huang S, Sposato K, Wang D, Claypool M, Agu E. The exergame enjoyment questionnaire (EEQ): an instrument for measuring exergame enjoyment. Presented at: Hawaii International Conference on System Sciences; Jan 7, 2020. [doi: 10.24251/HICSS.2020.416]
67. Gil-Gómez JA, Manzano-Hernández P, Albiol-Pérez S, Aula-Valero C, Gil-Gómez H, Lozano-Quilis JA. USEQ: a short questionnaire for satisfaction evaluation of virtual rehabilitation systems. Sensors (Basel) 2017 Jul 7;17(7):1589. [doi: 10.3390/s17071589] [Medline: 28686174]
68. Jennett C, Cox AL, Cairns P, et al. Measuring and defining the experience of immersion in games. Int J Hum Comput Stud 2008 Sep;66(9):641-661. [doi: 10.1016/j.ijhcs.2008.04.004]
69. Chin JP, Diehl VA, Norman LK. Development of an instrument measuring user satisfaction of the human-computer interface. Presented at: Conference on Human Factors in Computing Systems - Proceedings; May 1, 1988; Washington, DC, United States p. 213-218. [doi: 10.1145/57167.57203]
70. Yip MP, Chang AM, Chan J, MacKenzie AE. Development of the Telemedicine Satisfaction Questionnaire to evaluate patient satisfaction with telemedicine: a preliminary study. J Telemed Telecare 2003;9(1):46-50. [doi: 10.1258/135763303321159693] [Medline: 12641893]
71. Bakken S, Grullon-Figueroa L, Izquierdo R, et al. Development, validation, and use of English and Spanish versions of the telemedicine satisfaction and usefulness questionnaire. J Am Med Inform Assoc 2006;13(6):660-667. [doi: 10.1197/jamia.M2146] [Medline: 16929036]
72. Hirani SP, Rixon L, Beynon M, et al. Quantifying beliefs regarding telehealth: Development of the whole systems demonstrator service user technology acceptability questionnaire. J Telemed Telecare 2017 May;23(4):460-469. [doi: 10.1177/1357633X16649531] [Medline: 27224997]
73. Bol N, van Weert JCM, de Haes H, et al. Using cognitive and affective illustrations to enhance older adults' website satisfaction and recall of online cancer-related information. Health Commun 2014;29(7):678-688. [doi: 10.1080/10410236.2013.771560] [Medline: 24160560]
74. Haydon HM, Major T, Kelly JT, et al. Development and validation of the Digital Health Acceptability Questionnaire. J Telemed Telecare 2023 Dec;29(10_suppl):8S-15S. [doi: 10.1177/1357633X231202279] [Medline: 38007698]
75. O'Cathain A, Hoddinott P, Lewin S, et al. Maximising the impact of qualitative research in feasibility studies for randomised controlled trials: guidance for researchers. Trials 2015 Dec;16(S2):1-13. [doi: 10.1186/1745-6215-16-S2-O88]
76. Cavanagh K. Turn on, tune in and (don't) drop out: engagement, adherence, attrition, and alliance with internet-based interventions. In: Oxford Guide to Low Intensity CBT Interventions: Oxford University Press; 2010:227-234. [doi: 10.1093/med:psych/9780199590117.003.0021]
77. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52. [doi: 10.2196/jmir.1772] [Medline: 21821503]
78. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
79. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. Int J Hum Comput Stud 2018 Apr;112:28-39. [doi: 10.1016/j.ijhcs.2018.01.004]
80. Petrie H, Carmien S, Lewis A. Assistive technology abandonment: research realities and potentials. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Springer Verlag; 2018:532-540. [doi: 10.1007/978-3-319-94274-2_77]
81. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11. [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

XSL•FO
RenderX

82. ISO 9241-11:2018 - Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. URL: https://www.iso.org/standard/63500.html [accessed 2023-04-10]
83. Arthanat S, Bauer SM, Lenker JA, Nochajski SM, Wu YWB. Conceptualization and measurement of assistive technology usability. Disabil Rehabil Assist Technol 2007 Jul;2(4):235-248. [doi: 10.1080/17483100701343665] [Medline: 19263540]
84. Tuena C, Pedroli E, Trimarchi PD, et al. Usability issues of clinical and research applications of virtual reality in older people: a systematic review. Front Hum Neurosci 2020;14:93. [doi: 10.3389/fnhum.2020.00093] [Medline: 32322194]
85. Nielsen J. Usability Engineering.: Google Books URL: https://books.google.ca/books?hl=en&lr=&id=95A2OF67OC&oi=fnd&pg=PR9&dq=usability+engineering&ots=3tGydpo&sig=9hNXDVq4iTcumB3X3aF9Hqlfv=onepage&q=usability%20engineering&f=false [accessed 2023-04-10]
86. Quesenbery W. Balancing the 5Es: usability. Research Gate. URL: https://www.researchgate.net/publication/327937198_Balancing_the_5Es_Usability [accessed 2023-04-10]
87. Medina JLP, Acosta-Vargas P, Rybarczyk Y, Medina JLP, Acosta-Vargas P, Rybarczyk Y. A systematic review of usability and accessibility in tele-rehabilitation systems. Assistive and Rehabilitation Engineering IntechOpen 2019 Apr 15. [doi: 10.5772/INTECHOPEN.85869]
88. Simor FW, Brum MR, Schmidt JDE, Rieder R, De Marchi ACB. Usability evaluation methods for gesture-based games: a systematic review. JMIR Serious Games 2016 Oct 4;4(2):e17. [doi: 10.2196/games.5860] [Medline: 27702737]
89. Rohrer C. When to use which user-experience research methods. : Nielsen Norman Group; 2014 URL: https://www.xdstrategy.com/wp-content/uploads/2018/08/When-to-Use-Which-User-Experience-Research-Methods-2014-10-12-Print.pdf [accessed 2026-03-25]
90. SUS: a “quick and dirty” usability scale. In: Usability Evaluation In Industry: CRC Press; 1996:207-212. [doi: 10.1201/9781498710411-35]
91. Gao M, Kortum P, Oswald F. Psychometric Evaluation of the USE (Usefulness, Satisfaction, and Ease of USE) Questionnaire for Reliability and Validity: SAGE Publications; 2018:1414-1418. [doi: 10.1177/1541931218621322]
92. Parmanto B, Lewis, Jr. AN, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehab 2016;8(1):3-10. [doi: 10.5195/ijt.2016.6196]
93. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
94. Arthanat S, Wu YWB, Bauer SM, Lenker JA, Nochajski SM. Development of the Usability Scale for assistive technology-wheeled mobility: a preliminary psychometric evaluation. TAD 2009 Jan 1;21(3):79-95. [doi: 10.3233/TAD-2009-0275]
95. Gosselin LR, Arthanat S. Reliability and validity of the usability scale for assistive technology for computer access: a preliminary study using video-based evaluation. Assist Technol 2021 Nov 2;33(6):350-356. [doi: 10.1080/10400435.2019.1661316]
96. Accessible usability scale - test usability with assistive technology users. Fable. URL: https://makeitfable.com/accessible-usability-scale/ [accessed 2023-04-14]
97. Pellichero A, Best KL, Routhier F, Viswanathan P, Wang RH, Miller WC. Blind spot sensor systems for power wheelchairs: obstacle detection accuracy, cognitive task load, and perceived usefulness among older adults. Disability and Rehabilitation: Assistive Technology 2023 Oct 3;18(7):1084-1092. [doi: 10.1080/17483107.2021.1983654]
98. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237. [doi: 10.1016/j.pmrj.2018.07.004]
99. Zander V, Gustafsson C, Landerdahl Stridsberg S, Borg J. Implementation of welfare technology: a systematic review of barriers and facilitators. Disability and Rehabilitation: Assistive Technology 2023 Aug 18;18(6):913-928. [doi: 10.1080/17483107.2021.1938707]
100. Moorcroft A, Scarinci N, Meyer C. A systematic review of the barriers and facilitators to the provision and use of low-tech and unaided AAC systems for people with complex communication needs and their families. Disabil Rehabil Assist Technol 2019 Oct;14(7):710-731. [doi: 10.1080/17483107.2018.1499135] [Medline: 30070927]
101. Liu L, Miguel Cruz A, Rios Rincon A, Buttar V, Ranson Q, Goertzen D. What factors determine therapists’ acceptance of new technologies for rehabilitation – a study using the Unified Theory of Acceptance and Use of Technology (UTAUT). Disabil Rehabil 2015;37(5):447-455. [doi: 10.3109/09638288.2014.923529] [Medline: 24901351]
102. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350(mar19 6):h1258. [doi: 10.1136/bmj.h1258] [Medline: 25791983]
103. Chakrabarti S. What’s in a name? Compliance, adherence and concordance in chronic psychiatric disorders. World J Psychiatry 2014 Jun 22;4(2):30-36. [doi: 10.5498/wjp.v4.i2.30] [Medline: 25019054]
104. Tuazon JR, Jahan A, Jutai JW. Understanding adherence to assistive devices among older adults: a conceptual review. Disability and Rehabilitation: Assistive Technology 2019 Jul 4;14(5):424-433. [doi: 10.1080/17483107.2018.1493753]
105. Tuazon JR, Jutai JW. Toward guidelines for reporting assistive technology device outcomes. Disabil Rehabil Assist Technol 2021 Oct;16(7):702-711. [doi: 10.1080/17483107.2019.1697384] [Medline: 31795783]

XSL•FO
RenderX

106. Fournier AK, Wasserman MR, Jones CF, et al. Developing AHRQ's feasibility assessment criteria for wide-scale implementation of patient-centered outcomes research findings. J Gen Intern Med 2021 Feb;36(2):374-382. [doi: 10.1007/s11606-020-06247-6] [Medline: 33051835]
107. Van Stan JH, Whyte J, Duffy JR, et al. Rehabilitation treatment specification system: methodology to identify and describe unique targets and ingredients. Arch Phys Med Rehabil 2021 Mar;102(3):521-531. [doi: 10.1016/j.apmr.2020.09.383] [Medline: 33065124]
108. Lenker JA, Fuhrer MJ, Jutai JW, Demers L, Scherer MJ, DeRuyter F. Treatment theory, intervention specification, and treatment fidelity in assistive technology outcomes research. Assist Technol 2010;22(3):129-138. [doi: 10.1080/10400430903519910] [Medline: 20939422]
109. Luke DA, Calhoun A, Robichaux CB, Elliott MB, Moreland-Russell S. The Program Sustainability Assessment Tool: a new instrument for public health programs. Prev Chronic Dis 2014 Jan 23;11:130184. [doi: 10.5888/pcd11.130184] [Medline: 24456645]
110. Vourganas I, Stankovic V, Stankovic L, Kerr A. Factors that contribute to the use of stroke self-rehabilitation technologies: a review. JMIR Biomed Eng 2019;4(1):e13732. [doi: 10.2196/13732]
111. Sriram V, Jenkinson C, Peters M. Informal carers' experience of assistive technology use in dementia care at home: a systematic review. BMC Geriatr 2019 Jun 14;19(1):160. [doi: 10.1186/s12877-019-1169-0] [Medline: 31196003]
112. Shin HR, Um SR, Yoon HJ, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. J Med Internet Res 2023 Apr 10;25:e41935. [doi: 10.2196/41935] [Medline: 37036760]
113. Bridgeman PJ, Bridgeman MB, Barone J. Burnout syndrome among healthcare professionals. Am J Health Syst Pharm 2018 Feb 1;75(3):147-152. [doi: 10.2146/ajhp170460] [Medline: 29183877]
114. Mortenson WB, Demers L, Fuhrer MJ, Jutai JW, Lenker J, DeRuyter F. Development and preliminary evaluation of the caregiver assistive technology outcome measure. J Rehabil Med 2015 May;47(5):412-418. [doi: 10.2340/16501977-1952] [Medline: 25783142]
115. Tenforde AS, Hefner JE, Kodish‐Wachs JE, Iaccarino MA, Paganoni S. Telehealth in physical medicine and rehabilitation: a narrative review. PM R 2017 May;9(5S). [doi: 10.1016/j.pmrj.2017.02.013]
116. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J NeuroEngineering Rehabil 2012 Dec;9(1). [doi: 10.1186/1743-0003-9-21]
117. Laut J, Porfiri M, Raghavan P. The present and future of robotic technology in rehabilitation. Curr Phys Med Rehabil Rep 2016 Dec;4(4):312-319. [doi: 10.1007/s40141-016-0139-0] [Medline: 28603663]
118. Sundgren S, Stolt M, Suhonen R. Ethical issues related to the use of gerontechnology in older people care: A scoping review. Nurs Ethics 2020 Feb;27(1):88-103. [doi: 10.1177/0969733019845132]
119. Zwijsen SA, Niemeijer AR, Hertogh C. Ethics of using assistive technology in the care for community-dwelling elderly people: an overview of the literature. Aging Ment Health 2011 May;15(4):419-427. [doi: 10.1080/13607863.2010.543662] [Medline: 21500008]
120. Ienca M, Kressig RW, Jotterand F, Elger B. Proactive ethical design for neuroengineering, assistive and rehabilitation technologies: the Cybathlon Lesson. J NeuroEngineering Rehabil 2017 Dec;14(1). [doi: 10.1186/s12984-017-0325-z]
121. Wang RH, Tannou T, Bier N, Couture M, Aubry R. Proactive and ongoing analysis and management of ethical concerns in the development, evaluation, and implementation of smart homes for older adults with frailty. JMIR Aging 2023 Mar 9;6(1):e41322. [doi: 10.2196/41322] [Medline: 36892912]
122. Beauchamp T, Childress J. Principles of biomedical ethics: marking its fortieth anniversary. Am J Bioeth 2019 Nov;19(11):9-12. [doi: 10.1080/15265161.2019.1665402] [Medline: 31647760]
123. Human values, ethics, and design. In: The Human-Computer Interaction Handbook: CRC Press; 2007:1267-1292. [doi: 10.1201/9781410615862-78]
124. Hofmann B, Droste S, Oortwijn W, Cleemput I, Sacchini D. Harmonization of ethics in health technology assessment: a revision of the Socratic approach. Int J Technol Assess Health Care 2014 Jan;30(1):3-9. [doi: 10.1017/S0266462313000688] [Medline: 24499630]
125. Ethische fragen im bereich altersgerechter assistenzsysteme. Miteinander durch Innovation [Web page in German]. URL: https://www.interaktive-technologien.de/service/publikationen/ethische-fragen-im-bereich-altersgerechter-assistenzsysteme [accessed 2023-04-13]
126. Kernebeck S, Holle D, Pogscheba P, et al. A tablet app- and sensor-based assistive technology intervention for informal caregivers to manage the challenging behavior of people with dementia (the insideDEM Study): protocol for a feasibility study. JMIR Res Protoc 2019 Feb 26;8(2):e11630. [doi: 10.2196/11630] [Medline: 30806626]
127. O'Donnell JC, Pham SV, Pashos CL, Miller DW, Smith MD. Health technology assessment: lessons learned from around the world—an overview. Value Health 2009 Jun;12 Suppl 2:S1-S5. [doi: 10.1111/j.1524-4733.2009.00550.x] [Medline: 19523179]
128. Mathes T, Jacobs E, Morfeld JC, Pieper D. Methods of international health technology assessment agencies for economic evaluations--a comparative analysis. BMC Health Serv Res 2013 Sep 30;13(1):371. [doi: 10.1186/1472-6963-13-371] [Medline: 24079858]

XSL•FO
RenderX

129. Albala SA, Kasteng F, Eide AH, Kattel R. Scoping review of economic evaluations of assistive technology globally. Assist Technol 2021 Dec 1;33(sup1):50-67. [doi: 10.1080/10400435.2021.1960449]
130. Turner HC, Archer RA, Downey LE, et al. An introduction to the main types of economic evaluations used for informing priority setting and resource allocation in healthcare: key features, uses, and limitations. Front Public Health 2021;9(1236):722927. [doi: 10.3389/fpubh.2021.722927] [Medline: 34513790]
131. Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. J Health Serv Res Policy 1997 Jan;2(1):26-30. [doi: 10.1177/135581969700200107] [Medline: 10180650]
132. Draborg E, Gyrd-Hansen D, Poulsen PB, Horder M. International comparison of the definition and the practical application of health technology assessment. Int J Technol Assess Health Care 2005;21(1):89-95. [doi: 10.1017/s0266462305050117] [Medline: 15736519]

Abbreviations

EMA: Ecological momentary assessment
FASTER: Framework for Accelerated and Systematic Technology-based intervention development and Evaluation Research
ISO: International Organization for Standardization
MRC: Medical Research Council
RCT: randomized controlled trial
TAM: Technology Acceptance Model
tDCS: transcranial direct current stimulation
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by J Sarvestan; submitted 13.Jun.2025; peer-reviewed by H Heppner, R Stockley; revised version received 18.Feb.2026; accepted 19.Feb.2026; published 14.Apr.2026.

Please cite as:
Shore J, Hwang AS, Wang RH
Evaluating the Feasibility of Technology-Based Interventions in Disability and Rehabilitation: Definitions, Considerations, and Dimensions
JMIR Rehabil Assist Technol 2026;13:e79026
URL: https://rehab.jmir.org/2026/1/e79026
doi:10.2196/79026

© Josh Shore, Amy S Hwang, Rosalie H Wang. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 14.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Satisfaction With Assistive, Rehabilitation, and Training Technologies: Questionnaire Development and Content Validation Study

Oskar Cvjetičanin[1*], MSc; Ivana Marić[1*], MSc; Nejc Šarabon[1,2,3,4], PhD

[1]Faculty of Health Sciences, University of Primorska, Polje 42, Izola, Slovenia

[2]Andrej Marušič Institute, University of Primorska, Koper, Slovenia

[3]Laboratory for Motor Control and Motor Behavior, S2P, Science to Practice, Ljubljana, Slovenia

[4]Ludwig Boltzmann Institute for Rehabilitation Research, St. Pölten, Austria

[*]these authors contributed equally

Corresponding Author:
Nejc Šarabon, PhD
Faculty of Health Sciences, University of Primorska, Polje 42, Izola, Slovenia

Abstract

Background: According to human-centered design principles, user experience and user satisfaction are key aspects when developing or providing patients and clients with assistive, rehabilitation, or training technologies. Existing tools for assessing user experience and satisfaction are often either too narrow (developed for specific technologies) or too broad, failing to capture all relevant aspects of the technology being evaluated.

Objective: The main objective of this study was to present the development process and content validation of a questionnaire that, following further validation steps, could be useful for researchers and developers in assessing the satisfaction of direct end users with assistive, rehabilitation, and training technologies.

Methods: The development of the User Satisfaction Questionnaire for Assistive, Rehabilitation, and Training Technologies (U-ART-Q) followed a multistep process: (1) literature review, (2) domain definition and item generation, (3) expert consultation, (4) linguistic validation, and (5) content validation. Initially, 72 items were generated and organized into 7 domains. After expert review and linguistic validation, the questionnaire underwent content validation by 8 experts in physiotherapy, kinesiology, and ergonomics. Each item was rated for relevance using a 4-point Likert-type scale. The content validity ratio (CVR) and item-level content validity index (I-CVI) were calculated, and items below the accepted thresholds were removed or further revised.

Results: Expert consultation resulted in the removal or merging of 20 items to improve clarity and reduce redundancy. Of the 52 items evaluated during content validation, 44 met the CVR and I-CVI criteria (CVR >0.75 and I-CVI >0.78). Eight items required additional review, and 2 items were ultimately removed. The final questionnaire comprises 10 initial screening questions and 50 content-valid items across 7 domains: design, comfort, effectiveness and usefulness, service, treatment or exercise, overall impressions, and safety. The overall average content validity index was 0.915, providing evidence of strong content validity. A macro-enabled Excel tool was developed to automate item selection and generate a PDF version of the questionnaire based on initial screening responses.

Conclusions: The U-ART-Q is a newly developed instrument for assessing user satisfaction with assistive, rehabilitation, and training technologies. Further research will focus on psychometric evaluation, reliability testing, and multilingual validation to determine whether the U-ART-Q can provide a standardized method for assessing user satisfaction. Once fully evaluated, the questionnaire could support both developmental usability testing and postmarket evaluation, facilitating evidence-based decision-making and guiding future product improvement.

(JMIR Rehabil Assist Technol 2026;13:e88182) doi:10.2196/88182

KEYWORDS

user satisfaction; human-centered design; user-centered design; assistive technology; rehabilitation technology; training technology; questionnaire development

XSL•FO
RenderX

Introduction

The field of assistive, rehabilitation, and training technologies is rapidly advancing, presenting both opportunities and challenges [1]. User satisfaction plays a critical role in the adoption, adherence, and effectiveness of these technologies, thereby directly influencing the user's quality of life [2]. Consequently, the reliable assessment of user satisfaction and experience is essential for the successful development, provision, and implementation of such technologies. Existing assessment tools frequently rely on generic questionnaires that may not adequately capture the specific characteristics of technologies associated with clinical and training applications. This underscores the need for instruments specifically designed to evaluate user satisfaction within the domains of assistive, rehabilitation, and training technologies. The objective of this study was to describe the development and content validation of a new questionnaire for assessing user satisfaction. The intended respondents include individuals who wear, use, or train with the technology, as well as professionals who operate or provide it, such as physiotherapists, personal trainers, kinesiologists, and other health care professionals. The proposed questionnaire aims to fill an important gap in the existing literature by providing a new instrument for assessing and providing feedback on the design, usability, and effectiveness of these technologies, ultimately enhancing overall user experience and satisfaction.

Assistive, rehabilitation, and training technologies comprise tools, devices, and systems designed to maintain or enhance individuals' functioning, participation, health, and well-being. Rehabilitative and assistive technologies facilitate the engagement of individuals with disabilities in daily activities and support caregivers by improving independence and task performance [1]. Rehabilitation technologies primarily aim to help individuals regain lost function following illness or injury, whereas assistive technologies support individuals in performing tasks that might otherwise be difficult or impossible [3]. Assistive technologies, within the scope of this study, are hardware technologies classified according to the ISO 9999:2022 classification and range from low-tech aids, such as communication boards, to advanced computer-based systems, including prosthetics, specialized keyboards, exoskeletons, and others [4,5]. Assistive technologies are often categorized by complexity and user needs as low-, medium-, or high-level systems [5]. Recent advances have expanded these categories to include wearable technologies, virtual reality, AI, and robotics, offering new opportunities to enhance independence and quality of life [1,6,7]. Some of these devices, with variations in design and intended use, are also classified as rehabilitation technologies. Examples of rehabilitation technologies include exoskeletons, robotic-assisted therapy, virtual reality rehabilitation, and electrotherapy devices, while common assistive technologies include wheelchairs, canes, hearing aids, and speech recognition software. Rehabilitation technologies can also encompass assessment and monitoring tools, such as motion capture systems and musculoskeletal modeling software, which are used in both rehabilitation and training contexts. Training technologies, on the other hand, include devices designed to maintain or improve health, well-being, and performance in individuals with or without disabilities. Examples include treadmills, stationary bicycles, rowing machines, isokinetic devices, and sport-specific systems aimed at optimizing movement patterns. Additionally, occupational training devices are used to simulate job tasks, train, and improve specific skills.

Human-centered design (HCD) is a key approach in device and system development that ensures products are usable, useful, and aligned with users' needs by integrating principles of human factors, ergonomics, and usability [8,9]. HCD enhances product effectiveness, efficiency, user satisfaction, and accessibility while reducing potential adverse effects and promoting overall well-being [9]. According to the ISO, HCD is an iterative, cyclic process comprising 4 main phases: understanding and specifying the context of use, specifying user requirements, designing solutions, and evaluating against requirements [9]. In the development of assistive, rehabilitation, and training technologies, applying an HCD approach is essential to ensure that the final product meets the needs and preferences of end users. The ultimate goal of HCD is to create technologies that are necessary, functional, and enjoyable to use.

Usability testing is a fundamental component of the HCD process used to evaluate a product's ease of use, efficiency, and user satisfaction through testing with representative users [10]. It ensures that systems are not only functional but also intuitive, efficient, and aligned with user needs [11]. By identifying usability issues early in development, usability testing helps reduce later design costs and supports the development of technologies that provide positive user experiences [12,13]. A well-designed usability study involves participant recruitment, task design, data collection, and systematic analysis to obtain both qualitative and quantitative insights that inform design decisions and improve user experience [10,14]. Well-structured usability testing is therefore essential for understanding user interactions, identifying relevant user experience factors and potential pain points, and detecting barriers to implementation and opportunities for improvement [15,16]. However, in practice, HCD and usability testing are often carried out without the iterative approach needed to fully follow the principles of the HCD process [17]. Insufficient involvement of end users in the design process can lead to poor user acceptance and approval, emphasizing the need for developers to effectively integrate user needs throughout the design cycle. Similarly, the absence of systematic usability evaluation and adequate consideration of users' needs may contribute to reduced acceptance or abandonment of assistive technologies [18-20].

When planning and conducting usability testing, developers must determine the appropriate timing and frequency of testing within the design process, identify target end users, and establish methods for evaluating effectiveness, efficiency, and user satisfaction. Common approaches include structured and semistructured interviews, focus groups, surveys, scenario-based testing, expert evaluations, and controlled user observations [12]. Both qualitative and quantitative methods are essential for comprehensive usability evaluation and are often combined depending on the development stage and testing objectives. However, reliance on existing quantitative tools such as surveys

XSL•FO
RenderX

and standardized questionnaires can sometimes result in narrow or incomplete insights, potentially overlooking user needs and opportunities for design improvement. Conversely, qualitative methods such as focus groups and observational studies, while rich in detail, can be time-consuming and expensive. In software usability testing, common quantitative measures include task completion metrics (with or without video or screen recording) and standardized instruments such as the System Usability Scale (SUS), Usefulness, Satisfaction, and Ease of Use (USE), or NASA Task Load Index (NASA-TLX). The literature review was conducted to identify existing instruments suitable for evaluating user experience and satisfaction with hardware and devices that interact with the human body. The identified questionnaires were systematically examined with respect to their intended purpose, target technology, and domain coverage to determine if the development of the new tool was justified. The literature review yielded similar findings to previous studies: there are a large number of custom-made instruments used for satisfaction evaluation, and the identified instruments are either too general to identify specific design issues or too specific for broader application [21-23]. To overcome these limitations, a quantitative assessment tool was designed specifically to evaluate user satisfaction with assistive, rehabilitation, and training technologies. The instrument is intended to complement qualitative methods by providing a structured and potentially cost-efficient means of assessing satisfaction across individual items, broader domains, and overall device use. Following further validation steps, the questionnaire may be suitable not only for researchers and developers during the design process but also for routine commercial or clinical use of the device after development, as well as for postmarket evaluation. In these contexts, it could support the ongoing evaluation of user experience and satisfaction throughout the product life cycle, thereby informing continuous improvement and future design iterations. Additionally, it could assist researchers and clinicians in comparing user satisfaction across devices, settings, and user populations, thereby informing evidence-based decision-making in device selection and provision.

During the development process of the User Satisfaction Questionnaire for Assistive, Rehabilitation, and Training Technologies (U-ART-Q), it was considered essential to account for the distinct characteristics, intended purposes, and needs of end users of assistive, rehabilitation, and training technologies.

As emphasized in previous studies, comprehensive coverage, item clarity, and precise definition of measured constructs are crucial for creating valid and usable measurement instruments [2]. While detailed psychological definitions of satisfaction are beyond the scope of this study, user satisfaction in this context is conceptualized as a component of user experience, in line with the ISO definition: “Users’ perceptions and responses that result from the use and/or anticipated use of a system, product or service.” Satisfaction itself is defined as “the extent to which the user’s physical, cognitive and emotional responses that result from the use of a system, product or service meet the user’s needs and expectations” [9].

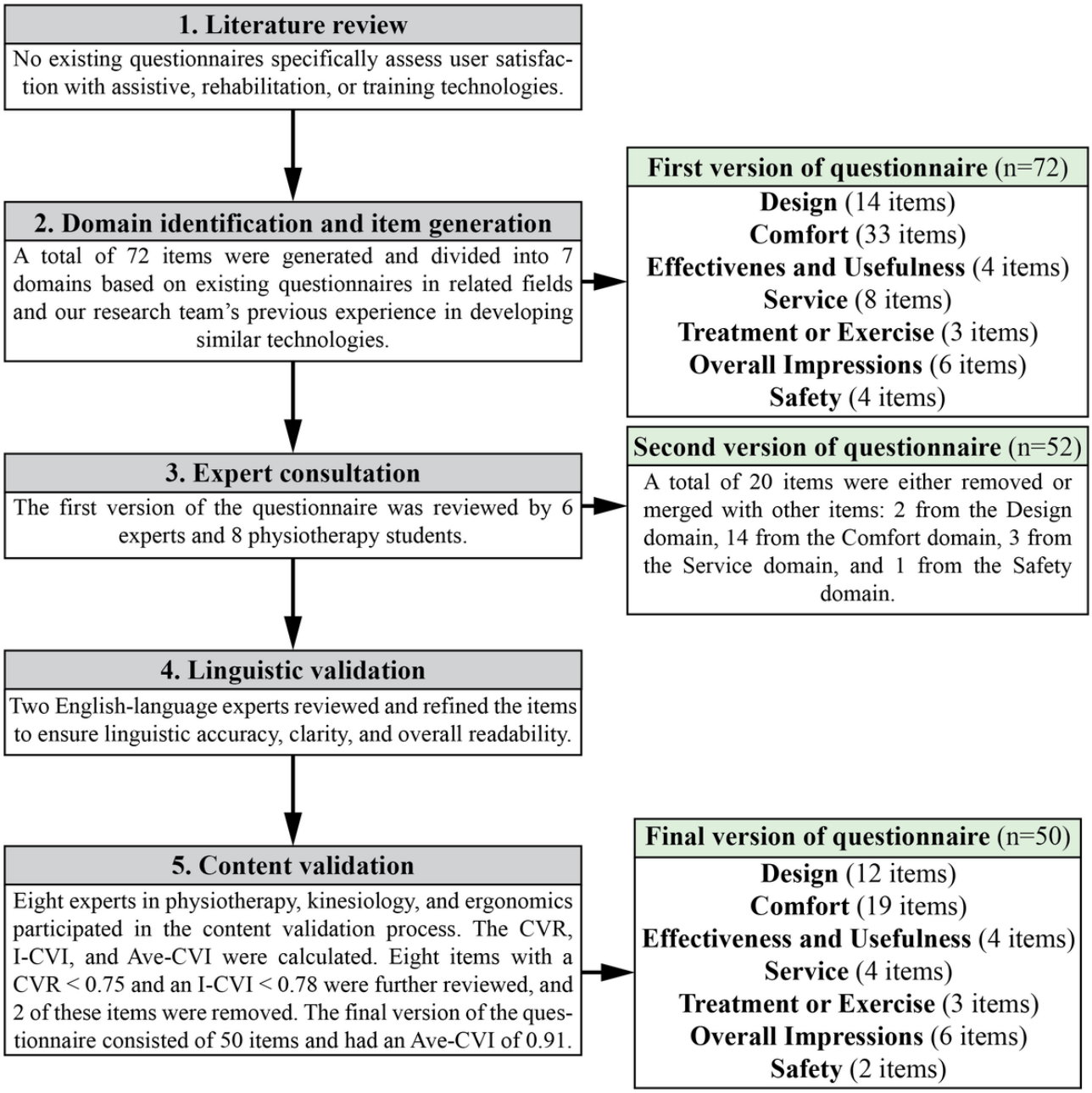
Methods

Overview

The development process of the U-ART-Q involved several distinct phases, including (1) literature review, (2) domain definition and item generation, (3) expert consultation, (4) linguistic validation, and (5) content validation. The detailed development process of this questionnaire is presented in Figure 1.

XSL•FO
RenderX

Figure 1. Development process of the User Satisfaction Questionnaire for Assistive, Rehabilitation, and Training Technologies (U-ART-Q). Ave-CVI: average content validity index; CVR: content validity ratio; I-CVI: item-level content validity index.



The first step in developing the U-ART-Q was to conduct a literature review to identify existing questionnaires assessing user satisfaction with assistive, rehabilitation, and training technologies. The search was conducted in the PubMed/MEDLINE, Scopus, and Web of Science databases on September 10, 2025. Database-specific search strategies were developed using combinations of terms related to satisfaction, user experience, questionnaires, development, rehabilitation, assistive technology, robotic rehabilitation, exoskeletons, prostheses, orthoses, and training technologies. The complete search strategies for each database are presented in Multimedia Appendix 1.

All retrieved records were exported to a Mendeley Reference Manager (Elsevier), and duplicate records were removed. Titles and abstracts were independently screened according to the predefined eligibility criteria. Potentially relevant studies underwent full-text review to determine final eligibility. Reasons for exclusion during the full-text assessment were documented in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) recommendations [24].

Studies were considered eligible if they reported the development, validation, adaptation, or application of questionnaires intended to evaluate user satisfaction with relevant technologies. Furthermore, studies were excluded if they focused on technologies outside the scope of rehabilitation, such as industrial exoskeletons, surgical robotics, dental technologies, and other unrelated medical devices. As our search was based only on hardware solutions, publications evaluating software or mobile apps without other hardware components were also excluded. Similarly, user interfaces without accompanying hardware components and virtual reality systems

XSL•FO
RenderX

were excluded. Additionally, studies that evaluated other constructs without measuring satisfaction (eg, quality of life, functional performance, technology acceptance without satisfaction outcomes, or clinical effectiveness alone) were excluded. Conference abstracts without full text available, editorials, letters, study protocols, dissertations, non-English publications, and studies with insufficient methodological information were also excluded.

Following full-text screening, studies were excluded for the following reasons: published study not in English language (n=9); technology outside the scope of this review (n=16); no satisfaction-related outcome measures reported (n=64); and study that does not use a questionnaire (n=3).

The complete study selection process is illustrated in the PRISMA flow diagram. All included studies are presented in Multimedia Appendix 2 (Microsoft Excel worksheet titled Studies for Extraction).

After searching the literature, we noted the existence of 117 questionnaires used to measure user experience, usability, and satisfaction with various technologies and systems. The most frequently used questionnaires were general questionnaires, such as the Quebec User Evaluation of Satisfaction With Assistive Technology (QUEST 2.0), USE, and SUS, as well as questionnaires specific to certain technologies, such as the Client Satisfaction With Device Module of the Orthotics and Prosthetics Users' Survey (CSD-OPUS), the Questionnaire for Usability Evaluation of orthopaedic shoes (QUE), the Trinity Amputation and Prosthesis Experience Scale (TAPES), the prosthesis-satisfaction subscale, and others [23,25-31]. The list of all identified questionnaires used in the evaluation of satisfaction with technologies is available in Multimedia Appendix 2 (Excel worksheet titled Identified Questionnaires).

Although a considerable number of questionnaires were identified (Multimedia Appendix 2), several limitations were noted. The identified questionnaires were either designed for a specific type of technology or failed to cover all important aspects necessary to provide detailed information on user satisfaction with assistive, rehabilitation, and training technologies. Forty-six out of 117 identified questionnaires were completely custom questionnaires, intended for the specific device or population in the study, limiting their broader applicability. Eighty-nine out of 117 questionnaires were developed specifically for the needs of the study, including custom questionnaires, modified instruments for usability testing of specific devices, newly developed questionnaires, and translations of existing questionnaires. Only 23 out of 117 questionnaires were reported in more than a single publication. Among the identified questionnaires, 58 primarily evaluated usability, user experience, technology acceptance, or overall system performance rather than user satisfaction, with satisfaction represented by only a small number of items instead of a comprehensive multidimensional construct. Additionally, none of the questionnaires appeared to cover all identified satisfaction domains. Item-level analysis demonstrated that domains with similar labels (eg, design) frequently assessed different underlying constructs and therefore did not capture equivalent aspects of user satisfaction. The identified questionnaires were also developed and used for different technology categories: 84 for assistive, 23 for rehabilitation, 9 for assistive and rehabilitation, 1 for nontechnology-specific questionnaire, and none for training technologies (Multimedia Appendix 2). Although several questionnaires have subsequently been applied beyond their original context, they were not specifically developed for these broader applications. In addition, widely used generic instruments, such as the SUS, were designed as technology-independent usability measures rather than for evaluating hardware-based technologies that directly interact with the human body and therefore do not capture several device-specific aspects identified as important in this review. Overall, while the identified questionnaires provide valuable information on specific aspects of user evaluation, no single instrument was found to simultaneously (1) assess user satisfaction as its primary outcome; (2) comprehensively cover all identified satisfaction domains at the item level; and (3) be specifically developed for application across assistive, rehabilitation, and training technologies.

Therefore, we aimed to develop a questionnaire to comprehensively assess user satisfaction across different types of assistive, rehabilitation, and training technologies. The questionnaire was specifically designed to capture the perspectives of direct end users, including both individuals using or training with the technology and professionals such as physiotherapists, kinesiologists, personal trainers, and other health care providers involved in its operation and delivery.

The second phase of the development process focused on identifying key domains and generating items to evaluate user experience and satisfaction with assistive, rehabilitation, and training technologies. The initial version of the questionnaire consisted of 72 items divided into 7 thematic domains. Item generation was guided by the preceding literature review and further informed by the research team's prior experience in developing and evaluating similar technologies. The list of all identified items and related domains during the literature review is available in Multimedia Appendix 2 (Excel worksheet titled Items and Domains). Items were generated with the intention that each item should apply to at least one type of existing assistive, rehabilitation, or training technology. However, given that such technologies can differ significantly in their fundamental characteristics, not all items are relevant for every device. To address this, a set of initial screening questions was developed to be completed by the researcher or developer prior to administering the questionnaire. The responses to these screening questions are used to generate a tailored set of questionnaire items, selected according to the characteristics of the technology and the role of the end user (eg, patient, client, therapist, or trainer). Furthermore, the researcher or developer may manually adjust the recommended item set by adding or removing questions to ensure optimal relevance for the specific technology.

After the domains were identified and the items generated, a round of expert revision was conducted. Six experts with extensive experience in the design and evaluation of assistive, rehabilitation, and training technologies and 8 master's students in physiotherapy were recruited to evaluate each item independently as either "relevant" or "not relevant." Additional

XSL•FO
RenderX

information about the experts included in the expert consultation is presented in Table 1. For any item marked as not relevant, the recruited experts were asked to provide a brief explanation for their decision. Based on the results obtained, the questionnaire was revised, and items marked as not relevant were either removed or modified. Following the revision process, the questionnaire was evaluated by 2 English-language experts to ensure linguistic accuracy, clarity, and overall readability.

Table . Additional information about the expert panel included in the expert consultation.

ID	Years of work experience	Employer	Field of research/work
E1	23	National research institute	Collaborative robotics
E2	16	National research institute	Collaborative robotics
E3	9	National research institute	Humanoid and cognitive robotics
E4	20	Clinical institution	Rehabilitation medicine
E5	27	Higher education institution	Engineering design
E6	12	Industry/private sector	Sports products development

The final part of the development process was the evaluation of content validity. Content validity relates to the degree to which a sample of items constitutes an adequate operational definition of a concept [32]. To assess item content validity, another 8 experts in physiotherapy, kinesiology, and ergonomics were recruited. These experts were independent of the 6 experts involved in the previous stage of questionnaire development and had at least 5 years of experience in the development or use of assistive, rehabilitation, or training technologies. The participating experts were informed of the purpose of the questionnaire and its intended use to evaluate user experience and satisfaction with the assistive, rehabilitation, and training technologies. Each expert independently evaluated all items of the questionnaire on a 4-point Likert-type scale (1="the item is not relevant to the measured domain," 2="the item is somewhat relevant to the measured domain," 3="the item is quite relevant to the measured domain," and 4="the item is highly relevant to the measured domain"). The content validity ratio (CVR) was calculated for each item alongside the content validity index for each item (I-CVI) using equations 1 and 2, where n_4 represents the number of experts grading an item as highly relevant (grade 4), $n_{3,4}$ represents the number of experts grading an item as quite or highly relevant (grade 3 or 4), and N represents the total number of experts. Additionally, the average content validity index (Ave-CVI) was calculated as the average of the I-CVIs, representing the content validity of the overall instrument [33].

(1)CVR=(ne−N2)(N2)

(2)I-CVI=n3,4N

Items with a CVR below 0.75 and an I-CVI below 0.78 were reviewed and either revised or removed in accordance with the content validity criteria described by Lawshe [34] and Lynn [35].

Ethical Considerations

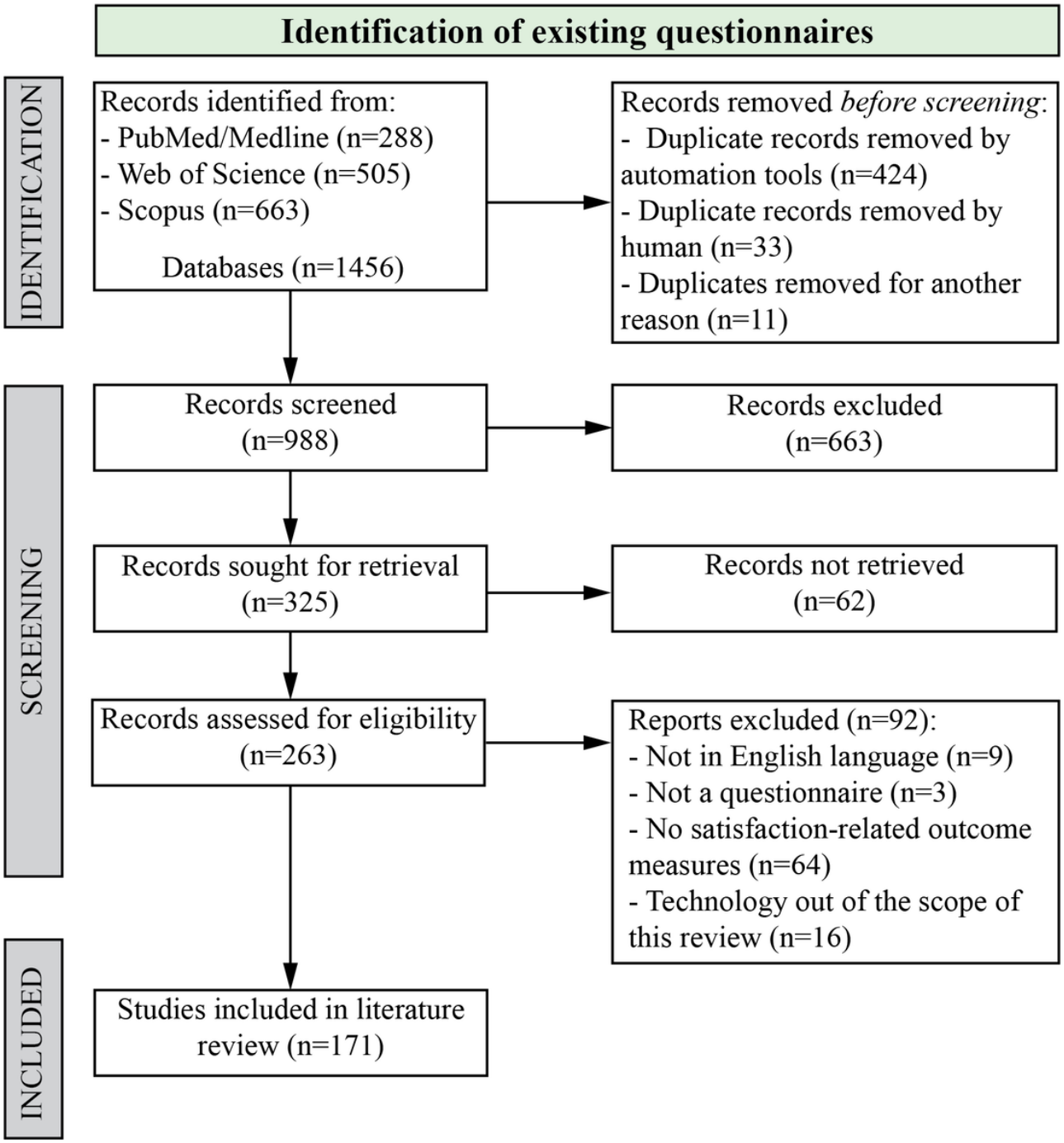
All participants received oral and written information about the study, were informed that participation was voluntary, and provided informed consent prior to participation. Participant confidentiality was ensured throughout the study. The study was approved by the University of Primorska Commission for Ethics in Research Involving Human Subjects (approval number: 4264-19-6/23) and was conducted in accordance with the principles of the Declaration of Helsinki.

Results

Following an initial literature review (the complete study selection process is illustrated in the PRISMA flow diagram Figure 2) and drawing on the research team's experience with similar technological developments, the first version of the questionnaire was developed, comprising 72 items divided into 7 domains. The identified domains were as follows: design (D), comfort (C), effectiveness and usefulness (EU), service (SE), treatment or exercise (TE), overall impressions (OI), and safety (SA).

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the literature review.



Once the first version of the U-ART-Q was prepared, it was reviewed by 6 experts with extensive experience in the design and evaluation of assistive, rehabilitation, and training technologies and 8 master's students in physiotherapy. Based on their feedback, 2 items from design, 14 from comfort, 3 from service, and 1 from safety domain were either removed or merged with other items to reduce redundancy and improve clarity.

The refined version of the questionnaire was further reviewed and revised by 2 English-language experts to improve accuracy, clarity, and overall readability. Content validity was then assessed based on an evaluation by 8 experts in physiotherapy, kinesiology, and ergonomics. Of the 52 items, 44 met the predefined CVR values >0.75 and I-CVI values >0.78 and were therefore considered suitable for inclusion in the final questionnaire. The remaining 8 items, with CVR values <0.75 and I-CVI values <0.78, were subjected to additional review and are presented in Table 2.

XSL•FO
RenderX

Table . Items with CVR[a] or I-CVI[b] values below predefined thresholds and corresponding actions taken.

Item	CVR	I-CVI	Action
D9[c]: I am satisfied with the battery life of the assistive/rehabilitation/training technology.	0.5	0.75	Retained
D12[d]: I am satisfied with the design and appearance of the assistive/rehabilitation/training technology.	0.5	0.75	Retained
C6[e]: I am satisfied with the comfort of the contact point/points in the upper arm area.	0.5	0.75	Retained
C10[f]: I am satisfied with the comfort of the contact point/points in the wrist area.	0.5	0.75	Retained
SE2[g]: I am satisfied with the time needed to understand how the assistive/rehabilitation/training technology is supposed to operate/function.	0.5	0.75	Retained
SE4[h]: I am satisfied with the performance of the assistive/rehabilitation/training technology in the specified weather conditions.	0.25	0.625	Removed
OI3[i]: The assistive/rehabilitation/training technology is pleasant to use.	0.5	0.75	Retained
SA2[j]: I have experienced no anxiety when using the assistive/rehabilitation/training technology.	0.25	0.625	Removed

[a]CVR: content validity ratio.

[b]I-CVI: item-level content validity index.

[c]D9: item number 9 from the design domain.

[d]D12: item number 12 from the design domain.

[e]C6: item number 6 from the comfort domain.

[f]C10: item number 10 from the Comfort domain.

[g]SE2: item number 2 from the service domain.

[h]SE4: item number 4 from the service domain.

[i]OI3: item number 3 from the overall impressions domain.

[j]SA2: item number 2 from the safety domain.

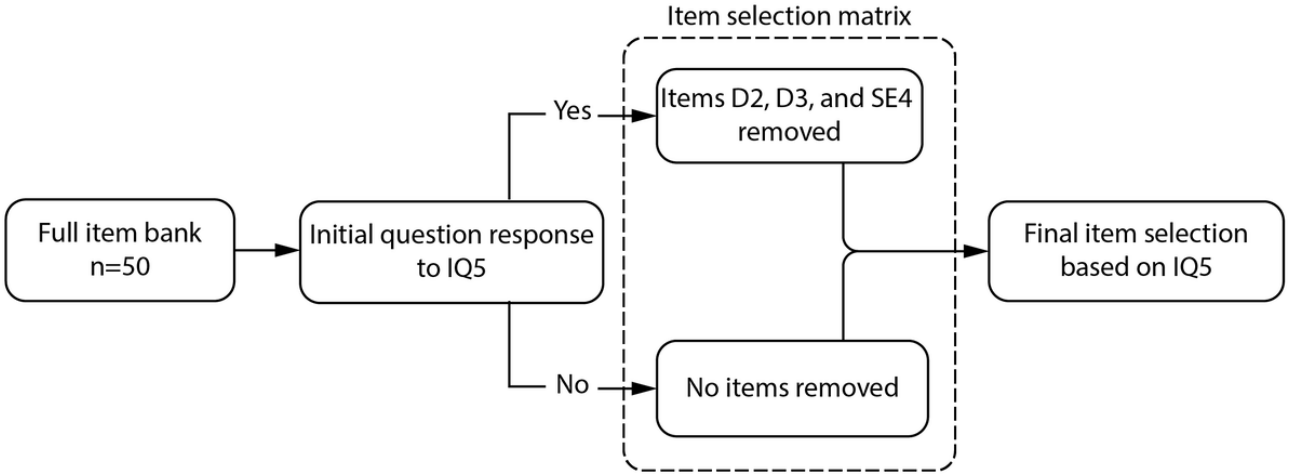
Following this review, 6 of these items were retained because they represented important aspects of the construct that were not adequately covered by other items. For example, although item D9, which refers to battery life, had CVR and I-CVI values below the predefined thresholds, it was retained because battery life was considered a highly relevant aspect of user satisfaction with technologies that operate on battery power [36]. Item D12, related to satisfaction with the appearance of the device, was retained, as satisfaction with appearance was identified in the literature as an important factor influencing user satisfaction, product abandonment, adoption, and long-term use [37-39]. Additionally, items regarding the appearance of the device were identified in 31 questionnaires during the literature review; therefore, this item was retained as part of the design domain. Items C6 and C10 were retained because all other items related to comfort at the contact points were rated as relevant. Because the upper arm or wrist may be the only contact points for some technologies, eliminating these items would limit the applicability of the questionnaire across different types of technologies. Regarding the retention of item OI3, the rationale was that user satisfaction encompasses the user's emotional response arising from interacting with a product or system. Therefore, perceived pleasantness represents an important affective dimension of satisfaction and was considered essential to retain in the questionnaire. The current definition of satisfaction in ISO 9241-11:2018 [40] explicitly encompasses users' physical, cognitive, and emotional responses, while ISO/IEC 25019:2023 [41] identifies user satisfaction as an emotional and hedonic outcome in which a pleasant, positive subjective experience is essential. In contrast, item SA2 was removed, even though anxiety was included as an item in some satisfaction questionnaires. The decision to remove this item was based on the fact that anxiety is a distinct psychological construct for which dedicated instruments exist, whereas

XSL•FO
RenderX

U-ART-Q is intended to assess satisfaction. There are many other standardized tools, such as the Abbreviated Technology Anxiety Scale (ATAS), Computer Anxiety Rating Scale (CARS), Technology Assisted Rehabilitation Patient Perception Questionnaire (TARPP-Q), and the General Anxiety Disorder 7-item screening tool, developed for the comprehensive assessment of technophobia and anxiety symptoms [42-45]. The item assessing satisfaction with the performance of the training technology under the specified weather conditions (SE4) was identified in only a small number of questionnaires, and insufficient evidence was found in the literature to justify its retention. Therefore, the item was removed. After removing these 2 items (SA2 and SE4), the final Ave-CVI, calculated across the remaining 50 items, was 0.915, which is above the acceptable standard reported in the literature. A complete list of all 10 initial screening questions and the final 50 items of the U-ART-Q is provided in Multimedia Appendix 3.

An additional set of 10 initial screening questions was created to determine the applicability of specific items based on the type of technology being evaluated. These initial questions are completed by the researcher who intends to use the questionnaire and are designed to identify which items are relevant, depending on the technology and the end user. Final questionnaire items are selected using the question selection matrix that is available in Multimedia Appendix 4. Each column of the matrix represents one possible response to an initial screening question, and each row represents 1 final questionnaire item. For each selected response, the corresponding matrix column is applied as a filter: items coded as 1 are retained, and items coded as 0 are excluded from the final questionnaire. Figure 3 illustrates the selection process for one of the screening questions. When the response to IQ5 is "Yes," the final questions Q4, Q5, Q39, and Q40 are removed, while all remaining items are retained.

Figure 3. Example of the item selection process for initial screening question number 5 (IQ5).



Items in the developed questionnaire are formulated in such a way that they can be graded on a 5-point Likert scale. The items can be graded as 1="strongly disagree," 2="disagree," 3="unsure," 4="agree," and 5="strongly agree."

To support the practical implementation of the item-selection process, a macro-enabled Excel tool (.xlsm) was developed and is provided in Multimedia Appendix 5. The purpose of this tool is to automatically generate the final set of questionnaire items based on responses to the initial screening questions. The tool is organized into separate worksheets. The worksheet named "Instructions" provides a step-by-step guide on how to use the Excel file to generate the final item selection. The "Initial_Questions" worksheet allows the user to select the responses to the initial questions from a drop-down list, thereby minimizing the possibility of incorrect entries. After all responses to the initial questions have been selected, the user can click the "Create_Final_Question Selection" button. The final questionnaire is then generated in the "Final_Item_Selection" worksheet. Following review, the generated U-ART-Q questionnaire can be exported as a PDF, printed, and used in practical applications.

Discussion

Principal Findings

HCD is a vital component of the product development process, ensuring that technologies are designed around the actual needs, abilities, and expectations of end users. In assistive, rehabilitation, and training technologies, HCD promotes usability, accessibility, and satisfaction by actively involving end users throughout the design, testing, and implementation phases. A key component of HCD is usability testing, which provides critical insight into how users interact with a product and identifies barriers to effective use. Failing to include usability testing at different stages of development can result in technically advanced products that do not meet user expectations, leading to poor acceptance, limited adoption, and ultimately unsuccessful implementation.

Although several relevant questionnaires were identified in the literature, none was considered fully suited to the specific purpose and scope of this study. Therefore, this study introduces the newly developed U-ART-Q and evaluates its content validity. It serves as a complementary tool for evaluating user satisfaction with assistive, rehabilitation, and training technologies. The items and domains for the initial version of the questionnaire were identified through a comprehensive

XSL•FO
RenderX

literature review. The preliminary set of items and domains underwent expert validation, which resulted in the removal or merging of 20 items to reduce redundancy and improve clarity. To further enhance linguistic accuracy, clarity, and overall readability, the revised questionnaire was reviewed and refined by 2 English-language experts. Finally, 8 additional experts with backgrounds in physiotherapy, kinesiology, and ergonomics participated in the content validation process. Out of 52 items, 44 met the predefined CVR and I-CVI thresholds and therefore met the predefined criteria for content validity. The remaining 8 items with CVR and I-CVI values below the recommended thresholds were reviewed further, resulting in the removal of 2 items from the item bank. Finally, the remaining 50-item questionnaire achieved a high Ave-CVI of 0.915, demonstrating strong evidence of content validity based on expert evaluation.

The proposed U-ART-Q represents a newly developed questionnaire with evidence of content validity for assessing user satisfaction within the domains of assistive, rehabilitation, and training technologies. To ensure its relevance across a broad range of technologies, the researcher or developer administering the questionnaire first completes a set of initial screening questions that guide item selection. This approach allows the instrument to be tailored to the specific characteristics of the technology being evaluated, making it both concise and practical for use across diverse assistive, rehabilitation, and training technologies.

The proposed questionnaire offers several important contributions to both research and practice. For researchers and clinicians, it is intended to serve as a structured and efficient evaluation tool for existing technologies, allowing user satisfaction to be compared across different devices, applications, and user groups. Its use can contribute to a more comprehensive understanding of user experience, leading to more informed decisions regarding device selection, clinical implementation, and future product improvement. For developers, the questionnaire provides a structured and standardized tool for collecting user feedback throughout the design process. It helps identify usability issues early and supports evidence-based design decisions. By systematically capturing satisfaction across multiple domains, the instrument can help optimize the development process by reducing redesign costs and improving the functionality of technologies under development. As user satisfaction and experience are closely linked to successful technology implementation, the results may also help clinical managers and other stakeholders compare available market options and support the decision-making, selection, and purchasing of technologies.

Future work will focus on further validation and practical application of the questionnaire. Pilot testing will be conducted to evaluate reliability, internal consistency, and test-retest reliability, and validated translations of the questionnaire into Serbian and Slovenian will be completed. Although a macro-enabled Excel tool (.xlsm) has been developed to support automated item selection, which is provided in Multimedia Appendix 5, the development of a mobile app or a web-based version of the questionnaire is planned to facilitate simple, time-efficient, and cost-effective administration in clinical and research settings.

Limitations

The development of the U-ART-Q has several limitations. First, the existing literature did not provide a comprehensive conceptual framework that specifically describes the multidimensional structure of user satisfaction with assistive, rehabilitation, and training technologies. Consequently, the domains and items of the U-ART-Q were informed by findings from the literature, existing questionnaires, expert consultations, and the research team's experience in developing similar technologies. Although the experts involved had substantial experience in relevant fields, expert-based evaluation is inherently subject to some degree of subjectivity. Different expert panels could have reached different conclusions regarding the relevance, wording, or importance of individual items.

Additionally, it is important to point out that the U-ART-Q is a newly developed questionnaire with initial evidence of content validity. No actual end users were included in the development or content-validation process, which represents an important limitation, particularly because the questionnaire is intended to assess user satisfaction. As such, the U-ART-Q should undergo further evaluation before it can be considered a fully validated and standardized instrument for assessing user satisfaction with assistive, rehabilitation, and training technologies. In line with recommended questionnaire-development procedures, important next steps include face validity assessment with intended users, pilot testing, evaluation of construct validity (through exploratory and confirmatory factor analysis), reliability testing, and establishment of a clear scoring system and interpretation of the results [46].

Conclusions

In conclusion, the U-ART-Q is a newly developed questionnaire with evidence of content validity. Its main objective is to support researchers and developers applying HCD principles in the evaluation and development of assistive, rehabilitation, and training technologies. The questionnaire is intended to complement existing methods of assessing user experience, including interviews, focus groups, and observational approaches. By integrating this quantitative measure with qualitative methods, the U-ART-Q can contribute to more informed, iterative, and end user–driven design processes. Future work will focus on further psychometric evaluation needed to determine whether the U-ART-Q can be established as a validated and standardized questionnaire for assessing user satisfaction with assistive, rehabilitation, and training technologies.

Acknowledgments

The authors used a generative AI tool, ChatGPT (GPT-5.2; OpenAI), to assist with language refinement and text editing. The tool was not used for literature searching, study selection, data extraction, analysis, interpretation of the literature, or formulation

XSL•FO
RenderX

of the study findings or conclusions. All content was subsequently reviewed and edited by the authors. The authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Funding

The authors declare that financial support was received for this work and its publication. This research was financially supported by the Slovenian Research and Innovation Agency (research program KINSPO—Kinesiology for the Effectiveness and Prevention of Musculoskeletal Injuries in Sports [P5-0443] and VIRTUES—Scientific Basis for the Use of Virtual Reality in Modulating Biopsychosocial Factors That Limit Engagement and Participation in Exercise Therapy [J5-70192]).

Authors' Contributions

Conceptualization: NŠ, OC, IM
Data curation: OC, IM
Formal analysis: OC, IM
Methodology: NŠ, OC, IM
Supervision: NŠ
Writing – original draft: OC, IM
Writing – review & editing: NŠ, OC, IM

Conflicts of Interest

None declared.

Multimedia Appendix 1
Literature review search strategies for each database.
[DOCX File, 21 KB - rehab_v13i1e88182_app1.docx]

Multimedia Appendix 2
Excel file containing additional information about identified studies and questionnaires.
[XLSX File, 181 KB - rehab_v13i1e88182_app2.xlsx]

Multimedia Appendix 3
Complete item bank containing all items developed for the Questionnaire for Assistive, Rehabilitation, and Training Technologies (U-ART-Q).
[DOCX File, 26 KB - rehab_v13i1e88182_app3.docx]

Multimedia Appendix 4
Question selection matrix.
[PDF File, 112 KB - rehab_v13i1e88182_app4.pdf]

Multimedia Appendix 5
Excel-based questionnaire generator (.xlsm) that uses initial responses to produce a tailored list of items, which can be viewed and exported to PDF.
[ZIP File, 74 KB - rehab_v13i1e88182_app5.zip]

References

1. Demiral DG. Emerging assistive technologies and challenges encountered. In: Canfi HM, Butuner R, editors. Current Studies in Technology, Innovation and Entrepreneurship: ISRES Publishing; 2023:1-21.
2. McNamara N. The psychometric approach to user satisfaction measurement. Interact Comput 2013 Jul 1;25(4):294-298. [doi: 10.1093/iwc/iwt014]
3. Cordier R. The technological revolution, occupational therapy practice and research: how are we responding? Aust Occup Ther J 2016 Dec;63(6):361-363. [doi: 10.1111/1440-1630.12344] [Medline: 27981633]
4. ISO 9999:2022: assistive products—classification and terminology. : International Organization for Standardization; 2022 URL: https://cdn.standards.iteh.ai/samples/72464/3f3608ed0bff4545bd53c02373f8cdbb/ISO-9999-2022.pdf [accessed 2026-09-09]
5. Anson DK. Assistive Technology for People with Disabilities, 1st edition: Greenwood; 2018. [doi: 10.5040/9798400615559]
6. Thakur N, Han CY. A review of assistive technologies for activities of daily living of elderly. In: Eckstein SM, editor. Elderly Care: Current Issues and Challenges: Nova Science Publishers; 2020:61-84.

XSL•FO
RenderX

7. Nierling L, Maia M. Assistive technologies: social barriers and socio-technical pathways. Societies 2020;10(2):41. [doi: 10.3390/soc10020041]
8. Harb A. Framework to assist in selecting human centred design tools and methods. J Arts Archit Res Stud 2020;1(2):276-287. [doi: 10.47436/jaarsfa.v1i2.72]
9. ISO 9241-210:2019: ergonomics of human-system interaction—part 210: human-centred design for interactive systems. : International Organization for Standardization; 2019 URL: https://cdn.standards.iteh.ai/samples/77520/33cc3cf17e5e4f1894ee872ed4f54fb6/ISO-9241-210-2019.pdf [accessed 2026-09-09]
10. Nielsen J. Usability Engineering: Academic Press; 1993.
11. Lyon AR, Munson SA, Renn BN, et al. Use of human-centered design to improve implementation of evidence-based psychotherapies in low-resource communities: protocol for studies applying a framework to assess usability. JMIR Res Protoc 2019 Oct 9;8(10):e14990. [doi: 10.2196/14990] [Medline: 31599736]
12. Maguire M. Methods to support human-centred design. Int J Hum Comput Stud 2001 Oct;55(4):587-634. [doi: 10.1006/ijhc.2001.0503]
13. Davis K, Douglas D. Translating usability test data into client-ready findings and recommendations. Proc Hum Factors Ergon Soc Annu Meet 2015 Sep;59(1):1764-1768. [doi: 10.1177/1541931215591381]
14. Garcia SE. Usability testing: creative techniques for answering your research questions. In: CHI EA '20: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2020:1-2. [doi: 10.1145/3334480.3375064]
15. Dumas JS, Redish JC. A Practical Guide to Usability Testing: Ablex Publishing Corporation; 1993.
16. Drungilas D, Ramašauskas I, Kurmis M. Emotion recognition in usability testing: a framework for improving web application UI design. App Sci 2024;14(11):4773. [doi: 10.3390/app14114773]
17. Persson J. A review of the design and development processes of simulation for training in healthcare: a technology-centered versus a human-centered perspective. Appl Ergon 2017 Jan;58:314-326. [doi: 10.1016/j.apergo.2016.07.007] [Medline: 27633228]
18. Masbernat-Almenara M, Rubi-Carnacea F, Opisso E, Duarte-Oller E, Medina-Casanovas J, Valenzuela-Pascual F. Developing an assistive technology usability questionnaire for people with neurological diseases. PLoS ONE 2023;18(1):e0281197. [doi: 10.1371/journal.pone.0281197] [Medline: 36719908]
19. Phillips B, Zhao H. Predictors of assistive technology abandonment. Assist Technol 1993;5(1):36-45. [doi: 10.1080/10400435.1993.10132205] [Medline: 10171664]
20. Wang J, Zhang J. Adapting a unified electronic health record usability framework for evaluation of connected health care technologies linking mobile data. iProc 2015;1(1):e20. [doi: 10.2196/iproc.4477]
21. Caronni A, Ramella M, Arcuri P, et al. The Rasch analysis shows poor construct validity and low reliability of the Quebec User Evaluation of Satisfaction with Assistive Technology 2.0 (QUEST 2.0) questionnaire. Int J Environ Res Public Health 2023 Jan 6;20(2):1036. [doi: 10.3390/ijerph20021036] [Medline: 36673791]
22. Koumpouros Y. A systematic review on existing measures for the subjective assessment of rehabilitation and assistive robot devices. J Healthc Eng 2016;2016:1048964. [doi: 10.1155/2016/1048964] [Medline: 27196802]
23. Bettoni E, Ferriero G, Bakhsh H, Bravini E, Massazza G, Franchignoni F. A systematic review of questionnaires to assess patient satisfaction with limb orthoses. Prosthet Orthot Int 2016 Apr;40(2):158-169. [doi: 10.1177/0309364614556836] [Medline: 25428901]
24. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160. [doi: 10.1136/bmj.n160] [Medline: 33781993]
25. Demers L, Weiss-Lambrou R, Ska B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): an overview and recent progress. TAD 2002;14(3):101-105. [doi: 10.3233/TAD-2002-14304]
26. Gao M, Kortum P, Oswald F. Psychometric evaluation of the USE (Usefulness, Satisfaction, and Ease of USE) questionnaire for reliability and validity. Proc Hum Factors Ergon Soc Annu Meet 2018 Sep;62(1):1414-1418. [doi: 10.1177/1541931218621322]
27. Jannink MJA, de Vries J, Stewart RE, Groothoff JW, Lankhorst GJ. Questionnaire for usability evaluation of orthopaedic shoes: construction and reliability in patients with degenerative disorders of the foot. J Rehabil Med 2004 Nov;36(6):242-248. [doi: 10.1080/16501970410033569] [Medline: 15841599]
28. Peres SC, Pham T, Phillips R. Validation of the System Usability Scale (SUS). Proc Hum Factors Ergon Soc Annu Meet 2013 Sep;57(1):192-196. [doi: 10.1177/1541931213571043]
29. Schrepp M, Hinderks A, Thomaschewski J. Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). Int J Interact Multimed Artif Intell 2017;4(6):103-108. [doi: 10.9781/ijimai.2017.09.001]
30. Sorrentino G, Vercelli S, Salgovic L, Ronconi G, Bakhsh HR, Ferriero G. Psychometric properties of the Client Satisfaction with Device module of the Orthotics and Prosthetics Users' Survey (OPUS): a scoping review. Int J Rehabil Res 2021 Sep 1;44(3):193-199. [doi: 10.1097/MRR.0000000000000484] [Medline: 34356037]
31. van Netten JJ, Hijmans JM, Jannink MJA, Geertzen JHB, Postema K. Development and reproducibility of a short questionnaire to measure use and usability of custom-made orthopaedic shoes. J Rehabil Med 2009 Nov;41(11):913-918. [doi: 10.2340/16501977-0447] [Medline: 19841843]

XSL•FO
RenderX

32. Allvin R, Ehnfors M, Rawal N, Svensson E, Idvall E. Development of a questionnaire to measure patient-reported postoperative recovery: content validity and intra-patient reliability. J Eval Clin Pract 2009 Jun;15(3):411-419. [doi: 10.1111/j.1365-2753.2008.01027.x] [Medline: 19366398]
33. Almanasreh E, Moles R, Chen TF. Evaluation of methods used for estimating content validity. Res Social Adm Pharm 2019 Feb;15(2):214-221. [doi: 10.1016/j.sapharm.2018.03.066] [Medline: 29606610]
34. Lawshe CH. A quantitative approach to content validity. Pers Psychol 1975;28(4):563-575. [doi: 10.1111/j.1744-6570.1975.tb01393.x]
35. Lynn MR. Determination and quantification of content validity. Nurs Res 1986;35(6):382-385. [doi: 10.1097/00006199-198611000-00017] [Medline: 3640358]
36. Lee BS, Lee JA, Choi H, et al. Development of the usability scale for power wheelchairs and its application in power wheelchairs users in Korea. J Mech Med Biol 2020 Dec;20(10):2040031. [doi: 10.1142/S021951942040031X]
37. Dos Santos ADP, Ferrari ALM, Medola FO, Sandnes FE. Aesthetics and the perceived stigma of assistive technology for visual impairment. Disabil Rehabil Assist Technol 2022 Feb;17(2):152-158. [doi: 10.1080/17483107.2020.1768308] [Medline: 32501732]
38. David BG. The role of aesthetics in medical device design. Res Inven J Biol Appl Sci 2024;4(1):26-30. [doi: 10.59298/RIJBAS/2024/412629]
39. Jeong J, Kim Y, Roh T. Do consumers care about aesthetics and compatibility? The intention to use wearable devices in health care. SAGE Open 2021 Jul;11(3):21582440211040070. [doi: 10.1177/21582440211040070]
40. ISO 9241-11:2018: ergonomics of human-system interaction—part 11: usability: definitions and concepts. : International Organization for Standardization; 2018 URL: https://preview.sist.si/sist-preview/63500/33c267a5a7564f298f02bbd65721a181/ISO-9241-11-2018.pdf [accessed 2026-09-09]
41. ISO/IEC 25019:2023: systems and software engineering—systems and software quality requirements and evaluation (square)—quality-in-use model. : International Organization for Standardization; 2023 URL: https://cdn.standards.iteh.ai/samples/78177/7c8fe3ed9fdc4c06a8c9a14a8cdae2ec/ISO-IEC-25019-2023.pdf [accessed 2026-09-09]
42. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
43. Fundarò C, Casale R, Maestri R, et al. Technology Assisted Rehabilitation Patient Perception Questionnaire (TARPP-Q): development and implementation of an instrument to evaluate patients' perception during training. J Neuroeng Rehabil 2023 Mar 24;20(1):35. [doi: 10.1186/s12984-023-01146-3] [Medline: 36964543]
44. Heinssen RK Jr, Glass CR, Knight LA. Assessing computer anxiety: development and validation of the Computer Anxiety Rating Scale. Comput Human Behav 1987 Jan;3(1):49-59. [doi: 10.1016/0747-5632(87)90010-0]
45. Wilson ML, Huggins-Manley AC, Ritzhaupt AD, Ruggles K. Development of the Abbreviated Technology Anxiety Scale (ATAS). Behav Res Methods 2023 Jan;55(1):185-199. [doi: 10.3758/s13428-022-01820-9] [Medline: 35338456]
46. Kishore K, Jaswal V, Kulkarni V, De D. Practical guidelines to develop and evaluate a questionnaire. Indian Dermatol Online J 2021;12(2):266-275. [doi: 10.4103/idoj.IDOJ_674_20] [Medline: 33959523]

Abbreviations

ATAS: Abbreviated Technology Anxiety Scale
CARS: Computer Anxiety Rating Scale
CSD-OPUS: Client Satisfaction With Device Module of the Orthotics and Prosthetics Users' Survey
CVR: content validity ratio
HCD: human-centered design
I-CVI: item-level content validity index
NASA-TLX: NASA Task Load Index
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUE: Questionnaire for Usability Evaluation of shoes
QUEST 2.0: Quebec User Evaluation of Satisfaction With Assistive Technology
SUS: System Usability Scale
TAPES: Trinity Amputation and Prosthesis Experience Scale
TARPP-Q: Technology Assisted Rehabilitation Patient Perception Questionnaire
U-ART-Q: User Satisfaction Questionnaire for Assistive, Rehabilitation, and Training Technologies
USE: Usefulness, Satisfaction, and Ease of Use

XSL•FO
RenderX

Edited by S Munce; submitted 20.Nov.2025; peer-reviewed by I Markoc, N Layton; revised version received 12.Aug.2026; accepted 20.Aug.2026; published 25.Sep.2026.

Please cite as:
Cvjetičanin O, Marić I, Šarabon N
User Satisfaction With Assistive, Rehabilitation, and Training Technologies: Questionnaire Development and Content Validation Study
JMIR Rehabil Assist Technol 2026;13:e88182
URL: https://rehab.jmir.org/2026/1/e88182
doi:10.2196/88182

© Oskar Cvjetičanin, Ivana Marić, Nejc Šarabon. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 25.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Analysis of Health Care Providers' Attitudes and Expectations on Intradialytic Exergaming: Qualitative SWOT Study

Aurel Zelko[1], PhD; Marianna Antoniadou[2], MSc; Jaroslav Rosenberger[3,4,5,6], MD, PhD; Zuzana Dankulincova Veselska[1], Prof Dr; Andrea Madarasova Geckova[1,3,4,7], Prof Dr; Leonie Klompstra[2], PhD; Tiny Jaarsma[2], Prof Dr

[1]Department of Health Psychology and Research Methodology, Faculty of Medicine, University of Pavol Jozef Šafárik, Trieda SNP 1, Košice, Slovakia

[2]Department of Health, Medicine and Caring Sciences, Linköping University, Växthuset, House 511-001, Linköping, Östergötland, Sweden

[3]Department of Community and Occupational Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[4]Olomouc University Social Health Institute, Palacký University Olomouc, Olomouc, Czech Republic

[5]Department of Internal Medicine, Faculty of Medicine, University of Pavol Jozef Šafárik, Košice, Slovakia

[6]Fresenius Medical Care Dialysis Services Košice, Košice, Slovakia

[7]Institute of Applied Psychology, Faculty of Social and Economic Sciences, Comenius University Bratislava, Bratislava, Slovakia

Corresponding Author:
Tiny Jaarsma, Prof Dr
Department of Health, Medicine and Caring Sciences, Linköping University, Växthuset, House 511-001, Linköping, Östergötland, Sweden

Abstract

Background: Physical inactivity significantly contributes to the rapid decline in clinical and health-related outcomes among patients undergoing hemodialysis. Exergaming may offer a safe, effective, and sustainable solution for maintaining adequate physical activity levels in these patients.

Objective: This study aims to support its clinical applicability; our aim was to examine the attitudes and expectations of health care providers regarding intradialytic exergaming (IEG) in hemodialysis care in Slovakia.

Methods: We conducted 21 individual, in-person, semistructured interviews with health care providers. Using purposive sampling, we included training assistants, nurses, nephrologists, and managers from 2 dialysis centers in Slovakia. The verbatim transcripts of interviews were qualitatively analyzed using an inductive thematic analysis approach. The identified codes and themes were further deductively categorized into the strengths, weaknesses, opportunities, and/or threats quadrants of the Strengths, Weaknesses, Opportunities, and Threats (SWOT) framework.

Results: We identified 2 strengths, 2 weaknesses, 3 opportunities, and 3 threats of IEG. Strengths included health care providers' enthusiasm and belief in the safety of IEG. Weaknesses involved concerns about the complexity of IEG and low digital literacy among patients and providers. Opportunities were seen in improving information availability and specificity, adjusting clinical culture, and tailoring exercises. Threats included skepticism about the availability of financial resources for IEG, resistance to changes in clinical practice, and issues related to patient behavior and privacy.

Conclusions: Health care providers perceived IEG as a safe and appropriate intervention for clinical practice. To improve implementation strategies, it is important to consider specific expectations regarding resource availability, as well as the technological and organizational aspects of IEG.

Trial Registration: ClinicalTrials.gov NCT03511924; https://clinicaltrials.gov/study/NCT03511924

International Registered Report Identifier (IRRID): RR2-10.1097/MD.0000000000015570

(JMIR Rehabil Assist Technol 2026;13:e106283) doi:10.2196/106283

KEYWORDS

hemodialysis; exergaming; qualitative research; intradialytic exergaming; health care providers; physical activity

Introduction

Maintaining adequate physical activity is one of the key challenges in the self-management of patients with chronic conditions. Kidney Disease Improving Global Outcomes Clinical Practice Guideline for the evaluation and management of chronic kidney disease recommends that all patients with chronic kidney disease be advised to undertake moderate-intensity physical activity for a cumulative duration of at least 150 minutes per week, or to a level compatible with their cardiovascular and physical tolerance [1]. Insufficient physical activity in patients with chronic diseases is associated with higher rates of

XSL•FO
RenderX

hospitalization, disability, and mortality, as well as increased financial and clinical burdens on both individuals and the health care system [2-5]. Health-related complications and behavioral limitations associated with primary chronic diseases, coexisting comorbidities, and the side effects of therapeutic procedures negatively affect patients' physical activity levels [6-11]. To address these complications and limitations, various exergaming strategies have been introduced as innovative behavioral interventions aimed at reducing sedentary lifestyles and promoting physical activity among individuals with chronic diseases [12-18].

Current evidence on the clinical application of exergaming, defined as the interactive experience of playing video games that require physical exertion or movement [19], indicates beneficial effects on physical activity behaviors and overall physical functioning in patients with chronic diseases. Benefits include greater adherence to recommended daily physical activity levels, enhanced muscle strength, endurance, physical capacity, and improved postural control, balance, and motor functions [20-27]. Maintaining regular physical activity, muscle function, and overall physical conditioning is a fundamental aspect in clinical management of chronic kidney disease, particularly for patients undergoing hemodialysis [28-30]. Low levels of physical activity and their progressive decline over time in patients undergoing hemodialysis are associated with increased morbidity, higher risks of all-cause and cardiovascular mortality, and poorer disease prognosis [31-36]. Given that many of these patients experience limitations and insufficiency in performance of daily physical activity and face a high risk of protein-energy wasting, malnutrition, and dynapenia or sarcopenia [37-41], intradialytic exergaming (IEG) may serve as a well-suited and effective component of clinical behavioral interventions [42-45].

The development and implementation of clinical behavioral interventions that primarily target physical activity and exercise through exergaming technologies present a highly challenging task [46-48]. Planning, execution, and continuation of such interventions for patients undergoing hemodialysis demand a comprehensive understanding of the disease's clinical and therapeutic background, the physiological effects of exercise, and current clinical guidelines for exercise prescription [28,49]. Additionally, effective integration of technological solutions into clinical practice requires consideration of logistical, financial, legislative, and organizational factors [50,51]. From an organizational perspective, it is necessary to understand and consider the qualitative dimensions of patients' and health care providers' awareness, knowledge, expectations, skills, competencies, cognition, and attitudes toward IEG [52-57].

Therefore, the aim of our study was to explore attitudes and expectations of health care providers regarding IEG in hemodialysis care in Slovakia. As participating health care providers had no prior experience with IEG, our study did not seek to evaluate actual experiences with the intervention. Rather, it focused on health care providers' hypothetical perceptions, attitudes, and expectations concerning the potential implementation of IEG in hemodialysis care. To achieve this objective, we used the Strengths, Weaknesses, Opportunities, and Threats (SWOT) framework to analyze collected data and guide the formulation of results. The SWOT framework provides a practical approach for evaluating the factors that may influence the success or failure of a procedure by examining both internal capabilities (Strengths and Weaknesses) and external challenges (Opportunities and Threats). It organizes these factors into 4 categories: Strengths are internal assets that enhance performance indicators and support the achievement of objectives. Weaknesses refer to internal limitations that may reduce the acceptability or effectiveness of the procedure. Opportunities are external conditions or emerging trends that can be leveraged to improve outcomes or broaden impact. Threats encompass external risks or challenges that may undermine the procedure's success or sustainability. As a strategic management tool, SWOT analysis supports informed decision-making by identifying key internal and external factors, thereby facilitating the effective introduction and integration of innovative procedures into clinical practice [58-62].

Methods

Study Design

This study was part of a project investigating the acceptability and effectiveness of a 12-week progressive, individualized, supervised intradialytic resistance training intervention on muscle function among patients undergoing hemodialysis. The study was an ancillary qualitative substudy conducted within the overarching framework of project NCT03511924 (ClinicalTrials.gov) [63]. Although the intervention resulted in moderate to large improvements in muscle function, the acceptability of intradialytic exercise and the sustainability of its benefits remained low [64,65]. In the current study, we used a 2-step qualitative design to explore and categorize health care providers' attitudes and expectations regarding IEG among patients undergoing hemodialysis. In the first step, we explored health care providers' perspectives using inductive thematic analysis (guided by [66] of data collected from participants' interviews. In the second step, we deductively classified the identified codes and themes using a SWOT analysis approach (as outlined [59].

Participants

We recruited participants from 2 dialysis centers located in Slovakia. Training assistants, nurses, nephrologists, and managers were eligible to participate if they had been employed or actively involved at the dialysis center for the entire duration of the 3-month intradialytic exercise intervention and were at least 18 years of age. This criterion ensured that participants had direct experience with the implementation of intradialytic exercise in routine clinical practice and were able to provide hypothetical insights into factors influencing the initiation and implementation of IEG. Furthermore, nurses, nephrologists, and managers were required to have a minimum of 2 years of professional experience in their respective roles. All available health care providers from 2 dialysis centers were prescreened based on their age, sex, and professional experience, including the total number of years worked in both current and previous health care positions and roles. While all participants had experience with the previously implemented intradialytic exercise intervention, none had prior experience with IEG. To

XSL•FO
RenderX

ensure diversity within the sample, individuals of different ages and professional backgrounds were purposefully selected and invited to participate in the study via an email sent by 2 members of the research team (AZ and IMS). Health care providers who accepted the invitation received oral and written information about the study, provided written informed consent, and had their interview date and time scheduled. We continued to include participants until we reached sufficiency of information for all studied topics. This occurred when responses to questions from successive participants became repetitive and no longer provided new, unique insights to the outcome of the study [67].

Data Collection and Procedures

Individual, in-person, semistructured interviews with health care providers were conducted using an interview guide (Multimedia Appendix 1). The interviews began in June 2019, approximately 1 month after the completion of the intradialytic exercise intervention study. Members of the research team (AZ, IMS, and DK) developed the interview guide using the Patient-Centered Access to Health Care framework [68] and the Biopsychosocial Model of Responsiveness to Physical Activity [69] as conceptual frameworks. Three senior members of the research team (JR, ZDV, and AMG) evaluated the initial interview guide and proposed adjustments to enhance its clarity. Further minor adjustments were made following a pilot interview with one dialysis center health care provider. During the interview process, the researchers continually adjusted the guide to delve into areas of specific interest, aligning with best practices for qualitative data collection [70].

The final interview guide covered 4 scope areas (detailed in Multimedia Appendix 1). The first 3 areas explored (1) the perceived benefits of physical activity for dialysis patients in general, (2) the perceived benefits and risks of intradialytic exercise, and (3) experiences with implementing exercise interventions in clinical settings. Findings from these 3 areas were presented in a separate article [65]. The fourth area focused on participants' perspectives on variations in intradialytic exercise interventions, with particular emphasis on IEG. Participants were asked about their opinions and suggestions on alternatives to intradialytic exercise interventions, as well as their attitudes and expectations regarding IEG. Data from this area served as the basis for the current article. The open-ended questions in the interview guide allowed participants to share their attitudes and expectations regarding IEG, enabling interviewers to probe deeper into individual perspectives. Participants were encouraged to elaborate and highlight issues they deemed important. The research team facilitated verbal expression and logical thought development, taking field notes when necessary.

Three members of the research team (AZ, IMS, and DK) conducted interviews from June 2019 through November 2019. Prior to the interviews, written informed consent was obtained from participants, along with a brief personal questionnaire summarizing their sociodemographic characteristics (age and sex) and professional experiences (years in previous and current job positions). The interviews were held in Slovak language, either in private meeting rooms at dialysis centers (for nurses and nephrologists) or in an office at the authors' workplace (for assistants and management staff). All interviews were conducted individually, involving only the participant and 1 or 2 members of the research team (AZ, IMS, and DK). The researchers conducting the interviews had no prior personal or professional relationships with the participants. Before the interviews began, ZDV instructed the researchers on optimal design and conduct of qualitative research interviews and assessed their professional backgrounds, potential biases, and personal interests related to the interview topics. To enhance confirmability, the research team comprised members with diverse theoretical and professional backgrounds, including Sports Sciences, Occupational Therapy, Health Psychology, Nursing, Digital Health, and Clinical Nephrology. Furthermore, team members were actively involved in all stages and levels of the data analysis process. This facilitated critical reflection on emerging interpretations, ensured consideration of multiple perspectives, and helped maintain a close connection between the findings and the participants' accounts while minimizing the influence of individual researcher biases. All interviews were audio-recorded with the participants' consent. The interview transcripts were not returned to participants for comments or corrections, and no follow-up interviews were conducted.

Data Handling and Analysis

During data handling, AZ and IMS transcribed all interview audio recordings verbatim and anonymized them, ensuring the removal of all personal participant data. AZ and IMS manually reviewed all transcriptions, comparing them to the audio recordings to verify the accuracy and authenticity of the transcripts. The authenticated transcripts and audio recordings were subsequently uploaded to the MAXQDA® standard platform (VERBI Software).

We analyzed the collected evidence of health care providers' perspectives regarding IEG in 2 steps. In the first step, we used Braun and Clarke [66] thematic analysis to inductively explore and identify codes, themes, and their patterns, specifically related to participants' attitudes and expectations. Two members of the research team (AZ and IMS) repeatedly read the original Slovak transcripts to familiarize themselves with the data (Phase 1) and then independently generated codes for the entire dataset (Phase 2). Quotations included in the paper were translated from Slovak into English by 2 members of the research team (AZ and IMS), and translation accuracy was reviewed by a senior member of the research team (ZDV). Outputs of initial coding were exchanged between 2 members (AZ and IMS) for coding quality cross-check. Four coding meetings (2 online, 2 in-person) were initiated to discuss problematic codes, disagreements in coding, and appeals to recode until consensus in the coding output was achieved. Through iterative discussion and reexamination of the transcripts, codes were refined and, where necessary, recoded until consensus was reached. Both researchers then associated thematically similar codes into groups and subgroups regarding attitudes and expectations on IEG (Phase 3). These themes were subsequently cross-referenced with the raw data to ensure they accurately represented the participants' experiences (Phase 4). The final output of the coding process and thematic code grouping was reviewed, discussed, validated, and approved by 3 senior members of the research team (JR, ZDV, AMG; Phase 5). At

XSL•FO
RenderX

the end of the first step of data analysis, the entire research team documented their findings and collectively summarized health care providers' attitudes and expectations regarding IEG (Phase 6). Using this procedure, we identified 139 codes and categorized them into 54 subthemes and 5 overarching themes.

In the second step, we used a SWOT analysis approach to organize emerging themes into strengths and weaknesses [potential internal factors], and opportunities and threats [potential external factors]. We conducted a deductive analysis using a revised SWOT model specifically adapted for health care research. The model is built on 3 foundational pillars: addressing stakeholder expectations, evaluating resource management, and considering contextual developments [59]. One member of the research team (AZ) extracted relevant parts from the results of the inductive thematic analysis and categorized them into the 4 quadrants of the SWOT framework. During this process, all codes and subthemes derived from the inductive thematic analysis were systematically examined and categorized as either internal or external factors. Internal factors were defined as those originating within the dialysis centers and directly related to their organizational structures, personnel, operational routines, and implementation processes. External factors were defined as those arising outside the dialysis centers and beyond their direct control, including health care policies, reimbursement frameworks, technological advancements, and broader societal and patient-related influences. They were then evaluated in terms of their hypothetical positive or negative effects on the initiation and implementation processes. Strengths were identified as internal benefits positively related to the IEG at dialysis centers. Weaknesses were identified as internal limitations that might negatively affect or impede the exergaming during dialysis. Opportunities were considered external factors that could positively enhance the applicability and sustainability of IEG. Lastly, threats were identified as external factors that could pose barriers to or adversely affect the implementation of IEG [71]. Afterward, members of the research team (JR, ZDV, AMG; and MA, LK, TJ serving as independent experts not involved in data collection) reviewed and validated whether the content extracted through inductive thematic analysis was appropriate and whether its deductive categorization aligned with the quadrants of the SWOT framework. Using this approach, 43 categories were generated, resulting in 2 strengths, 2 weaknesses, 3 opportunities, and 3 threats regarding IEG. The process of mapping the results of the inductive thematic analysis onto the SWOT framework is presented in the analytical flow table (Multimedia Appendix 2).

Ethical Considerations

The study protocol was reviewed and approved by the Ethics Committee of the Pavol Jozef Šafárik University in Košice, Slovakia [approval no. 14N/2017]. Written informed consent was obtained from all study participants. We report this qualitative study according to the Standards for Reporting Qualitative Research and the Consolidated Criteria for Reporting Qualitative Research guidelines (Checklist 1) [72,73]. The research team comprised experts from various fields, including Sports Sciences (AZ, PhD), Occupational Therapy (MA, MSc), Health Psychology (IMS, PhD; DK, MSc; ZVD, prof; AMG, prof), Nursing and Digital Health (LK, PhD; TJ, prof), and Clinical Nephrology (JR, PhD and MD). All team members had prior experience in researching older people with chronic diseases, as well as expertise in designing, analyzing, and writing qualitative studies.

Results

Characteristics of the Sample (Demographic Data)

We invited 24 health care providers from 2 cooperating dialysis centers to participate in our study. Three care providers (1 nurse and 2 nephrologists) declined to participate due to high workloads and privacy concerns. We conducted semistructured, individual interviews with 3 training assistants, 11 dialysis nurses, 5 nephrologists, and 2 members of dialysis center management. Of the 21 participating health care providers, 11 were recruited from the first dialysis center (6 dialysis nurses, 2 nephrologists, 2 training assistants, and 1 dialysis center manager) and 10 from the second dialysis center (5 dialysis nurses, 3 nephrologists, 1 training assistant, and 1 dialysis center manager). Interviews lasted between 20 and 60 minutes, with an average duration of 39 minutes. One-fourth of the participants were over 50 years old, one-third were male, and half had over 15 years of professional experience. Table 1 provides the demographic and working experience characteristics for each participant's professional position.

Table . Participants' demographic and working experience characteristics by professional position.

Professional position (n)	Sex (male/female), n	Age (years), mean (SD)	Working experience (years), mean (SD)
Dialysis nurse (n=11)	0/11	45.2 (8.1)	18.2 (10.0)
Nephrologist (n=5)	4/1	46.8 (12.6)	20.6 (11.8)
Training assistant (n=3)	2/1	21.7 (0.6)	1.0 (0.0)
Dialysis center manager (n=2)	2/0	55.5 (2.1)	16.0 (1.4)
Total (n=21)	8/13	43.2 (12.5)	16.1 (11.0)

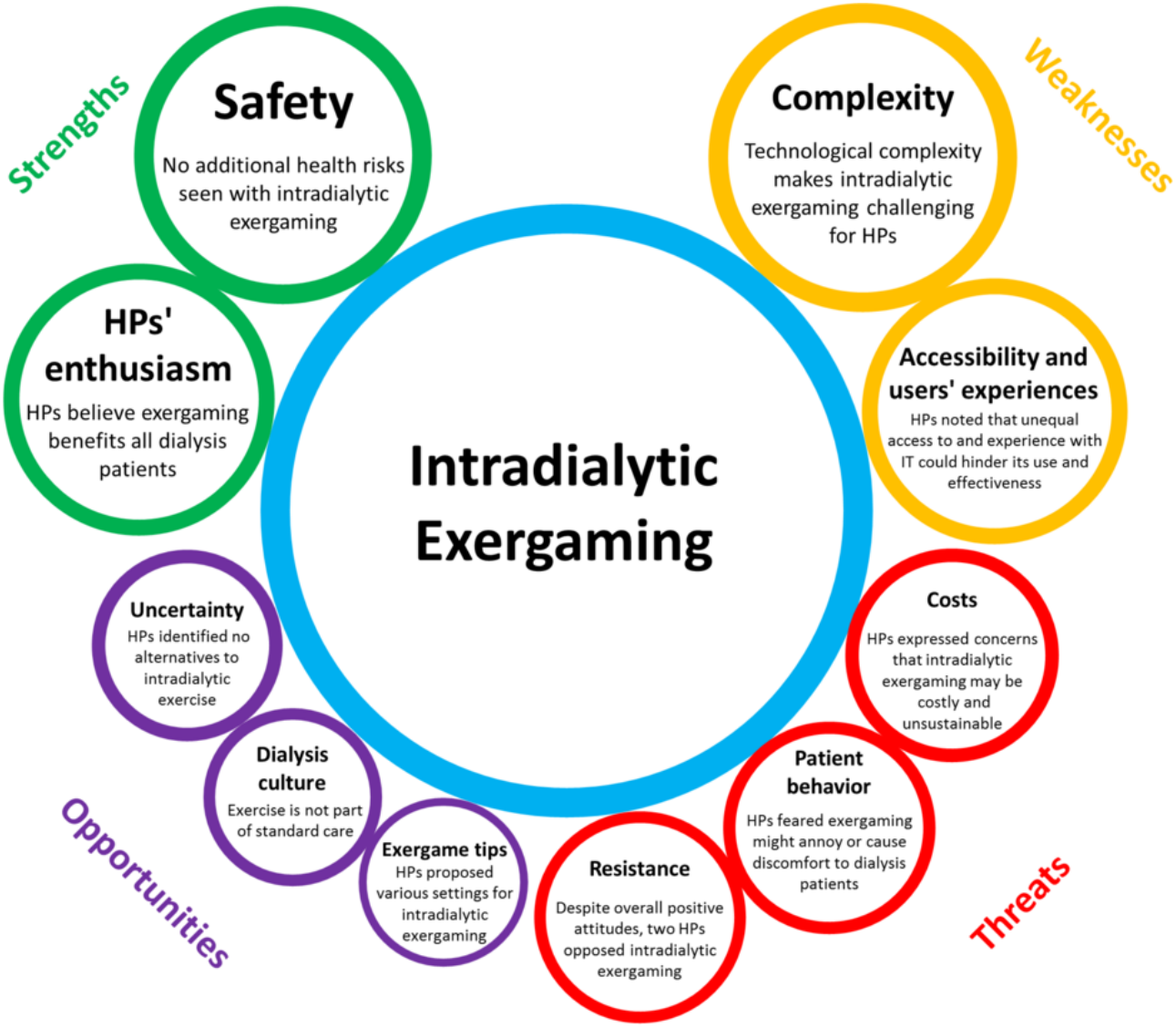
XSL•FO
RenderX

Overview of Findings From Inductive Thematic Analysis and SWOT Analysis

Using inductive thematic analysis, we identified 139 codes, which we then grouped into 54 subthemes and further organized into 5 overarching themes. These themes reflect a range of perspectives, including (1) mixed attitudes toward IEG: balancing optimism and skepticism, (2) attitudes on the acceptability and long-term sustainability of IEG interventions for patients, (3) expectations for addressing patients' individual needs in IEG interventions, (4) expected benefits of IEG, and (5) expectations of the potential risks and barriers associated with IEG.

The SWOT analysis led to the identification and formulation of 2 strengths, 2 weaknesses, 3 opportunities, and 3 threats concerning IEG (Figure 1). (1) Strengths included (a) health care providers' beliefs that exergaming is a safe activity to perform during dialysis, along with their (b) enthusiastic attitudes toward its integration into clinical practice. (2) Weaknesses reflected concerns about (a) the technological complexity of exergaming devices and platforms, and (b) inadequacies in patients' access to and experiences with digital technologies. (3) Opportunities reported by health care providers included (a) uncertainty regarding alternative forms of exercise that could be implemented during dialysis, (b) calls for changes in dialysis culture, particularly to ensure better availability and greater specificity of information and support for IEG, and (c) suggestions and tips regarding the settings and modifications of exergaming interventions to facilitate their implementation during hemodialysis therapy. (4) Threats included (a) health care providers' skepticism about the availability of financial resources, (b) factors related to patients' behavior, personal space, and privacy, and (c) the resistance to IEG interventions among some health care providers. The overall categorization of the outputs derived from the inductive thematic analysis into the SWOT framework is presented in the analytical flow table (Multimedia Appendix 2).

Figure 1. Strengths, Weaknesses, Opportunities and Threats framework (SWOT) analysis template of identified and formulated themes regarding health care providers' attitudes and expectations on Intradialytic Exergaming (IEG). HP: health care provider.



XSL•FO
RenderX

Strengths of IEG

Safety

Health care providers regarded exergaming as a safe activity to be undertaken during dialysis. They expressed that, when properly introduced, clinically monitored, and technologically supervised, exergaming does not pose any additional health-related risks for patients undergoing hemodialysis. Other participants stated that they were either unaware of or not qualified to comment on any potential health risks that might be affected by IEG.

Health care providers mentioned that they had no prior experience with IEG and limited knowledge and familiarity with technology-based exercise interventions. They anticipated that, prior to the commencement of IEG, both patients and the hemodialysis care team would be adequately introduced to, and familiarized with, the exergaming environment, its functions and controls. Care providers stated that a comprehensive introduction to IEG could help minimize the risks of misuse, underuse, or overuse of the system.

> *No, I'm not sufficiently educated and experienced in this area [of exergaming]. So, I can't say exactly what types of exercises are suitable. Because not all stretching or resistance exercise can be done during dialysis… they should definitely be designed by someone who is a muscle [kinesiology] specialist. So, not just like, 'Here's a spring or band, exercise with it up and down, and it will go better.* [Nephrologist, male, 44 years old, 20 years of professional experience]

Health care providers emphasized the importance of clinical monitoring in managing potential health-related risks associated with IEG. During hemodialysis, patients must maintain a stable body position to ensure effective extracorporeal blood circulation. Sudden or uncontrolled movements of the torso or limbs could compromise vascular access and impair the efficiency of the dialysis process. Care providers emphasized the necessity of monitoring patients' activity during IEG to safeguard the adequacy and efficacy of the therapy. They also noted that end-stage kidney disease and the hemodialysis procedure itself may contribute to clinical risk conditions, such as dialysis-induced hypotension, hypertension, dysrhythmias, infections, or bleeding, which could potentially be influenced by exergaming during treatment. As such, these conditions must be closely monitored and evaluated. Once appropriate clinical monitoring is in place, health care providers did not anticipate an increased risk of injury or complications related to the patient's clinical profile.

> *We have quite an aged dialysis population. So, they generally have quite osteoporotic bones. It really needs to be individualized based on who the patient is, right? Because we have elderly women who have recently fallen. So, it's really necessary to approach each patient individually, considering their diagnoses, their age, and essentially what can be done with them, just like you were assessing at the beginning, I believe.* [Nephrologist, male, 40 years old, 10 years of professional experience]

> *The safety risk could be that the patient becomes so immersed in the game that they forget they are connected to the dialysis machine. They might get agitated during gameplay, causing their blood pressure to rise, or they could accidentally pull out the needle with a sudden movement or if they get startled. Such consequences would need to be addressed.* [Nephrologist, male, 40 years old, 10 years of professional experience]

> *No [risks for IEG], I don't think so. Only if it could be that his [patient's] blood pressure increases [during exergaming], right? And then the consequences would have to be dealt with.* [Nurse, female, 38 years old, 16 years of professional experience]

Regular inspections of exergaming technology used in dialysis health care settings are essential to ensure patient safety. Health care providers noted that undetected defects, faults, or failures in exergaming systems could increase the risk of injury and health-related complications. However, they stated that these risks could be mitigated through regular, properly conducted technical inspections, including evaluations of both the physical integrity and functional safety of the exergaming technology.

> *I would say the technical aspect of such things [exergaming]. It is related to some devices or controllers, and it is electronics, you know, which is not that simple. Technical side of dialysis is my area and everything technical here requires certain maintenance and also inspection when used. Within a certain time frame during use. All of this must be well controlled.* [Manager, male, 54 years old, 9 years of professional experience]

> *The question is, who is supposed to technologically secure all of this, the technology. We were distributing headphones to patients for listening to the television program… they got damaged.* [Nephrologist, 68 years old, 40 years of professional experience]

Health Care Providers' Enthusiasm

Patients undergoing hemodialysis often spend treatment sessions engaging in passive activities such as watching television, browsing the internet, reading, or listening to music. However, health care providers have noted that these activities often become monotonous and repetitive, leaving patients feeling bored, as they are disconnected from their real-life interests and hobbies. Care providers anticipated that IEG could serve as a more stimulating and varied alternative, an unconventional activity that captures attention, is personalized, and offers greater enjoyment through content that is engaging, distracting, and fun.

> *It might be fun and maybe the dialysis would pass more quickly for them... That would be perfect because they would have visuals, maybe play something, and just exercise along with it. Would that be one way to motivate the patient? Even someone*

XSL•FO
RenderX

who is shy might get into it. [Nephrologist, male, 43 years old, 19 years of professional experience]

Yes, because many of them watch things that are completely pointless or repetitive. There are people, yes, I can say, there are people here who watch educational programs, which is great. They are constantly learning. But there are also younger ones, for example, a 30-year-old who brings a series and basically watches it for 4 hours, which I think he doesn't need to do, he could be doing something else. He's too young for this. [Nurse, female, 38 years old, 16 years of professional experience]

Weaknesses of IEG

Complexity

Participants identified the technological complexity of using exergaming devices and platforms in clinical settings as a significant limitation. Several noted that limited space in some dialysis centers could hinder the installation and positioning of such equipment. They emphasized the importance of maintaining unobstructed access around the patient, particularly the area between the patient's vascular access site and the dialysis machine. Regarding design and installation setup, health care providers expressed a preference for exergaming devices and platforms to be wireless and equipped with internal or built-in power supplies. For positioning, they favored designs that could be securely attached to the external structure of the dialysis bed using a fastening system, allowing for quick and easy detachment by health care staff when necessary.

Because those beds are really designed for the patient to fit in, and maybe have something next to them. So, depending on how big the device would be, I'm afraid it would be a bit restrictive in terms of placement, right? [Nephrologist, male, 45 years old, 16 years of professional experience]

Accessibility and Users' Experiences

With regard to other limitations, health care providers expressed 2 broad groups of factors that may influence patients' engagement with IEG. The first group relates to technological accessibility and digital experience. Participants noted that patients undergoing hemodialysis, particularly those with limited access to digital technologies and minimal previous experience using them in daily life, may be skeptical, dismissive, or uninterested in IEG. They also noted that socioeconomic background and prior digital experience, particularly the frequency and quality of that experience, were strongly associated with patients' capacity to accept and adopt digital technologies in health care settings. The second group comprises individual patients' characteristics affecting technology use. Participants identified impaired visual, auditory, and cognitive functions as significant barriers to engaging with digital technologies, potentially hindering patients' ability to participate in exergaming during dialysis sessions. Interestingly, there was inconsistency among care providers regarding the role of age in patients' acceptance of exergaming technologies. Some mentioned that IEG is unsuitable for older patients undergoing hemodialysis, while others perceived no age-related disadvantages. The remaining participants acknowledged that older age could present a barrier but emphasized that many older patients are fully capable of using and benefiting from digital tools.

In my opinion, there are older patients who will find it harder to handle [exergaming device/platform] than a mobile phone. So, this will be more difficult for them. But sometimes we are surprised that even an 80-year-old grandmother has her [own] phone and can use it well. So, it would probably need to be assessed individually first. I can't say definitively yes or no for everyone. [Nephrologist, female, 34 years old, 8 years of professional experience]

Well, we have a computer available for patients if they are interested. We have internet in the hall [dialysis room] if patients are interested, but almost no one used it.... But I don't know if it would work like 'here you go, play.' And of course, they have to be... physically and mentally fit to do something like that [IEG]. [Nurse, female, 54 years old, 36 years of professional experience]

Opportunities of IEG

Uncertainty

Most of the health care providers were unable to suggest specific types and settings that could be used to tailor exergaming interventions to better suit the clinical environment, as well as the needs, expectations, and limitations of both health care providers and patients undergoing hemodialysis. The limited ability of health care providers to identify external modalities for tailoring exergaming interventions to the hemodialysis setting indicated an opportunity for future research, technological innovation, and stakeholder engagement to support the implementation of IEG. To enhance the usage of intradialytic exercise and exergaming, care providers identified several opportunities: increasing health care providers' awareness and knowledge related to the planning and organization of such interventions; ensuring either permanent or temporary access to exercise specialists for patients undergoing hemodialysis; and fostering international networks to support the exchange of best practices and shared experiences in this field.

You are asking me a difficult question, and I don't know, at the moment I don't know, I haven't thought about it [variations of intradialytic exercise], what and it would probably be the same with everyone [of participants]… because basically, when we think about it, we are always dealing only with the patient's health condition. [Nurse, female, 38 years old, 16 years of professional experience]

Dialysis Culture

Intradialytic exercise is not currently integrated into routine hemodialysis care in Slovakia. According to health care providers, exercise during dialysis, regardless of its form, is not recognized as a standard component of health care services. They reasoned that the health care system, particularly its legislative and organizational frameworks, has not yet been

XSL•FO
RenderX

adapted to support the integration of such services into clinical practice. Drawing on experiences shared by participants during their visits to dialysis centers abroad, they provided suggestions and proposed strategies for implementing effective IEG interventions. Opportunities to enhance the use of IEG include changes in dialysis culture, including creating a more family-like and relaxed clinical environment within dialysis centers; providing direct legislative support for such activities; and establishing complementary services such as nutritional and psychological counseling in routine hemodialysis care.

> *But this is not just a local problem of our [clinical] workplace; it is a systemic error. Simply put, the entire healthcare system [in Slovakia]. So we cannot sustain [nutrition status among] patients and do not exercise with them… In the [other areas of the] world, it works exactly like that: they know how many kilojoules a patient needs, what kind of diet they need, and which day after acute treatment they can start exercising.* [Nephrologist, female, 34 years old, 8 years of professional experience]

> *I can't imagine it [routine intradialytic exercise]. Because, for example, when you look at dialysis abroad, they dance and exercise with the patients, and there's a more relaxed atmosphere. Here, it's just so clinical, the dialysis, connecting and disconnecting, needles, everything sterile. It seems to me that everything has its rules, everything must be on time.* [Nurse, female, 34 years old, 10 years of professional experience]

Exergame Tips

After receiving a verbal introduction to intratherapeutic exercise interventions from members of the research team during the interview, health care providers were able to offer several suggestions regarding the setting and potential variations for implementing IEG. They proposed that patients could engage in endurance exercises such as bed bike pedaling, resistance exercises using rehabilitation equipment, and relaxation or breathing exercises during dialysis. According to them, these activities could be delivered in the form of exergaming and might be enhanced with accompanying music. Other participants suggested incorporating exergaming as a routine component of health care but recommended scheduling exergaming sessions at dialysis centers immediately after the completion of dialysis treatments, rather than during the sessions themselves.

> *Maybe they could actually pedal on those bikes, for example. I don't mean the ones used by athletes, but there are some, I don't know what they're called, where the patient sits upright, and they're actually used during rehabilitation.* [Training assistant, male, 23 years old, 1 year of professional experience]

Threats of IEG

Costs

Health care providers expressed uncertainty when identifying the entity or authority that might be responsible for funding IEG services. They expressed skepticism about the applicability and long-term sustainability of such interventions within the Slovak health care system, primarily due to the anticipated high initial financial investment and limited funding opportunities for supplementary health services. Several participants were unwilling to discuss or chose not to clearly articulate their views on the financial aspects of implementing and sustaining IEG. One participant suggested that demonstrating both the health-related benefits for patients and the cost-effectiveness of the intervention could serve as a foundation for further discussions with health care management organizations and stakeholders about securing additional financial support for such intratherapeutic services.

> *Well, for them it might actually be fun, and maybe it would make the dialysis go by faster for them, but who would fund it? Yeah, like for him [patient] it might actually be [fun], yeah, just unfortunately, who would fund it, who would take care of it?* [Nephrologist, male, 43 years old, 19 years of professional experience]

Patient Behavior

The health care providers expressed the view that due to the presence of kidney diseases and the complexity of associated health complications, patients often experience periods of frustration, demotivation, resignation, negativism, and oppositional behavior. Furthermore, care providers who participated in our study mentioned that each patient presents a unique set of personality traits, with an individual trajectory of development and expression. As a result, care providers anticipated that the introduction, familiarization, and adoption of the IEG concept would require a long-term, sustained effort and a personalized approach. They recommended avoiding strict, formal, task-oriented communication, and instead emphasized the importance of a sensitive, familiar, and pressure-free approach during the introduction and initiation of IEG. With regard to the specific nature of exergaming, health care providers mentioned that some patients might feel embarrassed due to the visibility of the activity within the shared space of dialysis rooms, as well as unfamiliarity with the exergaming technology among patients. Two participants specifically noted that differences in levels of game immersion could influence patients' engagement with exergaming. They also emphasized that psychological predispositions toward gaming addiction should be considered a potential risk associated with IEG.

Health care providers expressed concerns that patient behavior-related factors, which are largely external to the intervention and beyond the direct control of health care providers, could significantly influence patients' attitudes and willingness to engage with exergaming during hemodialysis therapy, and, as a consequence, could threaten the acceptability, usability, and sustainability of IEG.

> *As for barriers [for exergaming], what comes to my mind is addiction. Because after all, these [exergaming interventions] are games, and addiction to games is quite easy to develop, so from that perspective.* [Training assistant, female, 22 years old, 1 years of professional experience]

XSL•FO
RenderX

With regard to other potential risks, participants noted that violations of patients' personal space, privacy, and comfort during dialysis caused by exergaming activity could compromise the usage of intradialytic exercise interventions. Such disruptions could also affect patients who are not directly participating in exergaming during their treatment. Health care providers identified several sources of these intrusions, including changes to the timing of dialysis connection and disconnection procedures, alterations to the spatial layout of the clinical environment during hemodialysis, and increased noise levels or other external factors that may be perceived by patients as negative distractions during therapy.

> *I think these are some patients who... how to put it? They're kind of offended by the whole world, and they would seriously cause problems and refuse [exergaming], and maybe they would even cause problems for the other patients [whose exergaming during their dialysis] if they didn't feel comfortable in some way, and if it somehow limited them. I mean the exercise of the other patients, not their own.* [Training assistant, male, 21 years old, 1 year of professional experience]

> *If there were such interest, then the conditions would have to be created for it [IEG]. Because we have a large hall [dialysis room]. There are a lot of people there. Right? The question is how the others would react, whether it [exergaming] would disturb them.* [Nurse, female, 54 years old, 36 years of professional experience]

Resistance

We identified critical, negative, and mixed attitudes and expectations regarding IEG, including resistance to the use of digital tools. Two health care providers expressed strong opposition to the integration of exergaming into dialysis therapy. They expressed a sense of technological resistance and frustration, or overwhelm, with the increasing reliance on digital tools. In the context of intradialytic exercise, they demonstrated a clear preference for traditional, patient-centered approaches to health care, placing high value on direct, personal interaction over technology-mediated experiences. Additionally, 2 other caregivers anticipated initial resistance to IEG and raised concerns about the added responsibilities it could place on health care providers, as well as potential disruptions to the routines of dialysis nurses.

> *Well, I already think this is a bit over the top with these computers. Everyone should maybe step away from it a little, I would say. You know, it's [exergaming] not the same as when you're physically with the patient. It's not that kind of contact. It's just the computer again, and then some pedals again.* [Nurse, female, 60 years old, 40 years of professional experience]

> *I would probably be lying if I said it [IEG] would go smoothly. I'm expecting initial resistance, definitely. [It means] A load of additional work [for healthcare providers]. That's just how it is. A burden, simply because it's another [work]load.* [Nephrologist, female, 34 years old, 8 years of professional experience]

Discussion

Principal Findings

This study aimed to explore the attitudes and expectations of health care professionals regarding the concept of IEG as a strategy to support patient self-management in the domain of physical activity. It offers novel insights into the perceived strengths, weaknesses, opportunities, and threats associated with IEG, as reported by training assistants, nurses, nephrologists, and dialysis center managers.

Health care providers identified the perceived safety of IEG, along with generally positive and enthusiastic attitudes toward its use, as key strengths. However, they also highlighted the complexity of the exergaming concept, as well as closely related issues of accessibility and user experience, as potential weaknesses that could hinder its integration into health care services. In terms of opportunities, participants expressed uncertainty regarding alternative approaches to the organization and delivery of intradialytic exercise. They emphasized the need for cultural adjustments within dialysis settings to facilitate the implementation of various proposed IEG modalities in the clinical environment. The main threats identified included skepticism about the financial and resource implications of IEG, potential patient reluctance or refusal to participate due to behavioral factors, and resistance to change among health care professionals in adopting exergaming as part of routine care.

Patients' perceptions of safety play a key role in their decision to initiate and continue using technology-based physical activity interventions during dialysis therapy [74,75]. Our findings regarding perceived strengths suggest that health care providers generally view exergaming as a safe activity for patients undergoing hemodialysis. This aligns with existing evidence from studies monitoring patient safety, which consistently report no serious adverse events attributable to IEG [42,76-80]. From a broader perspective, research on exergaming interventions in patients with heart failure, stroke, or cancer similarly reports either no adverse events or only rare, exceptional cases [81-83]. Paradoxically, evidence from older adults and patients with multiple sclerosis, mild cognitive impairment, other neurocognitive disorders, and hypertension suggests that exergaming may raise concerns about physical safety. These concerns include fear of falls and general injuries, as well as occurrences of knee or back pain [16,84-89]. Such safety issues have been attributed to reduced awareness of the physical environment and insufficient guidance and support during exergaming activity [84,85,90]. Our participants expressed that, when properly introduced, clinically monitored, and technologically supported, IEG does not pose additional health risks to patients. They also expressed enthusiastic attitudes toward its use during hemodialysis therapy, which is consistent with existing evidence on positive perceptions among health care providers working with patients with dementia, cardiovascular conditions, stroke, and cancer [53,55,82,91,92]. Taken together, these insights into perceived strengths underscore the importance of promoting IEG by prioritizing

XSL•FO
RenderX

patient safety and leveraging the enthusiastic support of health care providers. Their positive perceptions can play a pivotal role in enhancing patient engagement and long-term adherence to technology-supported physical activity during dialysis therapy.

Technological limitations and a lack of design flexibility in exergaming devices and platforms may undermine their ecological validity and hinder their transferability into clinical practice [93-95]. Our participants identified several weaknesses related to the design, installation setup, and positioning of exergaming technologies within dialysis environments, collectively referred to as technological complexities. These included limited space in dialysis centers, the need to maintain unobstructed access between the patient and the dialysis machine, and preferences for wireless, self-powered devices that can be securely yet easily attached to dialysis beds. These findings are consistent with existing evidence on the general limitations and challenges of exergaming among patients undergoing hemodialysis, particularly those arising from environmental and organizational constraints [96-99]. Importantly, other authors have noted that such complexities may present challenges not only during the integration of these technologies into clinical settings but also for health care providers and end-users [55,100-105]. Our participants also identified additional barriers to IEG, including patients' digital literacy, sensory and cognitive impairments, socioeconomic factors, and differing views on the influence of age on user experience. These factors were seen as key challenges that may limit patients' accessibility and engagement with exergaming technologies [16,56,97,98,100,101,103,106]. To improve the quality of implementation, particularly the transferability and accessibility of IEG technologies in clinical practice, developers and promoters should be aware of and take into account the specific conditions of dialysis environments and patient-related limitations from the earliest stages of design and development.

Health care providers are central figures in the development, initiation, and long-term sustainability of innovative therapeutic approaches in clinical practice [107,108]. In relation to opportunities for IEG, participants expressed limited ability to tailor such interventions to clinical settings and individual patient needs, highlighting gaps in their awareness, knowledge, and experience. Nevertheless, they proposed several key opportunities to enhance implementation, including raising awareness among health care professionals, integrating exercise specialists into routine dialysis care, and fostering international collaboration to share best practices. Although intradialytic exercise is not yet a standard component of dialysis care in Slovakia, participants suggested feasible strategies to support its integration. These included legislative backing, cultural shifts within dialysis centers, and the availability of complementary services. Our evidence on health care providers' expectations and proposed opportunities aligns with existing literature, which emphasizes the importance of health care providers' competencies, the presence of support teams and services, and the need for legislative and leadership support within dialysis networks [107,109-114]. Enhancing the availability of professional competencies and providing institutional and legislative support may serve as foundational drivers for cultural change in dialysis care. Such improvements could help meet the needs and expectations of health care providers and enable broader implementation of IEG in clinical practice. A recent study highlights that implementation of innovative exercise interventions, such as exergaming, depends on addressing multilevel factors, such as patient needs, health care professional engagement, and supportive policy frameworks [105].

Underestimating external factors, particularly potential threats, during the planning of behavioral interventions can compromise their usage in clinical practice. Our participants expressed uncertainty and skepticism regarding the financial feasibility of IEG, emphasizing the need for robust evidence of its health benefits and cost-effectiveness. Patient behavior was also identified as a potential threat, with participants highlighting the importance of a personalized and sensitive approach to introducing exergaming, given the psychological, social, and environmental challenges that may influence initial engagement and long-term adherence. Additionally, some health care providers expressed resistance to the concept of IEG, favoring traditional care methods and raising concerns about technological burden, workflow disruptions in dialysis settings, and potential patient discomfort. Existing literature reflects similar concerns, with reports of uncertainty and skepticism about the availability of financial resources for exergaming [89,92,115,116], although other studies highlight its favorable cost-effectiveness and cost-efficiency [117-119]. However, cost analyses specific to IEG are currently lacking and are critically needed to inform planning and development. Patient behavior, particularly motivation, and social and emotional factors has been significantly associated with the effectiveness and adherence to exergaming interventions across diverse populations, and is recognized as both a key determinant and a potential threat [120-124]. A particularly salient concern raised by health care providers was the potential for gaming addiction, reflecting a critical awareness of unintended psychological consequences that may arise from IEG. The evidence highlighting resistant, reserved, and oppositional attitudes, as well as expectations surrounding IEG, represents a challenging yet essential insight for effective planning and development. Recommendations to overcome initial resistance include the provision of education and support from specialized professionals (eg, physiotherapists), and the active involvement of health care providers in the planning and management of the intervention [65,109,113,125]. To effectively mitigate potential risks, it is essential to conduct a thorough assessment of the resources required, prioritize individualization by addressing each patient's specific needs, and actively involve health care providers in the planning process prior to implementation of IEG.

Our study provides comprehensive evidence on Slovak health care providers' attitudes and expectations regarding the implementation of IEG in clinical practice, systematically organized into thematic quadrants of the SWOT framework. Future research exploring perceptions of IEG, including those of patients, exergame developers, and stakeholders (particularly patients' organizations, governmental and regulatory bodies), could enhance understanding of critical factors for strategic planning. Evidence on attitudes and expectations toward IEG

XSL•FO
RenderX

from respondents representing diverse health care systems, geographical regions, and socioeconomic and cultural backgrounds could strengthen the generalizability of these findings.

This study presents several key strengths. First, the sample included professionals from all levels of the organizational hierarchy within dialysis centers, from training assistants and nurses to nephrologists and senior management, providing a well-rounded and comprehensive perspective on health care providers' attitudes and expectations regarding IEG. Second, the use of prescreening enabled the recruitment of a highly diverse group of participants, varying in age, gender, and professional background. Third, all participants contributed relevant expertise through their direct involvement in implementing intradialytic exercise interventions in clinical practice. Fourth, we used an inductive approach to analyze health care providers' attitudes and expectations, followed by deductive categorization using the SWOT framework. We believe this methodological strategy yielded a practical, systematically organized, and transferable overview of the SWOT associated with IEG interventions, offering valuable insights to support strategic planning and informed decision-making.

This study has several limitations that should be acknowledged. First, despite our efforts to ensure a balanced representation across professional roles, we were unable to achieve the optimal sample size of 6 to 8 participants in the groups representing dialysis center managers and training assistants. However, the strong thematic alignment between managers, training assistants, and other professional groups suggests that their perspectives were largely reflected. Second, all participants were recruited from 2 dialysis centers affiliated with leading nephrology networks in Slovakia. As a result, the findings may be shaped by context-specific factors such as socioeconomic, demographic, and cultural characteristics, which could limit their generalizability to other dialysis settings with different health care structures. Third, our inclusion criteria required varying minimum levels of professional experience across roles. This decision was made to address the limited availability of experienced training assistants and to avoid potential bias from insufficiently experienced nurses, nephrologists, and managers. Nonetheless, this approach may have excluded early-career perspectives from some professional groups. Fourthly, our study is limited by the fact that participating health care providers had no direct experience with IEG. Therefore, the findings are based on hypothetical perceptions and expectations regarding its implementation in hemodialysis care and may not fully reflect experiences or challenges associated with its actual use in clinical settings.

Conclusions

This study offers valuable insights into health care providers' views on the implementation of IEG in clinical practice. They regarded IEG as safe and potentially beneficial for patients undergoing hemodialysis when supported by clinical monitoring and technological oversight. They highlighted the potential of IEG to enhance patient engagement and alleviate boredom during hemodialysis while noting weaknesses such as technological complexity and limitations on accessibility and users' experiences. Health care providers identified opportunities such as clarifying alternative exergame options, enhancing the availability and quality of information and support for IEG, and refining available IEG settings and modifications to facilitate implementation during hemodialysis therapy.

Although IEG is not yet standard practice in Slovakia, our findings indicate that, with appropriate planning, training, and policy adjustments, it could become a sustainable component of hemodialysis care. Health care providers acknowledged its potential benefits for engagement and well-being but raised concerns about financial feasibility, long-term sustainability, and availability of support for integration into the health care system. Common threats included uncertainty over funding and skepticism about health care system readiness, alongside worries about patients' privacy, clinical routines, staff workload, and resistance. In addition, our participants emphasized the need for a personalized approach to address patients' psychological variability and possible discomfort in shared spaces. The findings highlight the need to address both human and systemic factors in the design, development, and implementation of IEG, indicating that successful integration will require not only technological innovation but also cultural, organizational, and legislative support.

Health care providers regarded IEG as a safe and promising option for patients on hemodialysis, particularly for boosting engagement and reducing treatment-related boredom. To make this approach feasible in practice, clear guidelines, staff training, and organizational support are essential to overcome technological and accessibility challenges. With thoughtful planning and adequate resources, IEG could become an effective addition to hemodialysis care, improving patient experience and overall well-being.

Acknowledgments

We sincerely appreciate the valuable cooperation and support provided by Ivana Meszaros Skoumalova, PhD (IMS), and Denisa Kravcova (DK) during the data collection, management, and analysis phases of this study.

Funding

This work was supported by the Swedish Research Council for Healthcare, Working Life and Welfare under Grant 2023-01586, the Slovak Research and Development Agency under Grant APVV-16-0490 and the Internal Research Grant System of Pavol Jozef Šafárik University under Grant VVGS-2019-1069. The funders had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

XSL•FO
RenderX

Data Availability

The data that support the findings of this study are available from the first author, AZ, upon reasonable request. The data are not publicly available due to privacy and ethical restrictions, as they contain information that could potentially identify study participants.

Authors' Contributions

Conceptualization: AZ, MA, JR, ZDV, AMG, LK, TJ
Data curation: AZ
Formal analysis: AZ, MA, JR, ZDV, AMG, LK, TJ
Funding acquisition: AZ, JR, TJ
Investigation: AZ
Methodology: AZ, ZDV
Project administration: AZ, JR
Supervision: ZDV, AMG, LK, TJ
Validation: JR, ZDV
Visualization: AZ
Writing – original draft: AZ
Writing – review & editing: AZ, MA, JR, ZDV, AMG, LK, TJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide used in this study.
[DOCX File, 19 KB - rehab_v13i1e106283_app1.docx]

Multimedia Appendix 2
Analytical flow table for mapping and categorizing inductive thematic analysis findings within the Strengths, Weaknesses, Opportunities, and Threat (SWOT) framework.
[XLSX File, 15 KB - rehab_v13i1e106283_app2.xlsx]

Checklist 1
Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist
[DOC File, 64 KB - rehab_v13i1e106283_app3.doc]

References

1. Stevens PE, Ahmed SB, Carrero JJ, et al. KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int 2024 Apr;105(4):S117-S314. [doi: 10.1016/j.kint.2023.10.018]
2. Ding D, Lawson KD, Kolbe-Alexander TL, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. The Lancet 2016 Sep;388(10051):1311-1324. [doi: 10.1016/S0140-6736(16)30383-X]
3. Lavie CJ, Ozemek C, Carbone S, Katzmarzyk PT, Blair SN. Sedentary behavior, exercise, and cardiovascular health. Circ Res 2019 Mar;124(5):799-815. [doi: 10.1161/CIRCRESAHA.118.312669] [Medline: 30817262]
4. Kuo CP, Tsai MT, Lee KH, et al. Dose-response effects of physical activity on all-cause mortality and major cardiorenal outcomes in chronic kidney disease. Eur J Prev Cardiol 2022 Mar 25;29(3):452-461. [doi: 10.1093/eurjpc/zwaa162] [Medline: 33704426]
5. Zaghi A, Holm H, Korduner J, et al. Physical inactivity is associated with post-discharge mortality and re-hospitalization risk among swedish heart failure patients-The HARVEST-malmö study. Front Cardiovasc Med 2022;9:843029. [doi: 10.3389/fcvm.2022.843029] [Medline: 35265689]
6. Newsom JT, Huguet N, McCarthy MJ, et al. Health behavior change following chronic illness in middle and later life. J Gerontol B Psychol Sci Soc Sci 2012 May;67(3):279-288. [doi: 10.1093/geronb/gbr103] [Medline: 21983040]
7. Williams K, Steptoe A, Wardle J. Is a cancer diagnosis a trigger for health behaviour change? Findings from a prospective, population-based study. Br J Cancer 2013 Jun 11;108(11):2407-2412. [doi: 10.1038/bjc.2013.254] [Medline: 23695026]
8. Barile JP, Mitchell SA, Thompson WW, et al. Patterns of chronic conditions and their associations with behaviors and quality of life, 2010. Prev Chronic Dis 2015 Dec 17;12:E222. [doi: 10.5888/pcd12.150179] [Medline: 26679491]
9. Barker J, Smith Byrne K, Doherty A, et al. Physical activity of UK adults with chronic disease: cross-sectional analysis of accelerometer-measured physical activity in 96 706 UK biobank participants. Int J Epidemiol 2019 Aug 1;48(4):1167-1174. [doi: 10.1093/ije/dyy294] [Medline: 30721947]

XSL•FO
RenderX

10. Zhang F, Yin X, Huang L, Zhang H. The “adult inactivity triad” in patients with chronic kidney disease: a review. Front Med (Lausanne) 2023;10:1160450. [doi: 10.3389/fmed.2023.1160450] [Medline: 37064039]
11. Kistler BM, Kirkman DL, Kusni D, et al. Physical activity and exercise for cardiometabolic health and fitness in CKD: an overview by exercise type. Clin J Am Soc Nephrol 2026 Jan 1;21(1):144-153. [doi: 10.2215/CJN.0000000784] [Medline: 40504623]
12. Verheijden Klompstra L, Jaarsma T, Strömberg A. Exergaming in older adults: a scoping review and implementation potential for patients with heart failure. Eur J Cardiovasc Nurs 2014 Oct;13(5):388-398. [doi: 10.1177/1474515113512203] [Medline: 24198306]
13. Höchsmann C, Schüpbach M, Schmidt-Trucksäss A. Effects of exergaming on physical activity in overweight individuals. Sports Med 2016 Jun;46(6):845-860. [doi: 10.1007/s40279-015-0455-z] [Medline: 26712512]
14. Mat Rosly M, Mat Rosly H, Davis Oam GM, Husain R, Hasnan N. Exergaming for individuals with neurological disability: a systematic review. Disabil Rehabil 2017 Apr;39(8):727-735. [doi: 10.3109/09638288.2016.1161086] [Medline: 27108475]
15. Viana RB, de Oliveira VN, Dankel SJ, et al. The effects of exergames on muscle strength: a systematic review and meta-analysis. Scand J Med Sci Sports 2021 Aug;31(8):1592-1611. [doi: 10.1111/sms.13964] [Medline: 33797115]
16. Rytterström P, Strömberg A, Jaarsma T, Klompstra L. Exergaming to increase physical activity in older adults: feasibility and practical implications. Curr Heart Fail Rep 2024 Aug;21(4):439-459. [doi: 10.1007/s11897-024-00675-9] [Medline: 39023808]
17. Selles WL, Santos EC, Romero BD, Lunardi AC. Effectiveness of gamified exercise programs on the level of physical activity in adults with chronic diseases: a systematic review. Disabil Rehabil 2024 Dec;46(26):6231-6239. [doi: 10.1080/09638288.2024.2323614] [Medline: 38444153]
18. Hsieh RL, Chen YR, Lee WC. Short-term effects of exergaming on patients with chronic low back pain: a single-blind randomized controlled trial. Musculoskelet Sci Pract 2025 Feb;75:103248. [doi: 10.1016/j.msksp.2024.103248] [Medline: 39746279]
19. Oh Y, Yang S. Defining exergames & exergaming. 2010 Presented at: Meaningful Play 2010; Oct 21-23, 2010 URL: https://meaningfulplay.msu.edu/proceedings2010/mp2010_paper_63.pdf [accessed 2026-10-19]
20. Klompstra L, Jaarsma T, Piepoli MF, et al. Objectively measured physical activity in patients with heart failure: a sub-analysis from the HF-Wii study. Eur J Cardiovasc Nurs 2022 Jun 30;21(5):499-508. [doi: 10.1093/eurjcn/zvab133] [Medline: 34993536]
21. Wu J, Chen Z, Zheng K, et al. Benefits of exergame training for female patients with fibromyalgia: a systematic review and meta-analysis of randomized controlled trials. Arch Phys Med Rehabil 2022 Jun;103(6):1192-1200. [doi: 10.1016/j.apmr.2021.10.022] [Medline: 35033538]
22. da Silva Alves R, de Carvalho JM, Borges JBC, Nogueira DA, Iunes DH, Carvalho LC. Effect of exergaming on quality of life, fatigue, and strength and endurance muscle in cancer patients: a randomized crossover trial. Games Health J 2023 Oct;12(5):358-365. [doi: 10.1089/g4h.2022.0161] [Medline: 37155685]
23. Da Silva KG, Nuvolini RA, Bacha JMR, et al. Comparison of the effects of an exergame-based program with conventional physiotherapy protocol based on core areas of the european guideline on postural control, functional mobility, and quality of life in patients with parkinson’s disease: randomized clinical trial. Games Health J 2023 Jun;12(3):228-241. [doi: 10.1089/g4h.2022.0039] [Medline: 36206023]
24. Gilani M, Torkaman G, Bahrami F, Bayat N. Virtual reality exergaming capability to change muscle strategy during the limits of stability test and reduce fear of falling in primary osteoporotic women. Games Health J 2023 Aug;12(4):310-322. [doi: 10.1089/g4h.2022.0172] [Medline: 36940295]
25. Jin X, Jin M, Zhang B, Niu M, Han Y, Qian J. The association of conventional therapy associated with somatosensory interactive game enhances the effects of early pulmonary rehabilitation for patients with acute exacerbation of chronic obstructive pulmonary disease: a randomized controlled trial. Games Health J 2025 Apr;14(2):127-135. [doi: 10.1089/g4h.2023.0095] [Medline: 39207252]
26. Kang X, Zhang Y, Sun C, et al. Effectiveness of virtual reality training in improving outcomes for dialysis patients: systematic review and meta-analysis. J Med Internet Res 2025 Jan 8;27:e58384. [doi: 10.2196/58384] [Medline: 39773859]
27. Klompstra L, Hägglund E, Jaarsma T, Kato NP, Strömberg A. Effects of exergaming and yoga on exercise capacity and physical and mental health in heart failure patients: a randomized sub-study. Eur J Cardiovasc Nurs 2025 Apr 11;24(3):389-398. [doi: 10.1093/eurjcn/zvae155] [Medline: 39743240]
28. Baker LA, March DS, Wilkinson TJ, et al. Clinical practice guideline exercise and lifestyle in chronic kidney disease. BMC Nephrol 2022 Feb 22;23(1):75. [doi: 10.1186/s12882-021-02618-1] [Medline: 35193515]
29. Battaglia Y, Baciga F, Bulighin F, et al. Physical activity and exercise in chronic kidney disease: consensus statements from the physical exercise working group of the Italian society of nephrology. J Nephrol 2024 Sep;37(7):1735-1765. [doi: 10.1007/s40620-024-02049-9] [Medline: 39269600]
30. Wilkinson TJ, Tarca B, Lightfoot CJ, et al. Prescribing physical activity and exercise for people with CKD: a practical guide by the global renal exercise network. Clin J Am Soc Nephrol 2025 Mar 14;20(6):876-888. [doi: 10.2215/CJN.0000000708] [Medline: 40085159]

XSL•FO
RenderX

31. Shimoda T, Matsuzawa R, Yoneki K, et al. Changes in physical activity and risk of all-cause mortality in patients on maintence hemodialysis: a retrospective cohort study. BMC Nephrol 2017 May 8;18(1):154. [doi: 10.1186/s12882-017-0569-7] [Medline: 28482880]
32. Matsuzawa R, Roshanravan B, Shimoda T, et al. Physical activity dose for hemodialysis patients: where to begin? Results from a prospective cohort study. J Ren Nutr 2018 Jan;28(1):45-53. [doi: 10.1053/j.jrn.2017.07.004] [Medline: 28893466]
33. Bernier-Jean A, Wong G, Saglimbene V, et al. Self-reported physical activity and survival in adults treated with hemodialysis: a DIET-HD cohort study. Kidney Int Rep 2021 Dec;6(12):3014-3025. [doi: 10.1016/j.ekir.2021.09.002] [Medline: 34901570]
34. Suzuki Y, Matsuzawa R, Hoshi K, et al. Physical activity and its trajectory over time and clinical outcomes in hemodialysis patients. Int Urol Nephrol 2022 Aug;54(8):2065-2074. [doi: 10.1007/s11255-021-03098-w] [Medline: 35038083]
35. Zhang F, Wang H, Wang W, Zhang H. The role of physical activity and mortality in hemodialysis patients: a review. Front Public Health 2022;10:818921. [doi: 10.3389/fpubh.2022.818921] [Medline: 35252096]
36. Hull KL, Abell L, Adenwalla SF, et al. Impact of physical activity on surrogate markers of cardiovascular disease in the haemodialysis population. Clin Kidney J 2024 Jul;17(7):sfae198. [doi: 10.1093/ckj/sfae198] [Medline: 39050864]
37. Souweine JS, Pasquier G, Kuster N, et al. Dynapaenia and sarcopaenia in chronic haemodialysis patients: do muscle weakness and atrophy similarly influence poor outcome? Nephrol Dial Transplant 2021 Sep 27;36(10):1908-1918. [doi: 10.1093/ndt/gfaa353] [Medline: 33306128]
38. Workeneh BT, Kalantar-Zadeh K, Moore LW. Progress in the identification and management of protein-energy wasting and sarcopenia in chronic kidney disease. J Ren Nutr 2021 Jul;31(4):335-339. [doi: 10.1053/j.jrn.2021.04.004] [Medline: 34053819]
39. Yoshikoshi S, Yamamoto S, Suzuki Y, et al. Associations between dynapenia, cardiovascular hospitalizations, and all-cause mortality among patients on haemodialysis. J Cachexia Sarcopenia Muscle 2022 Oct;13(5):2417-2425. [doi: 10.1002/jcsm.13039] [Medline: 35916353]
40. Piccoli GB, Cederholm T, Avesani CM, et al. Nutritional status and the risk of malnutrition in older adults with chronic kidney disease - implications for low protein intake and nutritional care: a critical review endorsed by ERN-ERA and ESPEN. Clin Nutr 2023 Apr;42(4):443-457. [doi: 10.1016/j.clnu.2023.01.018] [Medline: 36857954]
41. Slee A, Reid J. Exercise and nutrition interventions for renal cachexia. Curr Opin Clin Nutr Metab Care 2024 May 1;27(3):219-225. [doi: 10.1097/MCO.0000000000001022] [Medline: 38386361]
42. Martínez-Olmos FJ, Gómez-Conesa AA, García-Testal A, et al. An intradialytic non-immersive virtual reality exercise programme: a crossover randomized controlled trial. Nephrol Dial Transplant 2022 Jun 23;37(7):1366-1374. [doi: 10.1093/ndt/gfab213] [Medline: 34245292]
43. Cavusoglu I, Tarakci D, Zenginler Yazgan Y, Yavuz S. Comparison of effectiveness of Nintendo Wii-based exergaming and home-based fun video exercises in pediatric patients with chronic kidney disease. Int J Rehabil Res 2023 Mar 1;46(1):26-34. [doi: 10.1097/MRR.0000000000000554] [Medline: 36416089]
44. Gurz D, Coimbatore Dada K, Naga Nyshita V, et al. The impact of Virtual Reality (VR) gaming and casual/social gaming on the quality of life, depression, and dialysis tolerance in patients with chronic kidney disease: a narrative review. Cureus 2023 Sep;15(9):e44904. [doi: 10.7759/cureus.44904] [Medline: 37814762]
45. Martins AR, Ferreira MC, Fernandes CS. Emerging technologies for supporting patients during hemodialysis: a scoping review. Int J Med Inform 2025 Jan;193:105664. [doi: 10.1016/j.ijmedinf.2024.105664] [Medline: 39481176]
46. Clarkson P, Stephenson A, Grimmett C, et al. Digital tools to support the maintenance of physical activity in people with long-term conditions: a scoping review. Digit Health 2022;8:20552076221089778. [doi: 10.1177/20552076221089778] [Medline: 35433017]
47. Pong C, Tseng R, Tham YC, Lum E. Current implementation of digital health in chronic disease management: scoping review. J Med Internet Res 2024 Dec 12;26(1):e53576. [doi: 10.2196/53576] [Medline: 39666972]
48. Wang W, Grundy J, Khalajzadeh H, Madugalla A, Obie HO. Designing adaptive user interfaces for mHealth application targeting chronic disease: a user-centered approach. ACM Trans Softw Eng Methodol 2026 Feb 28;35(2):1-57. [doi: 10.1145/3731750]
49. Zheng X, Yang Z, Liu S, Li Y, Wang A. Digital symptom management interventions for people with chronic kidney disease: a scoping review based on the UK medical research council framework. BMC Public Health 2024 Dec 19;24(1):3534. [doi: 10.1186/s12889-024-20871-5] [Medline: 39702218]
50. Bennett PN, Kohzuki M, Bohm C, et al. Global policy barriers and enablers to exercise and physical activity in kidney care. J Ren Nutr 2022 Jul;32(4):441-449. [doi: 10.1053/j.jrn.2021.06.007] [Medline: 34393071]
51. Ribeiro HS, Andrade FP, Leal DV, Oliveira JS, Wilund KR, Viana JL. How is exercise being prescribed for patients on hemodialysis? A scoping review. J Nephrol 2023 Jun;36(5):1307-1319. [doi: 10.1007/s40620-022-01513-8] [Medline: 36418777]
52. Li Y, Muñoz J, Mehrabi S, Middleton L, Cao S, Boger J. Multidisciplinary iterative design of exergames (MIDE): a framework for supporting the design, development, and evaluation of exergames for health. In: Lecture Notes in Computer Science: Springer; 2020:128-147. [doi: 10.1007/978-3-030-50164-8_9]
53. Bond S, Laddu DR, Ozemek C, Lavie CJ, Arena R. Exergaming and virtual reality for health: implications for cardiac rehabilitation. Curr Probl Cardiol 2021 Mar;46(3):100472. [doi: 10.1016/j.cpcardiol.2019.100472] [Medline: 31606141]

XSL•FO
RenderX

54. Chu CH, Biss RK, Cooper L, Quan AML, Matulis H. Exergaming platform for older adults residing in long-term care homes: user-centered design, development, and usability study. JMIR Serious Games 2021 Mar 9;9(1):e22370. [doi: 10.2196/22370] [Medline: 33687337]
55. Klompstra L, Jaarsma T, Olsson M, et al. Health care professionals expectations on mobile exergaming to decrease sedentary time in patients with heart failure. Eur J Cardiovasc Nurs 2024 Jul 17;23(Supplement_1). [doi: 10.1093/eurjcn/zvae098.121]
56. Klompstra L, Stromberg A, Jaarsma T, Rytterstrom P. Feasibility exergaming as a physical activity for older adults: a systematic literature review. Eur J Prev Cardiol 2024 Jun 13;31(Supplement_1). [doi: 10.1093/eurjpc/zwae175.048]
57. Hu H, Wu C, Chau PH, Choi EPH. Availability of exercise program, perceived exercise benefits and barriers, and exercise habits in maintenance hemodialysis patients: a multicenter cross-sectional study. J Ren Nutr 2025 Sep;35(5):568-577. [doi: 10.1053/j.jrn.2025.03.009] [Medline: 40280376]
58. Johnson G, Scholes K, Whittington R. Exploring Corporate Strategy: Prentice Hall; 2008:119-120 URL: https://www.academia.edu/35139191/EXPLORING_CORPORATE_STRATEGY [accessed 2026-10-19]
59. van Wijngaarden JDH, Scholten GRM, van Wijk KP. Strategic analysis for health care organizations: the suitability of the SWOT-analysis. Int J Health Plann Manage 2012;27(1):34-49. [doi: 10.1002/hpm.1032] [Medline: 20603842]
60. Hussein R. Medical informatics in the digital personalized health and medicine era: a SWOT analysis and actionable strategies. Stud Health Technol Inform 2020 Jun 16;270:869-873. [doi: 10.3233/SHTI200285] [Medline: 32570506]
61. Benzaghta MA, Elwalda A, Mousa M, Erkan I, Rahman M. SWOT analysis applications: an itegrative literature review. JGBI 2021;6(1):55-73. [doi: 10.5038/2640-6489.6.1.1148]
62. Khoshsirat S, Soleimanpour H, Rezaei-Hachesu P. Strengths, Weaknesses, Opportunities, and Threats (SWOT) of implementing teleconsultation: a systematic review. Health Sci Rep 2025 May;8(5):e70645. [doi: 10.1002/hsr2.70645] [Medline: 40330766]
63. Zelko A, Skoumalova I, Kolarcik P, et al. The effects of intradialytic resistance training on muscle strength, psychological well-being, clinical outcomes and circulatory micro-ribonucleic acid profiles in haemodialysis patients: protocol for a quasi-experimental study. Medicine (Baltimore) 2019 May;98(19):e15570. [doi: 10.1097/MD.0000000000015570] [Medline: 31083229]
64. Zelko A, Rosenberger J, Skoumalova I, et al. The effects of an intradialytic resistance training on lower extremity muscle functions. Disabil Rehabil 2022 Jan;44(2):275-281. [doi: 10.1080/09638288.2020.1766581] [Medline: 32448017]
65. Zelko A, Skoumalova I, Kravcova D, et al. Perceptions of healthcare providers on benefits, risks and barriers regarding intradialytic exercise among haemodialysis patients. Int J Qual Stud Health Well-being 2024 Dec;19(1):2287597. [doi: 10.1080/17482631.2023.2287597] [Medline: 38055756]
66. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
67. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
68. Levesque JF, Harris MF, Russell G. Patient-centred access to health care: conceptualising access at the interface of health systems and populations. Int J Equity Health 2013 Mar 11;12(1):18. [doi: 10.1186/1475-9276-12-18] [Medline: 23496984]
69. Thiel A, Sudeck G, Gropper H, et al. The iReAct study - a biopsychosocial analysis of the individual response to physical activity. Contemp Clin Trials Commun 2020 Mar;17:100508. [doi: 10.1016/j.conctc.2019.100508] [Medline: 31890988]
70. Forbes M. Thematic analysis: a practical guide. Eval J Australas 2022 Jun;22(2):132-135. [doi: 10.1177/1035719X211058251]
71. Leigh D. SWOT analysis. In: Watkins R, Leigh D, editors. Handbook of Improving Performance in the Workplace: Selecting and Implementing Performance Interventions: John Wiley & Sons, Ltd; 2009:115-140. [doi: 10.1002/9780470587102]
72. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
73. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
74. Altorfer P, Adcock M, de Bruin ED, Graf F, Giannouli E. Feasibility of cognitive-motor exergames in geriatric inpatient rehabilitation: a pilot randomized controlled study. Front Aging Neurosci 2021;13:739948. [doi: 10.3389/fnagi.2021.739948] [Medline: 34912206]
75. Qiu Y, Luximon Y. Analysis of postures, perceived physical safety, and technology acceptance of immersive exergames among older adults. Appl Sci 2025;15(7):3711. [doi: 10.3390/app15073711]
76. Maynard LG, de Menezes DL, Lião NS, et al. Effects of exercise training combined with virtual reality in functionality and health-related quality of life of patients on hemodialysis. Games Health J 2019 Oct;8(5):339-348. [doi: 10.1089/g4h.2018.0066] [Medline: 31539293]
77. Segura-Ortí E, García-Testal A. Intradialytic virtual reality exercise: increasing physical activity through technology. Semin Dial 2019 Jul;32(4):331-335. [doi: 10.1111/sdi.12788] [Medline: 30916415]
78. Zhou H, Al-Ali F, Kang GE, et al. Application of wearables to facilitate virtually supervised intradialytic exercise for reducing depression symptoms. Sensors (Basel) 2020 Mar 12;20(6):1571. [doi: 10.3390/s20061571] [Medline: 32178231]

XSL•FO
RenderX

79. Omonaiye O, Smyth W, Nagle C. Impact of virtual reality interventions on haemodialysis patients: a scoping review. J Ren Care 2021 Sep;47(3):193-207. [doi: 10.1111/jorc.12362] [Medline: 33491276]
80. Finco MG, Najafi B, Zhou H, Hamad A, Ibrahim R, Al-Ali F. Game-based intradialytic non-weight-bearing exercise training on gait speed and balance in older adults with diabetes: a single-blind randomized controlled trial. Sci Rep 2023 Aug 30;13(1):14225. [doi: 10.1038/s41598-023-41290-3] [Medline: 37648695]
81. Klompstra L, Jaarsma T, Strömberg A. Exergaming to increase the exercise capacity and daily physical activity in heart failure patients: a pilot study. BMC Geriatr 2014 Nov 18;14(1):119. [doi: 10.1186/1471-2318-14-119] [Medline: 25407612]
82. Tough D, Robinson J, Gowling S, Raby P, Dixon J, Harrison SL. The feasibility, acceptability and outcomes of exergaming among individuals with cancer: a systematic review. BMC Cancer 2018 Nov 21;18(1):1151. [doi: 10.1186/s12885-018-5068-0] [Medline: 30463615]
83. Moan ME, Vonstad EK, Su X, Vereijken B, Solbjør M, Skjæret-Maroni N. Experiences of stroke survivors and clinicians with a fully immersive virtual reality treadmill exergame for stroke rehabilitation: a qualitative pilot study. Front Aging Neurosci 2021;13:735251. [doi: 10.3389/fnagi.2021.735251] [Medline: 34795576]
84. Pirovano M, Surer E, Mainetti R, Lanzi PL, Alberto Borghese N. Exergaming and rehabilitation: a methodology for the design of effective and safe therapeutic exergames. Entertain Comput 2016 May;14:55-65. [doi: 10.1016/j.entcom.2015.10.002]
85. Stamm O, Vorwerg S, Müller-Werdan U. Exergames in augmented reality for older adults with hypertension: a qualitative study exploring user requirements. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Social Media, Games and Assistive Environments: Springer International Publishing; 2019:232-244. [doi: 10.1007/978-3-030-22015-0_18]
86. Ning H, Jiang D, Du Y, et al. Older adults' experiences of implementing exergaming programs: a systematic review and qualitative meta-synthesis. Age Ageing 2022 Dec 5;51(12):afac251. [doi: 10.1093/ageing/afac251] [Medline: 36571772]
87. Seinsche J, de Bruin ED, Carpinella I, et al. Older adults' needs and requirements for a comprehensive exergame-based telerehabilitation system: a focus group study. Front Public Health 2022;10:1076149. [doi: 10.3389/fpubh.2022.1076149] [Medline: 36711352]
88. Chen X, Jiang D, Ning H, et al. The experience of and needs for exergames in older adults with mild cognitive impairment: qualitative interview study. JMIR Serious Games 2025 Jun 17;13(1):e53631. [doi: 10.2196/53631] [Medline: 40526915]
89. Moeinzadeh AM, Calder A, Petersen C, Hoermann S. Views of people with MS regarding VR-exergaming to improve physical function and cognition: a qualitative study. Disabil Rehabil 2025 Oct;47(20):5348-5357. [doi: 10.1080/09638288.2025.2461265] [Medline: 39976253]
90. Rüth M, Schmelzer M, Burtniak K, Kaspar K. Commercial exergames for rehabilitation of physical health and quality of life: a systematic review of randomized controlled trials with adults in unsupervised home environments. Front Psychol 2023;14:1155569. [doi: 10.3389/fpsyg.2023.1155569] [Medline: 37333591]
91. Nguyen AV, Ong YLA, Luo CX, et al. Virtual reality exergaming as adjunctive therapy in a sub-acute stroke rehabilitation setting: facilitators and barriers. Disabil Rehabil Assist Technol 2019 May;14(4):317-324. [doi: 10.1080/17483107.2018.1447608] [Medline: 29529895]
92. van Santen J, Dröes RM, Blanson Henkemans O, et al. Implementation of exergaming for people with dementia: facilitators, barriers, and recommendations. Aging Ment Health 2024 Feb;28(2):244-253. [doi: 10.1080/13607863.2023.2238259]
93. Graham TCN, King N, Coo H, Zabojnikova P, Gurd BJ, Samdup D. Design and evaluation of an exergaming system for children with autism spectrum disorder: the children's and families' perspective. Front Virtual Real 2022;3:817303. [doi: 10.3389/frvir.2022.817303]
94. Manser P, Herold F, de Bruin ED. Components of effective exergame-based training to improve cognitive functioning in middle-aged to older adults - a systematic review and meta-analysis. Ageing Res Rev 2024 Aug;99:102385. [doi: 10.1016/j.arr.2024.102385] [Medline: 38914262]
95. Manser P, de Bruin ED, Temprado JJ, Bherer L, Herold F. Beyond "just" fun: The role of exergames in advancing health promotion and disease prevention. Neurosci Biobehav Rev 2025 Sep;176:106260. [doi: 10.1016/j.neubiorev.2025.106260] [Medline: 40518089]
96. Li T, Lv A, Xu N, et al. Barriers and facilitators to exercise in haemodialysis patients: a systematic review of qualitative studies. J Adv Nurs 2021 Dec;77(12):4679-4692. [doi: 10.1111/jan.14960] [Medline: 34258784]
97. Weber MB, Ziolkowski S, Bootwala A, Bienvenida A, Anand S, Lobelo F. Perceptions of physical activity and technology enabled exercise interventions among people with advanced chronic kidney disease: a qualitative study. BMC Nephrol 2021 Nov 10;22(1):373. [doi: 10.1186/s12882-021-02591-9] [Medline: 34758729]
98. Graham-Brown MPM, Smith AC, Greenwood SA. Digital health interventions in chronic kidney disease: levelling the playing field? Clin Kidney J 2023 May;16(5):763-767. [doi: 10.1093/ckj/sfac259] [Medline: 37151422]
99. Hu S, Yang Q, Qi S, et al. Maintenance hemodialysis patients participate in a virtual reality rehabilitation training experience: a qualitative study. Disabil Rehabil 2025 Jun;47(12):3151-3156. [doi: 10.1080/09638288.2024.2417029] [Medline: 39440851]

XSL•FO
RenderX

100. Wiemeyer J, Deutsch J, Malone LA, et al. Recommendations for the optimal design of exergame interventions for persons with disabilities: challenges, best practices, and future research. Games Health J 2015 Feb;4(1):58-62. [doi: 10.1089/g4h.2014.0078] [Medline: 26181682]
101. Kojic T, Sirotina U, Moller S, Voigt-Antons JN. Influence of UI complexity and positioning on user experience during VR exergames. Presented at: 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX); Jun 5-7, 2019. [doi: 10.1109/QoMEX.2019.8743273]
102. Meulenberg CJW, de Bruin ED, Marusic U. A perspective on implementation of technology-driven exergames for adults as telerehabilitation services. Front Psychol 2022;13:840863. [doi: 10.3389/fpsyg.2022.840863] [Medline: 35369192]
103. Manser P, Adcock-Omlin M, de Bruin ED. Design considerations for an exergame-based training intervention for older adults with mild neurocognitive disorder: qualitative study including focus groups with experts and health care professionals and individual Semistructured In-depth Patient Interviews. JMIR Serious Games 2023 Jan 5;11(1):e37616. [doi: 10.2196/37616] [Medline: 36602851]
104. Ravani E, Ali F, Albuainain S, et al. Exploring the barriers and facilitators experienced by patients with heart failure when using popular exergaming platforms for self-management-a systematic review. Digit Health 2024;10:20552076241249823. [doi: 10.1177/20552076241249823] [Medline: 39070896]
105. Antoniadou M, Zelko A, Strömberg A, Jaarsma T, Klompstra L. Implementing exergames into healthcare for chronic conditions – insights from stakeholders: a qualitative study. Health Res Policy Sys 2025;23(1):144. [doi: 10.1186/s12961-025-01416-7]
106. Ringgenberg N, Mildner S, Hapig M, et al. ExerG: adapting an exergame training solution to the needs of older adults using focus group and expert interviews. J NeuroEngineering Rehabil 2022;19(1):89. [doi: 10.1186/s12984-022-01063-x]
107. van Kessel R, Roman-Urrestarazu A, Anderson M, et al. Mapping factors that affect the uptake of digital therapeutics within health systems: scoping review. J Med Internet Res 2023 Jul 25;25(1):e48000. [doi: 10.2196/48000] [Medline: 37490322]
108. Carrera A, Lettieri E, Lietti G, Martignoni S, Sgarbossa C, Cafazzo J. Therapies go digital. What drives physicians' acceptance? PLoS ONE ;19(5):e0303302. [doi: 10.1371/journal.pone.0303302]
109. Thompson S, Tonelli M, Klarenbach S, Molzahn A. A qualitative study to explore patient and staff perceptions of intradialytic exercise. Clin J Am Soc Nephrol 2016 Jun 6;11(6):1024-1033. [doi: 10.2215/CJN.11981115] [Medline: 27026522]
110. Rye M, Friborg O, Skre I. Attitudes of mental health providers towards adoption of evidence-based interventions: relationship to workplace, staff roles and social and psychological factors at work. BMC Health Serv Res 2019 Feb 8;19(1):110. [doi: 10.1186/s12913-019-3933-4] [Medline: 30736786]
111. Cowie J, Nicoll A, Dimova ED, Campbell P, Duncan EA. The barriers and facilitators influencing the sustainability of hospital-based interventions: a systematic review. BMC Health Serv Res 2020 Jun 28;20(1):588. [doi: 10.1186/s12913-020-05434-9] [Medline: 32594912]
112. Taberna M, Gil Moncayo F, Jané-Salas E, et al. The Multidisciplinary Team (MDT) approach and quality of care. Front Oncol 2020;10:85. [doi: 10.3389/fonc.2020.00085] [Medline: 32266126]
113. Castillo G, Presseau J, Wilson M, et al. Addressing feasibility challenges to delivering intradialytic exercise interventions: a theory-informed qualitative study. Nephrol Dial Transplant 2022 Feb 25;37(3):558-574. [doi: 10.1093/ndt/gfab228] [Medline: 34415351]
114. Pereira VC, Silva SN, Carvalho VKS, Zanghelini F, Barreto JOM. Strategies for the implementation of clinical practice guidelines in public health: an overview of systematic reviews. Health Res Policy Sys 2022 Dec;20(1):13. [doi: 10.1186/s12961-022-00815-4]
115. Farič N, Yorke E, Varnes L, et al. Younger adolescents' perceptions of physical activity, exergaming, and virtual reality: qualitative intervention development study. JMIR Serious Games 2019 Jun 17;7(2):e11960. [doi: 10.2196/11960] [Medline: 31210135]
116. Choi C, Kim DK, Lee I. Virtual golf, "Exergaming", using virtual reality for healthcare in older adults: focusing on leisure constraints, participation benefits, and continuous participation intention. Health Care (Don Mills) 2024;12(10):962. [doi: 10.3390/healthcare12100962]
117. Stanmore EK, Mavroeidi A, de Jong LD, et al. The effectiveness and cost-effectiveness of strength and balance exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: a multi-centre, cluster randomised controlled trial. BMC Med 2019 Feb 28;17(1):49. [doi: 10.1186/s12916-019-1278-9] [Medline: 30813926]
118. Cieślik B, Mazurek J, Wrzeciono A, et al. Examining technology-assisted rehabilitation for older adults' functional mobility: a network meta-analysis on efficacy and acceptability. NPJ Digit Med 2023 Aug 24;6(1):159. [doi: 10.1038/s41746-023-00907-7] [Medline: 37620411]
119. Klompstra L, Mourad G, Jaarsma T, Strömberg A, Alwin J, On behalf of the HF-Wii study team. Costs of an Off-the-Shelf Exergame Intervention in patients with heart failure. Games Health J 2023 Jun;12(3):242-248. [doi: 10.1089/g4h.2022.0013] [Medline: 36473172]
120. Kaos MD, Rhodes RE, Hämäläinen P, Graham TCN. Social play in an exergame: how the need to belong predicts adherence. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2019. [doi: 10.1145/3290605.3300660]

XSL•FO
RenderX

121. van Beek JJW, van Wegen EEH, Bohlhalter S, Vanbellingen T. Exergaming-based dexterity training in persons with Parkinson disease: a pilot feasibility study. J Neurol Phys Ther 2019 Jul;43(3):168-174. [doi: 10.1097/NPT.0000000000000278] [Medline: 31136450]
122. Labib M, Bohm C, MacRae JM, et al. An international delphi survey on exercise priorities in CKD. Kidney Int Rep 2021 Mar;6(3):657-668. [doi: 10.1016/j.ekir.2020.12.001] [Medline: 33732980]
123. Gunst M, De Meyere I, Willems H, Schoenmakers B. Effect of exergaming on wellbeing of residents in a nursing home: a single blinded intervention study. Aging Clin Exp Res 2022 Jan;34(1):151-157. [doi: 10.1007/s40520-021-01903-1] [Medline: 34156650]
124. Ho LYW, Tse JHY, Chan WLS, Cheung DSK. Effects of exergaming on frailty: a systematic review and meta-analysis. J Clin Nurs 2025 May;34(5):1913-1930. [doi: 10.1111/jocn.17672] [Medline: 39871670]
125. Wodskou PM, Reinhardt SM, Andersen MB, Molsted S, Schou LH. Motivation, barriers, and suggestions for intradialytic exercise-a qualitative study among patients and nurses. Int J Environ Res Public Health 2021 Oct 6;18(19):10494. [doi: 10.3390/ijerph181910494] [Medline: 34639794]

Abbreviations

IEG: intradialytic exergaming
SWOT: Strengths, Weaknesses, Opportunities, and Threats

Edited by A Mastropietro; submitted 05.Jul.2026; peer-reviewed by F Niknam, I Widyawati; revised version received 19.Sep.2026; accepted 05.Oct.2026; published 04.Nov.2026.

Please cite as:
Zelko A, Antoniadou M, Rosenberger J, Dankulincova Veselska Z, Madarasova Geckova A, Klompstra L, Jaarsma T
Analysis of Health Care Providers' Attitudes and Expectations on Intradialytic Exergaming: Qualitative SWOT Study
JMIR Rehabil Assist Technol 2026;13:e106283
URL: https://rehab.jmir.org/2026/1/e106283
doi:10.2196/106283

© Aurel Zelko, Marianna Antoniadou, Jaroslav Rosenberger, Zuzana Dankulincova Veselska, Andrea Madarasova Geckova, Leonie Klompstra, Tiny Jaarsma. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 4.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Appropriateness and Impact of a Vocal Cord Vibration Switch for Children with Complex Communication Needs: Case Series

Leslie Mumford[1], MHSc; Denise Guerriere[2], PhD; Tom Chau[1,3], PhD

[1]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, 150 Kilgour Road, Bloorview Research Institute, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Tom Chau, PhD
Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, 150 Kilgour Road, Bloorview Research Institute, Toronto, ON, Canada

Abstract

Background: Communication is an essential component of participation. Communication impairment restricts full participation for children who have unintelligible speech. A vocal cord vibration switch offers an avenue for meaningful interaction to children who cannot rely on speech or voluntary limb movement but have some control of their vocal cords. Previous evaluations of the vocal cord vibration switch have been conducted primarily with adults and adolescents. However, implementation of a vocal cord vibration switch with younger, school-aged children in their natural environmental contexts can potentially foster the development of early communication skills.

Objective: This case series evaluated the appropriateness and impact of a vocal cord vibration switch, the "Hummingbird" (Holland Bloorview) with school-aged children who have complex communication needs and their mothers and teachers, using an individualized, collaborative, and iterative assistive device implementation protocol.

Methods: The Hummingbird was evaluated with 3 school-aged children, across educational and health-related contexts, over a 2-year period. Baseline, midterm, and final assessments took place at home or school in the first year with a follow-up assessment in the second year. In addition to field observations and device performance assessments by the research team, feedback from mothers and teachers was collected via questionnaires (Pictorial Children's Effort Rating Table, Quebec User Evaluation of Satisfaction with Assistive Technology, and the Family Impact of Assistive Technology Scale) to ascertain the Hummingbird's appropriateness and its impact.

Results: Appropriateness data indicated the suitability of the Hummingbird across settings. Compared to the child's prestudy devices, mother and teacher participants reported that the physical effort required by all 3 children to use the Hummingbird was lower (scores on the Pictorial Effort Rating Table decreased from 8 and 11 to scores of 4 - 6). The switch efficacy assessment of the Hummingbird indicated moderate-to-high specificity and high sensitivity at midterm and high sensitivity (0.91 - 0.94) and specificity (0.92 - 1) at final assessments. Total satisfaction scores increased from baseline (prestudy device) to the 2-year assessment for all 3 children. While data on the impact on family and communication were incomplete for 1 participant, generally favorable effects were reported. The field notes underscored the value of an individualized protocol, where the implementation and evaluation phases were adapted to accommodate the health-related characteristics (eg, seizure disorder and sleep deprivation), evolving school contextual factors (new school and teacher), and unique family environments (involving the child participant's toddler-sibling in Hummingbird sessions).

Conclusions: Overall, the Hummingbird was appropriate across home and school settings for our case study participants, all of whom had complex communication needs. The device was well-received by children and their mothers or teachers, providing an effective, setting-agnostic option for communication support. Modifications to both the device and its implementation process were required to address unanticipated health, family, and school challenges.

(JMIR Rehabil Assist Technol 2026;13:e75626) doi:10.2196/75626

KEYWORDS

complex communication needs; augmentative and alternative communication; vocal cord vibration switch; children with disabilities; access technology

XSL•FO
RenderX

Introduction

Communication impairment is associated with limited participation in children with complex disabilities [1,2]. Participation, defined as the engagement in all facets of life, has been identified as an essential outcome for children with disabilities [3,4]. Full and effective participation demands being on an equal basis with others and transcends home, school, health, and community environments. This engagement facilitates quality of life [5] and development [6], allowing children to realize their potential [4]. While communication is one of the core components of participation, communicative impairment has been associated with participation restrictions for children with unintelligible speech [1,3,4].

Augmentative and alternative communication (AAC) systems can enhance communication and, ultimately, participation for children with complex needs [7,8]. Electronic AAC systems produce speech via technology external to the child's body [9,10]. For children who have physical limitations that restrict direct access to electronic AACs via touch, an accompanying access technology is necessary. An access technology consists of a sensor that captures the child's intent from a physical movement, vocalization, or physiological indicator and a processing unit that translates the detected intent into a corresponding control signal [11]. A vocal cord vibration switch, one type of access technology, provides an avenue for meaningful interaction to children who cannot rely on speech or voluntary limb movement but have some control of their vocal cords. This switch consists of a sensor, positioned over the vocal cords, to detect periodic vibrations resulting from the child's vocalizations, which are translated into an output signal for computer control [12].

Previous evaluations of the vocal cord vibration switch have been conducted primarily with adolescents, adults, and children older than 8 years [12-15]. Intervening with AAC as early as possible allows for the opportunity to foster children's communication skills and establish their participation potential earlier in their development [10]. School-aged children, particularly those with complex disabilities, rely heavily on their family members and teachers to facilitate social engagement. As parents and teachers play an integral part in these children's participation by defining the opportunities to participate [3], engaging them in the implementation is essential [16]. Furthermore, in previous studies, the vocal cord vibration switch was evaluated over trajectories of 4 months or less. Offering children and their communication partners more time to use the vocal cord switch with support may encourage device adoption [17,18] and enable an evaluation of their experiences later in their trajectory. Finally, these assessments focused on the function of the Hummingbird and did not address the environments in which it would be used.

Accordingly, the purpose of this longitudinal case series was to evaluate the appropriateness and impact of the implementation of a vocal cord vibration switch with school-aged children who have complex communication needs in their natural environments (home and school), based on a collaborative, iterative, child- and family-centered protocol. Mothers' and teachers' perceptions of effort exertion, satisfaction, engagement, and psychosocial impact were measured. Device appropriateness was also assessed by the research team, and observations of process and progress were recorded in field notes. Temporal changes were measured across contexts (home and school) over a 2-year assessment period to allow for ongoing support and assessment after the initial introduction of the vocal cord vibration switch.

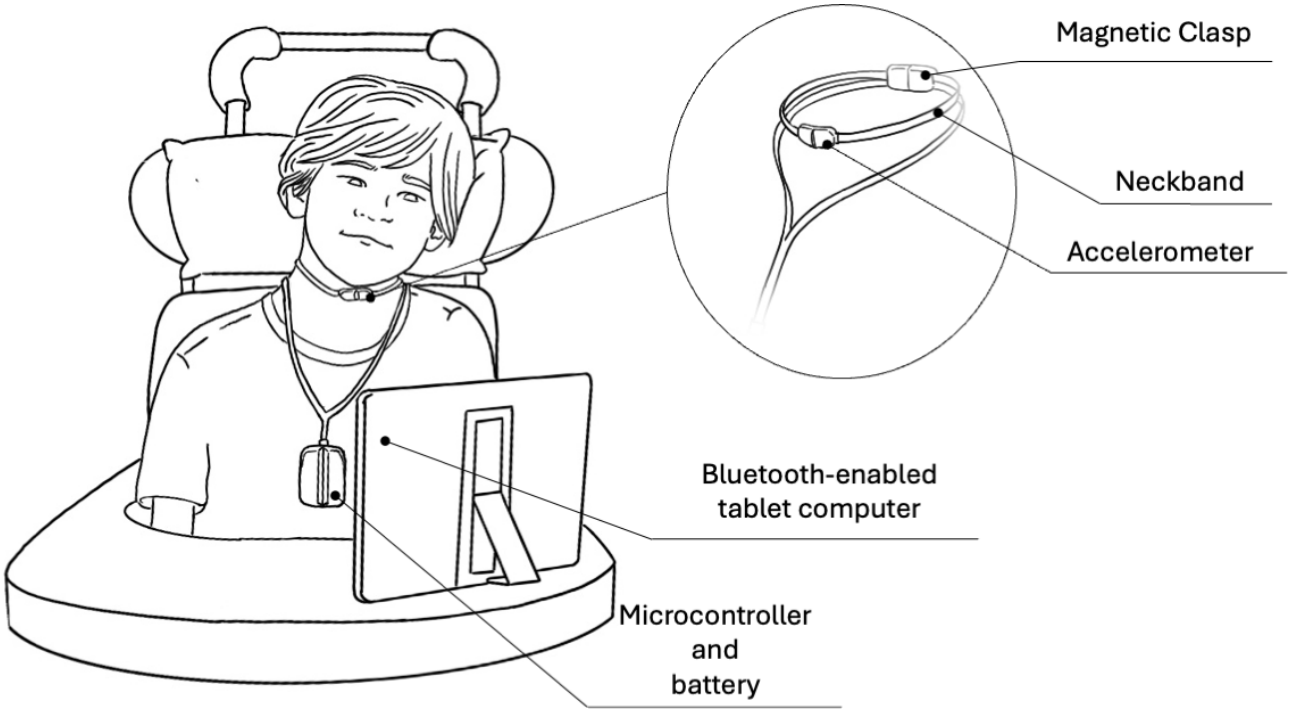
Methods

Overview

In this 2-year multiple case series, a vocal cord vibration switch, the "Hummingbird," was introduced to 3 school-aged children, guided by the Assistive Technology Delivery Protocol (ATDP) [19,20]. The ATDP informs the introduction and adoption of assistive technology for children who have severe disabilities and encourages device adoption by focusing on the compatibility of the technology with the needs of the users, while ensuring standardized training protocols, multidisciplinary inclusion, and an iterative assessment process. The Hummingbird consists of a dual-axis accelerometer, which is held in place on the child's neck proximal to the vocal cords using a neckband. The sensor is connected to a microcontroller box, where input vocal cord vibration signals are monitored and analyzed to detect short- or long-duration intended sounds (ie, voiced sounds or hums; Figure 1).

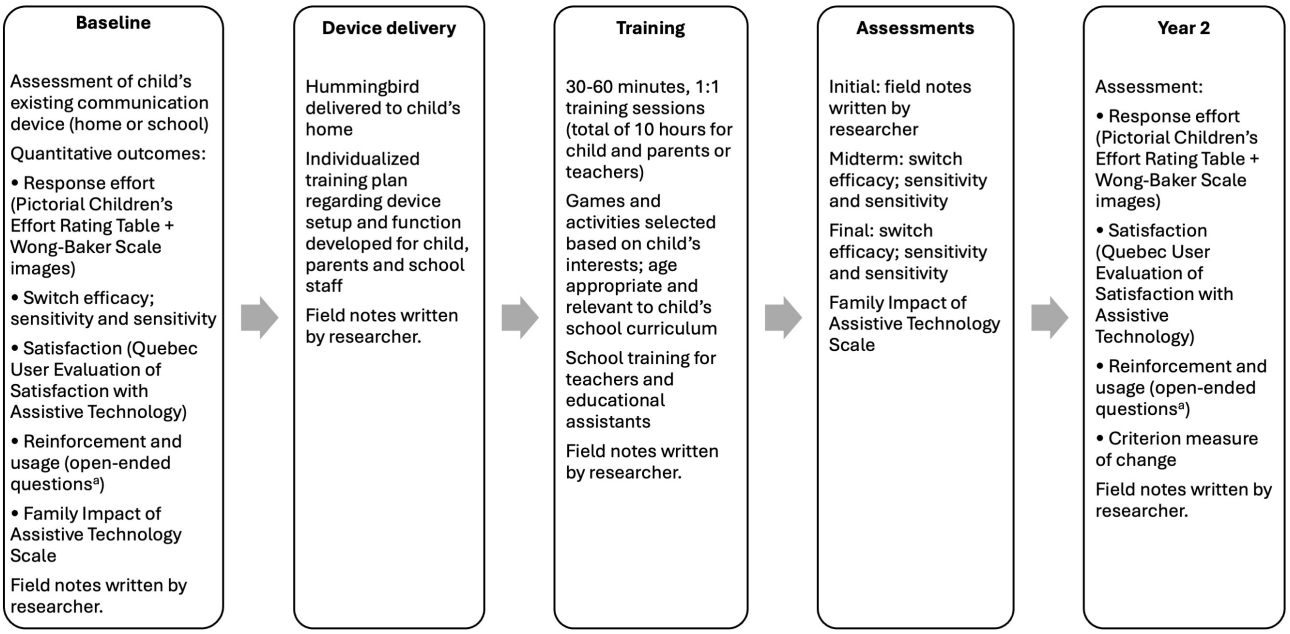
XSL•FO
RenderX

Figure 1. The Hummingbird device worn by a child, interacting with a computer through intentional vocal fold vibrations. The neckband is fastened by a magnetic clasp around the child's neck. A control signal is transmitted to the tablet computer whenever an intentional vocal fold vibration is detected.



For this implementation, assessments across multiple time points were conducted, including a follow-up assessment (year 2; Figure 2). In the baseline session, assessments of the child's existing communication device were conducted at home or the participant's school, according to the preference of the family. Following the baseline assessment, the Hummingbird was delivered to families at home, and an individualized training plan was developed by the research team to include the children's teachers and parents. The training plan was informed by the child's family, caregivers (personal support workers, nurses, or additional family members), and hospital-based occupational therapists or speech-language pathologists. Device training for the children occurred in 30 to 60 minute, one-on-one sessions for a total of 10 hours over a 10 to 20-week time frame and involved games and activities that were age-appropriate and of interest to the child. A school integration phase consisted of training for teachers and/or educational assistants. No researcher involvement occurred in year 2 except for a check-in to determine if the children were still using the Hummingbird and to complete a subset of the outcome measures.

Figure 2. Overview of implementation and assessment sessions. [a]Open-ended questions for reinforcement: On average, how many attempts does the child make to activate the Hummingbird before he or she is successful? Is there any delay between when the Hummingbird is activated and when the child perceives a response from the technology? Is there any delay or lack of response from the communication partner?



XSL•FO
RenderX

Ethical Considerations

The study was approved by the research ethics board of the Holland Bloorview Kids Rehabilitation Hospital (eREB0291) and each of the participating school boards. To be eligible for the study, a participant had to meet the following criteria: (1) aged at least 4 years, (2) did not have an access pathway or had an existing access pathway that was deemed inappropriate as per International Organization for Standardization (ISO) 9241‐9, (3) understand cause and effect, and (4) have voluntary control over their vocal cords in a manner that generated a measurable vibration.

Additionally, one of the parents (or guardians) of the participant as well as his or her teacher had to possess English competency at no less than a grade 6 level and had to commit to the training workshops.

Participants were recruited by self-referral via the hospital website or through a therapist or teacher. A participant was deemed competent to provide consent if the participant had a reliable access pathway associated with a communication interface (AAC device, iPad with AAC application, or computer with AAC software) that allowed him or her to express his or her complete understanding or could use partner-assisted scanning to express his or her full understanding.

If the potential participant was deemed competent to provide consent, at the initial meeting, the principal investigator or co-investigator explained the study to the potential participant and his or her legal guardian. The potential participant was subsequently asked to explain to the researcher the purpose of the study and his or her involvement in his or her own words using his or her current mode of communication. If the potential participant appeared to understand and was able to provide responses with reasonable accuracy for all questions, he or she was asked to sign consent forms for participation in the study. If the potential participant could not provide answers that were accurate with his or her current mode of communication, he or she was asked to provide assent, and his or her legal guardian was asked to provide written consent. Assent was obtained by both a researcher and the family in order to ensure that the participant understood that he or she could opt out at any time. The participant was also reminded on a weekly basis that he or she did not have to participate if he or she did not want to.

If the potential participant did not meet competency criteria, the consent process followed as described above except that the participant was asked to provide assent, and his or her legal guardian was asked to provide written consent. Where applicable, the participant's teacher and educational assistant or assistants provided written informed consent to participate in structured communication partner training. All participants were compensated for their time according to institutional research ethics board guidelines.

Appropriateness and Impact

Four concepts were measured to evaluate the appropriateness of the vocal cord switch: (1) response effort; (2) switch efficacy; (3) mother or teacher satisfaction; and (4) reinforcement and usage. Response effort refers to the amount of physical and cognitive effort required to use the Hummingbird [19]. It was measured using a modified Pictorial Children's Effort Rating Table (PCERT), consisting of effort descriptors, and a 6-level scale ranging from 0 (very easy) to 5 (very hard), with total scale scores ranging from 0 to 20 [21] and the Wong-Baker scale images [22]. The PCERT has yielded valid and reliable measures of effort with typically developing children [21,23]. In previous studies, scale scores have been moderately to highly correlated with objective measures of effort (Pearson r=0.54‐0.87), demonstrating concurrent validity of the scale at various levels of effort intensity [21], and the scale has demonstrated sufficient test-retest reliability (intraclass correlation coefficient=0.77) [23]. The PCERT was modified to use the images from the Wong-Baker pain scale, as cartoon faces ranging from happy to sad were considered more meaningful to children in wheelchairs than depictions of a child climbing stairs.

Switch efficacy represents a measure of effectiveness as defined by ISO 9241‐9; it was assessed using a switch-accessible game that required low cognitive effort and had unambiguous correct and incorrect responses for both the mother or teacher and the child [18]. The game also provided clear auditory and visual feedback to the child. Each child had 3 training sessions with the game to familiarize themselves with the task. Switch efficacy was based on researchers manually recorded data during the game and calculated using two equations: (1) specificity=true negatives/(true negatives+false positives); and (2) sensitivity=true positives/(true positives+false negatives). The specificity of a vocal cord vibration switch reflects the extent to which the device rejects the user's unintended vocal sounds (eg, a cough) while the sensitivity is the device's ability to correctly identify intended vocalizations. Based on previous evaluations of the efficacy of the Hummingbird with comparatively older children (aged >9 y), we anticipated that the 3 school-aged children in our series would achieve similar performances of specificity between 90% and 100% and sensitivity between 65% and 80% [13-15].

The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) was used to assess caregivers' satisfaction with assistive technology on a 5-point Likert scale ranging from 1 (not satisfied at all) to 5 (very satisfied). A total scale score and 2 subscale scores (Assistive Device and Services) are generated [24,25]. The QUEST 2.0 has been used extensively with clients of all ages, and it has been used to elicit parents' [26] and health professionals' [27] perceptions of their satisfaction with their child's assistive technology. Psychometric evaluations of the QUEST 2.0 have yielded acceptable levels of internal consistency (Cronbach α=0.74‐0.79) [28]. Furthermore, moderate correlations between subscales of the QUEST 2.0 and the Psychosocial Impact of Assistive Devices Scale [29] have been reported (Pearson r=0.34‐0.45), supporting the concurrent validity of the scale [30].

To evaluate the performance of the Hummingbird, the rate of reinforcement (attempts required to activate the Hummingbird) and the immediacy of reinforcement (delay from technology or communication partner) were assessed by asking mothers and teachers open-ended questions (Figure 2). In addition, technology usage was determined by asking them to estimate

XSL•FO
RenderX

the frequency of usage of the Hummingbird. Responses were recorded by the researcher.

The impact of the Hummingbird was assessed using two questionnaires. The Family Impact of Assistive Technology Scale (FIATS-AAC) detects the multidimensional effect of assistive device use on families with children with disabilities [31]. The purpose of this parent-reported questionnaire is to evaluate the functional effects of AAC systems on the lives of children and their families. It consists of 89 items, comprising 13 domains (7 child-related domains and 6 parent-related domains). Item scores are averaged to yield a domain score between 1 and 7. Higher scores suggest positive functional effects within a specific domain, and overall total scale scores range between 13 and 91. It has high internal consistency (Cronbach α=0.91) and test-retest reliability (intraclass correlation coefficient =0.95) [32]. Furthermore, its concurrent validity was supported as demonstrated by moderate agreement with a measure of family functioning (Pearson r=0.54) [33].

The Criterion Measure of Change Questionnaire [34] measures respondents' rating of the change in their child's communication since the commencement of Hummingbird with 3 ordinal categories (worse, no change, and better), as well as the extent and importance of this change on 7 levels ranging from a score of 1 ("a tiny bit") to 7 ("a very great deal").

Field Notes

In addition to the administration of the quantitative measures, field notes were written. These notes, which were constructed by the researchers after each interaction with the children, mothers, and teachers, were unstructured and included observations of: the performance of the technology and the child, study and child or mother or teacher progress, communication with the child's mothers, teaching staff and clinicians, and protocol implementation. The notes were compiled and analyzed using a manifest content analysis [35,36], which relies on the face value of the words written by the researcher and requires the use of preconceived categories for initiating and guiding the analysis. These categories included the following: the child's health and communication characteristics and needs, the benefits and disadvantages of the child's existing communication technology at study entry, the child's experiences with using the Hummingbird (child's likes and dislikes, successful strategies for encouraging use, and activities performed), challenges in using the Hummingbird specific to the child and mother or teacher, problem-solving, and device training and Hummingbird use at year 2. The field notes were grouped into units (sentences or phrases written by the researcher) within each of these categories, and then further organized by concepts (eg, successful strategies) within the category (eg, experiences with using the Hummingbird). This deductive approach allowed for a rich description and provided context to the child's use of the Hummingbird at their home and school, over time.

The purpose of the field notes was to provide context, that is, to identify aspects of the technology or the environment that can improve and support the child's use of the technology. They are useful beyond the study context in that they provide perspective on the aspects of the implementation and evaluation process that may need to be customized for individual children in future studies.

The ATDP [19,20] was used as a framework to design and guide the implementation and assessment process for this case series. The choice of questionnaires and the conceptual anchors for the field notes were guided by the ATDP, specifically the extent to which the Hummingbird meets the overall communication needs of the children. According to this protocol, successful Hummingbird adoption relies on the extent to which the technology is appropriate for each individual child. Appropriateness is defined as the efficiency, effectiveness, and satisfaction associated with the Hummingbird. Furthermore, the impact of the Hummingbird on the child and the family is essential to its success. The extent to which the Hummingbird is aligned with the characteristics of the children, mothers, and teachers is ultimately dependent on achieving a good match between the child and the Hummingbird within their given environment (school and home). The likelihood of this match is strengthened by the guidance of the ATDP.

Results

Herein, the field notes are presented among the results from the quantitative measures to provide context and detail. The clinical and demographic characteristics of the 3 child participants appear in Table 1.

Table . Demographic characteristics of child participants.

Child	Age[a] (y)	Sex	Diagnoses	School	Sensory function (vision; hearing)	Cognitive function
Eve	7	Female	Hydrocephalus	Special education classroom	Some impairment; no impairment	Some impairment
Jameel	5	Male	Cerebral palsy spastic quadriplegia	Home-schooled	No impairment; no impairment	No impairment
Zach	6	Male	Global developmental delay and epilepsy	School	Cortical visual impairment; no impairment	Significant impairment

[a]At study enrollment.

XSL•FO
RenderX

Field Notes: Participant Description and Rationale for Hummingbird Introduction

At the start of the study, Eve was aged 7 years and attended a congregated public school for children with physical disabilities. In school, she used button-type switches (BIGmack button device) [31] with her hands, requiring a lot of assistance, demonstrating little accuracy, and experiencing a lot of fatigue. Her fatigue increased as the day progressed. Eve could control her head movements but had difficulty maintaining a steady head position, which resulted in her head dropping forward or to the side. She would vocalize when excited, happy, or upset. Accompanied by facial expressions, these vocalizations were used by parents and teachers to attempt to interpret what she was trying to communicate. She was able to verbalize the words "mama," "papa," and "hi," and communicate with facial expressions and sounds that only familiar communication partners could understand.

Jameel enrolled in the study at the age of 5 years and was home-schooled. At that time, he had been attempting to use a Tobii eye-gaze system [32], but the high tone of his muscles made it difficult for him to hold his head in one position, even with physical head support. As such, he found it challenging to maintain a dwell long enough for the eye-gaze system to activate the desired choice on the screen. Jameel had good voluntary control of his voice. The Hummingbird was considered a potentially appropriate device for him, as it does not require head stability, and he possessed the vocal control for activation. It was hoped by his parents that he would be able to combine the eye-gaze system and the Hummingbird.

Zach was aged 6 years and attended a specialized school for children with disabilities at the start of the study. It was difficult for Zach to control his limbs and head, and he experienced frequent seizures. He had previously tried to use button-type switches (BIGmack button device) by controlling them with his hand and head, but he found them too tiring. He was able to indicate 'yes' using a simple non-word vocalization ('ah'). This vocalization had been used to elicit answers using partner-assisted scanning and a communication book. However, his family reported that the pictures in the communication book were challenging for him to discern due to his cortical visual impairment. Zach's family expressed a desire for an appropriate technology for him so that he could demonstrate his comprehension and answer questions in the classroom, as it was otherwise difficult for him to display his capacity for learning and understanding class material. Zach did not have the first-year final assessment for the FIATS because he was not using the Hummingbird at home, only at school during the first year. The Criterion of Change was also missing for this participant.

Field Notes: Observations During Sessions

Eve really enjoyed music and as such, it was used as the initial basis for her switch training. She began with errorless activities where she had to activate the Hummingbird to hear the next verse of a song and moved on to timed options where she would activate her Hummingbird to sing a repeating line from a song at the correct time. She often needed prompting to "wait" for the correct time to activate, but this need for prompting reduced over time. There were several interruptions to Eve's participation in the study as she changed schools and school boards multiple times throughout the study. After consultation with one teacher, choice-making during circle time activities was chosen as the focus for classroom training for Eve.

Jameel progressed very rapidly from simple cause and effect exercises to games that required precise timing. At the start of the study, he was working with his mother (who was also his home-school teacher) and an educational assistant, with a home visit approximately once per week from a member of the research team. One of the main challenges for Jameel was finding activities that motivated him; he could easily become bored with an activity that was presented in a repetitive fashion. This made evaluation of the switch efficacy quite challenging—after just 1 or 2 trials of a game, he sometimes refused to play more. Despite his excellent performance with the Hummingbird, the study progressed more slowly for Jameel due to personal and family illnesses, leading to a hiatus of approximately 4 months. Even after the break, however, Jameel quickly regained proficiency in the use of the Hummingbird. In addition, Jameel's family had been limiting their use of the Hummingbird to prevent Jameel's toddler sibling from pulling and/or breaking the device wires. Jameel started using a wireless version of the Hummingbird, which became available midway through the study, and his use of the Hummingbird increased after this change. With the new wireless Hummingbird, Jameel was playing games in addition to controlling switch-activated toys that also engaged his sibling.

During Hummingbird training, Zach was motivated by music, astronauts and space, and any activities involving his family members. He was also motivated by helping others, so the research team would reframe simple games as him "helping" the character move through the story or activity. Auditory cues were used as often as possible to reduce his need to rely on his vision; bright colors on a dark background were used as these were easiest for him to see. Zach had sleep difficulties, which often required him to sleep during the day. Occasionally, he would fall asleep during his sessions or struggle to stay awake. On the days when he was well-rested, the time required for him to coordinate a response to scanning choices was 2 to 3 seconds, compared to 15 seconds when fatigued. In addition, he had frequent absent seizures. At times, the research team was unsure if he was processing his answer or temporarily unaware of the task due to seizure activity. Zach would begin a session in an alert state but could then become fatigued partway through. On good days, a quick "warm-up" using a short errorless game, followed by the actual task, seemed to give shorter response times. When he was fatigued at the start, the researchers tried to preserve his energy as much as possible for the actual task.

Appropriateness and Impact: Response Effort (PCERT)

Table 2 presents effort levels, as reported by a mother or teacher, for the baseline assessment when the child was using their baseline or non-Hummingbird device and those for the year 2 assessment. The reported perceived effort required to use the Hummingbird by all three of the children was lower than the effort required for their previous device. The largest temporal

XSL•FO
RenderX

change was for the physical effort required by Jameel's mother, decreasing from the maximum effort of 5 to 2.

Table . Response effort scores[a].

	Eve		Jameel		Zach	
	Baseline[b]	Year 2	Baseline[c]	Year 2	Baseline[b]	Year 2
Respondent	Teacher	Teacher	Mother	Mother	Mother	Teacher
Item						
Physical effort for user (child)	3	1	3	1	4	2
Cognitive effort for user (child)	2	1	0	1	4	2
Physical effort for caregiver (mother or teacher)	2	2	5	2	0	1
Cognitive effort for caregiver (mother or teacher)	1	0	3	1	0	1
Total (0 - 20)	8	4	11	5	8	6

[a]By assessment time and item: 0 (very easy) to 5 (very hard), with total scale scores ranging from 0 to 20.

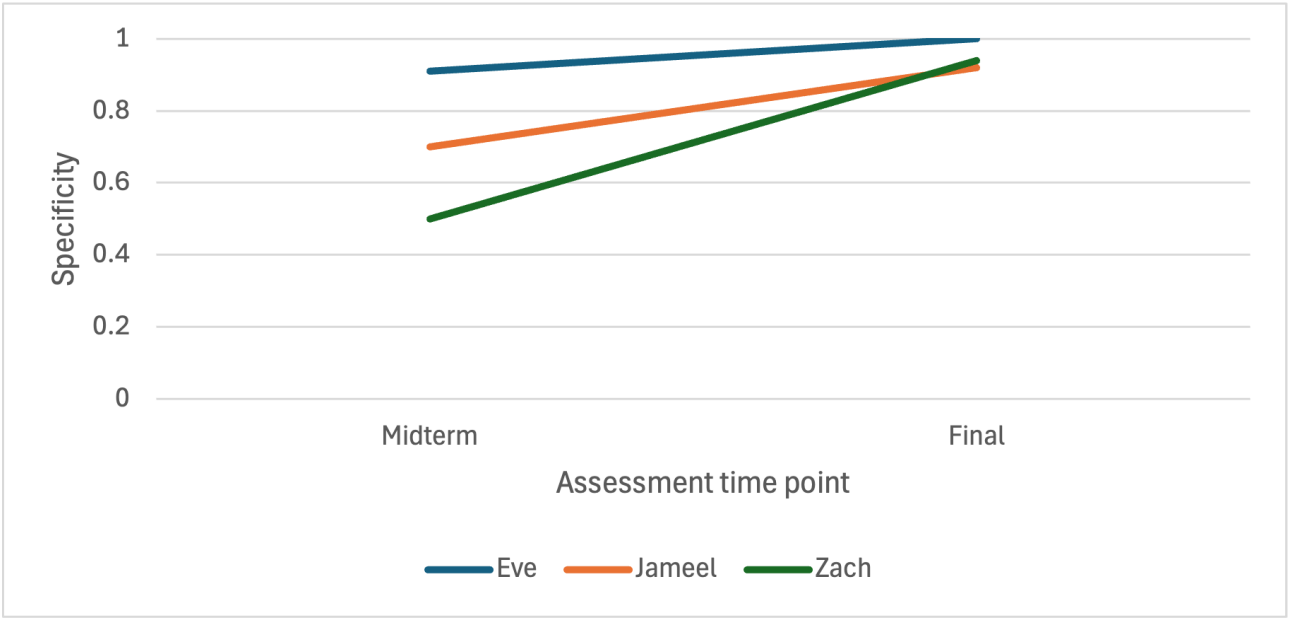
[b]BIGmack button device [37].

[c]Tobii eye-gaze [38].

Hummingbird-Switch Efficacy

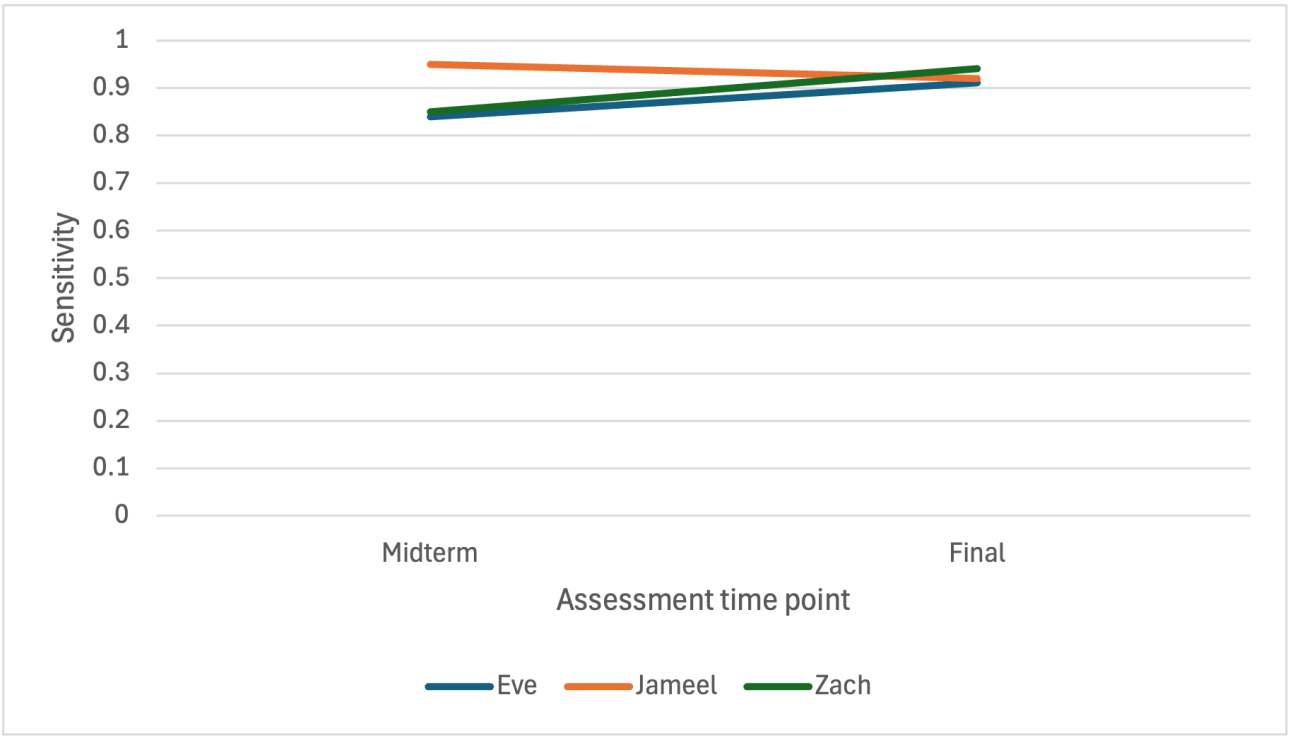
Figures 3 and 4 display switch efficacy at the midterm and final assessments for all 3 children. The Hummingbird demonstrated very high specificity (0.9 and 1) and sensitivity (0.8 and 0.9) for Eve at both the midterm and final assessments. For Jameel, specificity at the midterm was moderate (0.7) but increased in the final assessment (0.9). Finally, very high sensitivity at both midterm (0.9) and final (0.9) assessments was obtained for Zach, and a relatively low specificity at the midterm (0.5), which improved to a very high level for the final assessment (0.9).

Figure 3. Switch specificity at midterm and final assessments.



XSL•FO
RenderX

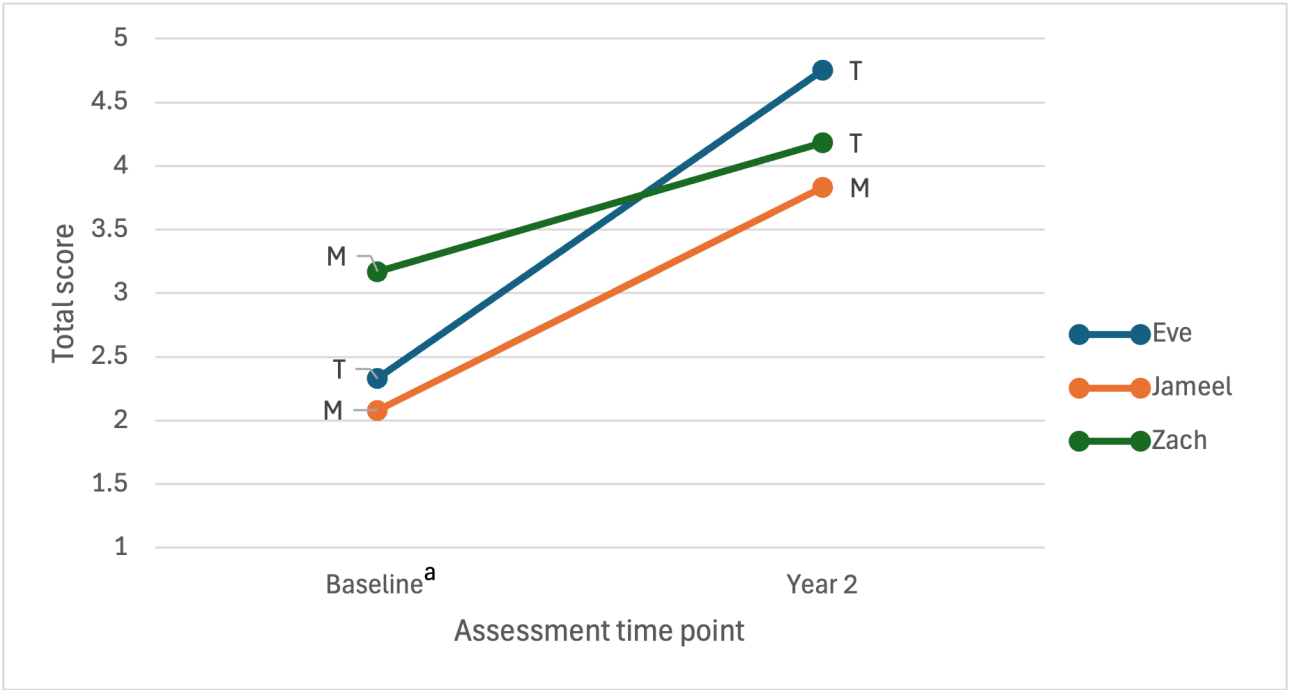
Figure 4. Switch sensitivity at midterm and final assessments.



Quebec User Evaluation of Satisfaction With Assistive Technology

Satisfaction ratings by teacher or mother participants increased from baseline (using a non-Hummingbird device) to year 2 for all 3 participants (Figure 5). All 4 respondents (2 teachers and 2 mothers) rated "easy to use" and "effectiveness" as being one of the top 3 most important items. "Professional service," "comfort," and "dimensions" were each reported by 2 respondents.

Figure 5. Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) total scores for baseline and year 2 (Hummingbird; 1="not satisfied" to 5="very satisfied"). [a]Eve and Zach using the BIGmack button device, and Jameel using the Tobii eye-gaze at baseline. T: teacher report; M: mother report.



Rate and Immediacy of Reinforcement

The immediacy of reinforcement with the Hummingbird was consistent across the 3 participants, with no delay from the technology during their year 2 evaluation (Table 3). The usage of the Hummingbird, as reported by mothers and teachers, was lower than that of her previous device for Eve. Due to inconsistency in reporting, comparative usage is difficult to

XSL•FO
RenderX

determine for Jameel and Zach. In Table 3, immediacy of response *from the technology* refers to whether there was a delay between activating the device and the intended output (ie, the Hummingbird responding via message output). On the other hand, immediacy of response *from the communication partner* refers to any delay between the time the child used the Hummingbird to communicate (eg, made a choice or requested something) and the time at which their communication partner acknowledged the message. In this latter context, no delay meant that the communication partner (eg, mother or teacher) acknowledged the message and desired action as soon as the child "says" it.

Table . Reinforcement and device usage at baseline and year 2 assessment.

	Eve		Jameel		Zach	
	Baseline (BIGmack)	Year 2 (Hummingbird)	Baseline (Tobii)	Year 2 (Hummingbird)	Baseline (BIGmack)	Year 2 (Hummingbird)
Respondent	• Teacher	• Teacher	• Mother	• Mother	• Mother	• Teacher
Usage	• 2 times per day for 30 minutes each • Weekly total: 420 minutes	• 3 days per week for 25 - 40 minutes each • Weekly range: 75 - 120 minutes	• 5 days per week for 10 - 30 minutes • Weekly range: 10 - 150 minutes	• Varied from: once daily for 5 days per week for months to not being used for months • Weekly rate: 5	• 2 times per day at home and school • Weekly rate: 14	• Minimum 1 time per day for a few minutes to share a short message • Weekly minimum: 35 minutes
Rate of reinforcement	• On some days, 2 attempts and other days, 3 - 4 attempts to hit switch—depends on mood+health	• Activates every time	• 1 attempt for playing games • Many attempts for communication	• With practice, 2 attempts	• Needs prompting or hand over hand to start, then uses without difficulty	• 1 attempt when device is working properly (ie, tuned to the user's vocal fold vibration), otherwise several attempts
Immediacy of reinforcement	• No delay	• No delay from either technology or communication partner	• Delay from technology when used for communication • No delay from communication partner	• No delay from technology	• Delay in activating—quicker for the child to vocalize than hit the button	• No delay

Impact

Table 4 presents scores for Eve and Jameel for baseline and final assessments, as reported by their mothers. An increase in scores over time indicates gains (positive functional effects) in dimensions from baseline to final for Jameel. The "face-to-face communication" domain increased from 2.4/7 at baseline to 5.1/7 for the final assessment. The "doing activities" domain also increased from 2.6 to 3.8. No temporal change in the overall score was observed for Eve. Only 2 domains increased for Eve over time, namely, scores from 3.6 to 4.3 for "contentment" and scores from 3.8 to 4.3 for "family roles," otherwise, her domain scores decreased (doing activities and security) or remained the same.

In year 2, the Criterion Measure of Change Questionnaire, which rates a child's communication on 3 ordinal categories (worse, no change, and better), indicated that both Eve's and Jameel's communication was rated "better" than when they entered the study (baseline). The extent of this improvement was quantified as 2/7 and 4/7, and the importance of this change as 3/7 and 7/7 for Eve and Jameel, respectively.

XSL•FO
RenderX

Table . Family Impact of Assistive Technology Scale mean domain scores for Jameel and Eve over time, as reported by their mothers.

Domain	Jameel[a]		Eve[b]	
	Baseline	Final	Baseline	Final
Child-related				
Behavior	4.5	4.7	6	6
Caregiver relief	2.2	2.8	2.8	2.9
Contentment	3.4	4.9	3.6	4.3
Doing activities	2.6	3.8	4.2	3.8
Education	4.7	6	6	5.7
Energy	1.7	3	3.6	3.6
Face-to-face communication	2.4	5.1	2.8	2.6
Family-related				
Family roles	2.6	3.3	3.8	4.3
Finances	3.6	3	3.4	3
Security	3	3.9	2.3	3.1
Self-reliance	2.9	4.9	3.6	3
Social versatility	3.4	4.6	5	4.7
Supervision	2.1	3	2	2.6
Total score (range 13 - 91)	39.1	53	49.1	49.6

[a]Home-schooled.

[b]Special education classroom.

Field Notes: Observations at Year 1 Completion

By the end of the training phase, both Eve and her teacher were comfortable with choice-making activities. Eve really enjoyed vocalizing to talk or sing along, so the Hummingbird switch was only vigilant (ie, turned on) when she was being asked to make a choice. At those times, she could calm herself and wait for the option she wanted. This selective use of the Hummingbird was adopted to avoid the frustration of excessive accidental activations when she wanted to expressively communicate through her vocalizations. Eve would become excited for sessions and wanted to vocalize a lot. If she was in an excited state, the researcher would give her time at the beginning of the session to just play, activating her switch as much as she wanted, before asking her to relax and calm herself to do the activity. The teacher also waited patiently for her to reach her calm state before presenting her with choices.

By the end of the first year, Jameel was using the Hummingbird with the eye-gaze system and appeared to find it easier than using the Hummingbird alone with a scanning-based communication grid on a tablet. His mother reported that it was easier for him to combine the 2 access methods rather than relying solely on switch scanning with his communication grid on his tablet. He was exploring spelling and sentence building using predictive text, with the assistance of his educational assistant. Depending on his level of motivation for an activity, he could often use his set-up to select from multiple choices independently. At this point, Jameel's toddler sibling had also spontaneously begun to model for him, attempting to get his attention by humming and pointing at the Hummingbird while saying, "My turn!"

Zach's sleep difficulties and seizures made some of the evaluations difficult. When the research team was unsure whether Zach was processing his answer or temporarily unaware of the task due to seizure activity, more frequent indirect prompting was used when the time to produce a response seemed especially prolonged. Indirect prompting included presenting Zach with a visual cue or picture, telling him that something was expected (eg, "Now what?" or "What's next?") and using body language (eg, questioning hand motion). The team learned to carefully observe signs of increased effort, as Zach could start a session alert but become fatigued partway through. During some of the switch efficacy testing sessions, he struggled with fatigue and falling asleep; his behavior during those sessions did not reflect the team's observations in other sessions.

Discussion

Principal Findings

This case series evaluated the appropriateness and impact of a vocal cord vibration switch with children aged 5 to 7 years across home and school settings, according to the ATDP. Despite some challenges related to the children's health and environment such as fatigue, school changes, and illness, the results from this evaluation lend support for the use of the Hummingbird for school-aged children with complex communication needs. Overall, the implementation of the

XSL•FO
RenderX

Hummingbird with 3 school-aged children was successful, and mothers and teachers responded positively to supporting its use at home and school.

Appropriateness and Impact

In our case series, we observed higher sensitivity rates at both midterm and year 2, and similar percentages of specificity compared to 2 previous evaluations of the Hummingbird that were conducted over a shorter evaluation period with children who were primarily older. In 1 study with 2 children aged 5 and 9 years, the average specificity of the Hummingbird was 98.1% and 98.8%, respectively, and the average sensitivity was 73.3% and 80%, respectively [15]. Furthermore, previously reported sensitivity percentages for a child aged 12 years using the Hummingbird were 60% and 65% at 8 and 16 weeks, respectively, and specificity percentages similar to ours were reported at >90% [20]. Switch efficacy levels with children aged 5 to 7 years in our study are similar to those observed with older users, thus supporting the use of the Hummingbird with a variety of age groups. For the youngest participant, Jameel, who was aged 5 years at study entry, the Hummingbird correctly identified his intended vocalizations at a higher or similar rate as older users. As with any efficacy evaluation, the results can be affected by dynamic factors specific to the child or user at the time of any given assessment; for example, fatigue was such a factor for 1 participant in our case series as well as for 1 child in a previous implementation [15].

We observed that mothers' and teachers' response effort scores on the PCERT decreased for all 3 children from baseline, indicating that the perceived physical effort exerted by the child was lower when using the Hummingbird compared to their prestudy communication device. It is noteworthy that this decrease in effort was observed during the year 2 assessment when regular support and contact with the research team had concluded. Ease of use of an AAC has been reported in previous studies as an important determinant of parent satisfaction [39] and device use [40]. Furthermore, literature indicates that parents report device abandonment when the effort to use the device is too demanding [17].

Measuring parents' and teachers' satisfaction is essential, given that dissatisfaction with an AAC system is 1 factor leading to device abandonment [17]. Our satisfaction results suggest that the mothers and teachers in our evaluation had a higher level of satisfaction with the Hummingbird compared to the child's previous device. Teachers' and mothers' satisfaction ratings increased from baseline to the year 2 assessment for all 3 children (scores ranging from 3.8-4.8). Similarly, in a previous case study, an individual aged 19 years also reported higher satisfaction with the Hummingbird (score=4.6) compared to a sound-based device [12]. Furthermore, in another assessment of the Hummingbird with a girl aged 12 years with cerebral palsy, similar QUEST 2.0 scores (4.0 and 3.5 at 8- and 16-wk postdevice introduction, respectively) were reported by her occupational therapist [20]. Our data thus support the applicability and usability of the Hummingbird with younger children than previously documented.

Device reliability is an important factor in preventing device abandonment [41]. In this case series, the quality of reinforcement evaluation indicated that, overall, the children were able to activate the Hummingbird more reliably, with fewer attempts and no delay compared to their previous prestudy device. The assessments during year 2 of the study revealed that all 3 children had continued to use the Hummingbird after the main study period.

Family impact scores on some of the FIATS domains for Eve and Jameel did not change from baseline, using their non-Hummingbird device, to the final assessment after months of using the Hummingbird. Overall, Jameel's total scale scores indicated a moderate increase in the positive functional effects with the Hummingbird (from 40/91 to 53/91) and no change in total functional effects for Eve. Perhaps because Jameel was home-schooled, his mother may have observed more of an effect on family functioning. In addition, Jameel's sibling started to play games with him while he was using the Hummingbird, which may have resulted in a greater overall impact on family functioning.

It is critical that AAC systems are designed for the individual to optimally facilitate communication and promote sustained user engagement [18]. In this study, personalization involved changes to the Hummingbird itself as well as to the process aspects of the protocol. The field notes provided insight into the characteristics of the children and their environments and permitted an iterative approach whereby adjustments to the training, implementation, and evaluation processes were made during the study. For example, the researcher (or researchers) explored different activities with each child; for example, Zach excelled when engaging in activities that involved helping people. A wireless version replaced Jameel's original Hummingbird to address the needs of his family environment. These customizations encouraged engagement in subsequent sessions with the three children as well as offering a guide to the implementation process in future work.

Previous research in the broader AAC literature has supported the importance of training communication partners; effective partner training promotes the communication abilities of individuals with complex communication needs [18,42]. Despite this important element, most studies have excluded AAC training for parents and siblings of children with multiple disabilities.

Value of Multisite Inclusion and Individualized Implementation Approach

Challenges are inherent in the implementation process of a device for children who have complex, co-existing conditions that compound their existing activity limitations. In our case series, the child participants differed from one another as they had a variety of health, social, family, and school context characteristics. The successful longitudinal implementation of the Hummingbird with these 3 school-aged children resulted from a comprehensive, individualized approach, including adjusting and modifying both the technology and the protocol, to meet the needs of the child and the child's environment. For example, changing to a wireless device for Jameel was essential to encourage continued use of the technology when his brother was present.

XSL•FO
RenderX

As purported by the ATDP, training and support by the team ensure a child-specific approach that recognizes and adapts to the technology user, including elements of their physical and social or family environment. Of priority to our implementation was the inclusion of key players from each of the child's environments. Furthermore, based on the ATDP, device training and implementation were individualized and embedded into the participants' daily lives and natural environments, while evaluation was semipersonalized, such as affording the child the choice of a game. Inclusion of the school setting in the implementation, attention to training, and quality of service delivery are indicators of users' and caregivers' satisfaction as well as device adoption [42-46].

Contextualization Within Existing Literature

The development and initial evaluation of the vocal cord vibration switch have been described previously [13-15]; these previous evaluations were conducted over shorter time periods (0.5 - 4.0 mo), primarily with adolescents, adults, and children older than 9 years and focused on the physical aspects of the Hummingbird, its reliability, and the ease of set-up for parents and teachers. The results from these earlier assessments informed a more formalized switch training process, the design of the neckband and its positioning, and the decision to use a questionnaire assessing family impact in the case series. In contrast, this case series measured the impact of long-term use in the child's natural environment.

Fundamentally, our findings indicate that the Hummingbird can afford a pathway through which a child can access switch-controllable technology, inclusive of communication, mobility, and environmental devices. From the perspective of the International Classification of Functioning, Disability and Health for Children and Youth framework [3], our results suggest that via the Hummingbird (environmental facilitator), children's vocal cords (body structures) can be accessed by amplifying vocal fold vibrations (body function) to expand independent, intentional switch control (activities) within their given setting (environmental participation conditions). Environmental conditions included barriers such as a child's fatigue during sessions and school changes, as well as motivators such as music or a sibling joining in on a game. In turn, such technology-facilitated switch activations can augment children's participation in communication, interpersonal interactions, education, and directing their own care (specific participation domains).

Limitations

The assessment time points did not always follow the intended trajectory. For reasons of illness and family health issues, which reflect the reality of many families with children with multiple disabilities, some of the assessments were not conducted at the intended times and/or were not completed. Furthermore, some of the questionnaires were not completed at certain time points, resulting in missing data for the FIATS and the Criterion of Change for Zach. However, given the case series design, the results were interpreted individually rather than by making quantitative comparisons across participants. Furthermore, in instances where 1 assessment was missed, data from other time points were available, permitting nonetheless a temporal view of children's and families' experiences. Future studies may consider a different measure of family impact catered more toward children who are developing basic communication skills, such as expressing preferences.

For some data collection questionnaires, 2 different respondents (ie, mother and teacher) provided the assessment for a given child participant, based on who was present in the assessment session and where it took place (home or school). For both the response effort and QUEST 2.0 measures, scores were provided by a child's mother at 1 session and a teacher at the subsequent session. This may have confounded the potential temporal changes in an outcome. Nonetheless, we arrived at a comprehensive assessment of the Hummingbird by deploying a variety of measures to capture the perspectives of multiple key informants in the training, use, and assessment phases of our protocol. Subsequent studies may consider mixed methods designs and associated integrated analyses that allow for informant-specific data sources [47].

Usage reporting was varied. Mothers' reports regarding the usage of both the child's prestudy device and the Hummingbird were unstructured (Table 3). As a result, the usage data varied in terms of how mothers reported use by units of time. To overcome this limitation, future research may consider invoking automatic logging of Hummingbird usage to objectively capture time and duration of use, the number of activations, and the functional application.

Conclusions

Across home and school environments, the Hummingbird was appropriate for 3 school-aged children with complex communication needs as indicated by their low response effort, high scores on switch efficacy and immediacy of reinforcement, and strong caregiver satisfaction. The impact was generally favorable in familial and communication realms for the 2 participants with complete data. Participants encountered unanticipated health, family, and school challenges during the study, which necessitated modifications to the device and the implementation protocol. Future studies should consider a more expansive set of impact measures in light of these real-world circumstances.

Acknowledgments

The authors would like to thank the families and children for generously participating in the protracted protocol of this study.

Funding

This research was supported by the Accessible Technology Program, Ministry of Innovation, Science & Economic Development Canada (Grant #511898). The study was funded for a 2-year period for CAD $311,111.

XSL•FO
RenderX

Authors' Contributions

TC and LM designed the protocol. TC secured funding, oversaw all aspects of the study, and provided detailed edits to the manuscript and its revisions. LM recruited participants, trained them, collected all the data, managed day-to-day study-related tasks, and revised various versions of the manuscript. DG wrote the manuscript and performed data analysis.

Conflicts of Interest

None declared.

References

1. Allison KM, Doherty KM, Cerebral Palsy Research Network. Relation of speech-language profile and communication modality to participation of children with cerebral palsy. Am J Speech Lang Pathol 2024 Mar 7;33(2):1040-1050. [doi: 10.1044/2023_AJSLP-23-00267] [Medline: 38215219]
2. Connaghan KP, Baylor C, Romanczyk M, Rickwood J, Bedell G. Communication and social interaction experiences of youths with congenital motor speech disorders. Am J Speech Lang Pathol 2022 Nov 16;31(6):2609-2627. [doi: 10.1044/2022_AJSLP-22-00034] [Medline: 36215658]
3. International Classification of Functioning, Disability and Health: Children & Youth version: ICF-CY. : World Health Organization; 2007 URL: https://iris.who.int/server/api/core/bitstreams/81e04699-2930-4da2-ac97-76725c7dcc07/content [accessed 2026-03-04]
4. van Schie PEM, Siebes RC, Dallmeijer AJ, et al. Development of social functioning and communication in school-aged (5-9 years) children with cerebral palsy. Res Dev Disabil 2013 Dec;34(12):4485-4494. [doi: 10.1016/j.ridd.2013.09.033] [Medline: 24139717]
5. Davis E, Shelly A, Waters E, et al. Quality of life of adolescents with cerebral palsy: perspectives of adolescents and parents. Dev Med Child Neurol 2009 Mar;51(3):193-199. [doi: 10.1111/j.1469-8749.2008.03194.x] [Medline: 19191833]
6. Chantry J, Dunford C. How do computer assistive technologies enhance participation in childhood occupations for children with multiple and complex disabilities? A review of the current literature. Br J Occup Ther 2010 Aug;73(8):351-365. [doi: 10.4276/030802210X12813483277107]
7. Crowe B, Machalicek W, Wei Q, Drew C, Ganz J. Augmentative and alternative communication for children with intellectual and developmental disability: a mega-review of the literature. J Dev Phys Disabil 2022;34(1):1-42. [doi: 10.1007/s10882-021-09790-0] [Medline: 33814873]
8. O'Neill T, Light J, Pope L. Effects of interventions that include aided augmentative and alternative communication input on the communication of individuals with complex communication needs: a meta-analysis. J Speech Lang Hear Res 2018 Jul 13;61(7):1743-1765. [doi: 10.1044/2018_JSLHR-L-17-0132] [Medline: 29931287]
9. Iacono T, Trembath D, Erickson S. The role of augmentative and alternative communication for children with autism: current status and future trends. Neuropsychiatr Dis Treat 2016;12(12):2349-2361. [doi: 10.2147/NDT.S95967] [Medline: 27703354]
10. Romski MA, Sevcik RA. Augmentative communication and early intervention: myths and realities. Infant Young Child 2005 Jul;18(3):174-185. [doi: 10.1097/00001163-200507000-00002]
11. Tai K, Blain S, Chau T. A review of emerging access technologies for individuals with severe motor impairments. Assist Technol 2008;20(4):204-219. [doi: 10.1080/10400435.2008.10131947] [Medline: 19160907]
12. Chan J, Falk TH, Teachman G, Morin-McKee J, Chau T. Evaluation of a non-invasive vocal cord vibration switch as an alternative access pathway for an individual with hypotonic cerebral palsy - a case study. Disabil Rehabil Assist Technol 2010 Jan;5(1):69-78. [doi: 10.3109/17483100903278107] [Medline: 19941442]
13. Lui M, Falk TH, Chau T. Development and evaluation of a dual-output vocal cord vibration switch for persons with multiple disabilities. Disabil Rehabil Assist Technol 2012;7(1):82-88. [doi: 10.3109/17483107.2011.557467] [Medline: 21345001]
14. Falk TH, Chan J, Duez P, Teachman G, Chau T. Augmentative communication based on realtime vocal cord vibration detection. IEEE Trans Neural Syst Rehabil Eng 2010 Apr;18(2):159-163. [doi: 10.1109/TNSRE.2009.2039593] [Medline: 20071275]
15. Lu EC, Falk TH, Teachman G, Chau T. Assessing the viability of a vocal cord vibration switch for four children with multiple disabilities~!2009-10-06~!2010-02-19~!2010-04-02~!. Open Rehabil J 2010 Apr 15;3(1):55-61. [doi: 10.2174/1874943701003010055]
16. Marshall J, Goldbart J. "Communication is everything I think." Parenting a child who needs augmentative and alternative communication (AAC). Int J Lang Commun Disord 2008;43(1):77-98. [doi: 10.1080/13682820701267444] [Medline: 17852533]
17. Moorcroft A, Scarinci N, Meyer C. "I've had a love-hate, I mean mostly hate relationship with these PODD books": Parent perceptions of how they and their child contributed to AAC rejection and abandonment. Disabil Rehabil Assist Technol 2021 Jan;16(1):72-82. [doi: 10.1080/17483107.2019.1632944] [Medline: 31250678]

XSL•FO
RenderX

18. Brittlebank S, Light JC, Pope L. A scoping review of AAC interventions for children and young adults with simultaneous visual and motor impairments: clinical and research implications. Augment Altern Commun 2024 Sep;40(3):219-237. [doi: 10.1080/07434618.2024.2327044] [Medline: 38578299]
19. Mumford L, Lam R, Wright V, Chau T. An access technology delivery protocol for children with severe and multiple disabilities: a case demonstration. Dev Neurorehabil 2014 Aug;17(4):232-242. [doi: 10.3109/17518423.2013.776125] [Medline: 23869969]
20. Mumford L, Chau T. Application of an access technology delivery protocol to two children with cerebral palsy. Disabil Rehabil Assist Technol 2016 Feb;11(2):166-175. [doi: 10.3109/17483107.2015.1063017] [Medline: 26171580]
21. Yelling M, Lamb KL, Swaine IL. Validity of a pictorial perceived exertion scale for effort estimation and effort production during stepping exercise in adolescent children. Eur Phy Educ Rev 2002 Jun;8(2):157-175. [doi: 10.1177/1356336X020082007]
22. Wong-baker FACES® pain rating scale. Wong-Baker FACES Foundation. URL: https://wongbakerfaces.org/ [accessed 2024-07-03]
23. Marinov B, Mandadjieva S, Kostianev S. Pictorial and verbal category-ratio scales for effort estimation in children. Child Care Health Dev 2008 Jan;34(1):35-43. [doi: 10.1111/j.1365-2214.2007.00767.x] [Medline: 18171442]
24. Demers L, Weiss-Lambrou R, Ska B. Item analysis of the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST). Assist Technol 2000;12(2):96-105. [doi: 10.1080/10400435.2000.10132015] [Medline: 11508406]
25. Demers L, Weiss-Lambrou R, Ska B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): an overview and recent progress. Technol Disabil 2002;14(3):101-105. [doi: 10.3233/TAD-2002-14304]
26. Field DA, Livingstone RW. Parents' and therapists' satisfaction with four early childhood power mobility devices. Can J Occup Ther 2022 Dec;89(4):364-375. [doi: 10.1177/00084174221098879] [Medline: 35656731]
27. Eriksson M, Jylli L, Villard L, Kroksmark AK, Bartonek Å. Health-related quality of life and orthosis use in a Swedish population with arthrogryposis. Prosthet Orthot Int 2018 Aug;42(4):402-409. [doi: 10.1177/0309364618774059] [Medline: 29775129]
28. Aledda S, Galeoto G, Fabbrini G, et al. A systematic review of the psychometric properties of Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST). Disabil Rehabil Assist Technol 2024 May;19(4):1228-1235. [doi: 10.1080/17483107.2022.2161648] [Medline: 36645802]
29. Jutai J, Day H. Psychosocial Impact of Assistive Devices Scale (PIADS). Technol Disabil 2002;14(3):107-111. [doi: 10.3233/TAD-2002-14305]
30. Demers L, Monette M, Lapierre Y, Arnold DL, Wolfson C. Reliability, validity, and applicability of the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) for adults with multiple sclerosis. Disabil Rehabil 2002;24(1-3):21-30. [doi: 10.1080/09638280110066352] [Medline: 11827151]
31. Ryan SE, Shepherd TA, Renzoni AM, et al. Responsiveness of a parent-reported outcome measure to evaluate AAC interventions for children and youth with complex communication needs. Augment Altern Commun 2018 Dec;34(4):348-358. [doi: 10.1080/07434618.2018.1520296] [Medline: 30369273]
32. Delarosa E, Horner S, Eisenberg C, Ball L, Renzoni AM, Ryan SE. Family Impact of Assistive Technology Scale: Development of a measurement scale for parents of children with complex communication needs. Augment Altern Commun 2012 Sep;28(3):171-180. [doi: 10.3109/07434618.2012.704525] [Medline: 22946992]
33. Kron AT, Kingsnorth S, Wright FV, Ryan SE. Construct validity of the Family Impact of Assistive Technology Scale for augmentative and alternative communication. Augment Altern Commun 2018 Dec;34(4):335-347. [doi: 10.1080/07434618.2018.1518993] [Medline: 30369255]
34. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. Phys Ther 1996 Oct;76(10):1109-1123. [doi: 10.1093/ptj/76.10.1109] [Medline: 8863764]
35. Downe-Wamboldt B. Content analysis: method, applications, and issues. Health Care Women Int 1992;13(3):313-321. [doi: 10.1080/07399339209516006] [Medline: 1399871]
36. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]
37. AbleNet Inc. URL: https://www.ablenetinc.com/ [accessed 2024-06-20]
38. Tobii Dynavox. URL: https://ca.tobiidynavox.com/ [accessed 2024-06-20]
39. Starble A, Hutchins T, Favro MA, Prelock P, Bitner B. Family-centered intervention and satisfaction with AAC device training. Commun Disord Q 2005 Dec;27(1):47-54. [doi: 10.1177/15257401050270010501]
40. Baxter S, Enderby P, Evans P, Judge S. Barriers and facilitators to the use of high-technology augmentative and alternative communication devices: a systematic review and qualitative synthesis. Int J Lang Commun Disord 2012;47(2):115-129. [doi: 10.1111/j.1460-6984.2011.00090.x] [Medline: 22369053]
41. Howard J, Fisher Z, Kemp AH, Lindsay S, Tasker LH, Tree JJ. Exploring the barriers to using assistive technology for individuals with chronic conditions: a meta-synthesis review. Disabil Rehabil Assist Technol 2022 May;17(4):390-408. [doi: 10.1080/17483107.2020.1788181] [Medline: 32663110]

XSL•FO
RenderX

42. Elmquist M, Crowe B, Wattanawongwan S, et al. Caregiver-implemented AAC interventions for children with intellectual or developmental disabilities: a systematic review. Rev J Autism Dev Disord 2025 Jun;12(2):290-310. [doi: 10.1007/s40489-023-00394-2] [Medline: 40893926]
43. Larsson Ranada Å, Lidström H. Satisfaction with assistive technology device in relation to the service delivery process—a systematic review. Assist Technol 2019;31(2):82-97. [doi: 10.1080/10400435.2017.1367737] [Medline: 28892461]
44. Lidström H, Almqvist L, Hemmingsson H. Computer-based assistive technology device for use by children with physical disabilities: a cross-sectional study. Disabil Rehabil Assist Technol 2012 Jul;7(4):287-293. [doi: 10.3109/17483107.2011.635332] [Medline: 22612787]
45. Craddock G, McCormack L. Delivering an AT service: a client-focused, social and participatory service delivery model in assistive technology in Ireland. Disabil Rehabil 2002;24(1-3):160-170. [doi: 10.1080/09638280110063869] [Medline: 11827150]
46. Anderson KL, Balandin S, Stancliffe RJ. Alternative service delivery models for families with a new speech generating device: perspectives of parents and therapists. Int J Speech Lang Pathol 2015 Apr;17(2):185-195. [doi: 10.3109/17549507.2014.979876] [Medline: 25472834]
47. Dariotis JK, Mabisi K, Jackson-Gordon R, et al. Implementing adolescent wellbeing and health programs in schools: insights from a mixed methods and multiple informant study. Prev Sci 2023 May;24(4):663-675. [doi: 10.1007/s11121-022-01481-2] [Medline: 36630022]

Abbreviations

AAC: augmentative and alternative communication
ATDP: Assistive Technology Delivery Protocol
FIATS: Family Impact of Assistive Technology Scale
ISO: International Organization for Standardization
PCERT: Pictorial Children's Effort Rating Table
QUEST: Quebec User Evaluation of Satisfaction with Assistive Technology

Edited by J Sarvestan, S Munce; submitted 07.Apr.2025; peer-reviewed by A Banayat, C Brum, D Ming; revised version received 09.Jan.2026; accepted 30.Jan.2026; published 31.Mar.2026.

Please cite as:
Mumford L, Guerriere D, Chau T
Appropriateness and Impact of a Vocal Cord Vibration Switch for Children with Complex Communication Needs: Case Series
JMIR Rehabil Assist Technol 2026;13:e75626
URL: https://rehab.jmir.org/2026/1/e75626
doi:10.2196/75626

© Leslie Mumford, Denise Guerriere, Tom Chau. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Technology-Driven Group Exercise Program Implementation in an Underserved Community: Multimethod Retrospective Evaluation Study

Whitney N Neal[1*], PhD; Laurie A Malone[2*], PhD, MPH

[1]Department of Health Behavior, School of Public Health, University of Alabama at Birmingham, 1665 University Boulevard, Birmingham, United States

[2]Department of Occupational Therapy, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States

[*]all authors contributed equally

Corresponding Author:
Whitney N Neal, PhD
Department of Health Behavior, School of Public Health, University of Alabama at Birmingham, 1665 University Boulevard, Birmingham, United States

Abstract

Background: Physical inactivity among people with disabilities and older adults is a persistent public health concern, particularly those from racial minority groups living in underserved communities where structural barriers limit access to exercise opportunities. Technology-driven exercise programs offer scalable solutions, but the contextual factors that influence their uptake, fidelity, and sustainability remain underexplored.

Objective: This study applied the Consolidated Framework for Implementation Research (CFIR) to identify barriers, facilitators, and lessons learned during the implementation of 2 technology-driven group exercise programs—synchronous online fitness and active virtual reality gaming—delivered to predominantly African American older adults with disabilities and/or chronic health conditions in an underserved community.

Methods: Using a multimethod design, we conducted semistructured interviews and focus groups with program participants, staff, and administrators (n=15) and administered implementation surveys to program deliverers (n=9). Qualitative data were analyzed deductively using CFIR domains and constructs, while survey data provided descriptive context.

Results: Both programs achieved high acceptability and appropriateness scores, with feasibility rated slightly lower for virtual reality. Themes mapped to 4 CFIR domains (innovation, outer setting, inner setting, and individuals). Qualitative analysis identified adaptability, affordability, engaging design, and community support as key facilitators, while barriers included technology and space constraints, resistance to new technology, and sustainability challenges.

Conclusions: Technology-driven exercise programs can expand access to physical activity in underserved communities when they are adaptable, affordable, and socially engaging. Addressing multilevel barriers, including resource limitations, technology hesitancy, and long-term sustainability, is critical to scaling and sustaining these interventions.

(JMIR Rehabil Assist Technol 2026;13:e79598) doi:10.2196/79598

KEYWORDS

physical activity; underserved communities; group fitness intervention; program implementation; exercise

Introduction

Despite the known health benefits of being physically active, people with disabilities report high levels of physical inactivity. Reports worldwide indicate that approximately 54% to 91% of adults without disabilities meet physical activity recommendations, compared with only 21% to 60% of people with disabilities [1]. In the United States, 1 in 4 adults is living with a disability [2], and sedentary people with disabilities are 3 times more likely to report primary health conditions such as heart disease, stroke, diabetes, or cancer compared to active people with disabilities [3]. Furthermore, physical inactivity exacerbates loneliness and social isolation, which are also associated with adverse health outcomes, including heart disease and stroke [4]. Of concern is that individuals with disabilities are not only physically inactive but often experience higher levels of social isolation and loneliness than persons without disabilities [5]. Together, these disparities underscore the urgent need for effective, inclusive strategies to promote physical activity among people with disabilities.

Barriers to physical activity for people with disabilities span multiple levels of the socioecological model, including individual (eg, lack of knowledge, physical limitations, pain,

XSL•FO
RenderX

and fatigue), interpersonal (eg, lack of social support), institutional or organizational (eg, lack of accessible fitness facilities and exercise equipment), community (eg, poorly designed outdoor spaces and street connectivity), and systemic (eg, limited or no transportation) factors [1,6-10]. These multifactorial barriers limit opportunities for people with disabilities to engage in effective and sustainable physical activity and are further compounded in underserved communities—those characterized by racial and ethnic health disparities, high poverty rates, low insurance coverage, and limited health care and physical activity resources—underscoring the need to address physical activity barriers within the socioecological framework [11,12].

Technology-driven physical activity [13] and health promotion programming can help eliminate many of these barriers by delivering accessible, scalable interventions that foster social engagement and a sense of community [14-17]. Recent systematic reviews and meta-analyses demonstrate that mobile and wearable technologies, telehealth-delivered exercise, online fitness programs, and interactive gaming platforms can increase physical activity levels and improve body composition and physical function, with consistent benefits shown across home-based and community settings [18-20]. Telehealth programs have shown improvements in physical functioning even among individuals with frailty, cognitive challenges, or mobility limitations, highlighting the value of remote delivery for populations facing barriers to traditional exercise access [21]. Adherence to technology-based exercise is generally favorable when programs are accessible, engaging, and designed to meet participant needs [22]. Interactive and game-based technologies further extend reach and appeal: virtual reality (VR) exergame programs have been shown to improve balance, mobility, and functional outcomes in older adults, including those living in long-term care, while enhancing enjoyment and motivation to be active [23-25]. Work by Malone et al [15,26] demonstrates that adults with mobility impairments can engage with adapted exergaming systems—such as modified Wii Fit balance boards and gaming mats—to achieve greater energy expenditure and report high enjoyment, suggesting these technologies can reduce barriers and support participation in meaningful movement. Additionally, online membership–based fitness programs have shown feasibility and preliminary effectiveness for increasing physical activity in people with mobility impairments [27]. Collectively, this body of evidence supports the use of technology-driven physical activity interventions as accessible, engaging, and adaptable options for promoting physical activity and supporting health in populations traditionally underserved by conventional exercise models.

Although technology-mediated exercise shows promise, evidence of efficacy and feasibility alone is insufficient to ensure successful deployment in resource-limited settings. To translate and sustain such programs in underserved settings, it is essential to examine multilevel determinants of implementation. Implementation science offers a structured approach to understanding and addressing contextual factors that influence the uptake, fidelity, and sustainability of health interventions [28]. The Consolidated Framework for Implementation Research (CFIR) provides a comprehensive, validated taxonomy for assessing contextual factors that influence implementation success across 5 domains: outer setting, inner setting, innovation, implementation process, and individuals [29,30]. Using a framework like the CFIR ensures that barriers and facilitators are identified and addressed in a structured, reproducible manner, strengthening both the science and practice of health promotion program delivery in underserved settings. The CFIR has been used to identify implementation barriers and facilitators of physical activity interventions for older adults and people with disabilities [31-33], and in a systematic review of community-based physical activity interventions [34]. However, research on technology-driven exercise interventions in underserved communities remains limited.

To address this gap, this study used the CFIR to identify facilitators, barriers, and lessons learned during the implementation of 2 technology-driven group exercise programs (synchronous online fitness and VR exergaming) delivered to older adults (aged >60 y) and individuals with disabilities and/or chronic health conditions living in an underserved community. Our primary aims were to (1) characterize the CFIR constructs that most strongly influenced each program's implementation outcomes and (2) compare implementation team member perspectives (program staff, administrators, and participants) on technology-specific and context-specific challenges and enablers. By leveraging a validated implementation framework, this work examines how technological and contextual factors shape program delivery in underserved communities and informs strategies to translate and sustain effective technology-based exercise in traditionally under-resourced settings.

Methods

Study Design and Setting

Overview

This study followed a multimethod research design, whereby the core of the research was qualitative, utilizing in-depth interviews and focus groups, and was supplemented by a minor quantitative component in the form of a brief survey provided to program staff and administrators. We chose both methods not for the purpose of triangulation or formal mixing of qualitative and quantitative paradigms, but rather for complementarity and expansion [35], allowing the survey data to offer a broader, descriptive overview of factors associated with the Lakeshore Online Fitness (LOF) and VR program implementation and the rich qualitative data to provide deep context.

This research was a retrospective evaluation of 2 technology-based interventions deployed between July 2022 and October 2022 at a neighborhood recreation center located in an underserved community (population <5800; ≈50% Black, ≈15% Hispanic or Latino; ≥25% living below the federal poverty line; 1 in 4 uninsured; 1 in 4 with a disability). The synchronous online fitness program was deployed in partnership with a local nonprofit organization that provides physical activity, recreation, and sport programs for people with physical disabilities and chronic health conditions (Lakeshore Foundation). For the VR program component, content- and age-appropriate (ie, limited

XSL•FO
RenderX

violence or foul language) active video games were identified by research staff. The LOF and Get Active with VR programs were delivered on-site at the municipal recreation facility building. Program recipients were current members of the recreation center's senior programs who met the following eligibility criteria: (1) aged 18 years or older, (2) self-reported a physical disability or health condition, (3) aged 60 years or older regardless of health condition, and (4) were not enrolled in any other physical activity program. All successfully screened participants provided informed consent prior to enrollment. Sessions for both programs were held twice a week, lasted 60 minutes, and included a warm-up, main activity, and cooldown. Participants completed 6 weeks of each program (LOF or VR) in a randomized crossover design using alternating assignments to either the LOF (Figure 1) or VR program (Figure 2).

Figure 1. Participants engaging in the "Lakeshore Online Fitness" program.



XSL•FO
RenderX

Figure 2. Participants engaging in the "Get Active with Virtual Reality" program.



Lakeshore Online Fitness

Online classes were delivered via Zoom (Zoom Communications, Inc) from Lakeshore to the municipal recreation facility [27,36]. Exercise dose was based on the predetermined length of online classes provided by Lakeshore, and frequency was based on scheduling needs at the recreation center. An instructor from Lakeshore taught all classes, and a community facilitator (staff or volunteer) and research team member oversaw the on-site group activity.

Get Active With Virtual Reality

Oculus Quest 2 headsets were used for the VR program. A community facilitator and research team member oversaw VR classes. For each session, program recipients would choose to play a predetermined game individually or competitively against other participants. Exercise dose and frequency were matched to those of the LOF program.

Study Participants

This research study used convenience sampling techniques to identify factors related to program implementation. Our sample size was determined by the availability of respondents and met minimum guidelines for thematic analysis [37,38]. We gathered input from individuals who (1) participated in the LOF and/or VR programs and (2) played a role in program implementation and/or delivery. Invitations for interviews were extended to program recipients who completed the LOF and VR programs (n=10). Invitations for focus groups were extended to 2 LOF instructors, 2 Lakeshore Foundation administrators, 1 recreation center administrator, 1 recreation center program staff member, and 2 community volunteers (n=9). The recreation center staff member's primary role was to help with the physical setup for LOF, assist participants during classes, and take down program equipment. Both community volunteers were recruited through the University of Alabama at Birmingham Center for the Study of Community Health to act as community liaisons, assist with classes, and identify potential new locations for implementation of the programs. These volunteers also helped with the physical setup for the VR program, assisted participants during classes, and took down program equipment.

Ethical Considerations

The University of Alabama at Birmingham Institutional Review Board approved this study (IRB-300008762), and all program participants provided written informed consent before taking part, which included agreement to participate in the study and permission for their images to appear in photos used in the study. Program participants received US $15 for completing the interviews. Data were stored on encrypted, password-protected servers accessible only to the research team.

Data Collection

Quantitative

Administrators, program staff, and volunteers involved in implementing the LOF and VR programs were invited to complete an online survey assessing accessibility, appropriateness, and feasibility of the two programs in October 2023. The Acceptability of Intervention Measure (AIM), Intervention Appropriateness Measure (IAM), and Feasibility of Intervention Measure (FIM) surveys were completed by all

XSL•FO
RenderX

9 relevant staff members and administrators. The AIM assesses the "perception that a given treatment, service, practice, or innovation is aggregable, palatable, or satisfactory," the IAM evaluates the "perceived fit of the innovation to address a particular issue or problem," while the FIM measures the ease at which a "new treatment or innovation can be successfully used or carried out within a given agency or setting." [39] Each measure consisted of 4 thematic questions, with 5 responses ranging from completely disagree to completely agree on a 5-point Likert scale [40]. Outcomes for the LOF and VR programs were assessed separately. Scores for each scale range from 4 to 20 and were classified as low (<12), medium (12-15), and high (>15) [40].

Qualitative

Program participants took part in one-on-one, semistructured interviews at the recreation center in January 2023 after completing the intervention. A CFIR-informed interview guide, developed in collaboration with the study principal investigator (LAM) and interviewer (WNN), was used to gauge perceptions of the most important factors influencing implementation of the 2 programs for older adults and people with disabilities in underserved communities. Interview guide questions focused on the individual, innovation, and inner setting domains of the CFIR (Multimedia Appendix 1) [29].

A separate CFIR-informed focus group guide was developed to capture staff members' opinions about barriers, facilitators, and lessons learned from implementation of the programs (Multimedia Appendix 2). Focus group questions covered the innovation, outer setting, inner setting, and implementation process domains of the CFIR [29]. Two semistructured, in-depth focus groups and 1 interview were carried out. The first focus group (n=3) included the recreation center program staff member and two community volunteers who helped set up and deliver the LOF and VR programs. The second focus group (n=3) comprised LOF fitness instructors who delivered the LOF program and 1 Lakeshore Foundation senior staff member. The focus group with recreation center staff and volunteers was conducted in person at the community center where participants completed the programs, while the focus group with LOF program deliverers and senior staff was carried out in person at the Lakeshore Foundation in November 2023. One Lakeshore Foundation senior staff member was unable to attend the focus group session, so they completed a follow-up, one-on-one, semistructured interview online in December 2023. Interviews and focus groups were completed with 79% of program participants and relevant staff and volunteers.

Data Analysis and Rigor

Survey data and participant characteristics are summarized using descriptive statistics. Means and SDs are reported for the continuous variables (ie, age and AIM, IAM, and FIM surveys), while categorical variables (ie, race and sex) are presented as frequencies and percentages.

All qualitative data were coded and organized using NVivo 14 (Lumivero). In the first stage of analysis, the Zoom recordings were transcribed verbatim using a professional transcription service (Landmark Associates, Inc), with time stamps at the beginning of each response and speakers labeled as interviewer or interviewee to ensure clarity and anonymity of the content. The first author became familiar with the data by listening to the audio recordings and repeatedly reading the transcripts. The data were then analyzed using a deductive content analysis approach, following the 3-phase guide outlined by Elo and Kyngäs [41], which includes preparation, organization, and reporting. During the preparation phase, it was decided that sentences, phrases, and themes from the transcripts would serve as the units of analysis to capture both detailed content and broader thematic patterns. Both manifest (explicit, surface-level content) and latent (underlying meanings) content were coded, allowing for a more nuanced understanding of factors influencing implementation. A deductive approach was chosen to align the identified themes with the CFIR domains and constructs, facilitating a theory-driven analysis of implementation factors.

In phase 2 of the qualitative analysis, a structured categorization matrix based on the 5 CFIR domains and 39 constructs was developed to guide organization and interpretation of themes. The first author then reviewed the data for content and coding based on correspondence with the predetermined domains and constructs. In phase 3, results were synthesized by identifying the most frequently reported themes. To ensure accurate alignment with CFIR constructs, the first and second authors reviewed and refined the synthesized results, confirming that each theme was appropriately categorized and all relevant data were considered. Integration occurred during the interpretation phase, where the survey results were used to provide supplementary context and illustration for the core qualitative themes. In this analysis, the main findings remain rooted in the qualitative evidence and are organized by facilitators and barriers, specific CFIR domains, and subthemes that correspond to the CFIR constructs [29].

To ensure quality and rigor, the evaluative markers coherence, transparency, and substantive contribution were selected from Smith and Caddick's [42] ongoing list and can be found in Table 1. At the time of data collection, the first author (WNN) was a predoctoral student in health behavior with formal training in qualitative research methods. With nearly a decade of experience working with individuals with chronic health conditions and no prior involvement in the study's program delivery portion, she brought an outsider perspective. The second author (LAM) has been involved in several qualitative studies related to physical activity and disability and served as a "critical friend" throughout the process. This role involved reviewing the first author's interpretations of the data and providing feedback on the accuracy and consistency of the coding and categorization process. Rather than reach an agreement, this collaborative, reflective approach was intended to generate nuanced and multifaceted interpretations of meaning.

XSL•FO
RenderX

Table . Evaluative markers chosen for quality and rigor.

Marker	Definition	Application
Coherence	The way that different parts of the interpretation create a complete and meaningful picture.	Demonstrated through the alignment of the research questions, methodology, and validated theoretical framework to guide the evaluation.
Transparency	Honesty about the research process.	Maintenance of an audit trail through detailed notes and critical friend throughout analytical process.
Substantive contribution	The extent to which a study makes a meaningful impact on our understanding of what is being researched theoretically, conceptually, practically, or methodologically.	Offers a theoretical understanding of factors influencing the implementation of technology-driven community programs.

Results

Implementation Surveys

All administrators, program staff, and volunteers involved in implementing the LOF and VR programs (n=9, 100%) completed the surveys. Results from the AIM, IAM, and FIM surveys are provided in Table 2. Overall, both programs attained high ratings across all 3 domains. However, one-third of respondents (n=3, 33.3%) rated the VR program's feasibility as medium (FIM score <15) rather than high.

Table . Acceptability, appropriateness, and feasibility measure results.

	LOF[a] program, mean (SD)	VR[b] program, mean (SD)
AIM[c]	16.89 (1.62)	17.56 (2.74)
IAM[d]	16.33 (1.58)	16.56 (2.70)
FIM[e]	16.78 (1.64)	16.22 (2.86)

[a]LOF: Lakeshore Online Fitness

[b]VR: virtual reality

[c]AIM: Acceptability of Intervention Measure

[d] IAM: Intervention Appropriateness Measure

[e]FIM: Feasibility of Intervention Measure

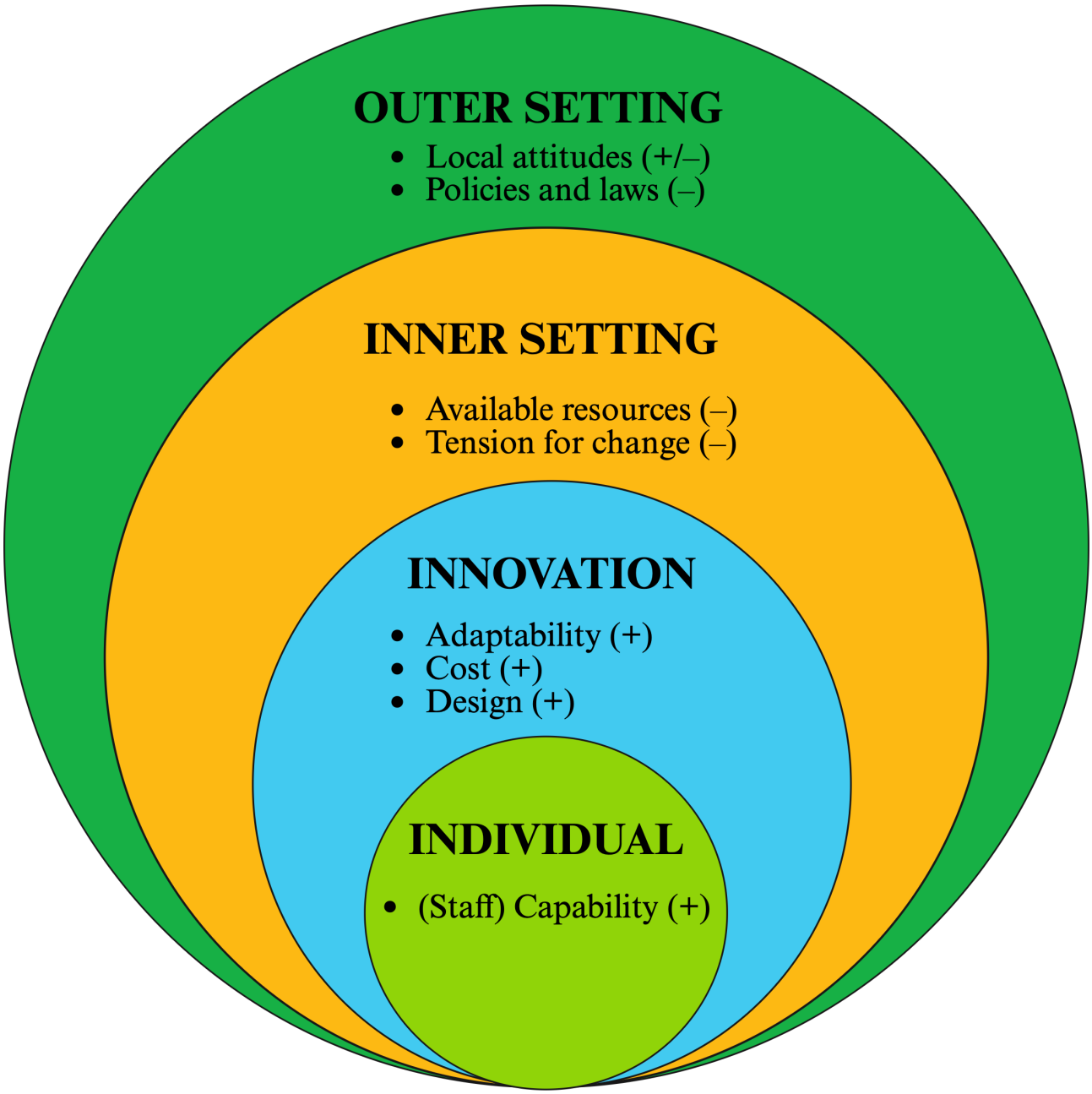
Interview and Focus Group Results

A total of 15 individuals took part in the interviews (mean 41.3 min, SD 16.1 min; range 24 - 63 min) or focus groups (mean 49.3 min, SD 7.1 min), including 8 of 10 program participants, 1 recreation center staff member, 2 community volunteers, 2 fitness instructors, and 2 administrators [1]. Interview participants (n=8) were primarily Black or African American (n=7, 87.5%) and female (n=5, 62.5%) participants with a mean age of 65.0 (SD 14.0; range 44 - 86) years. The majority of interview participants had hypertension (n=6, 75%) and arthritis (n=4, 50%), with 37.5% having both concurrently. Focus group participants were primarily White (5 /7, 71.4%) and male (4/7, 57.1%) individuals; age was not collected. Program participants completed 78.3% and 85.4% of LOF and VR sessions, respectively.

Qualitative data analysis identified 4 overarching domains and 8 constructs that represented key facilitators and barriers to implementing the LOF and VR programs. The results mapped to CFIR domains and constructs are depicted in Figure 3.

XSL•FO
RenderX

Figure 3. Implementation barriers and facilitators coded into Consolidated Framework for Implementation Research (CFIR) domains and constructs. "–" indicates barriers, "+" indicates facilitators. Created in BioRender [43].



Key Facilitators to LOF or VR Implementation

Six subthemes based on 5 CFIR constructs were identified as facilitators to program implementation. These constructs belonged to 3 CFIR domains ("innovation," "outer setting," and "individuals") and are described in more detail under the "Innovation," "Outer Setting," and "Individuals" subsections below.

Innovation

Overview

The innovation domain comprises features of an intervention that might influence implementation. Key facilitators related to program characteristics centered on the adaptability, cost, design quality, and packaging constructs. Implementation team members praised both the LOF and VR programs for their flexibility, allowing for modifications to accommodate diverse physical abilities. The low cost of equipment and free access at the community center made both programs accessible and affordable, while the socially interactive and engaging design enhanced motivation and enjoyment for program recipients.

Flexible LOF Exercises (Adaptability)

All program recipients and deliverers noted the flexibility of the LOF program for participants with diverse physical abilities, such as low mobility, dexterity limitations, low vision, or stroke-related impairments. Participants with dexterity limitations mentioned the ease of modifying the LOF program components to facilitate engagement, "Once I could pick up the weights, when the resistance ball isn't—none of it bothered me, it was just putting it in my hand" (program recipient 04). Instructors adjusted exercises to participants' needs without full

XSL•FO
RenderX

information on their abilities but noted that experience was essential for making these on-the-fly adaptations.

VR Program Technology Considerations (Adaptability)

Recreation center staff and volunteers emphasized the importance of the VR program's adaptability to diverse settings, such as standing versus seated sessions to support group activity and ensure participant safety:

> *I think if they were standing up, they would likely free roam. They're in their virtual world, and they can't see where they're going or what they're doing. When you start moving, you just movin'. They might be bumping into each other. I think sitting was the perfect thing to stay in the circle and not bump each other. I think that was good.* [Community volunteer 01]

Although some participants initially struggled with dexterity, this issue was resolved quickly by the research staff, who modified the VR equipment for one-handed use.

Program Access and Affordability (Cost)

Program deliverers felt that the LOF program's low-cost exercise equipment, such as balls, bands, and lightweight dumbbells, would make the program relatively simple to replicate in other locations:

> *In theory, the way this works is we're bringing a lot of the expense to them. We're bringing the TV. We're trying to cut away as much of the red tape as we can. What we really need is for them to say, yes, here's a room. Here's somebody to turn on the TV. Here's somebody who can hover in there to help with stuff, but you guys do the hard part of buying the TV and having the expertise to teach the class. It's just [a] partnership.* [Administrator 01]

Program recipients appreciated that both programs were free and available at the community center, citing affordability and ease of access as key facilitators of engagement. They acknowledged that many older adults and people with chronic health conditions in their community lacked transportation to attend paid physical activity programs, but that, "…they would be happy to do exercise like that, long as they don't have to leave the building'' (program recipient 09).

Socially Interactive and Engaging Design (Design)

This subtheme pertains to participant perceptions of the quality of the materials associated with the LOF and VR programs. Program recipients valued the socially supported, group-based nature of the LOF program, which enhanced motivation. "We did those stretch band things. We did the balls and the dumbbells…I liked it. I liked being on the equipment" (program recipient 02). They found the competitive and interactive elements of the VR program particularly motivating, as these elements allowed them to compete against others or themselves, "We had a competition going together shooting games and stuff like that, bowling, and stuff like that" (program recipient 03). These program components made physical activity more enjoyable and immersive:

> *I like the [VR] headset because you get the feeling of really being in there. That's the part I really like. The sceneries are so beautiful. You really did, you felt like you had actually stepped into a space age, or a tennis court or baseball diamond.* [Program recipient 08]

Outer Setting: Social and Community Engagement (Local Attitudes)

This subtheme reflects the degree to which sociocultural values (eg, shared responsibility in helping recipients) and beliefs (eg, convictions about the worthiness of recipients) encourage the outer setting (eg, the city the program was implemented in) to support implementation and/or delivery of the intervention." [30,31] Those involved in program implementation and delivery felt that, "Having a community champion like the assistant mayor" (administrator 01) was crucial to the program's success, as their advocacy facilitated connections with community leaders and strengthened local engagement. Positive reactions from program recipients further affirmed the programs' value and energized program deliverers. These groups noted that this enthusiasm made the program more enjoyable for the staff and was a key indicator of its success:

> *Just seeing them enjoying made me feel good, because they [were] enjoying something. It's not like they [were] forced to do it or "Okay, we're in this program here at the rec, so this is part of it. We have to do it." "No, this is a program that I enjoy. I'm glad when it start[s]. I hate when it end[s]. We just want to have fun." They all got along. That was a good thing.* [Community volunteer 01]

> *...it was fun to coach, because you could see them move and you could see 'em doing things and it seemed like they really enjoyed it and then when it ended, they would clap and say thank you, and all that and give you that reassurance.* [Fitness instructor 01]

Individuals: Staff Training and Role Clarity (Capability)

Effective implementation required clear role distribution and adequate staffing to assist participants, manage technology, and troubleshoot issues onsite. Program deliverers felt that staff with strong interpersonal and problem-solving skills and at least some knowledge of exercise were beneficial to LOF and VR program success:

> *So, ideally, it would be someone with fitness experience…[but] minimally, it has to be, in my opinion, someone who is consistent, dependable, and someone that we could train.* [Administrator 02]

> *It was a very easy setup and once the class was going, you were hands off. If someone didn't really seem like they knew what they were doing, you would try and mimic what the instructor was doing. I don't really feel like there was too much training needed for that portion of it.* [Recreation center staff 01]

Key Barriers to LOF or VR Implementation

There were 4 barriers identified from the interviews and focus groups, with subthemes divided evenly between the inner and outer setting domains.

XSL•FO
RenderX

Inner Setting

Overview

Barriers to LOF and VR program implementation within the inner setting domain (eg, the organization in which the intervention is being implemented) [30,31] centered on the available resources and tension for change constructs. The LOF and VR programs at the recreation center faced technical and spatial challenges, and program recipients expressed dissatisfaction with the limited accessible physical activity options in their area, highlighting the need for sustainable community resources to address underlying infrastructure deficits.

Technology and Space Challenges (Available Resources)

This subtheme reflects "the degree to which resources are available to implement and deliver the intervention" [30,31] Addressing technical and spatial barriers, including technology setup and space constraints, was essential to successful implementation of the LOF and VR programs at the recreation center. Stable Wi-Fi, reliable equipment, and adequate space were critical to overcoming these challenges and ensuring successful program delivery. Program users reported minimal technical issues:

> *The only [LOF program] technical issue I think we had—well, then it wasn't an issue—was that we could not hear what she was sayin', but we could see what she was doin', so we still continued because we could see her doin' the exercises.* [Program recipient 05]

> *One day the [VR program] computer wasn't acting right for some. … [But] we never had to miss a day because of the technical difficulties.* [Program recipient 08]

However, program providers felt that LOF program delivery was complicated by spatial constraints at the recreation center:

> *I think the biggest complaint with the [LOF] class was we just didn't have enough space… Sometimes we had 10 people, and so we had all the chairs. It was like they couldn't fully put out their arms, or [they'd] be running into each other and stuff like that. … I feel like just the size of the room was the biggest issue.* [Recreation center staff 01]

> *You could see 'em moving and everything, but you weren't able to read people's eyes, you saw gross movements, but you didn't see the finer movements.* [Fitness instructor 01]

LOF instructors noted that participant monitoring could be enhanced with larger screens and/or additional cameras, "…so you can see people a little bit closer, 'cause we were at a distance, and you really couldn't see their faces" (fitness instructor 01). Instructors also emphasized the importance of standardized and simplified technological setups on their end to minimize the burden of adjusting to new rooms, equipment, and lighting conditions.

Need for Infrastructure and Resource Support (Tension for Change)

This subtheme aligns with the CFIR construct "tension for change," where implementation team member recognition of significant gaps creates momentum for implementation efforts [30,31]. Program recipients expressed a strong desire for more accessible physical activity options in their area, noting that the absence of such resources created what they perceived as an urgent community need: "There aren't no other [free physical activity opportunities] here...People, they go out down there to the community store and buy 'em a can of beer or a pack of cigarettes" (program recipient 04). Participants highlighted both the immediate value of the LOF and VR programs, "Well, me and my other two siblings, we got high blood pressure, so we need somethin' to motivate us" (program recipient 02), and emphasized the need for additional community resources to support these programs long-term:

> *...Cause we don't have nothin' there. I needed it. ...Those people need it. They really need somethin' like that…* [Program recipient 09]

The emphasis on sustainability concerns further reflected program recipients' insight that temporary interventions, while beneficial, would not address the fundamental infrastructure deficits underlying their community's physical activity challenges.

Outer Setting

Key barriers to LOF and VR program implementation within the outer setting domain centered on the local attitudes and policies and laws constructs. Resistance to new technology and sustainability challenges were identified as barriers to implementing and sustaining the LOF and VR programs.

Resistance to New Technology and Approaches (Local Attitudes)

Program recipients, administrators, recreation center staff, and volunteers all mentioned resistance to adopting new technology and program approaches. Some program recipients expressed concerns that others in the community might resist the VR program due to discomfort with wearing headsets, influenced by cultural and personal preferences: "...I don't think they would like it 'cause just puttin' stuff on their face…" (program recipient 09). Similarly, recreation center administration initially doubted participants' willingness to engage with the LOF and VR programs, reflecting skepticism about the programs' acceptance and feasibility:

> *When we said, well, how about this group take the class, [senior staff] at the rec center were like, y'all are out of your minds. They are never going to participate in this* [Administrator 01]

Sustainability Challenges (Policies and Laws)

This subtheme reflects the "degree to which legislation, regulations, professional group guidelines and recommendations, or accreditation standards support implementation and/or delivery of an intervention" [30,31]. One administrator felt that relying solely on grant funding made long-term sustainability of the LOF program uncertain, especially without a financial

XSL•FO
RenderX

model for continued implementation, "So how do we deliver it when it's not through a grant, and how do we try to align as closely to the way that we would onboard members now, but on the online, which is different" (administrator 02). This administrator also noted liability and membership issues, including the need for waivers and consistent liability protocols, as barriers to continuing the LOF program at the community center, "...Because we can't deliver any product without some kinda safeguard set up...and I don't believe we ever figured that out" (administrator 02). Similarly, program deliverers expressed concerns about the VR program's long-term feasibility due to equipment maintenance, cultural unfamiliarity, and staff burden—factors that align with the relatively lower feasibility scores reported in the FIM survey.

Discussion

Key Findings and Contributions

This study applied the CFIR to explore barriers and facilitators influencing implementation of two technology-driven group exercise programs—an online platform (LOF) and VR gaming—delivered to predominantly African American older adults living in an underserved community. Through semistructured interviews, focus groups, and implementation surveys, we identified factors across 4 CFIR domains (innovation, outer setting, inner setting, and individuals) that collectively influenced the programs' uptake, delivery, and sustainability. Our findings highlight that successful implementation was shaped by a combination of program adaptability, affordability, social engagement, and community support, while sustainability and resource constraints emerged as persistent barriers.

A key contribution of this study lies in demonstrating how technology-enabled interventions can create structured exercise opportunities in community centers that previously lacked formal fitness programming. Unlike settings where technology supplements existing exercise classes, the recreation center in this study had no structured physical activity programs prior to implementation. Technology served as the catalyst for introducing accessible, engaging, and adaptable exercise options, addressing a critical gap in underserved communities where traditional programs are often unavailable due to resource limitations. This approach aligns with growing evidence that technology-based interventions can overcome environmental and logistical barriers by delivering scalable, customizable programs that promote physical activity and social engagement among older adults and individuals with disabilities [18-20]. By leveraging online platforms and VR gaming, the programs provided not only exercise opportunities but also novel experiences that enhanced enjoyment and motivation—factors shown to improve adherence in populations facing mobility and transportation challenges [15,23,26].

Facilitators of Implementation

Adaptability emerged as a key facilitator across both programs. The ability to adjust exercises and technology to accommodate diverse physical abilities allowed for broad participation, even among individuals with mobility limitations and dexterity challenges. These findings align with previous research highlighting the importance of customizable interventions that support engagement among populations with heterogeneous physical limitations [44,45]. Affordability and ease of access were also critical enablers, extending beyond equipment costs to include transportation and membership fees, which are common deterrents among people with disabilities [1]. By eliminating these costs, the programs created equitable access to structured exercise opportunities, reinforcing prior evidence that cost-sensitive solutions are vital for program sustainability in resource-constrained settings [46].

The socially interactive and gamified design of both programs further enhanced motivation and enjoyment. LOF's group-based format fostered camaraderie, while the VR program's competitive elements introduced novelty and fun, supporting the value of incorporating social and gamified components into technology-based group physical activity programs for older adults and people with disabilities to sustain engagement [47]. Future implementation of technology-driven exercise interventions should leverage these design principles to strengthen both physical and psychosocial outcomes in this population.

Finally, community support amplified program success. Advocacy from community champions (eg, city leaders), enthusiastic endorsement from community volunteers and recreation center administrators, and visible program recipient enthusiasm strengthened buy-in among staff and volunteers, creating a sense of shared ownership. This is consistent with previous research emphasizing the role of community advocacy in successful implementation of health promotion programs in underserved areas [45] and suggests that early identification of community champions should be integral to implementation planning.

Barriers to Implementation

Despite these strengths, several barriers impeded implementation. Technology and space limitations were the most immediate obstacles. Unstable Wi-Fi, suboptimal camera setups, and inadequate physical space complicated LOF program delivery and monitoring, while the VR sessions required careful configuration to ensure safety, highlighting the need for robust infrastructure and standardized protocols [45,46,48]. Resistance to new technology reflected cultural and personal discomfort, particularly with VR headsets, and skepticism among staff about program feasibility. While program recipients spoke highly of the VR program, initial hesitancy underscores the importance of tailored program orientation and gradual exposure to technology to build confidence and acceptance [45,49,50].

Additionally, sustainability challenges stemming from reliance on grant funding and unresolved liability protocols raised concerns about program continuity. Administrations cited liability and membership protocols as unresolved issues for the LOF program, while survey responses from program deliverers reflected reservations about the VR program's feasibility. This underscores the need for alternative funding strategies (eg, partnerships with local health departments and sliding-scale fees), infrastructure development, and policy alignment to support long-term program continuation. Without such planning,

XSL•FO
RenderX

even well-received programs risk discontinuation once initial funding ends.

Implications

Our findings suggest that technology-driven exercise programs can reduce health disparities in underserved communities when they are adaptable, affordable, socially engaging, and supported by community champions. However, implementation efforts must anticipate and address resource constraints, technology resistance, and sustainability challenges. Multilevel implementation strategies, including investment in digital infrastructure, community advocacy, targeted orientation to reduce technology hesitancy, and sustainable financing models, are essential to scale and sustain such interventions. Future research should explore strategies for scaling these programs, such as partnerships with local organizations, integration into existing community services, or development of cost-sharing models. Programs should engage community champions from inception, build standardized technical protocols, and test these programs across diverse geographic and demographic contexts.

Strengths and Limitations

Study strengths include its use of a validated implementation framework (CFIR), integration of multiple implementation team member perspectives, and use of both qualitative and quantitative methods to identify barriers and facilitators to program implementation. However, the current study has limitations that warrant consideration. Our sample, though achieving 79% participation among relevant implementation team members, was relatively small and drawn from a single, predominantly African American community, which may limit generalizability. Additionally, racial distribution differed between data collection formats: interview participants were predominantly African American, while focus group participants were primarily White. This imbalance may have influenced the range of perspectives captured. Implementation team member perspectives were collected several months after program delivery, which may not have captured immediate perspectives of program sustainability or technology acceptance. Nonetheless, the rich qualitative insights underscore the value of incorporating ongoing implementation team member evaluation and feedback into program design.

Conclusions

Technology-enhanced group exercise interventions hold considerable promise for reducing physical inactivity among underserved older adults and people with disabilities. The results of our CFIR-guided analysis suggest that, while adaptability, low cost, and social engagement facilitated uptake, the LOF and VR programs presented technical, infrastructural, and policy-related barriers. The findings of this study contribute to the growing literature on technology-enabled health interventions for aging populations and people with disabilities and offer insights into implementation strategies in underserved, resource-limited settings.

Acknowledgments

The authors wish to acknowledge Eric Evans for his initiative on the virtual reality portion of this project and to thank Ashley Wright for her expertise and assistance with program delivery. The authors also extend our appreciation to the Lakeshore Foundation fitness instructors and senior leadership, community volunteers, and staff at the Parks and Recreation Department. Without their time, dedication, commitment, and resources, this project would not have been possible. The authors thank Emily Delzell for her editorial review and assistance in the preparation of this manuscript.

Funding

This work was supported by the University of Alabama at Birmingham Center for the Study of Community Health University-Wide Interdisciplinary Research Center and the University of Alabama Center for Clinical and Translational Science.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Lakeshore Online Fitness or Get Active with Virtual Reality intervention participant interview guide.
[DOCX File, 34 KB - rehab_v13i1e79598_app1.docx]

Multimedia Appendix 2
Lakeshore Online Fitness or Get Active with Virtual Reality intervention focus group guide.
[DOCX File, 25 KB - rehab_v13i1e79598_app2.docx]

References

1. Martin Ginis KA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. Health Psychol Rev 2016 Dec;10(4):478-494. [doi: 10.1080/17437199.2016.1198240] [Medline: 27265062]

XSL•FO
RenderX

2. Okoro CA, Hollis ND, Cyrus AC, Griffin-Blake S. Prevalence of disabilities and health care access by disability status and type among adults—United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Aug 17;67(32):882-887. [doi: 10.15585/mmwr.mm6732a3] [Medline: 30114005]
3. Carroll DD, Courtney-Long EA, Stevens AC, et al. Vital signs: disability and physical activity—United States, 2009-2012. MMWR Morb Mortal Wkly Rep 2014 May 9;63(18):407-413. [Medline: 24807240]
4. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016 Jul 1;102(13):1009-1016. [doi: 10.1136/heartjnl-2015-308790] [Medline: 27091846]
5. Emerson E, Fortune N, Llewellyn G, Stancliffe R. Loneliness, social support, social isolation and wellbeing among working age adults with and without disability: cross-sectional study. Disabil Health J 2021 Jan;14(1):100965. [doi: 10.1016/j.dhjo.2020.100965] [Medline: 32843311]
6. Rimmer JH, Padalabalanarayanan S, Malone LA, Mehta T. Fitness facilities still lack accessibility for people with disabilities. Disabil Health J 2017 Apr;10(2):214-221. [doi: 10.1016/j.dhjo.2016.12.011] [Medline: 28143707]
7. Malone LA, Barfield JP, Brasher JD. Perceived benefits and barriers to exercise among persons with physical disabilities or chronic health conditions within action or maintenance stages of exercise. Disabil Health J 2012 Oct;5(4):254-260. [doi: 10.1016/j.dhjo.2012.05.004] [Medline: 23021736]
8. Kayes NM, McPherson KM, Taylor D, Schlüter PJ, Kolt GS. Facilitators and barriers to engagement in physical activity for people with multiple sclerosis: a qualitative investigation. Disabil Rehabil 2011;33(8):625-642. [doi: 10.3109/09638288.2010.505992] [Medline: 20695816]
9. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]
10. Vasudevan V, Vang PD, Fernandez-Baca D. An exploration of healthy eating and physical activity habits of Hmong high-school students by disability status: a pilot study. Disabil Health J 2019 Oct;12(4):694-698. [doi: 10.1016/j.dhjo.2019.06.012] [Medline: 31257033]
11. Cockerham WC, Hamby BW, Oates GR. The social determinants of chronic disease. Am J Prev Med 2017 Jan;52(1S1):S5-S12. [doi: 10.1016/j.amepre.2016.09.010] [Medline: 27989293]
12. Mahmoudi E, Meade MA. Disparities in access to health care among adults with physical disabilities: analysis of a representative national sample for a ten-year period. Disabil Health J 2015 Apr;8(2):182-190. [doi: 10.1016/j.dhjo.2014.08.007] [Medline: 25263459]
13. Evans EJ, Naugle KE, Owen T, Naugle KM. Active gaming: it is not just for young people. J Aging Phys Act 2020;28(5):731-739. [doi: 10.1123/japa.2019-0303] [Medline: 32422600]
14. Rowland JL, Malone LA, Fidopiastis CM, Padalabalanarayanan S, Thirumalai M, Rimmer JH. Perspectives on active video gaming as a new frontier in accessible physical activity for youth with physical disabilities. Phys Ther 2016 Apr;96(4):521-532. [doi: 10.2522/ptj.20140258] [Medline: 26316530]
15. Malone LA, Thirumalai M, Padalabalanarayanan S, Neal WN, Bowman S, Mehta T. Energy expenditure and enjoyment during active video gaming using an adapted Wii Fit balance board in adults with physical disabilities: observational study. JMIR Serious Games 2019 Feb 1;7(1):e11326. [doi: 10.2196/11326] [Medline: 30707098]
16. Lai B, Davis D, Narasaki-Jara M, et al. Feasibility of a commercially available virtual reality system to achieve exercise guidelines in youth with spina bifida: mixed methods case study. JMIR Serious Games 2020 Sep 3;8(3):e20667. [doi: 10.2196/20667] [Medline: 32880577]
17. Lindsay-Smith G, Eime R, O'Sullivan G, Harvey J, van Uffelen JGZ. A mixed-methods case study exploring the impact of participation in community activity groups for older adults on physical activity, health and wellbeing. BMC Geriatr 2019 Sep 2;19(1):243. [doi: 10.1186/s12877-019-1245-5] [Medline: 31477054]
18. Wang G, Xiang R, Yang X, Tan L. Digital technology empowers exercise health management in older adults: a systematic review and meta-analysis of the effects of mHealth-based interventions on physical activity and body composition in older adults. Front Public Health 2025 Oct 1;13:1661028. [doi: 10.3389/fpubh.2025.1661028] [Medline: 41103467]
19. Alley SJ, Waters KM, Parker F, et al. The effectiveness of digital physical activity interventions in older adults: a systematic umbrella review and meta-meta-analysis. Int J Behav Nutr Phys Act 2024 Dec 18;21(1):144. [doi: 10.1186/s12966-024-01694-4] [Medline: 39696583]
20. Dubbeldam R, Stemplewski R, Pavlova I, et al. Technology-assisted physical activity interventions for older people in their home-based environment: scoping review. JMIR Aging 2025 Sep 15;8:e65746. [doi: 10.2196/65746] [Medline: 40952779]
21. Dawson R, Oliveira JS, Kwok WS, et al. Exercise interventions delivered through telehealth to improve physical functioning for older adults with frailty, cognitive, or mobility disability: a systematic review and meta-analysis. Telemed J E Health 2024 Apr;30(4):940-950. [doi: 10.1089/tmj.2023.0177] [Medline: 37975811]
22. Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. J Geriatr Phys Ther 2018;41(1):49-61. [doi: 10.1519/JPT.0000000000000095] [Medline: 27362526]

XSL•FO
RenderX

23. Chen PJ, Hsu HF, Chen KM, Belcastro F. VR exergame interventions among older adults living in long-term care facilities: a systematic review with meta-analysis. Ann Phys Rehabil Med 2023 Apr;66(3):101702. [doi: 10.1016/j.rehab.2022.101702] [Medline: 36028201]
24. Donath L, Rössler R, Faude O. Effects of virtual reality training (exergaming) compared to alternative exercise training and passive control on standing balance and functional mobility in healthy community-dwelling seniors: a meta-analytical review. Sports Med 2016 Sep;46(9):1293-1309. [doi: 10.1007/s40279-016-0485-1] [Medline: 26886474]
25. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inform 2016 Jan;85(1):1-16. [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
26. Malone LA, Davlyatov GK, Padalabalanarayanan S, Thirumalai M. Active video gaming using an adapted gaming mat in youth and adults with physical disabilities: observational study. JMIR Serious Games 2021 Aug 26;9(3):e30672. [doi: 10.2196/30672] [Medline: 34435962]
27. Malone LA, Mehta T, Mendonca CJ, Mohanraj S, Thirumalai M. A prospective non-randomized feasibility study of an online membership-based fitness program for promoting physical activity in people with mobility impairments. Pilot Feasibility Stud 2024 Aug 2;10(1):104. [doi: 10.1186/s40814-024-01528-x] [Medline: 39095876]
28. Bauer MS, Kirchner J. Implementation science: What is it and why should I care? Psychiatry Res 2020 Jan;283:112376. [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]
29. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
30. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4(1). [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
31. Cardona MI, Monsees J, Schmachtenberg T, Grünewald A, Thyrian JR. Implementing a physical activity project for people with dementia in Germany-Identification of barriers and facilitator using consolidated framework for implementation research (CFIR): a qualitative study. PLoS One 2023;18(8):e0289737. [doi: 10.1371/journal.pone.0289737] [Medline: 37556503]
32. Bhardwaj A, FitzGerald C, Graham M, MacFarlane A, Kennedy N, Toomey CM. Barriers and facilitators to implementation of an exercise and education programme for osteoarthritis: a qualitative study using the consolidated framework for implementation research. Rheumatol Int 2024 Jun;44(6):1035-1050. [doi: 10.1007/s00296-024-05590-9] [Medline: 38649534]
33. Boxall C, Dennison L, Miller S, et al. Implementing a digital physical activity intervention for older adults: qualitative study. JMIR Aging 2025 Aug 21;8:e64953. [doi: 10.2196/64953] [Medline: 40840434]
34. Cooper J, Murphy J, Woods C, et al. Barriers and facilitators to implementing community-based physical activity interventions: a qualitative systematic review. Int J Behav Nutr Phys Act 2021 Sep 7;18(1):118. [doi: 10.1186/s12966-021-01177-w] [Medline: 34493306]
35. Anguera MT, Blanco-Villaseñor A, Losada JL, Sánchez-Algarra P, Onwuegbuzie AJ. Revisiting the difference between mixed methods and multimethods: is it all in the name? Qual Quant 2018 Nov;52(6):2757-2770. [doi: 10.1007/s11135-018-0700-2]
36. Mohanraj S, Malone LA, Mendonca CJ, Thirumalai M. Development and formative evaluation of a virtual exercise platform for a community fitness center serving individuals with physical disabilities: mixed methods study. JMIR Form Res 2023 Dec 15;7:e49685. [doi: 10.2196/49685] [Medline: 38100173]
37. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2021;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]
38. Braun V, Clarke V, Weate P. Using thematic analysis in sport and exercise research. In: Routledge Handbook of Qualitative Research in Sport and Exercise: Routledge; 2016:191-205.
39. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
40. Saxena S, Bhan A, Fleming W, Singh A, Tugnawat D. Acceptability, feasibility and appropriateness of care of patients with depression at Health and Wellness Centre in a district of India. Prev Med Res Rev 2024;1(1):41-46. [doi: 10.4103/PMRR.PMRR_77_23]
41. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
42. Smith B, Caddick N. Qualitative methods in sport: a concise overview for guiding social scientific sport research. Asia Pac J Sport Soc Sci 2012;1(1):60-73. [doi: 10.1080/21640599.2012.701373]
43. Neal W. Figure 3. Implementation barriers and facilitators coded into consolidated framework for implementation research (CFIR) domains and constructs. "–" indicates barriers, "+" indicates facilitators. bioRender. 2026. URL: https://BioRender.com/0rybw39 [accessed 2026-03-02]
44. Rimmer JH, Riley B, Wang E, Rauworth A, Jurkowski J. Physical activity participation among persons with disabilities: barriers and facilitators. Am J Prev Med 2004 Jun;26(5):419-425. [doi: 10.1016/j.amepre.2004.02.002] [Medline: 15165658]

XSL•FO
RenderX

45. Hung L, Mann J, Wallsworth C, et al. Facilitators and barriers to using virtual reality and its impact on social engagement in aged care settings: a scoping review. Gerontol Geriatr Med 2023;9:23337214231166355. [doi: 10.1177/23337214231166355] [Medline: 37020921]
46. Schroeder K, Deatrick JA, Klusaritz H, et al. Using a community workgroup approach to increase access to physical activity in an underresourced urban community. Health Promot Pract 2020 Jan;21(1):5-11. [doi: 10.1177/1524839919867649] [Medline: 31423845]
47. Zimmer C, McDonough MH, Hewson J, et al. Experiences with social participation in group physical activity programs for older adults. J Sport Exerc Psychol 2021 Jun 24;43(4):335-344. [doi: 10.1123/jsep.2020-0335] [Medline: 34167084]
48. Dal Bello-Haas VPM, O'Connell ME, Morgan DG, Crossley M. Lessons learned: feasibility and acceptability of a telehealth-delivered exercise intervention for rural-dwelling individuals with dementia and their caregivers. Rural Remote Health 2014;14(3):2715. [Medline: 25081991]
49. Khoshrounejad F, Hamednia M, Mehrjerd A, et al. Telehealth-based services during the COVID-19 pandemic: a systematic review of features and challenges. Front Public Health 2021;9:711762. [doi: 10.3389/fpubh.2021.711762] [Medline: 34350154]
50. Mitzner TL, Savla J, Boot WR, et al. Technology adoption by older adults: findings from the PRISM trial. Gerontologist 2019 Jan 9;59(1):34-44. [doi: 10.1093/geront/gny113] [Medline: 30265294]

Abbreviations

AIM: Acceptability of Intervention Measure
CFIR: Consolidated Framework for Implementation Research
FIM: Feasibility of Intervention Measure
IAM: Intervention Appropriateness Measure
LOF: Lakeshore Online Fitness
VR: virtual reality

Edited by S Munce; submitted 24.Jun.2025; peer-reviewed by C Idiong, E Dove, J Ambros-Antemate, O Abasilim; accepted 26.Jan.2026; published 24.Mar.2026.

Please cite as:
Neal WN, Malone LA
Technology-Driven Group Exercise Program Implementation in an Underserved Community: Multimethod Retrospective Evaluation Study
JMIR Rehabil Assist Technol 2026;13:e79598
URL: https://rehab.jmir.org/2026/1/e79598
doi:10.2196/79598

© Whitney N Neal, Laurie A Malone. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Adherence to a Digital Knee Rehabilitation Platform Among Patients With Knee Osteoarthritis and Anterior Cruciate Ligament Reconstruction in Hong Kong: Qualitative Study

Nipuna Cooray[1], PhD; Stacie Powell[1], PhD; Mingqian Yu[2], MSc; Michael Tim Yun Ong[2], BSc, MBChB, MSc; Julie Brown[1], PhD

[1]Injury Division, Faculty of Medicine and Health, University of New South Wales Sydney, The George Institute for Global Health, Level 8, Health Translation Hub, 55 Botany Street, Randwick, New South Wales, Australia

[2]Department of Orthopaedics & Traumatology, Chinese University of Hong Kong, Hong Kong SAR, China (Hong Kong)

Corresponding Author:
Nipuna Cooray, PhD
Injury Division, Faculty of Medicine and Health, University of New South Wales Sydney, The George Institute for Global Health, Level 8, Health Translation Hub, 55 Botany Street, Randwick, New South Wales, Australia

Abstract

Background: Exercise therapy is fundamental to rehabilitation for knee osteoarthritis and anterior cruciate ligament (ACL) reconstruction, yet adherence to prescribed exercise typically declines once clinical supervision ends. Digital rehabilitation platforms offer a promising means of supporting sustained exercise adherence, but qualitative evidence on how patients experience these platforms in real-world clinical practice remains limited, particularly in non-Western health care contexts.

Objective: This study aimed to explore how patients with different knee conditions experienced the Healthy Knees digital rehabilitation platform in Hong Kong and to identify the factors shaping their platform engagement and exercise adherence.

Methods: A qualitative design was adopted using reflexive thematic analysis. Fifteen adults (9 with ACL, 6 with osteoarthritis) who had been prescribed the Healthy Knees web-based platform at Prince of Wales Hospital participated in semistructured, in-person interviews (30 - 45 min). Interviews were conducted in Cantonese or Mandarin, transcribed verbatim, translated into English, and analyzed inductively. Ethics approval was obtained from the Chinese University of Hong Kong and the University of New South Wales.

Results: Participants were aged 21 to 79 years, with most being male (11/15). Younger participants were predominantly patients with postoperative ACL, while older participants were predominantly patients with preoperative osteoarthritis. Three interrelated themes were identified, collectively describing the fit between the platform and participants' contexts. Content fit captured the alignment between exercise content and rehabilitation needs; participants across both groups perceived substantial overlap with existing physiotherapy, and content was often mismatched to their recovery stage. Motivational fit captured the alignment between platform support features and motivational needs; pain functioned as both a driver and a deterrent to exercise, and participants ranged from highly self-directed to reliant on external scaffolding, not following a simple age pattern. Access fit captured the alignment between the platform's delivery mechanism and participants' technological circumstances; QR code–dependent access, absence of a dedicated mobile app, and display issues created friction that led several participants to migrate to alternative resources, maintaining exercise adherence while abandoning platform engagement.

Conclusions: Adherence to digital knee rehabilitation was shaped by the degree of fit between the platform and users' contexts across content, motivational, and access dimensions. When access fit failed, participants often substituted alternative exercise resources rather than ceasing exercise entirely, highlighting a distinction between platform engagement and exercise adherence. As the sample's clinical and demographic characteristics were closely linked, these findings should not be interpreted as diagnostic comparisons between ACL and osteoarthritis populations but as patterns shaped by the recovery phase and age. These findings suggest that digital rehabilitation platforms should incorporate adaptive content aligned with the recovery stage, integrated feedback mechanisms, and reduced access friction to sustain platform engagement within an ecosystem of competing alternatives.

(JMIR Rehabil Assist Technol 2026;13:e87471) doi:10.2196/87471

KEYWORDS

digital rehabilitation; exercise adherence; knee osteoarthritis; anterior cruciate ligament; thematic analysis; Hong Kong

XSL•FO
RenderX

Introduction

Exercise is widely recognized as a key approach in managing knee osteoarthritis and improving recovery following knee surgeries, such as total knee replacement (TKR) and anterior cruciate ligament (ACL) reconstruction [1]. Therapeutic exercise programs are critical in reducing pain, improving joint function, and enhancing overall quality of life for individuals with knee impairments [2].

Despite the proven benefits of these programs, maintaining exercise for patients remains a significant challenge. Exercise adherence usually declines once clinical supervision tapers off [3]. A decline in exercise can lead to suboptimal recovery outcomes, prolonged rehabilitation periods, and increased health care costs [4,5].

Digital rehabilitation interventions—including web-based platforms and mobile apps—have emerged as promising tools to support sustained exercise adherence beyond the clinic [6]. These platforms can offer flexible access to exercise guidance, structured programs, and remote monitoring capabilities [7]. Yet, the effectiveness of such interventions depends on whether patients engage with and adhere to them in practice [8]. Adherence itself is a complex, multidimensional phenomenon: the World Health Organization has argued for a systems approach that considers socioeconomic, health system, condition-related, therapy-related, and patient-related factors, rather than treating nonadherence as a patient-driven problem [9]. Understanding how digital rehabilitation interventions are experienced in real-world settings—and why users persist with, adapt, or abandon them—is therefore essential for informing intervention design [10].

A growing body of quantitative research has examined digital rehabilitation for knee conditions, with evidence suggesting that digital adjuncts can improve short-term adherence, though benefits may diminish over time [11]. However, qualitative research exploring patients' lived experiences of digital knee rehabilitation remains limited, particularly in non-Western health care contexts. Existing qualitative work has identified broad themes such as the importance of personalization, usability, and motivational support [6], but it is not well understood how patients with differing clinical and demographic profiles experience digital knee rehabilitation platforms in practice, nor how their experiences shape adherence in practice. There is also limited evidence from contexts such as Hong Kong, where long public hospital waiting times for elective orthopedic surgery [12], significant physiotherapy workforce shortages, and the growing availability of alternative digital health resources create a distinctive landscape for digital rehabilitation adoption.

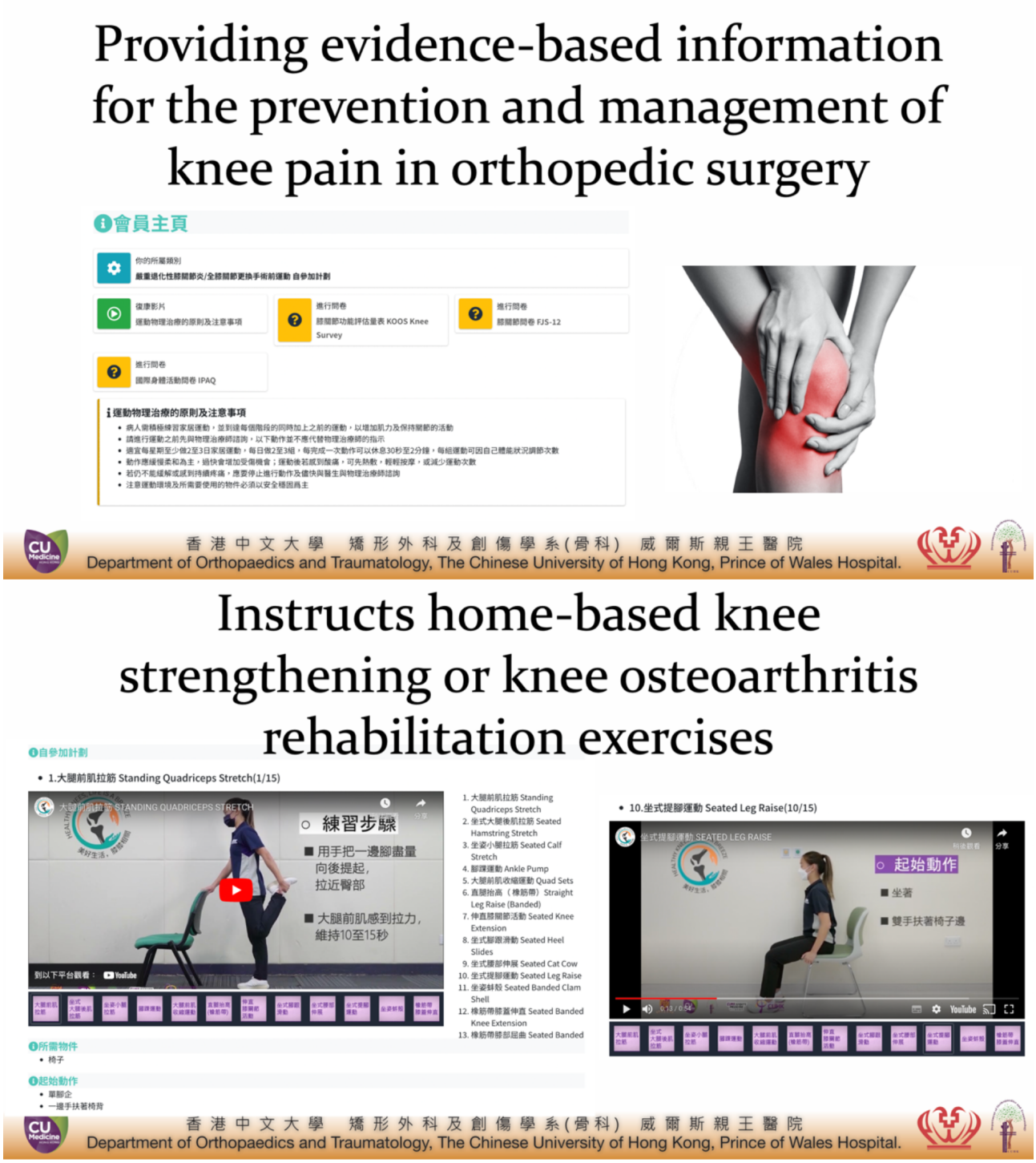
The Healthy Knees intervention (Figure 1) is a web-based rehabilitation platform developed by the Department of Orthopaedics and Traumatology at the Chinese University of Hong Kong. It provides evidence-based information for the prevention and management of knee pain in orthopedic settings and delivers structured, home-based rehabilitation exercises for patients with knee osteoarthritis and those recovering from ACL surgery. The platform includes video demonstrations of strengthening and mobility exercises, educational resources on knee health, and progress monitoring tools such as patient-reported outcome surveys. The platform uses online questionnaires to monitor at-risk patients. It is accessed via a web browser using a QR code provided by the treating clinician; no dedicated mobile app is available. The platform does not include automated adherence-support features, such as goal-setting tools, personalized feedback, or social comparison functions. It is intended to support remote rehabilitation by improving access to exercise guidance outside conventional in-person physiotherapy services.

In this study, adherence refers to the extent to which participants performed the prescribed exercises after accessing them through the Healthy Knees platform. This encompasses both engaging with the platform to access exercise content and subsequently completing exercises as intended. The study did not capture objective platform usage metrics such as login frequency or session duration; rather, adherence was explored through participants' self-reported accounts of their exercise behavior in qualitative interviews. The term "adherence" is used consistently throughout this paper; where a distinction is necessary, "platform engagement" refers specifically to interaction with the digital platform, while "exercise adherence" refers to the performance of prescribed exercises.

This study qualitatively examines adherence to the Healthy Knees digital rehabilitation platform among individuals with different knee conditions in Hong Kong. The aim is to explore how patients experience the platform in real-world clinical practice and to identify the factors that shape sustained engagement and exercise behavior across different clinical profiles.

XSL•FO
RenderX

Figure 1. Healthy knees digital intervention.



Methods

Overview

This qualitative study used reflexive thematic analysis following the 6-phase process described by Braun and Clarke [13,14]. Trustworthiness was addressed using the strategies outlined by Nowell et al [15], including prolonged engagement with the data and peer debriefing. The study was planned and designed by NC, SP, JB, and MTYO. NC and SP trained MY, who conducted the interviews. The researchers, except for MTYO, who was the consulting clinician, had no prior relationships with the participants. The study was reported using the COREQ (Consolidated Criteria for Reporting Qualitative Research). The study was conducted at Prince of Wales Hospital in Hong Kong.

The research team brought diverse disciplinary perspectives to the study. NC is a digital health researcher with no clinical relationship with the participants. SP contributed qualitative research expertise and had no prior involvement with the Healthy Knees platform or its users. JB provided senior research oversight for the project. MTYO, the consulting orthopedic clinician, introduced patients to the Healthy Knees platform as part of routine clinical care and was involved in study design but did not participate in recruitment, conduct interviews, or

XSL•FO
RenderX

contribute to the analysis. MY, who conducted the interviews, was trained by NC and SP and had no prior relationship with participants. Participants were existing users of the Healthy Knees platform who were invited to participate by MY during their clinic visit. The team acknowledges that participants' awareness of MTYO involvement in their clinical care may have influenced their willingness to participate or the candor of their responses. This should be considered when interpreting the findings.

Participant Selection and Recruitment

Participants were selected using a convenience sampling strategy from patients at Prince of Wales Hospital who had been introduced to the Healthy Knees app by author MO as their consulting clinician. Participants were recruited during their clinic visits. The sample included individuals diagnosed with knee osteoarthritis, those awaiting ACL surgery, and those who had undergone ACL surgery. Inclusion criteria focused on patients who had been introduced to the Healthy Knees app. No explicit exclusion criteria were defined for this study.

Data Collection

Data were collected through in-person, semistructured interviews. A total of 15 interviews were conducted, each lasting around 30 to 45 minutes. The interviews were designed to identify the participants' experiences with the Healthy Knees app.

The research team developed a semistructured interview guide (Multimedia Appendix 1) to facilitate the interviews. The guide began with questions to collect participant demographic and clinical background information, followed by open-ended questions exploring participants' experiences with the Healthy Knees platform, their perceptions of the exercises, barriers and facilitators to adherence, motivations for exercise, and suggestions for platform improvement. Probing questions were used to elicit detailed responses. Prior to data collection, the interview guide was pilot-tested within the research team to ensure clarity and relevance. In practice, the interviewer used the guide flexibly, adapting the order and phrasing of questions in response to participants' accounts, consistent with a semistructured approach.

All interviews were conducted in Cantonese or Mandarin by MY and were audio-recorded. The audio recordings were transcribed verbatim in the original language. The first transcript was translated into English by MY. The remaining transcripts were translated by a professional translation service at Guangdong University of Foreign Studies, a university recognized for its expertise in translation studies and accredited by the International Conference of University Institutes of Translators and Interpreters. No formal back-translation or independent verification of translations was conducted; this is acknowledged as a limitation.

Data Analysis

Data analysis followed the 6-phase process of reflexive thematic analysis [13,14]. NC familiarized with the data by reading the transcripts multiple times and then generated initial codes by identifying key patterns, concepts, and recurring ideas. These codes were grouped into broader categories in the search for potential themes, which were subsequently reviewed, refined, and named to ensure they accurately represented the data and were distinct from one another. Throughout this process, NC, SP, and JB engaged in recurring discussions to critically examine the developing codes and themes, ensuring they were grounded in the data and reflected the range of participant experiences. Where individual accounts diverged from emerging patterns, these were examined closely to assess whether theme boundaries required refinement. During the final phase of analysis, the team identified an overarching pattern connecting the 3 themes, which was developed inductively rather than applied from a pre-existing theoretical framework. The coding and analysis were carried out manually without using software tools. A completed COREQ checklist is provided as Checklist 1.

Ethical Considerations

Ethical approval for the study was obtained from the Chinese University of Hong Kong (2023.29), and external approval was registered with the University of New South Wales (iRECS5785) ethics committees. Informed consent was obtained from all participants through a signed written consent form before participation. Participants were informed of the study's purpose, that participation was voluntary, and that they could withdraw at any time without affecting their clinical care. All interview transcripts were deidentified prior to analysis, with participants assigned alphanumeric codes (eg, P01) in place of identifying information, and stored securely on password-protected institutional servers accessible only to the research team. Participants did not receive any financial or other compensation for taking part in the study.

This project was funded by the Chinese University of Hong Kong-The George Institute Joint Collaborative Research Fund.

Results

Participants

Fifteen participants took part in the study (Table 1). The sample comprised 9 participants with ACL conditions and 6 with knee osteoarthritis. Most were male (11/15). Notably, diagnosis, age, and surgical stage were closely intertwined in this sample: participants with ACL conditions were predominantly younger (age range 21 - 52 y) and postoperative (8/9), while participants with osteoarthritis were predominantly older (age range 63 - 79 y) and preoperative (5/6). This overlap means that differences observed between the 2 groups cannot be attributed to diagnosis alone and may reflect age, surgical stage, or their combination. Throughout the "Results" section, participants are described by their experiential profile (eg, "younger, postoperative participants") with the associated diagnosis noted in parentheses.

XSL•FO
RenderX

Table . Participant demographics[a].

Characteristic	ACL[b] (n=9)	Osteoarthritis (n=6)	Total (n=15)
Surgery status, n			
Presurgery	1	5	6
Postsurgery	8	1	9
Sex, n			
Male	8	3	11
Female	1	3	4
Age (y), range	21 - 52	63 - 79	21 - 79

[a]Detailed clinical variables, such as Kellgren-Lawrence grade for knee osteoarthritis and time since surgery for postoperative participants, were not systematically collected in this study.

[b]ACL: anterior cruciate ligament.

Analysis of the interview data identified 3 interrelated themes that collectively describe the fit between the Healthy Knees platform and participants' contexts. These themes—content fit, motivational fit, and access fit—capture the dimensions along which the platform aligned or misaligned with participants' rehabilitation needs, motivational profiles, and technological circumstances. Across the 3 themes, adherence was shaped not by any single factor but by the degree of alignment across all 3 dimensions.

Content Fit: Alignment Between Platform Exercises and Rehabilitation Needs

This theme captures the extent to which the Healthy Knees exercise content matched participants' perceived rehabilitation needs at their particular stage of recovery. Three patterns emerged: perceived redundancy with existing physiotherapy, temporal mismatch between platform content and recovery stage, and difficulty calibrating between exercises and physical capacity.

Perceived Redundancy With Existing Care

Across both groups, a recurring finding was that participants perceived substantial overlap between the Healthy Knees content and the exercises they were already receiving through physiotherapy or other sources. This perception reduced the platform's added value and contributed to disengagement:

> *I saw that the exercises were similar to those in physical therapy, so I didn't persist.* [P01, male, 76 years, osteoarthritis, preoperative]
>
> *They're the same by 70% to 80%. The physiotherapy exercises consist of stretching routines.* [P09, male, 45 years, ACL, preoperative]
>
> *I am not interested because they're very basic.* [P04, male, 26 years, ACL, postoperative]

Even among participants who found the content acceptable, the platform was often seen as confirming what they already knew rather than offering new guidance. One older participant noted:

> *I've received these conservative treatments and know how to perform these movements. However, I just do them depending on my own abilities.* [P05, male, 77 years, osteoarthritis, preoperative]

Similarly, 1 postoperative participant who had progressed well under the guidance of her personal coach found the platform content comparable to what she was already doing (P08, female, 63 years, osteoarthritis, post-TKR). This pattern was cross-cutting, appearing in younger postoperative and older preoperative participants alike, suggesting that redundancy with existing care was a common barrier regardless of age or diagnosis.

Temporal Mismatch Between Content and Recovery Stage

Among younger, postoperative participants (predominantly ACL), a distinct pattern emerged in which the platform content was seen as appropriate for a particular recovery stage—but not for the stage the participant was at when they accessed it.

> *These exercises are suitable for those who need rehabilitation training immediately after surgery. However, I was not at the rehabilitation stage when I accessed this platform.* [P13, male, 38 years, ACL, approximately 1 year postoperative]

One participant discovered the platform 6 months after surgery, by which time she had already resumed her preinjury physical activities daily and found the early-stage content no longer relevant: "I searched for related videos on YouTube...there weren't many videos in this respect on YouTube and were not so detailed as this platform" (P14, female, 21 years, ACL, postoperative). Another attempted the exercises immediately after injury but found them too painful:

> *I tried to follow your videos to do home exercises immediately after starting the rehabilitation at home. However, doing those movements was so painful that I failed.* [P02, male, 52 years, ACL, postoperative]

A notable counter-case was P07 (male, 35 years, ACL, postoperative), who reviewed later-stage content and found it genuinely challenging:

> *The jumping with the BOSU ball and the single-leg jump are challenging. But I think they are movements at the final stage of the rehabilitation program on your platform. I have not reached that stage currently.*

XSL•FO
RenderX

This suggests that the platform does contain progressive content, but participants who entered at the wrong recovery stage were unlikely to discover or reach it.

Difficulty Calibration

Perceptions of exercise difficulty diverged along the age and surgical-stage lines in this sample. Younger, postoperative participants (predominantly ACL) consistently found the exercises to be below their capacity and preinjury activity levels:

> *Your platform is good. Maybe as an athlete I might think the exercises on the platform are easy, but they are still useful for the elderly.* [P04, male, 26 years, ACL, postoperative]
>
> *Your exercises are too basic for them, but are adequate for me.* [P06, male, ACL, postoperative]

Older, preoperative participants (predominantly osteoarthritis) more frequently reported that specific exercises exceeded their physical capabilities:

> *This movement is challenging for me, but if I find it beneficial, I will strive to do it even despite difficulties. However, the third and fourth movements are beyond my capacity.* [P05, male, 77 years, osteoarthritis, preoperative]
>
> *There were a few exercises that I couldn't do at all, such as bending my feet, stilting my feet.* [P01, male, 76 years, osteoarthritis, preoperative]

However, this pattern was not universal. P11 (male, 79 years, osteoarthritis, preoperative) found all exercises within his capacity, and P08 (female, 63 years, osteoarthritis, post-TKR) noted that exercises initially difficult became manageable over time: "At the beginning, I found some movements difficult, but now I can do them." One participant's experience highlighted a potential safety concern: after stretching for one to two minutes rather than the recommended 10 to 30 seconds, he experienced adverse effects: "I've done them for a week and stopped after my muscles hurt. Maybe I didn't follow their recommended duration" (P09, male, 45 years, ACL, preoperative). This suggests that without adequate guidance on exercise dosage, some users may inadvertently overexert.

Motivational Fit: Alignment Between Platform Support and Motivational Needs

This theme captures how participants' motivational profiles interacted with the platform's support features (or lack thereof) to shape adherence. Two patterns emerged: the complex role of pain as both a driver and deterrent of exercise behavior, and the spectrum between self-directed motivation and reliance on external scaffolding.

Pain as a Double-Edged Motivational Driver

Across both groups, pain emerged not simply as a barrier to exercise but as a motivational force operating in 2 opposing directions. For some participants, the presence of pain drove continued engagement; for others, it triggered disengagement.

Among older, preoperative participants (predominantly osteoarthritis), pain served primarily as a motivator through fear of functional decline. Despite finding exercises difficult, P01 (male, 76 years, osteoarthritis, preoperative) persisted because he feared deterioration: "Not really fond of exercises, but if I don't exercise, my physical function may deteriorate soon." P11 (male, 79 years, osteoarthritis, preoperative) continued exercising because exercises directly alleviated his symptoms: "After doing these exercises, the pains of thighs can be alleviated to some extent." P05 (male, 77 years, osteoarthritis, preoperative) adopted a selective approach, performing exercises that were tolerable and skipping those that caused excessive discomfort:

> *I will skip the movements that are totally beyond my capacity. But I definitely will do these stretches.*

Among younger, postoperative participants (predominantly ACL), the relationship was more paradoxical. P06 (male, ACL, postoperative) explicitly described reduced pain as a reason to stop exercising: "When I feel the therapeutic effect without pain, I tend to skip exercises, otherwise I'm motivated to do more exercises." In this case, the absence of pain—which would typically signal recovery—actually undermined adherence by removing the perceived need for continued exercise. By contrast, P09 (male, 45 years, ACL, preoperative) stopped exercising because pain from overexertion deterred him: "I've done them for a week and stopped after my muscles hurt."

The Healthy Knees platform did not include features to help participants interpret pain in the context of their recovery—for instance, distinguishing expected postexercise discomfort from signals to reduce intensity. In the absence of such guidance, participants developed their own heuristics, with some persisting through pain and others withdrawing from exercise entirely.

Self-Direction Versus External Scaffolding

Participants varied considerably in the degree to which they could sustain exercise independently, and this variation did not follow a simple age or diagnosis pattern.

Several participants described strong internal motivation. P04 (male, 26 years, ACL, postoperative) reported no need for external prompts due to his athletic background. P10 (female, 74 years, osteoarthritis, preoperative) exercised 45 minutes to 1 hour daily, combining exercises sourced from a physiotherapy platform (HA Go) with a community exercise group led by a retired coach:

> *I think at least if you have some kind of this kind of platform or introduce some patient, at least they have the knowledge to know and then they have the freedom whether to follow is up to their decision.*

Her case challenges any simple association between older age and low self-direction.

Others, however, described difficulty maintaining routines without external structure. P03 (male, 21 years, ACL, postoperative) was candid:

> *I don't think I have that high level of self-consciousness. I will have greater drive to perform workouts with companions than alone.*

P06 (male, ACL, postoperative) attributed his nonadherence to "laziness," and P01 (male, 76 years, osteoarthritis, preoperative)

XSL•FO
RenderX

described a cycle of initial effort followed by abandonment: “I can do it for 10, 20 times, and then I gave up.” One participant (P08, female, 63 years, osteoarthritis, post-TKR) resolved this by building a structured external routine—weekly sessions with a coach supplemented by independent practice at home—effectively creating her own scaffolding outside the platform.

Two younger, postoperative participants independently raised the idea that progress tracking and feedback features could bridge the gap between self-direction and external support. P15 (male, 34 years, ACL, postoperative) described this in detail:

> *If I have the record about what exercises I do at what time, the results, are they good or bad, or easy to do? Then I will have more motivation. With something to feedback.*

P13 (male, 38 years, ACL, postoperative) similarly suggested staged content with expected milestones so users could gauge their progress. While this pattern was observed in only 2 participants, both arrived at a convergent suggestion: that the platform’s lack of feedback mechanisms left a motivational gap that neither purely self-directed nor externally prompted approaches could fill.

Access Fit: Alignment Between Platform Design and Users’ Technological Context

This theme captures how the design and delivery mechanism of the Healthy Knees platform interacted with participants’ technological circumstances to shape engagement. Two patterns emerged: access barriers created by the platform’s delivery format and competition from alternative exercise resources that participants found easier to use.

Platform Access Barriers

The Healthy Knees platform is accessed via a web browser using a QR code provided on a printed pamphlet during a clinic visit. It has no dedicated mobile app, no saved login, and no bookmark prompt. Several participants described how this access mechanism created friction that disrupted or prevented engagement.

The most direct account came from P08 (female, 63 years, osteoarthritis, post-TKR), who lost the pamphlet and was unable to return to the platform:

> *Sometimes it was impossible to access your platform without scanning the QR code … I lost the pamphlet; without the QR code, I couldn’t log into the website.*

Rather than abandoning exercise, she switched to YouTube videos and sessions with her coach.

P01 (male, 76 years, osteoarthritis, preoperative) faced a more fundamental barrier: he did not have a mobile data plan, relying on a pay-as-you-go card, and could only access the platform via home Wi-Fi. A nurse assisted him with initial access during his clinic visit, but sustained engagement required connectivity he did not routinely have:

> *My phone doesn’t have internet access. I didn’t sign a service contract.*

P03 (male, 21 years, ACL, postoperative)—a younger participant with no digital literacy difficulties—identified the format itself as the problem and suggested a solution:

> *Maybe it would be more convenient to change your platform into an app. Users can view video as soon as they access the app.*

He noted this would also benefit older users: "Some elderly users may not be able to find the QR code each time, but can always check videos on the app with one click.”

One participant reported a technical display issue that compromised usability:

> *I have an issue with your website on my mobile phone. Only the latter half is visible; the upper portion remains hidden … It’s an iPhone issue.* [P06, male, ACL, postoperative]

These barriers were not exclusively a function of age or digital literacy. While older participants faced additional challenges around connectivity and device familiarity, younger participants also encountered friction from the platform’s web-based, QR code–dependent format. The common thread was that the access mechanism placed the burden of re-entry on the user, and when that burden exceeded a threshold—whether through a lost pamphlet, absent connectivity, or a display bug—engagement ceased.

Competition With Alternative Exercise Resources

A distinctive finding was that when the platform’s access mechanism failed or when content was perceived as insufficient, participants did not simply stop exercising. Instead, many migrated to alternative resources they found more accessible or better suited to their needs.

The most frequently cited alternative was HA Go, a mobile app provided by the Hospital Authority that offers exercise content with daily reminders. P05 (male, 77 years, osteoarthritis, preoperative) described it as functionally comparable to Healthy Knees: “HA Go app has similar functions to your platform and can remind users to do exercises every day.” P10 (female, 74 years, osteoarthritis, preoperative) used HA Go as her primary exercise resource and had not engaged with Healthy Knees at all, supplementing the app with a community exercise group:

> *They gave me also another platform that is belonged to the physiotherapist. And then I follow that kind of exercise. I think there’s some improvement for me.*

P11 (male, 79 years, osteoarthritis, preoperative) similarly used HA Go rather than Healthy Knees for his daily exercise routine, responding to its SMS text messaging reminders.

YouTube was another common alternative. P14 (female, 21 years, ACL, postoperative) searched for ACL recovery exercises on YouTube during her home rehabilitation, though she noted the content was “not so detailed as this platform.” P08 (female, 63 years, osteoarthritis, post-TKR) turned to YouTube after losing access to Healthy Knees. P01 (male, 76 years, osteoarthritis, preoperative) took a different approach entirely, purchasing a stretching board and devising his own exercise routine based on prior physiotherapy knowledge.

XSL•FO
RenderX

This pattern of substitution has important implications. It suggests that the participants in this study were not, on the whole, lacking in motivation to exercise—rather, the Healthy Knees platform was competing for their engagement within an ecosystem of alternative resources, some of which offered lower access friction (HA Go's app format, YouTube's familiarity) or more personalized guidance (personal coaches, community groups). Nonengagement with the platform, in these cases, did not equate to nonadherence to exercise.

Table 2 summarizes the 3 themes describing the fit between the Healthy Knees platform and participants' rehabilitation contexts.

Table . Summary of themes describing the fit between the Healthy Knees platform and participants' rehabilitation contexts.

Theme	Description
Content fit	The degree to which the platform's exercise content matched participants' rehabilitation needs, including perceived redundancy with existing physiotherapy, temporal mismatch between content and recovery stage, and difficulty calibration relative to physical capacity.
Motivational fit	The alignment between the platform's motivational support features and participants' needs, including the dual role of pain as both a driver and deterrent of exercise, and the spectrum between self-directed motivation and reliance on external scaffolding.
Access fit	The alignment between the platform's delivery mechanism and participants' technological circumstances, including access barriers created by the QR code–dependent web format and competition from alternative exercise resources such as HA Go and YouTube.

Across the 3 themes, a common pattern emerged inductively during analysis: adherence to the Healthy Knees platform depended on the degree of fit between the platform and the participant's context across content, motivational, and access dimensions. When fit was adequate across all 3 dimensions, participants maintained both platform engagement and exercise adherence. When any single dimension was misaligned—whether through content that did not match the recovery stage, insufficient motivational scaffolding, or access friction—adherence was disrupted. Notably, the consequences of misfit differed by dimension: content and motivational misfit tended to reduce exercise adherence overall, while access misfit more often redirected exercise behavior toward alternative resources, disrupting platform engagement while preserving exercise adherence. Furthermore, fit was not static: participants' alignment with the platform shifted over the course of their recovery, with content that was initially appropriate becoming redundant or insufficient, and motivational needs evolving as pain and functional capacity changed. These patterns, including the temporal dimension of fit and its implications for platform design, are developed further in the Discussion section.

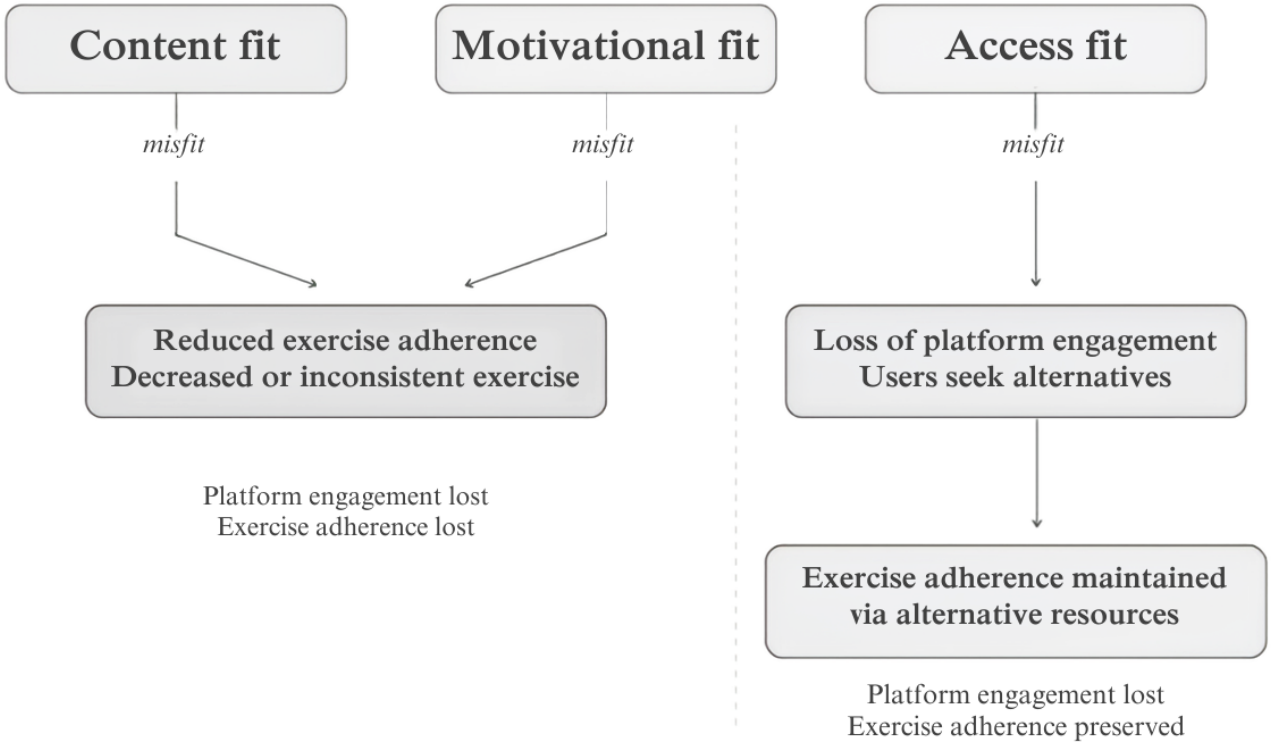
Discussion

Principal Findings

This qualitative study explored adherence to the Healthy Knees digital knee rehabilitation platform among 15 participants in Hong Kong. Three themes were identified—content fit, motivational fit, and access fit—which together describe the alignment between the platform and participants' rehabilitation needs, motivational profiles, and technological circumstances. The central finding is that adherence was not determined by any single factor but by the degree of fit across all 3 dimensions. When fit was adequate across content, motivation, and access, participants maintained both platform engagement and exercise adherence. When any single dimension was misaligned, adherence was disrupted—but the nature of that disruption differed: content and motivational misfit tended to reduce exercise adherence overall, while access misfit more commonly redirected exercise behavior toward alternative resources, preserving exercise adherence while disrupting platform engagement. These predominant patterns are depicted in Figure 2.

XSL•FO
RenderX

Figure 2. Relationship between the 3 fit dimensions and divergent adherence patterns.



Fit as an Organizing Framework for Understanding Digital Rehabilitation Adherence

The concept of fit that emerged from this analysis resonates with established adherence frameworks while offering a more specific account of how adherence operates in digital rehabilitation contexts. The World Health Organization's multidimensional adherence model identifies 5 interacting dimensions—socioeconomic, health system, condition-related, therapy-related, and patient-related factors—and emphasizes that nonadherence is a systemic issue rather than a patient-driven problem [9]. Our findings align with this principle: participants' nonadherence reflected mismatches in the intervention design, health care delivery context, and access mechanisms, not simply a lack of motivation. The fit framework extends this perspective by specifying the dimensions most salient in a digital rehabilitation context and by revealing that different types of misfits produce qualitatively different patterns of nonadherence.

Our findings can also be situated within the Capability, Opportunity, and Motivation Behavior (COM-B) model of behavior change [16], which posits that capability, opportunity, and motivation must all be present for a behavior to occur. Content fit maps broadly onto capability—participants needed exercises that were appropriate to their physical capacity and recovery stage. Motivational fit maps onto motivation—participants varied in the degree to which they could sustain exercise without external scaffolding. Access fit maps onto opportunity—the platform's delivery mechanism either enabled or constrained the possibility of engagement. Consistent with previous applications of COM-B to rehabilitation contexts [17], our findings suggest that interventions addressing only 1 component—for example, providing exercise content without motivational support or accessible delivery—are unlikely to sustain adherence. However, the fit framework adds a temporal dimension that COM-B does not explicitly address: fit is not static but shifts as users progress through recovery, meaning that a platform well-suited to a participant at one stage may become misaligned at another.

Content Fit and the Challenge of Perceived Redundancy

A striking cross-cutting finding was that participants across both groups perceived the Healthy Knees content as substantially overlapping with exercises they were already receiving through in-person physiotherapy or other digital resources. This finding suggests that providing evidence-based content alone may be insufficient to drive engagement; rather, as the person-based approach to digital intervention development argues, content must offer perceived added value from the user's perspective to sustain use [18]. The temporal mismatch finding—in which participants accessed the platform at a recovery stage for which the content was either too advanced or no longer relevant—has important design implications. Previous work on digital rehabilitation has emphasized the importance of tailoring exercise content to the individual [19], but our data suggest that tailoring must account not only for diagnosis and physical capacity but also for the timing of platform introduction relative to the patient's rehabilitation trajectory. The experience of participants who discovered the platform months after surgery, when its early-stage content was no longer relevant, illustrates how the window of optimal content fit can be missed entirely. Conversely, participants who attempted exercises immediately

XSL•FO
RenderX

postinjury found the content too demanding, suggesting that premature introduction can be equally problematic.

The safety concern raised by one participant, who overexerted during stretching exercises and experienced adverse effects, highlights the limitations of a platform that provides exercise demonstrations without guidance on dosage or adaptation. Digital rehabilitation platforms that lack real-time feedback or clear dosage parameters may inadvertently place responsibility for safe exercise progression on the user, which is consistent with concerns raised in the wider literature on unsupervised home-based exercise [20].

Motivational Fit and the Dual Role of Pain

The finding that pain operated simultaneously as a motivator and a deterrent adds nuance to existing understandings of adherence in musculoskeletal rehabilitation. Previous research has identified pain as a barrier to exercise adherence [21], but our findings suggest a more complex dynamic in which the presence or absence of pain served as an implicit indicator of need. For some participants, feeling pain signaled that exercise was necessary; for others, the absence of pain removed the perceived rationale for continued effort. This paradox has been observed in medication adherence research, where symptom improvement is a well-documented trigger for discontinuation [9], but it has received less attention in the digital rehabilitation literature.

The spectrum between self-directed motivation and reliance on external scaffolding also challenges simple demographic assumptions. Some of the most self-directed exercisers in the sample were older preoperative participants, while some younger postoperative participants described significant difficulty maintaining routines without companions or external prompts. This finding is consistent with research suggesting that health literacy, prior exercise experience, and self-efficacy may be more important predictors of sustained engagement than age alone [22]. The convergent suggestion from 2 participants that progress tracking and feedback features could support adherence aligns with evidence that self-monitoring is among the most effective behavior change techniques in digital health interventions [23].

Access Fit and the Ecosystem of Alternative Resources

The finding that participants migrated to alternative exercise resources rather than ceasing exercise entirely is perhaps the most distinctive contribution of this study. Recent work on engagement with digital health interventions has drawn a useful distinction between microengagement—interaction with the digital tool itself—and macroengagement—engagement with the target health behavior [8,24]. Our findings illustrate this distinction empirically: several participants maintained macroengagement (exercise adherence) while abandoning microengagement (platform engagement), substituting YouTube videos, the HA Go app, personal coaches, or self-devised routines.

This pattern of substitution reflects the competitive landscape in which digital rehabilitation platforms operate. In Hong Kong, the Hospital Authority's HA Go app provides a widely available, app-based alternative with built-in reminder features. YouTube offers a vast library of rehabilitation exercise content. Personal coaches and community exercise groups provide social support and accountability. The Healthy Knees platform, delivered via a QR code–dependent web browser with no dedicated app, no saved login, and no automated reminders, faced a significant access disadvantage relative to these alternatives. This is consistent with broader evidence that the usability and convenience of digital health tools are critical determinants of sustained engagement, particularly when users have readily available alternatives [25].

The Hong Kong Context

Several features of the Hong Kong health care system shaped the adherence patterns observed in this study. Waiting times for elective orthopedic surgery in the public system can extend to several years [12], meaning that participants with preoperative osteoarthritis in this study were managing their condition over prolonged periods with limited physiotherapy access. Participants described physiotherapy sessions conducted at long intervals that felt insufficient, with limited opportunities for 1-on-1 attention. This context makes digital rehabilitation particularly relevant as a means of bridging gaps in formal care—but it also means the platform must sustain engagement over long waiting periods, a challenge that the current static content model is unlikely to meet.

The cost differential between public and private physiotherapy in Hong Kong is substantial. Participants in this study were drawn from a public hospital where physiotherapy resources were described as limited. Digital platforms offer a potential solution to this resource constraint, but our findings suggest that a platform perceived as duplicating rather than extending existing care will not be adopted, regardless of its clinical quality.

While these health care system factors are specific to Hong Kong, several of the adherence patterns observed in this study may extend to other contexts. Context-specific findings include the reliance on a QR code–based web platform as the primary access mechanism and the availability of HA Go as a readily accessible alternative provided by the Hospital Authority. However, the challenge of perceived redundancy with existing physiotherapy, the temporal mismatch between static platform content and evolving rehabilitation needs, and the tendency for users to substitute alternative resources when access friction is high are likely relevant wherever digital rehabilitation platforms operate alongside established care pathways.

Strengths and Limitations

This study has several strengths. To the authors' knowledge, it is among the first qualitative studies to explore digital knee rehabilitation adherence in Hong Kong, providing context-specific insights into a health care system with distinctive resource constraints and cultural characteristics. The postimplementation design captured real-world experiences of participants who had used the platform in practice, rather than hypothetical responses to a prototype. The inclusion of both ACL and osteoarthritis participants allowed comparison across different experiential profiles, revealing both cross-cutting and group-specific patterns.

XSL•FO
RenderX

Several limitations must be acknowledged. Diagnosis, age, and surgical stage were closely intertwined in this sample: participants with ACL were predominantly younger and postoperative, while participants with osteoarthritis were predominantly older and preoperative. Consequently, differences observed between groups cannot be attributed to diagnosis alone, and may reflect age, perioperative stage, or their combination. Detailed clinical descriptors—such as osteoarthritis severity grading, time since surgery, or concomitant injuries—were not systematically collected, which limits the degree to which findings can be linked to specific clinical characteristics. Future research should include such variables to enable more precise stratification.

The study used a convenience sample drawn from a single hospital, potentially limiting transferability, although there is no reason to expect patients at this site to differ substantially from those at other public hospitals in Hong Kong. The platform did not capture objective usage data (eg, login frequency, session duration), so all engagement information is based on participants' self-reported accounts. Transcripts were translated from Cantonese or Mandarin into English, which may have introduced minor shifts in meaning despite efforts to ensure accuracy through the use of a professional translation service. Data collection did not pursue theoretical saturation, and with 15 participants, some perspectives may not have been captured. While our sample included diverse demographic and clinical profiles, future research with larger and more varied samples is needed to explore additional viewpoints.

Conclusions

Adherence to digital knee rehabilitation is best understood not as a property of the individual or the platform, but as a product of the ongoing fit between the 2 across content, motivational, and access dimensions. This fit is dynamic, shifting as users progress through recovery. Digital rehabilitation platforms that offer static content through high-friction access mechanisms risk losing users to more accessible alternatives—without necessarily losing them to inactivity. Designing for sustained fit, particularly through adaptive content, integrated feedback, and reduced access friction, may be more effective than targeting adherence as an isolated behavior. Future research should test this framework longitudinally and with objective usage data to identify when and how fit breaks down across recovery trajectories.

Acknowledgments

During the preparation of this manuscript, the authors used Grammarly and Anthropic's Claude to refine grammar, language, and clarity of the draft. The authors reviewed and edited all output, verified the accuracy of the content and references, and take full responsibility for the final manuscript. Neither tool is listed as an author.

Funding

This project was funded by the Chinese University of Hong Kong-The George Institute Joint Collaborative Research Fund.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview discussion guide.
[DOCX File, 20 KB - rehab_v13i1e87471_app1.docx]

Checklist 1
COREQ checklist.
[DOCX File, 12 KB - rehab_v13i1e87471_app2.docx]

References

1. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev 2015 Jan 9;1(1):CD004376. [doi: 10.1002/14651858.CD004376.pub3] [Medline: 25569281]
2. Zeng CY, Zhang ZR, Tang ZM, Hua FZ. Benefits and mechanisms of exercise training for knee osteoarthritis. Front Physiol 2021;12:794062. [doi: 10.3389/fphys.2021.794062] [Medline: 34975542]
3. Pisters MF, Veenhof C, Schellevis FG, Twisk JWR, Dekker J, De Bakker DH. Exercise adherence improving long‐term patient outcome in patients with osteoarthritis of the hip and/or knee. Arthritis Care Res 2010 Aug;62(8):1087-1094. [doi: 10.1002/acr.20182]
4. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021 Feb 19;18(4):2023. [doi: 10.3390/ijerph18042023] [Medline: 33669679]
5. Le TK, Montejano LB, Cao Z, Zhao Y, Ang D. Health care costs in US patients with and without a diagnosis of osteoarthritis. J Pain Res 2012;5:23-30. [doi: 10.2147/JPR.S27275] [Medline: 22328832]

XSL•FO
RenderX

6. Patten RK, Tacey A, Pile R, et al. Digital self-management interventions for osteoarthritis: a systematic scoping review of intervention characteristics, adherence and attrition. Arch Public Health 2022 Mar 31;80(1):103. [doi: 10.1186/s13690-022-00854-x] [Medline: 35361270]
7. Shah N, Costello K, Mehta A, Kumar D. Applications of digital health technologies in knee osteoarthritis: narrative review. JMIR Rehabil Assist Technol 2022 Jun 8;9(2):e33489. [doi: 10.2196/33489] [Medline: 35675102]
8. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
9. Adherence to long-term therapies: evidence for action. : World Health Organization; 2003 URL: https://iris.who.int/server/api/core/bitstreams/121c6b73-8651-442f-9560-e3b1d9c8a75c/content [accessed 2026-06-02]
10. Murray E, Hekler EB, Andersson G, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851. [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
11. Lang S, McLelland C, MacDonald D, Hamilton DF. Do digital interventions increase adherence to home exercise rehabilitation? A systematic review of randomised controlled trials. Arch Physiother 2022 Oct 3;12(1):24. [doi: 10.1186/s40945-022-00148-z] [Medline: 36184611]
12. Tsui YM, Fong BYF. Waiting time in public hospitals: case study of total joint replacement in Hong Kong. Public Adm Policy 2018 Nov 30;21(2):120-133. [doi: 10.1108/PAP-10-2018-009]
13. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
14. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
15. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1609406917733847. [doi: 10.1177/1609406917733847]
16. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
17. Paterson S, Dawes H, Winward C, et al. Use of the capability, opportunity and motivation behaviour model (COM-B) to understand interventions to support physical activity behaviour in people with stroke: an overview of reviews. Clin Rehabil 2024 Apr;38(4):543-557. [doi: 10.1177/02692155231224365] [Medline: 38192225]
18. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30. [doi: 10.2196/jmir.4055] [Medline: 25639757]
19. McHugh GA, Lavender EC, Bennell KL, et al. A qualitative evaluation of two electronic-rehabilitation programmes for managing persistent knee pain. Musculoskeletal Care 2025 Mar;23(1):e70051. [doi: 10.1002/msc.70051] [Medline: 39788887]
20. von Heideken J, Chowdhry S, Borg J, James K, Iversen MD. Reporting of harm in randomized controlled trials of therapeutic exercise for knee osteoarthritis: a systematic review. Phys Ther 2021 Oct 1;101(10):pzab161. [doi: 10.1093/ptj/pzab161] [Medline: 34180534]
21. Jack K, McLean SM, Moffett JK, Gardiner E. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228. [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
22. Shaw JF, Pilon S, Vierula M, McIsaac DI. Predictors of adherence to prescribed exercise programs for older adults with medical or surgical indications for exercise: a systematic review. Syst Rev 2022 Apr 29;11(1):80. [doi: 10.1186/s13643-022-01966-9] [Medline: 35488307]
23. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
24. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions. JMIR Form Res 2019 Oct 10;3(4):e14052. [doi: 10.2196/14052] [Medline: 31603427]
25. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]

Abbreviations

ACL: anterior cruciate ligament
COM-B: Capability, Opportunity, and Motivation Behavior
COREQ: Consolidated Criteria for Reporting Qualitative Research
TKR: total knee replacement

XSL•FO
RenderX

Edited by S Munce; submitted 09.Nov.2025; peer-reviewed by N Shah, Z Liu; revised version received 12.May.2026; accepted 13.May.2026; published 26.Jun.2026.

Please cite as:
Cooray N, Powell S, Yu M, Ong MTY, Brown J
Adherence to a Digital Knee Rehabilitation Platform Among Patients With Knee Osteoarthritis and Anterior Cruciate Ligament Reconstruction in Hong Kong: Qualitative Study
JMIR Rehabil Assist Technol 2026;13:e87471
URL: https://rehab.jmir.org/2026/1/e87471
doi:10.2196/87471

© Nipuna Cooray, Stacie Powell, Mingqian Yu, Michael Tim Yun Ong, Julie Brown. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Aids for Social Media in Adults With Traumatic Brain Injury: Longitudinal Evaluation Study

Dakota Sullivan[1], MS; Nirav N Patel[2], BS; Lisa Kakonge[3], PhD; Callie Y Kim[1], MS; Yaxin Hu[1], PhD; Hailey L Johnson[1], BS; Catalina L Toma[1], PhD; Melissa Duff[2], PhD; Lyn Turkstra[3], PhD; Bilge Mutlu[1], PhD

[1]Department of Computer Sciences, University of Wisconsin–Madison, Madison, WI, United States

[2]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, United States

[3]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Dakota Sullivan, MS
Department of Computer Sciences
University of Wisconsin–Madison
1205 University Avenue
Madison, WI, 53706
United States
Phone: 1 608 262 1204
Email: dsullivan8@wisc.edu

Abstract

Background: Digital social media aids are features that enhance accessibility to social media for users with disabilities. One group of adults with disabilities who have reported barriers to social media access is adults with traumatic brain injury (TBI). Results of previous studies showed that adults with TBI want to engage in social media but encounter significant barriers to doing so, due in large part to TBI-related cognitive and communication challenges. We previously collaborated with adults with TBI to develop 5 aids to increase social media access. Aids were implemented as Chrome extensions that modified user interface functions and visuals on Facebook. The present study was the next step: testing the use of the 5 aids by a novel group of adults with TBI longitudinally and in naturalistic settings.

Objective: This study aimed to determine (1) how participants perceived each aid (ie, in terms of satisfaction, usability, likes, and dislikes), (2) how aids influenced participants' well-being, and (3) which aid features participants rated as most and least useful.

Methods: We recruited 39 adult Facebook users with TBI to take part in an approximately 1-month-long within-subject evaluation study of Facebook browser aids. Each aid was used for 3 days, followed by an evaluation and a subsequent 2-day no-aid use period. Participants evaluated each of the 5 unique browser aids in a randomized order.

Results: Users reported the aids to be beneficial, especially aids that reduced visual clutter and information overload. Users also reported increased life satisfaction after using the aids (P=.02).

Conclusions: These results provide proof of concept for the types of digital aids that could improve access to social media for users with TBI. We advocate including aids like these across all social media platforms so users with TBI can access the benefits of engagement in social media with fewer barriers.

(JMIR Rehabil Assist Technol 2026;13:e83639) doi:10.2196/83639

KEYWORDS

accessibility; assistive technology; human factors; social media; TBI; traumatic brain injury

Introduction

Overview

Moderate or severe traumatic brain injury (TBI) can cause cognitive and communicative challenges that profoundly affect a survivor's activities of daily living and social networks [1-3]. Social media platforms have the potential to help individuals with TBI maintain and expand their social networks, join communities with people sharing similar experiences post injury, and acquire relevant information [4,5]. These platforms can also create spaces and opportunities for self-expression and identity

XSL•FO
RenderX

construction post injury [6]. Individuals with TBI, however, report significant challenges when using social media platforms.

Building upon the existing literature on TBI-related social media challenges [6,7], our team designed and developed an accessibility toolkit to provide cognitive and communication support for individuals with TBI [8]. The toolkit includes 5 browser extensions that modify the Facebook interface for greater accessibility and usability. Here, we build on this prior work by testing how these tools perform during real-world, longitudinal use.

Cognitive and Communication Impairments After TBI

TBI can cause a myriad of impairments in cognition, communication, and social interaction, as well as sensorimotor and psychological functions [1-3,9]. The most commonly impaired cognitive functions are executive functions and memory [2,10]. Executive function impairments can limit the individual's ability to manage thoughts and feelings and self-regulate behaviors [11]. Memory impairments are typically in working memory, the ability to hold and manipulate information temporarily, and declarative learning, the ability to consciously encode new information. The combination of these impairments often leads to difficulties in higher-order cognitive functions, such as reasoning and decision-making [12,13].

Social cognition impairments have been well documented in adults with moderate to severe TBI, including deficits on affect recognition and Theory of Mind tasks [14-22]. While researchers continue to debate whether these impairments are secondary to impairments in executive function (eg, if impaired cognitive flexibility limits the ability of a person with TBI to take another person's perspective) [23-25], there is consensus that performance on social cognition tasks is impaired and that these impairments contribute to negative social outcomes after TBI [26].

Cognitive impairments typically translate into communication challenges [25,27], including difficulty understanding long or complex verbal information, either written or spoken; failure to understand abstract or figurative language, like sarcasm or slang; and expressive language that is less informative, appropriate, efficient, or complete than would be expected in a given context [25,28]. Speed is also a factor in communication, as everyday communication requires rapid comprehension and expression. Speed of thinking and expressing thoughts is commonly affected by TBI [28], so even individuals with relatively strong cognitive function can have difficulty communicating at the pace of typical social interactions.

Use of inappropriate language and lack of social understanding can lead to limitations in family and friend support and cause social isolation and loneliness for individuals with TBI [1,29-33]. Psychological changes can also occur after TBI, causing negative self-perception, heightened stress and anxiety, and depressive symptoms [34,35]. Specifically, individuals with TBI can have a stronger emotional response to pessimistic information or negative experiences [8]. These TBI-related challenges can profoundly affect an individual's social connections, resulting in a lower quality of life [36].

Social Media Use After TBI

As a result of cognitive and communication impairments, individuals with TBI face significant barriers to using social media [4,8-42]. Complex user interfaces can be difficult to navigate and make some features of social media platforms less accessible [6,43]. Undesirable content in newsfeeds, such as arguments between users or posts pertaining to violence, can trigger heightened negative emotions and cause frustration for individuals with TBI, particularly individuals with impairments in emotion recognition. Advertisements in the newsfeed can easily be a source of distraction and prevent users from focusing on desirable content [8]. Because of their communication challenges, individuals with TBI can misunderstand social media posts, "misreading" the sentiment of the sender and thus responding with inappropriate comments [6,8]. These types of interactions can yield an overall worse user experience and limit the development of social connections [44].

Why Improve Social Media Access?

It could be argued that the costs of social media access for adults with TBI outweigh the benefits, as cognitive challenges associated with TBI can pose risks to users. Our view is that adults with TBI who want to access social media should have the same opportunity to benefit from that access as their uninjured peers. When accessible, social media platforms help users maintain connections with their social ties, form bonds with communities, obtain new information, and foster information exchange [4,6,38,40,45]. For individuals with TBI, social media platforms provide functions that can help them overcome TBI-related challenges [4,5,38,43], and there is evidence that individuals with TBI may even prefer to have social interactions online rather than face-to-face [39]. Specifically, social media platforms can mitigate the communication challenges caused by TBI by reducing time pressure through increased time for users to comprehend information and formulate responses [4]. Individuals with TBI can also form online support groups to exchange information related to TBI and share personal experiences [4,5]. Moreover, individuals with TBI have used social media as a space to rebuild their self-identity and social connections post injury, particularly those who face dramatic changes in their personal lives [6]. With these benefits in mind, we next discuss our team's prior work toward making social media more accessible to users with TBI.

Related Work

Overview

There is a growing body of work aimed at reducing social media–related challenges faced by individuals with TBI and improving their social media experience [6,7]. For example, Brunner et al [46] developed online training and an educational toolkit for users with acquired brain injury to ensure users' safety and privacy in social media use [46]. This important work has reduced barriers and risks to using social media but has focused on training persons with TBI rather than changing the platforms themselves. Our research team focused on the latter: developing tools that address these specific challenges and improve the social media experience for users with TBI.

XSL•FO
RenderX

Our research team sought to understand the challenges of social media use in users with TBI, with the long-term goal of reducing barriers to use and improving opportunities for communication, connection, and well-being through computer-mediated communication. In this section, we review our previous efforts toward the development and evaluation of social media aids, which have used a variety of research designs, including needs finding [5,39,47], participatory design [6], prototyping [48], and a formative study [8]. These studies culminated in our current longitudinal study assessing the effectiveness of the SMART-TBI extensions.

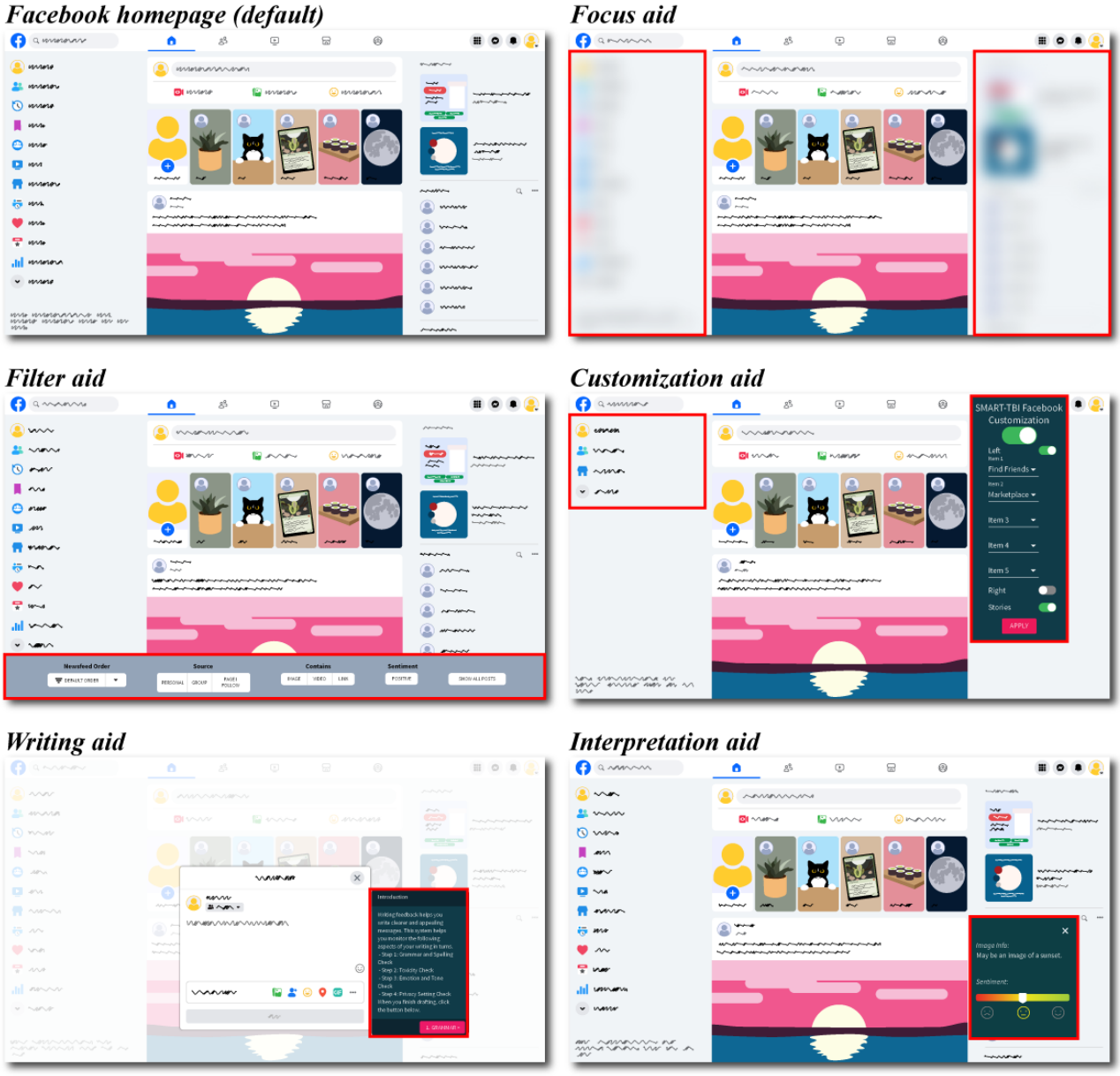
To begin, Zhao et al [48] conducted a preliminary study to assess individuals with TBI's perceptions of the acceptability of 4 prototype social media aids intended to support user attention, memory, social cue interpretation, and message production. These 4 support areas were chosen based on existing literature and the clinical experience of the research team. Individuals with TBI were asked to review images of the prototype aids, read descriptions of their functions, and rate the aids in terms of acceptability, ease of use, and utility. The findings of this study indicated that participants found the proposed aids easy to use and expressed interest in utilizing them regularly. This study offered foundational insights for the future development of technological support tools on social media platforms.

Building on this work, Lim et al [6] further explored the use of social media supports through a participatory design study that involved exploration, ideation, and sketching to create prototype supports. Through this study, several critical design features were identified to address challenges faced by users with TBI on social media. These features included a safe filtering mechanism for the newsfeed, a customizable interface, and prompts to support communication, among many others. The study emphasized the importance of addressing not only sensory-level support but also cognitive, emotional, and communication impairments, along with the altered sense of identity experienced by users with TBI when using social media. The generation of these prototypes allowed the research team to ground future aid development in the needs of real users with TBI.

Based on these participant-led designs, Hu et al [8] developed the Social Media Accessibility and Rehabilitation Toolkit (SMART-TBI). The toolkit consists of 5 aids: the focus aid, filter aid, customization aid, writing aid, and interpretation aid. These aids were specifically designed to address the cognitive and communication challenges faced by individuals with TBI on social media platforms. A visualization of each aid can be seen in Figure 1.

XSL•FO
RenderX

Figure 1. Social Media Accessibility and Rehabilitation Toolkit (SMART-TBI) extensions. Comparison of the default Facebook homepage and visualizations of each aid extension. Focus aid: the left and right side panels are blurred to increase focus on the newsfeed. Filter aid: a menu bar at the bottom of the interface allows users to tailor newsfeed content to their preferences. Customization aid: users can tailor the stories and the left and right side panels of the interface to meet their needs and preferences. Writing aid: users receive feedback on the posts they write, including grammar, spelling, toxicity, tone, and privacy checks. Interpretation aid: users are provided with an interpretation of a post's content. The interface differences associated with each aid are highlighted in red.



Focus Aid

This aid helps users focus on the newsfeed section of the Facebook interface by blurring other peripheral information present on the home page. This function reduces the cognitive challenge of information overload and creates a more manageable viewing experience. As a result, users with TBI are able to direct their attention toward meaningful content rather than extraneous details.

Filter Aid

Rather than limiting the amount of information on screen, this aid offers users the ability to reorganize newsfeed content in chronological order or view only positive posts. In doing so, this aid reduces cognitive challenges related to information overload and maintaining focus by allowing users to control the content they view and limit the appearance of undesirable content. IBM Watson NLP API [49] is used to analyze the sentiment of the posts and, in turn, enable positive post filtering.

Customization Aid

Similar to the focus aid, this aid helps users simplify the home page layout by offering the ability to customize which features are shown in the left column of Facebook's home page (eg, friends, groups, and events), hide the right column (eg, advertisements and contacts), or hide the stories sections of the home page. While the focus aid also minimizes extraneous information on the home page, the customization aid enables users to optimize their navigation experience by choosing which elements they prefer to view. This aid helps to address cognitive

XSL•FO
RenderX

challenges related to organization and planning, maintaining focus, and information overload.

Writing Aid

Unlike the prior 3 aids, this aid primarily focuses on communication by helping users refine their Facebook posts and comments through provided grammar suggestions, sentiment analysis (ie, tone and emotion type), toxicity analysis, and privacy settings (ie, private or public). This function allows users with TBI to better understand how others might interpret their posts and comments. As a result, this aid helps to address communication challenges related to social cognition, comprehension, word finding, and language production. Grammar errors, sentiment, and toxicity are detected using external APIs, including TextGears [50] for grammar checks, IBM Watson NLP API [49] for sentiment and emotion analysis, and Perspective API [51] for toxicity assessment. It is important to note that the writing aid runs entirely on the client side as a browser extension. To generate feedback, text is sent from the participant's browser to the aforementioned natural language processing services. These services return only numeric scores and category labels. Google APIs are used under terms that prohibit data retention and training. We also set Perspective API calls so that text is not stored. TextGears stores only the analysis result and the number of API requests.

Interpretation Aid

While the writing aid helps users with TBI understand how others might interpret their posts and comments, the interpretation aid allows users to understand the sentiment and meaning of posts written by others. This aid helps to address the communication challenge of comprehension by highlighting negative text in red and positive text in green, thereby clarifying the sentiment of the post. The same external APIs used within the writing aid are also used to analyze the sentiment of posts in the interpretation aid (ie, IBM Watson NLP API [49]).

These 5 aids are intended to directly address the specific challenges faced by users with TBI. As a result, we hope to see improvements in participant self-reports, including feelings of emotional support, social connection, self-esteem, and satisfaction with life.

Objective

We conducted a longitudinal in-the-wild study to evaluate the effectiveness of the 5 aids described above in supporting users' social and psychological well-being. A total of 39 users with TBI engaged in this study over the course of 4 to 7 weeks. Throughout the course of the study, all participants completed a prestudy interview, provided daily numerical and open-ended ratings of their satisfaction with each aid, completed evaluation surveys of each aid after each evaluation period, and completed a poststudy interview describing their experience throughout the study. Through these data, we sought to answer the following questions:

- How do participants perceive each of our 5 SMART-TBI extensions in terms of satisfaction, usability, likes, and dislikes?
- How do the aid extensions influence participants' well-being?
- What aid features are most and least useful to participants?

The contributions of this work are threefold:

- a longitudinal evaluation of our SMART-TBI extensions in participants' everyday use;
- insights into the social media experience of users with TBI, including their challenges and needs; and
- design implications for social media developers who may integrate the features of these aids into their applications.

Methods

Overview

We next describe our approach to evaluating the SMART-TBI extensions, including details on participants; measures; study materials, design, and procedure; and our analysis plan.

Participants

A total of 39 adults with moderate to severe TBI (female: n=25; male: n=14) were recruited at Vanderbilt University Medical Center (n=30) and McMaster University (n=9). At the time of the study, the participants' mean age was 41.85 (SD 11.55) years. All participants were from the continental United States and were native speakers of North American English. While some participants occasionally missed their daily Facebook usage, all 39 participants successfully completed all 5 evaluation phases of the study for a final retention rate of 100%.

Participants at the Vanderbilt site were recruited through the Vanderbilt Brain Injury Patient Registry [52], and participants at the McMaster site were recruited from among participants in previous studies who consented to be recontacted for future studies. All participants with TBI were in the chronic phase of injury (ie, >6 months post injury). All but 1 participant sustained their injuries in adulthood (ie, at or after age 18 years), and the average time since injury among participants was 8 (SD 8.37) years. Participants did not have a history of neurological or cognitive disabilities before the qualifying brain injury. TBI severity was determined using the Mayo Classification System [53]. By this system, participants were classified as having sustained a moderate to severe TBI if at least one of the following criteria was met: (1) a Glasgow Coma Scale (GCS) score of less than 13 within 24 hours of acute care admission, indicating moderate or severe injury; (2) positive neuroimaging findings; (3) loss of consciousness (LOC) greater than 30 minutes; or (4) posttraumatic amnesia (PTA) after 24 hours. Injury-related information was collected from available medical records and a semistructured interview with participants. GCS scores were available for 26 participants, with scores ranging from 3 to 15; LOC was available for 28 participants; and PTA information was available for 29 participants. Acute imaging information was available for 27 participants, with 26 containing positive findings. Causes of injury included motor vehicle accidents (n=16), falls (n=9), motorcycle accidents (n=6), being hit by a car as a pedestrian (n=3), assault (n=2), multiple concussions (n=1), or other (n=2). Multimedia Appendix 1 provides demographic and injury characteristics of our participant population.

XSL•FO
RenderX

Inclusion criteria for this study included self-reported English fluency, having an active Facebook account and using it within the past year, knowing the login information for the account, and daily access to Wi-Fi.

Study Materials

The research team provided each participant with a Samsung Chromebook 4 laptop (model #XE310XBA-K01US), which was used for all study-related activities (eg, Facebook usage and Zoom meetings). To prevent lack of familiarity with a Chromebook from being a barrier to participation, the research team prepared each Chromebook with a new Gmail account, Zoom for aid activation and prestudy or poststudy calls, all 5 SMART-TBI extensions, and aid demonstration videos for each aid. Additionally, each laptop contained browser bookmarks for Facebook and other relevant applications (eg, YouTube and Gmail) required throughout the study. Each participant also received a binder with their personalized schedule, a series of "cheat sheets" describing the functions of each aid, and login information for both the Chromebook and Gmail account. Each participant was also provided with a wireless mouse and keyboard for easier navigation of their laptop and a laptop bag to hold all the study materials.

Study Design

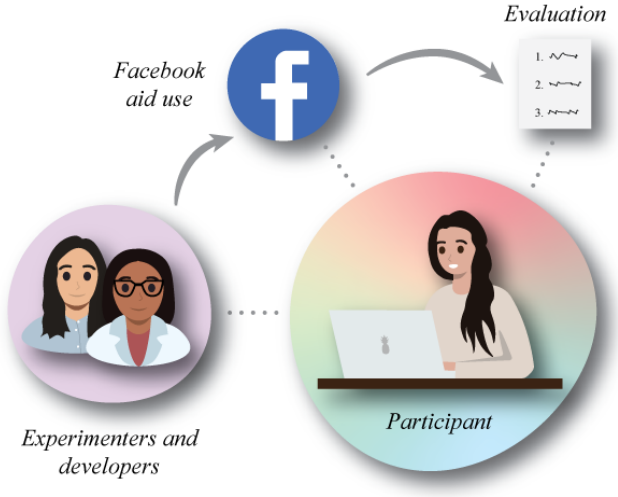
This study used a within-subjects, in-the-wild design. The study protocol was approved by the institutional review boards of Vanderbilt University and McMaster University, and all participants provided consent prior to engaging in the study.

Participants used the developed aid extensions for 4 to 7 weeks, with the duration of a given participant depending on external factors such as travel and holiday plans. Each participant received a personalized study schedule tailored to their individual availability and preferences (eg, time of day for Zoom calls). An example schedule is shown in Figure 2.

To begin, participants attended an in-person or online prestudy session where they were introduced to the study and provided with study materials. Throughout the course of this study, participants were asked to use and evaluate the 5 SMART-TBI extensions. Each evaluation included 3 days of using an individual aid, followed by a period of Facebook use with no aids to mitigate carryover effects. The number of baseline days varied depending on individual schedules and availability, but participants were always given at least 2 days between each aid evaluation phase. The aid extension order was randomized using a balanced Latin-square design to reduce order effects. At the beginning of each evaluation phase, participants were required to complete a Zoom call to activate the necessary aid for that period. After using all 5 aids, participants attended an online poststudy session to end the study.

Participants were compensated US $20 per hour for the prestudy session, poststudy session, and each of the 5 online aid introduction sessions. For the duration of the study, participants received US $10 per day that they used Facebook for at least 10 minutes. Compensation was provided in the form of a check or Amazon e-gift card. Finally, participants who completed the study were gifted the Chromebook laptop they used throughout the study.

Figure 2. Participant schedules. Participants engaged in the study following schedules with the same structure illustrated here. Participants initially met with experimenters to complete a prestudy interview and later met with experimenters at the beginning of each aid evaluation period to activate the relevant aid. Each evaluation period consisted of 3 days during which participants used 1 of the aid extensions. Following each evaluation period, participants completed an aid evaluation survey. After all 5 evaluation periods, participants completed a final poststudy survey. Participants additionally met with the development team to discuss any technical challenges as needed throughout the study.



Sun	Mon	Tue	Wed	Thu	Fri	Sat
	1	2	3	4	5	6
	Prestudy interview			Zoom call to switch aid Focus aid	Focus aid	Focus aid
7	8	9	10	11	12	13
Evaluation survey		Zoom call to switch aid Filter aid	Filter aid	Filter aid	Evaluation survey	
14	15	16	17	18	19	20
Zoom call to switch aid Customization aid	Customization aid	Customization aid	Evaluation survey		Zoom call to switch aid Writing aid	Writing aid
21	22	23	24	25	26	27
Writing aid	Evaluation survey		Zoom call to switch aid Interpretation aid	Interpretation aid	Interpretation aid	Evaluation survey
28	29	30	31			
Poststudy interview						

Procedure

Prestudy Session

Participants were given the option of completing the prestudy session online or in person. If a participant chose the in-person option, the research team met the participant in the laboratory and gave them the Chromebook laptop to take home for the rest of the study. If a participant chose the online option, the research team mailed them their Chromebook laptop and met with them on Zoom. During the prestudy session, participants were given an overview of the Chromebook (ie, login information and bookmarks) and their study schedule (ie, evaluation days, baseline days, and survey dates). Prior to or during this session,

XSL•FO
RenderX

participants provided informed consent by signing digital consent forms.

Once a participant logged into their assigned Chromebook, they completed a prestudy interview on REDCap and a prestudy questionnaire on Qualtrics. Next, participants logged into their Facebook accounts and were informed that their Facebook usage would only register if they used the provided Chromebook. Additionally, the experimenters informed participants that data collected throughout the study would never include the specific content of their posts and would be anonymized through the use of participant IDs.

Following this initial meeting, participants began the study by completing baseline days (ie, using Facebook with no aids active) until the first scheduled Zoom meeting, during which the first aid would be activated and the evaluation phase would begin.

Aid Evaluation Phases

Throughout the course of the study, participants completed a total of 5 aid evaluation phases. Each phase began with a Zoom call with an experimenter to (1) update the laptop background to the "cheat sheet" that matched the aid being evaluated, (2) assist the participant in activating the aid and a "helper window" extension for feedback, (3) watch the demonstration video for the relevant aid, and (4) guide participants through the process of using the aid prior to individual use. Participants were asked to use the aid for 3 days and for a minimum of 10 minutes per day. Outside of the minimum usage time, participants were encouraged to use Facebook as they would during normal use. Following their daily use of Facebook, participants were asked to complete a brief feedback survey to report their satisfaction with the active aid. This survey was made available through a "helper window" extension. After 3 days of use, the active aid extension automatically deactivated. Participants were then sent an email with an evaluation survey to provide detailed feedback on the aid extension they had evaluated. This full process was then repeated for each of the remaining aids until all had been evaluated.

Poststudy Session

Following their evaluation of all 5 aid extensions, participants completed an online poststudy session with an experimenter. During this session, participants completed an interview on REDCap and a questionnaire on Qualtrics to report their experiences with the aids and well-being outcomes. Additionally, experimenters assisted participants in deleting study materials and restoring the laptop to factory settings during this time.

Measures

Overview

Throughout the course of this study, the research team collected a considerable amount of data ranging from qualitative interviews and open-ended survey questions to quantitative perception measures and Facebook usage data.

Prestudy Interview

Within this initial interview, experimenters asked participants a series of 8 open-ended questions, which centered on current social media usage and experience with assistive technology for TBI challenges. A listing of these questions is provided in Multimedia Appendix 2.

Aid Evaluation Surveys

Following each aid evaluation phase, participants completed a survey concerning their experiences with and perceptions of the evaluated aid. This survey included the full 15-item Usefulness, Satisfaction, and Ease of Use (USE) Questionnaire [54], which captures perceptions of usefulness, ease of use, ease of learning, and satisfaction. This questionnaire is a widely used metric to evaluate system usability and was chosen to capture a multifaceted perspective on participant experience with the developed aids. Five open-ended questions concerning likes, dislikes, technical difficulties, aid challenges, and design suggestions were also included. These items allowed the research team to develop a thorough understanding of participant perceptions and experiences while also providing an opportunity for participants to give open-ended feedback. All items that were developed by the research team are provided in Multimedia Appendix 2.

Daily Ratings

After a participant had used Facebook on a given day, they completed a daily rating. Participants were prompted with the question, "How satisfied are you with the Facebook aid today?" followed by the optional question, "Do you have any other comments?" The first question was answered with a star rating of 1 to 5, while the second question was answered with a free-text response.

Poststudy Interview

During the final interview, experimenters asked participants a series of 3 open-ended questions centering on their experiences with the aids throughout the study. These questions included:

- How was your experience using the aids?
- Do you have a favorite aid? And if so, which one was it? And why?
- Is there anything else you'd like to share with us?

Prestudy and Poststudy Surveys

Within both the prestudy and poststudy surveys, participants rated their feelings of emotional support, social connection (adapted from Ledbetter [55]), self-esteem [56], and satisfaction with life [57]. These ratings were intended to capture any pre-post study changes. Additionally, the poststudy survey also contained 2 additional items regarding unaddressed TBI challenges (ie, "Are there any challenges related to your TBI that were not addressed by the aids you used during this study? If yes, please explain.") and endorsements of the aids for use in everyday life (ie, "Would you use any of the aids you have used during this study in your daily life? If so, which ones would you use?"). After having evaluated all aids, these questions allowed the research team to understand the main benefits and shortcomings of all aids holistically.

XSL•FO
RenderX

Analysis

Given the complexity of this study and the high volume of data collected throughout, quantitative data were first reviewed and cleaned to remove any erroneous data (ie, duplicate surveys and extraneous daily ratings). Following this step, aggregate daily ratings were generated for each aid per participant and then for each aid across all participants. Aggregates were also generated for each perception rating collected within the evaluation surveys across all participants. Finally, paired-samples 1-tailed *t* tests were used to determine whether there were positive pre-post changes in participant feelings of emotional support, social connection, self-esteem, and satisfaction with life.

We examined several types of qualitative data, including prestudy and poststudy REDCap interview data and multiple ratings of participants' perceptions of the extensions (Table 1). We conducted inductive thematic analysis for the survey feedback under each category for each aid (ie, likes, dislikes, technical difficulties, and aid challenges) [58,59]. The primary coder and a secondary coder each coded a subset of the dataset and developed the initial codebook. Disagreements were resolved through discussion. The primary coder then coded the remaining dataset, consolidated new codes, and updated the codebook iteratively. The primary coder performed the interview analysis independently. This analysis was aided by qualitative analysis software, specifically Dedoose [60], that facilitated data management and coding consistency. Based on the final codebooks of the survey responses and the interview data, our interdisciplinary research team discussed and generated the final themes and subthemes.

Table 1. Data analysis approach. An overview of the analysis for all data collected throughout the longitudinal user study, including collection method, measures, and analysis approach.

Collection method	Measures	Analysis approach
Prestudy interview	Prior social media usage and experience with assistive technology for TBI[a] challenges	Thematic analysis
Prestudy survey	Feelings of emotional support, social connection, self-esteem, and satisfaction with life	Descriptive statistics
Daily rating	Satisfaction with the evaluated aid	Descriptive statistics
Aid evaluation survey	Quantitative perceptions (ie, usefulness, ease of use, ease of learning, and satisfaction) and qualitative perceptions (ie, likes, dislikes, technical difficulties, aid challenges, and design suggestions) of aid extensions	Descriptive statistics and open coding
Poststudy interview	Experiences with all aids throughout the study	Thematic analysis
Poststudy survey	Feelings of emotional support, social connection, self-esteem, and satisfaction with life as well as aid endorsements	Descriptive statistics and *t* test

[a]TBI: traumatic brain injury.

Ethical Considerations

The study protocol was approved by the institutional review board (IRB) at Vanderbilt University (IRB 191489) and the Hamilton Integrated Research Ethics Board at McMaster University (Project 16271). All participants provided informed consent prior to participating. Participants were compensated at a rate of US $20 per hour for the prestudy session, poststudy session, and each of the 5 online aid introduction sessions. During the study period, they also received US $10 for each day they used Facebook for at least 10 minutes. Compensation was provided by check or Amazon e-gift card. In addition, participants who completed the study were permitted to keep the Chromebook laptop used throughout the study. Participant data were deidentified through the use of codes. All identifying information present in interview data was removed during the transcription process. Additionally, all of our social media aids ran on the client side as browser extensions, and no content from any Facebook posts or comments was collected. Collected data, such as consent forms and electronic data files, were stored in either secure physical storage or university-approved servers accessible only to study personnel. In the case of Vanderbilt University, research records will be retained for at least 3 years following closure of the research with the IRB, and any HIPAA (Health Insurance Portability and Accountability Act)-related documentation will be retained for at least 6 years following the last date of use. Deidentified data collected at McMaster University will be stored indefinitely.

Results

Overview

In this section, we present the results of our longitudinal study, including themes derived from our prestudy and poststudy interviews, findings from our multifaceted aid evaluation, and the overall impact the aids had on participants.

Prestudy Interview

Overview

To begin, a total of 39 prestudy interview transcripts were analyzed using thematic analysis. Through this analysis, we identified a total of 6 themes that illustrate the varying backgrounds of our participants.

Theme 1: Facebook Usage Changes Following TBI

Overview

Our analysis revealed several subthemes that reflected reported changes in patterns of Facebook usage in relation to TBI-related challenges.

XSL•FO
RenderX

Pattern of Usage

A total of 51% (20/39) of participants noted a decrease in Facebook use following their TBI, with varying patterns of reduced engagement. Some participants ended their use immediately after their TBI and only re-engaged in a limited manner sometime later. Participant 5 stated:

> *I just got back on Facebook about a year and a half ago. If I added anyone, it was just immediate family, a pool of ten people. These people don't have Instagram, so I have them on Facebook...I rejoined because I missed seeing the progression of close people's stories.*

Others reported decreasing or ending specific activities, such as posting. For example, participant 27 explained:

> *I have stopped posting a ton of things on Facebook. I limit it to just family things and not so much on where I am, like checking in anywhere.*

Some participants expressed a preference for re-engaging with visual-oriented platforms, such as Instagram, due to their TBI-related needs or personal interests. Participant 1 mentioned, "I switched over mostly to Instagram, with it being easier to take a look at pictures." Additionally, several participants indicated a desire to limit their time online or adapt their social media use to better suit their post-TBI lifestyle. For example, participant 10 noted:

> *Not specifically due to my TBI, but I do use social media significantly less than I did before. Now I primarily use Facebook for Facebook Marketplace and to remember birthdays.*

Participants 1 and 5 also noted a switch to Instagram from Facebook post TBI, as the picture-based nature of Instagram supports their engagement.

However, 10% (4/39) of participants noted an increase in their Facebook use post TBI. Participant 18 remarked, "My Facebook use has increased since the injury, partially due to COVID since I was home more." This participant also attributed their increased use to social pressure, stating:

> *It has increased due to social norms forcing the use of Facebook to be able to be involved in community events, etc. since that is how they notify you of times and changes in events.*

Additionally, 31% (12/39) of participants did not report any specific changes in their Facebook usage post TBI. For example, one participant noted that they "use it about the same" and described themself as "not a frequent user" (P31).

Sensory Challenges

Additionally, 21% (8/39) of participants noted sensory challenges while engaging with Facebook post TBI. Specifically, vision challenges were reported to impede Facebook navigation immediately following a sustained TBI:

> *The biological part, my eyes were so bothered by the bright light. Screens were just out, completely. Scrolling was impossible. I would close my eyes as I was scrolling and then look at posts, it was a nightmare. Trying to read people's updates, I found it was frustrating and a useless experience.* [P5]

Participants also described experiencing sensory overload while on Facebook, noting, "It just goes very fast; it takes me a while to understand" (P4).

Social Challenges

Further, 10% (4/39) of participants noted social challenges while engaging on Facebook post TBI. One such challenge was reduced confidence in posting on Facebook:

> *Actually, before I was confident using it. Now, I will start to put something on Facebook and then delete it because I'm not confident about what I've posted. It's wording, punctuation, that I'm not as confident with.* [P7]

Concerns with social comparison were also expressed: "I had deleted Facebook initially. I couldn't watch everyone's updates of getting married, having kids" (P5). This sentiment was echoed by other participants who felt overwhelmed by others' lives:

> *The way I see Facebook, it's more attracting families, moms and dads, posting their pictures—blasting all of their pictures, like birthday parties—It's too much.* [P1]

Cognitive Challenges

Beyond changes in Facebook usage, many participants also noted challenges following their TBI. Changes in cognitive abilities, such as attention, memory, and understanding, were specifically noted by 8% (3/39) of participants. They shared that their "attention span" was reduced when on Facebook (P33), that they relied on written reminders to support their memory (P6), and that understanding online information was difficult at times: "Some things I don't understand. Definitely. It's difficult sometimes" (P4).

Theme 2: Use of Technology, Apps, and Tools for TBI Management

Overview

Participants described the use of various tools, apps, and technology to manage their TBI symptoms in general and specifically to support their use of online communication. The following subthemes relate to the types of supports participants indicated they used prestudy.

Cognitive Supports

A total of 26% (10/39) of participants highlighted the use of various technologies and apps to address cognitive challenges following a TBI, both in their daily lives and on Facebook, where available. These tools were primarily used to support memory and attention difficulties. For instance, participant 3 mentioned:

> *I've attempted to use a lot of apps for research on symptom management, for example like a type of exercise to do to overcome the dizziness, the cognitive impairment aspect of it, and the balance of course.*

The types of tools and apps used included timers, calendars, reminders, alarms, note-taking apps, and meditation apps. These

XSL•FO
RenderX

cognitive supports were considered essential by many participants. One participant emphasized the importance of these supports, stating, "It's kinda' like my external hard drive. When I can't rely on my brain, it's helpful for storing and recalling information" (P8). Participants also discussed the use of various cognitive retraining apps, including brain games and specific apps (eg, Peak) "to improve brain function and stamina" (P12).

Participants also noted challenges with the consistent updates to and availability of these tools on social media. For example, participant 10 expressed difficulty in keeping up with changes on social media platforms: "I also like the birthday reminders but don't currently know where to find this on Facebook; things seem to move often."

Sensory Supports

A total of 21% (8/39) of participants discussed the types of apps, tools, and technology they use online to cope with visual, sensory overload, and hearing challenges. Visual supports included setting their screen to dark mode, dimming the screen, using a larger font, and orienting the screen to photo mode. Participant 19 noted, "I like darker screens now and less clutter, so we've adapted my preferences to be darker and less clutter[ed]; email, my search engine, etc." Similarly, participant 5 stated, "I use larger font—14-16 size font," to support their vision. Additionally, one participant mentioned post-TBI auditory sensory overload and the use of "hearing aids...for anacusis and tinnitus, with active noise canceling" (P5).

Assistive Technology

Additionally, 21% (8/39) of participants described an increased reliance on the use of assistive technology in the form of devices and assistive features post TBI. For example, some described using larger devices to access social media: "My iPad and computer have purposefully lots of pixels, big screens, etc., for when my eyes get tired and it's harder to process" (P19). Specific assistive features noted included speech-to-text, text-to-speech, and AssistiveTouch. Reasons participants reported for using these features centered around the need for cognitive, physical, and communication supports. For example, one participant noted:

> *With the voice-to-text, because of the affected side [of my body], sometimes it is hard for me to write or type. So I use the voice-to-text. It also helps my speech because the injury affected my speech. In the beginning, you get a lot of therapy but then it stops. I also use the touch screen and magnifying and closed captioning a lot.* [P3]

Written Communication Supports

Some participants (3/39, 8%) also reported using spelling and grammar check tools while writing online, specifically highlighting the use of apps like Grammarly to support post-TBI cognitive challenges with writing. Participant 36 shared:

> *When I forget a word or erroneously substitute, [for example I] meant 'baseball' but typed 'basement', I like that it highlights potential errors and allows me to decide if I want to 'correct' it.* [sic]

AI

AI search tools, such as Bard AI, were described by one participant as "helpful when sifting through large amounts of data" (P2) and were highlighted for their role in facilitating information processing and retrieval post TBI.

Lack of Support

A substantial number of participants (16/39, 41%) reported not using any specific apps or technology post TBI. Participants' rationale varied, ranging from a lack of awareness about available tools to a preference for managing their interactions with technology in other ways. For instance, participant 18 stated, "[I do] not know of any tools to use," indicating a lack of awareness as a barrier to using supportive technology. Some participants preferred to manage their technology use by controlling their exposure to overwhelming stimuli. For example, participant 27 shared their strategy for handling Facebook usage: "I got back onto Facebook once I could handle it again. If I became overwhelmed using it, I would stop and come back to it later." This approach highlights a preference for self-regulation and intermittent use to avoid sensory and cognitive overload.

Theme 3: Frequent Use of Social Media and Facebook

Out of 39 direct responses regarding the frequency of social media and Facebook use, 37 (95%) participants provided detailed insights. The majority of participants (34/39, 87%) rated their use as "5," indicating they use social media "very frequently" (eg, daily or several times a day). Example comments included: "[I] check stuff daily but more app-specific preferences" (P36). Some participants noted that they are on social media "constantly, I'm on it all the time" (P2), or "on a daily basis, 2-3 hours per day, this is where I get my news now" (P1). A total of 2 (5%) participants rated their use as "4," signifying that they use social media often (eg, once a week to several times a week). One participant rated their use as "2," indicating "rare" use of social media (eg, less than once a month), stating:

> *Maybe once a month or once every two months [I use Facebook]. Then I'm only on there for a few minutes to send a message or receive one. I find it overwhelming and confusing and too time consuming for me.* [P19]

As with the frequency of social media use, the majority of participants noted that they use Facebook either "very frequently" (27/39, 69%) or "often" (8/39, 21%). One participant (P35) noted that they use Facebook "rarely," specifically "once every 3 months." Participants discussed attempting to limit their use, noting, "[I am on] daily but [for a] limited amount of time; mostly to check on community events, not for personal people [sic] networking" (P18). Another participant mentioned a specific use of Facebook to keep up with a certain demographic of friends:

> *I check it once a week primarily to keep in contact with older sergeants from my time in the Army that only use Facebook, along with some members of my family.* [P23]

XSL•FO
RenderX

Theme 4: Preferred Social Media Platforms

Participants described using multiple social media platforms in addition to Facebook, with many engaging with 2 or 3 other platforms. Instagram was mentioned most frequently (30/39, 77%), followed by TikTok (9/39, 23%), Snapchat (5/39, 13%), and several others. A visual of the reported social media platforms is provided in Multimedia Appendix 3.

One participant explained their selective use of social media platforms:

> *I was on Instagram, but I don't like it. I just find that everything is about your picture and posting, and I'm getting overwhelmed with social media. So, I just choose which one is functional for me. And TikTok, but it's getting too addictive, I like the dancing and the funny stuff. Of course, YouTube—everything you need is on YouTube. And YouTube is educational for me. I like fixing stuff so I look for how to do stuff.* [P3]

Several participants (9/39, 23%) noted that Facebook was the only social media platform they actively used. One participant shared:

> *Facebook's the only one I use. I Google things and use email, but that's about it. I don't have Instagram or TikTok or Twitter. If there are other social media platforms, I'm not really aware of them...* [P19]

Theme 5: Perceived Benefits of Social Media

Overview

Participants were asked what they generally liked about engaging on social media, specifically on Facebook. Several subthemes of liked aspects of social media emerged.

Social Connection

Most participants (32/39, 82%) reported that engaging on social media provided them with an important context for social connection. They reported using social media to connect with friends and family who live far away, converse with new people with similar interests, and stay updated with others' activities. One participant noted about Facebook:

> *After my brain injury, I like the fact that I can, it can be two in the morning or afternoon, there's always somebody, a conversation, something to see. It's very stimulating.* [P9]

General Information

A total of 26% (10/39) of participants highlighted the ease of accessing information, such as staying current with the news and learning new things. For example, one participant mentioned, "It's great for staying up to date with the news" (P24), and another added, "I use it for learning new things" (P21).

Entertainment

Several participants (6/39, 15%) described the entertainment value of watching videos, sharing memes, and passing time with humor. One participant noted, "I follow very few things on Facebook, but the ones that I do are fun things" (P7).

Professional Development

Some participants (3/39, 8%) said they used specific sites for business purposes. One participant said that LinkedIn was a "powerful tool in the business environment" (P23), allowing them to build their businesses.

Theme 6: Concerns With Facebook

Overview

While about half of the participants (18/39, 46%) reported not having any concerns with their Facebook use, the others described concerns about Facebook in 4 categories: cognitive, social, privacy and safety, and skills-related.

Cognition

Cognitive concerns included feeling overloaded and overwhelmed by information and concerns about their own overuse of the platform. For example, participant 31 noted, "It takes up too much time, too much screen time, [and] feels like it makes my attention span worse." Another participant echoed this sentiment, stating, "I have ADHD and it's a dopamine hit, so I'm prone to get on social media too much. Especially TikTok" (P21). One participant reported concerns with their long-term use of Facebook, noting, "my brain has enough challenges, and it's hard to keep up with all the changes on Facebook" (P19).

Social

Social concerns included personal and social judgments about what they and others posted, as well as social comparison. Participant 1 said, "Sometimes I feel weird about posting and people asking too much [sic] questions," and participant 30 expressed the concern that Facebook "leads to comparison of others."

Privacy and Safety

Privacy and safety concerns included concerns about data tracking, identity and information theft, scammers, hacking, and a lack of anonymity. For example, participants reported knowing their information is not private and feeling "leery" about posting or experiencing heightened anxiety online. Participants noted, "There are so many unscrupulous people" (P2) and "Predatory people are out there, and [you] need to be careful" (P9).

Lack of Skills

Finally, skills-related concerns centered around familiarity with using Facebook. For example, participant 7 explained why they chose Facebook because of its familiarity:

> *Because I'm used to Facebook, that's why I go there the most. Others, like Instagram, I'm not used to moving around, navigating. With Facebook, I get a pop-up message that it's there.*

Other participants expressed concerns with frequent Facebook interface updates and the resulting reduced confidence with the platform. These reports showcase the importance of consistency in allowing for a manageable social media experience.

XSL•FO
RenderX

Aid Perceptions

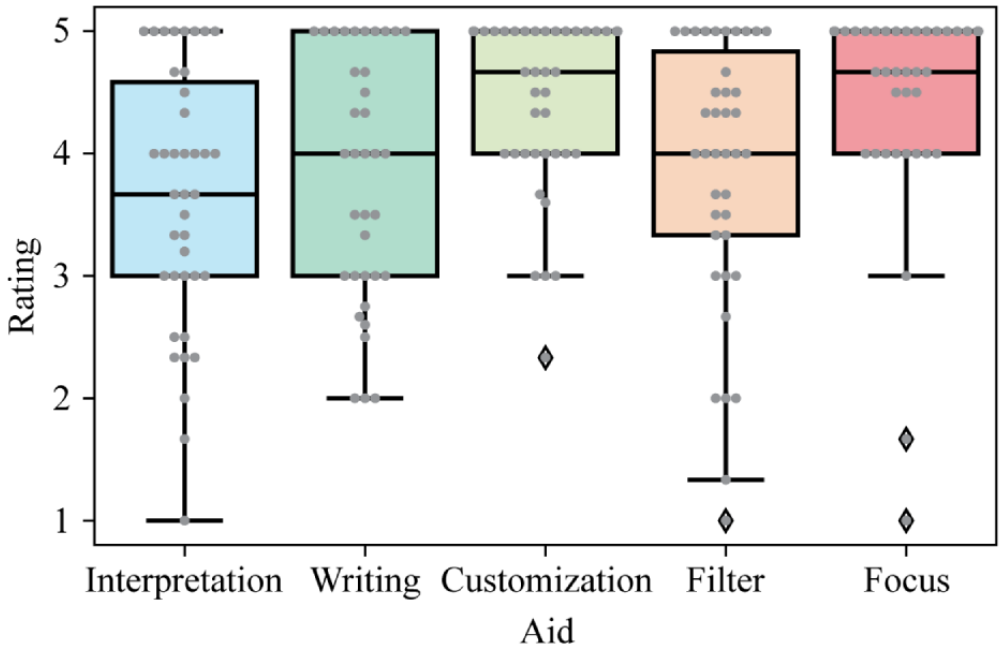
Overview

To understand how participants perceived each of our aid extensions, we investigated their daily ratings (ie, numerical and descriptive) while the aids were active and their perceptions of each aid holistically across the metrics of the USE Questionnaire (ie, usefulness, ease of use, ease of learning, and satisfaction). We also conducted open coding of participants' reported likes and dislikes, technical difficulties, and aid-specific challenges for each aid. Finally, we analyzed which aids participants reported they would like to continue using in daily life. These data allow us to understand which aid features were beneficial and where improvements can be made.

Daily Ratings

When evaluating each aid, participants were asked to provide a daily rating on each of the 3 evaluation days. While some participants provided exactly 3 ratings, others provided additional ratings or missed days. In total, 518 ratings were collected throughout the study. Due to the inconsistency in the number of ratings between participants, overall scores were calculated by first generating average scores per participant per aid and then averaging those aggregates to calculate per-aid average ratings. These ratings range from 1 to 5 stars, where 1 star indicates a poor experience and 5 stars indicate a highly positive experience. On average, participants rated their satisfaction with each aid as follows: focus, 4.45 (SD 0.89); customization, 4.41 (SD 0.71); filter, 3.86 (SD 1.09); writing, 3.83 (SD 1.01); and interpretation, 3.65 (SD 1.06). While these scores express an overall positive perception of the aids, there appear to be noteworthy divisions, with the focus and customization aids being rated similarly and most highly, followed by the filter and writing aids, and finally the interpretation aid. These scores are visualized in Figure 3.

Figure 3. Daily ratings. Aggregate daily ratings of all participants across days of aid use for each of the 5 Social Media Accessibility and Rehabilitation Toolkit (SMART-TBI) extensions.

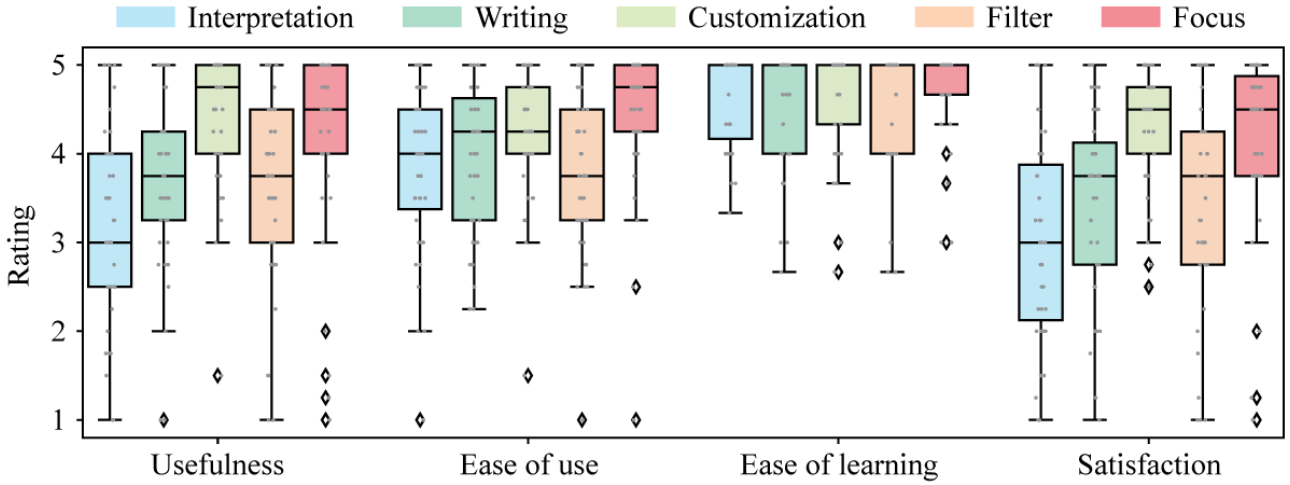


USE Questionnaire

Following each evaluation phase of the study, participants completed the USE Questionnaire. The average ratings per aid in terms of usefulness, ease of use, ease of learning, and satisfaction can be seen in Figure 4. While there is some variation between the daily ratings and holistic postevaluation ratings of satisfaction, we see a similar pattern of perception wherein the customization (mean 4.33, SD 0.68) and focus (mean 4.17, SD 0.98) aids are rated most highly, followed by the filter (mean 3.44, SD 1.19) and writing (mean 3.44, SD 1.05) aids, and finally the interpretation aid (mean 2.97, SD 1.17). The same pattern is also present in participant ratings of aid usefulness and ease of use. There is, however, a noteworthy difference in terms of ease of learning, for which all aids scored 4.56 or higher on average (focus: mean 4.76, SD 0.51; customization: mean 4.64, SD 0.61; interpretation: mean 4.63, SD 0.51; writing: mean 4.57, SD 0.66; filter: mean 4.56, SD 0.73).

XSL•FO
RenderX

Figure 4. Usefulness, Satisfaction, and Ease of Use (USE) Questionnaire ratings. Participant ratings of usefulness, ease of use, ease of learning, and satisfaction across all 5 Social Media Accessibility and Rehabilitation Toolkit (SMART-TBI) extensions. These subscales are shown in the order in which they are presented within the full scale.



Likes and Dislikes

Overview

Along with the USE Questionnaire, participants also provided open-ended responses regarding aspects of the aids that they most liked and disliked. We collected additional feedback on aid challenges, technical difficulties, and design suggestions, though these data were much sparser, as participants often provided responses such as “none” or “see previous.” Here, we present the most common sentiments shared by participants for each aid.

Focus Aid

Participants noted that the focus aid decreased distractions within the Facebook interface, allowed for improved focus, and made their use of Facebook less overwhelming. As one participant stated:

> *[I] loved how it helped me stay focused on the news feed. [I]t made me feel less overwhelmed and [I] would probably stay on [Facebook] longer due to this.* [P30]

Participants also expressed that they would have preferred the translucent blurring effect to instead be opaque so that moving content (eg, advertisements) would not be visible. Additionally, participants noted that while it was active, the focus aid made some elements of the interface inaccessible.

Filter Aid

The features of the filter aid that participants most liked were the positive filter, the chronological-ordering filter, and the overall ability to control and customize their Facebook experience. The benefits of the positive filter are well illustrated by participant 5, who stated:

> *I liked knowing that if I was having a bad day and wanted to use the filter...I had the option to engage with Facebook without the fear of encountering any content that could be interpreted as ‘too heavy’ to handle for that day.*

This statement captures the need for both limitations on certain types of content and the ability to tailor content based on day-to-day user preferences. Participants also expressed, however, that the filter aid caused lag when scrolling through the newsfeed and that the menu bar covered certain interface elements. The lag experienced while using the aid was caused by information loading and rendering on the web page.

Customization Aid

Like the focus aid, participants reported that the customization aid enabled greater control, decreased distractions, and simplified the Facebook interface by removing clutter. These benefits were expressed by participant 12, who stated:

> *I liked how much [the customization aid] decluttered my feed since I am so easily distracted and overwhelmed by information. I liked that I could choose certain tabs to keep on the left side and hide the rest.*

While the customization aid requires users to specify which elements they would like to keep and remove, which could be a cognitive challenge, the result is a more tailored Facebook experience that ultimately could reduce cognitive load and increase the ability to focus on desired content. Depending on the user, this granular selection might be more desirable than the broad blurring of interface elements with the focus aid. The primary negative feedback about the customization aid was the inability to use the aid beyond the Facebook home page, such as blocking advertisements in the Facebook Marketplace.

Writing Aid

Participants’ most-liked features were the spelling and grammar checks and the tone and toxicity checks. As stated by participant 27:

> *I’m glad that [the writing aid] checked the toxicity and tone of the post. It made me think twice about whether or not I wanted to post a comment. I really like that it checked grammar and spelling. There’s nothing worse than bad spelling.*

XSL•FO
RenderX

This quote illustrates that participants appreciated the review of their content before posting to limit errors in both communication of intention and technical aspects of writing. While the aid had a positive impact for some participants, others reported aid inaccuracies. For example, participant 19 expressed concern, stating:

> *It wasn't very accurate...and it made me anxious—if I think I can trust something like that I might let my guard down and depend on it. When I find it's not accurate, it's disturbing.*

Interpretation Aid

Participants reported that the interpretation aid allowed for easier understanding of posts by clarifying sentiment and summarizing content. Participant 27 reported a positive experience with the aid, stating:

> *I really liked that it colored the text. I found myself not reading the red text at all and skimming through the white text. It really made the whole experience more positive and more enjoyable.*

Not all participants felt they needed the aid. Participant 29 said:

> *[The interpretation aid] wasn't something that I would use, but it is also something that I don't need. For those who struggle with perception it could be useful.*

Some participants reported disagreeing with the interpretations. As with the writing aid, participant 19 said, "I had a hard time trusting [the interpretation aid] because it didn't always agree with my interpretations of what was positive or negative."

Poststudy Evaluation

Overview

Participants were asked 3 questions at the conclusion of the study about their experience using the aids and their overall favorite aid. The following themes emerged from these data.

Overall Positive Experience Using the Aids

A total of 95% (37/39) of participants reported positive experiences using the aids. Participants appreciated the ease of learning and using the aids, with participant 24 stating:

> *I had a very good experience using the aids. They were all easy to learn how to use and didn't have too many kinks. I definitely felt like some were more specific to the symptoms I struggle with than others.*

The aids were noted to enhance engagement and enjoyment on Facebook, as evidenced by participants' comments like, "I absolutely loved the aids. I loved that they kept me more focused and less distracted. I was able to enjoy Facebook more" (P27) and "It was very helpful, and I feel like I would use Facebook more" (P35). Participants also emphasized the aids' user-friendliness, noting, "All aids were easy to use" (P22) and "The tools were really useful based on my needs" (P1), though some aids were easier to use than others. Additionally, the aids were recognized for their potential value to individuals with TBI, with one participant stating, "I believe the aids are very helpful for the Brain Injury Community" (P37), and another stating, "They each could have value depending on the user's needs" (P22). Overall, the aids were well received, contributing to increased focus, enjoyment, and perceived value among participants.

A total of 21% (8/39) of participants also reported specific negative experiences while using the aids. Technical difficulties were a common issue, as evidenced by participant comments such as, "Sometimes [the aids] didn't work well enough to make them worth using" (P19) and "I found a lot of errors" (P14). The interpretation aid was particularly noted to have technical challenges: "The interpretation aid was glitchy. I think it could be easier if it was working. Ideally, it could be really great" (P8). Some participants questioned the usefulness of certain aids, with one stating, "Some of the aid[s] I was not able to really understand how it would be beneficial for someone with TBI" (P14). Technological difficulties with the equipment were also mentioned:

> *The aids...I don't know if it was the Chromebook that was making [sic], or the debugging of the aid, but the Facebook responses would come out slower or slow. It would be a long time waiting for something to happen while using the aid. The waiting was frustrating. Nowadays, everyone wants things to be done very quick. Waiting just for a Facebook post or feed, sometimes it's not worth the wait. A lot of times, I would just close the window because I thought it was crashing, and would have to close and restart.* [P1]

Issues with accuracy were highlighted, with one participant noting, "The writing tool was not as accurate, and I found it frustrating when there were so many technology glitches" (P2). Lastly, some participants raised concerns about the potential harm of using AI for functions such as sentiment analysis:

> *Everything else has the potential to be harmful for users, especially for those who want to engage in self-expression and just want help with navigating online interactions.* [P5]

Favorite Aids: Focus, Filter, and Facebook Customization

Overview

Participants were asked to name their favorite overall aid from the study. The focus, filter, and Facebook customization aids emerged as their favorites, with some participants ranking more than 1 aid as their favorite.

Focus Aid

A total of 49% (19/39) of participants appreciated the focus aid's straightforward and easy-to-use interface, noting that it significantly reduced distractions and kept the user interface clutter-free, thereby allowing for a more focused Facebook experience. For instance, participant 3 shared:

> *I liked everything! Like OMG, yesterday I was watching some cooking stuff, and the fact there is no distraction, nothing on the left that is taking my attention. Just to focus on one thing, it brought some*

XSL•FO
RenderX

> *peace of mind, less anxious. Yes, I love the focus mode.*

Similarly, participant 6 said:

> *I loved this one because when you're on Facebook, and all the apps disappear, then it's less distraction. And you're more focused on what you're doing.*

The focus aid also helped reduce feelings of being overwhelmed, as highlighted by participant 10, who said:

> *Focus Mode was my favorite. I felt less overwhelmed using Facebook and could pay more attention to what I was looking at, it felt easier to try to and write something without so much to pay attention to.*

Participant 7 echoed this sentiment, stating, "I was able to focus on what I was doing, and ignore the 'visual noise'."

Filter Aid

A total of 41% (16/39) of participants reported various features they enjoyed about the filter aid. They appreciated the ability to control what they see, such as filtering by positive posts: "I really enjoyed the ability to block negative posts" (P37) and to only show positive posts (P32 and P28). Others valued filtering by the most recent posts: "I like having the ability to see posts in chronological order!" (P11). Participants also noted the aid's ability to prioritize important posts over advertisements and irrelevant content. For example, participant 19 remarked:

> *I really liked the one that filtered my posts to where the pertinent ones were the ones that would show up—I would use this all the time. I get too easily sucked into ads and distracted by them, and it seems most of my feed is ads—it's not worth wading through all of them to find the more important posts. I don't like re-reading the same posts over and over either, so to have them in chronological order keeps them from repeating and saves my energy.*

Facebook Customization Aid

A total of 36% (14/39) of participants chose the Facebook customization aid as a favorite. One participant noted:

> *I liked the Facebook Customization aid the best in that it allowed me to turn off the various panels of FB [sic] which increased my desire to interact with FB [sic].* [P36]

Others noted that Facebook can be cluttered, and the aid helped with this problem:

> *Facebook can feel so cluttered with all the options of things to click on, that I rarely even click the options I like. To have the few options I would want to click on made Facebook look so much cleaner, not cluttered, and tailored to exactly what I would use.* [P15]

Participants appreciated the flexibility and simplicity of the aid and that it made navigating the Facebook page less busy (P29 and P22). Others noted that decluttering the page, similar to the Focus aid, helped minimize distractions, allowing them to "be able to focus on the content people posted rather than getting distracted" (P26). Participants also appreciated that the aid minimized unwanted content that could be potentially emotionally triggering:

> *The Facebook Customization was a game changer. It helped me a lot on my bad days, when I was feeling really emotional. I could just display exactly what I was able to view on that day.* [P2]

During the final poststudy survey, participants reported which of the aids they would use in daily life if given the opportunity. Of the 39 participants, 21 (54%) chose the filter and focus aids, 19 (49%) chose the customization aid, 11 (28%) chose the writing aid, and 9 (23%) chose the interpretation aid. These results largely comport with the other metrics we have reviewed; however, the filter aid was endorsed by more participants than expected. This misalignment with our other collected measures might be due to the fact that the aid endorsement questions asked which aids participants would like to use rather than which they thought worked the best. Participants may have believed the filter aid would be highly useful after further development, but that its current form did not quite meet that potential. A visualization of per-aid endorsement counts is provided in Multimedia Appendix 4.

Impact on Participants

Next, we attempted to understand how participants may have been impacted by each of our aids. Toward this effort, we conducted a pre-post study analysis to understand whether participants reported changes in feelings of emotional support, social connection, self-esteem, and life satisfaction (Figure 5). Further, 1-tailed paired-samples *t* tests with Bonferroni correction showed a statistically significant increase in participant ratings of satisfaction with life (P=.02) and nonsignificant increases across the remaining 3 metrics following the use of our aids. A full listing of the relevant statistics from these tests can be seen in Table 2.

XSL•FO
RenderX

Figure 5. Pre-post study metrics. Participants reported their perception of emotional support, social connection, self-esteem, and life satisfaction prior to and following their engagement in this study. A statistically significant increase was found between participants' life satisfaction before and after participation.

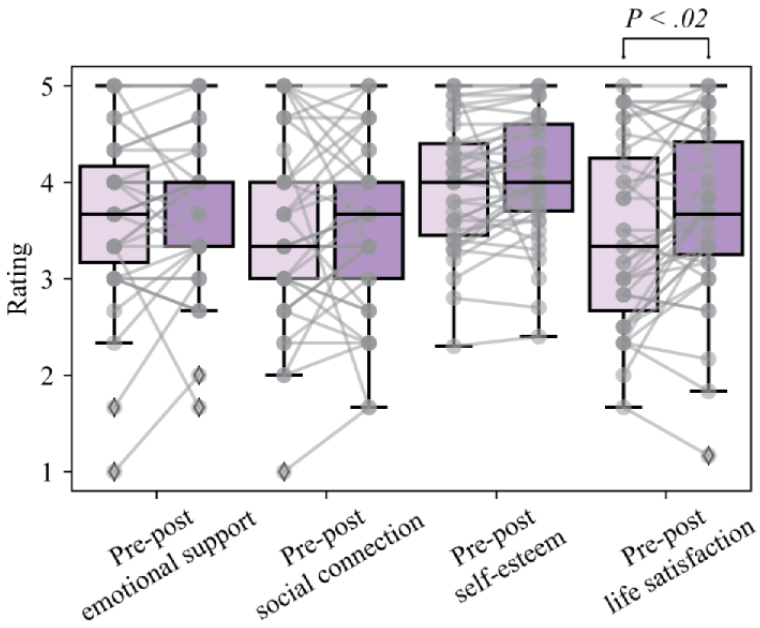


Table 2. Paired-samples *t* test statistics. The study used 1-tailed paired-samples *t* tests to assess statistical significance in pre-post differences between participants' reported perceptions of emotional support, social connection, self-esteem, and life satisfaction. A statistically significant increase was only found in perceptions of life satisfaction.

Outcome	Pretest, mean (SD)	Posttest, mean (SD)	*t* test (*df*)	Corrected *P* value	95% CI	Cohen *d*
Emotional support	3.684 (0.88)	3.761 (0.80)	−0.851 (38)	.80	−∞ to 0.075	−0.091
Social connection	3.496 (0.95)	3.547 (0.88)	−0.406 (38)	>.99	−∞ to 0.162	−0.055
Self-esteem	3.972 (0.66)	4.072 (0.64)	−1.441 (38)	.32	−∞ to 0.017	−0.153
Life satisfaction	3.449 (0.94)	3.705 (0.88)	−2.666 (38)	.02	−∞ to −0.094	−0.278

Discussion

Principal Findings

Participants with moderate to severe TBI have cognitive, sensorimotor, and communication challenges that can make social media engagement difficult or unmanageable. Our study illustrates that many of these challenges can be addressed through proper technological supports. Through this work, we have (1) demonstrated the feasibility of a longitudinal study design to examine the social media use of individuals with TBI, (2) gained novel insights into the social media experience of users with TBI and related design implications, and (3) identified new directions for researchers and developers working to promote online social connection among users with TBI.

SMART-TBI Perceptions

Through our results, we have shown that, on average, all of our aids were regarded positively by participants across all metrics. The filter, focus, and customization aids were rated by participants as their favorite aids and the aids they would have most liked to use beyond the study. These results are consistent with participants' reported sensory and cognitive challenges, as all 3 aids allowed users to configure what was displayed onscreen.

The writing and interpretation aids were regarded less positively overall. This result was unsurprising given that social challenges can be more complex and multidimensional than sensory limitations and, hence, more difficult to overcome with a technical solution. The sentiment analyses used in these aids also yielded inconsistent results, which meant they were regarded by some as untrustworthy. Participant perception may then be more reflective of our specific implementation than the core ideas the aids represent. Across daily ratings and all USE metrics, the writing and interpretation aids were consistently rated the lowest. Future implementations may be able to provide the aids' intended utility and be perceived as highly useful.

It is noteworthy that the findings did not apply to all participants, as some found the writing and interpretation aids very helpful in their current form. This variability in findings is common in TBI due to the heterogeneity across this clinical population in strengths and limitations, previous experience, and goals. It is expected that some users will benefit from certain aids, while others will not. If implemented in a commercially available product, the problems associated with these aids would likely be addressed, and users would be able to use them fully. Even with such an implementation, however, it is important to consider the impact that bias, subtle misclassifications, and overreliance may have on users. While our study did not capture

XSL•FO
RenderX

the nuance of these problems, they could introduce new challenges to social media users with TBI. Through proper research and development, these challenges may also be mitigated, but further evaluation will be needed.

Accessibility Design Insights

Beyond the aids themselves, we also identified the following three major insights from our study results: (1) participants preferred aids that assisted with sensory and cognitive challenges over social ones, (2) participants would like to tailor their social media experience to their specific needs, and (3) simple solutions to TBI challenges were preferred over complex ones.

Participants preferred aids that assist with TBI-related sensory and cognitive challenges over social ones. These findings are consistent with Zhao et al [48], who proposed that social deficits are difficult to recognize in oneself and therefore make an individual less likely to believe they may need related assistance. While the focus, filter, and customization aids all allowed users to make their Facebook experience more manageable through selection of interface elements and newsfeed content, the writing and interpretation aids attempted to compensate for social challenges by interpreting content for the user. While most interface and content changes are objective and clear, interpretations are subjective and more difficult to assess in terms of accuracy. When participants assess that some interpretations are not accurate, a challenging judgment in its own right, they may lose a sense of trust in all interpretations. As a result, these types of aids may be seen as fraught compared with those that address cognitive and sensory challenges. Additional validation will therefore be needed to improve these aids and understand the implications of their use prior to widespread implementation on social media platforms.

Consistent with the variability among people with TBI, participants appreciated the ability to tailor their social media interface to fit their needs. The ability to customize aids is essential in TBI to allow users the freedom to activate features as needed or select a baseline set of features that fit their needs. This point was well illustrated by participants' positive ratings of the filter and customization aids, which allow users to explicitly select which interface elements appear on their screen and what content appears in their newsfeed. Allowing users to choose specific accessibility features ensures that benefits are fully realized without making the resulting experience needlessly cumbersome.

While user choice allows specific needs to be met, it can also generate greater complexity to the detriment of a user's experience. If we consider, for example, the focus and writing aids, they each require different levels of effort from users. In the case of the focus aid, users simply activate the aid to blur the left and right panels of the Facebook interface. With the writing aid, however, participants activate the aid and then follow several steps while writing a post. While both of these aids have unique advantages and disadvantages, the writing aid adds much greater complexity to the baseline use of Facebook than does the focus aid. While we cannot disentangle the utility of a given aid from its cognitive demand, our results do suggest that participants prefer simpler aids (ie, the focus, filter, and customization aids) to more complex ones (ie, the interpretation and writing aids).

Limitations and Future Work

While this work shows great promise for the use of social media aids, some challenges have remained unaddressed. To begin, we had intended for participants to use our aids in their daily Facebook use as would occur naturally. However, it is clear from participants' daily ratings that participants would experiment with the aids to understand their limitations. For example, some participants wrote draft posts that included highly positive, negative, or subjective sentiments to evaluate the accuracy of the writing aid's sentiment analysis. While these participants' curiosity is understandable, these interactions illustrate the potential limitations inherent to an in-the-wild study. Additionally, we found that some participants had difficulty remembering whether they had used Facebook on a given day, leading to missed or duplicated participation. This difficulty is understandable and to be expected among participants who experience memory challenges. To address this challenge, experimenters provided friendly reminders to participants about daily participation, but challenges persisted. In the future, we believe that a usage indicator would be helpful for users to determine whether they had already participated on a given day. Such an indicator could take the form of a small overlay dashboard that contains an icon to signal whether the participant had completed their use for the day.

Furthermore, we saw that not all challenges participants reported were addressed by our aids. In particular, we expect that the writing and interpretation aids can be refined for a better overall user experience by improving their sentiment analysis features. Additionally, while the filter aid was positively regarded, we believe there are other filter options we can provide to users so they can fully tailor their display to their specific preferences. Beyond our existing aids, we also believe there are opportunities for new supports that can assist users on other social media platforms. While the current aids were developed specifically for Facebook, there may also be opportunities to apply these concepts to other platforms. Some supports, like the focus aid, may need to be tailored to the specific interface elements of a given platform, but others, like the filter aid, could easily be translated to platforms that use a feed. As social media evolves over time, similar adaptations may be needed, but, provided the core features of social media remain (eg, feed, posting, comments, and advertisements), the benefits these aids provide should continue to be relevant to individuals with TBI. Examining how these aids can be integrated into other platforms and exploring their potential impact on other user populations would substantially widen their impact.

Beyond the development and use of the aids, it is worth noting that the compensation participants received (ie, monetary compensation for participation and the gifted Chromebook) could have skewed their perceptions. While many participants did provide critical feedback about the aids, daily use and life satisfaction ratings could have been skewed by monetary gains and the novelty of using a new device. Additionally, while our results show a statistically significant increase in life satisfaction between the beginning and end of our study, we do not see a

XSL•FO
RenderX

significant increase in emotional support, social connection, or self-esteem. Given that changes of this nature may require greater exposure to our aids, further evaluation is needed to understand whether the use of such aids can achieve these outcomes.

Finally, it is important to consider the generalizability of our results. To be considered for our study, participants needed to be active users of Facebook who were willing to partake in a technology-evaluation study. Our results may therefore be limited to users with higher digital literacy and more social media experience than the average individual with TBI. Future work may be able to bridge this gap by evaluating such aids with a broader population of users with TBI, including those with more severe cognitive impairments or anosognosia. Additionally, our study does not use a comparison group or randomized controlled design. This step could be taken in similar future studies to clearly contextualize the perceptions of the developed aids and the broader impacts these aids have on participant well-being.

Conclusion

Users with TBI experience challenges related to sensory perception, cognition, communication, and social interaction. While social media platforms like Facebook can provide an outlet for users to maintain connections post injury, these platforms must provide TBI-specific accessibility features so that users may enjoy the full benefits of their social media experience. To that end, we created 5 digital aids that allowed users to modify their Facebook user interface and collected feedback on use of the aids in a longitudinal study with 39 Facebook users with TBI. While participants had varying needs and preferences for the aid extensions, overall, they had positive ratings for the aids' accessibility features. These findings suggest that aids like these have the potential to improve the social media experience of users with TBI.

Acknowledgments

We would like to thank Kim Walsh, Natalia Rivera, and Amy Koike for their contributions to our longitudinal study. This research was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (award number R01HD071089). Open-access publication was supported by the Sheldon B and Marianne S Lubar Professorship awarded to BM.

Authors' Contributions

Conceptualization: LT (lead), MD (equal), BM (equal), CLT (equal)
Data curation: DS (lead), LK (equal,), NNP (supporting), CYK (supporting)
Formal analysis: DS (lead), LK (equal), YH (supporting)
Funding acquisition: LT (lead), MD (equal), BM (equal), CLT (equal)
Investigation: NNP (lead), LK (equal), YH (supporting), CYK (supporting)
Methodology: HLJ (lead), NNP (supporting), LT (supporting), MD (supporting)
Project administration: MD (lead), LT (equal), CLT (equal), BM (equal)
Resources: MD (lead), LT (equal)
Software: YH (lead), CYK (supporting)
Supervision: MD (lead), LT (equal), BM (equal), CLT (supporting)
Validation: NNP (lead), LK (equal)
Visualization: DS
Writing – original draft: DS (lead), LK (supporting), HLJ (supporting), NNP (supporting), YH (supporting), CK (supporting)
Writing – review & editing: MD (lead), LT (equal), CLT (equal), BM (equal)

Conflicts of Interest

YH, LK, HLJ, LT, MD, CLT, and BM were involved in the development of the digital aids evaluated in this study (reported in Hu et al [8]. The authors have no financial or other competing interests to declare.

Multimedia Appendix 1
Demographic and injury characteristics of the 39 participants included in the longitudinal study. The "other" cause of traumatic brain injury includes highly specific circumstances that may otherwise identify participants. Although motorcycle collisions (MCCs) are encompassed within motor vehicle accidents (MVAs), they are categorized separately due to the unique risks and injury patterns associated with motorcycle-related injuries.
[XLSX File (Microsoft Excel File), 21 KB - rehab_v13i1e83639_app1.xlsx]

Multimedia Appendix 2
Questionnaires used during prestudy interviews and aid evaluation surveys. The former allowed the research team to assess participants' traumatic brain injury (TBI) and social media use histories, while the latter enabled evaluation of each Social Media Accessibility and Rehabilitation Toolkit (SMART-TBI) extension.

XSL•FO
RenderX

[PDF File (Adobe PDF File), 92 KB - rehab_v13i1e83639_app2.pdf]

Multimedia Appendix 3
Social media platform use—counts of participants who reported using major social media platforms.
[PNG File , 80 KB - rehab_v13i1e83639_app3.png]

Multimedia Appendix 4
Counts of participants who reported they would use each aid in their daily lives if given the opportunity. These data were collected during the poststudy survey, although additional endorsements were also captured during the poststudy interview.
[PNG File , 33 KB - rehab_v13i1e83639_app4.png]

References

1. Mazaux JM, Masson F, Levin HS, Alaoui P, Maurette P, Barat M. Long-term neuropsychological outcome and loss of social autonomy after traumatic brain injury. Arch Phys Med Rehabil 1997;78(12):1316-1320. [doi: 10.1016/s0003-9993(97)90303-8] [Medline: 9421984]
2. Sohlberg MM, Hamilton J, Turkstra L. Transforming Cognitive Rehabilitation: Effective Instructional Methods. New York, United States: Guilford Press; 2023.
3. Togher L, McDonald S, Coelho CA, Byom L. Cognitive communication disability following TBI: examining discourse, pragmatics, behaviour and executive function. In: Social and Communication Disorders Following Traumatic Brain Injury. Hove, England: Psychology Press; 2013:89-118.
4. Brunner M, Palmer S, Togher L, Hemsley B. 'I kind of figured it out': the views and experiences of people with traumatic brain injury (TBI) in using social media-self-determination for participation and inclusion online. Int J Lang Commun Disord 2019;54(2):221-233 [FREE Full text] [doi: 10.1111/1460-6984.12405] [Medline: 29873159]
5. Morrow EL, Zhao F, Turkstra L, Toma C, Mutlu B, Duff MC. Computer-mediated communication in adults with and without moderate-to-severe traumatic brain injury: survey of social media use. JMIR Rehabil Assist Technol 2021;8(3):e26586 [FREE Full text] [doi: 10.2196/26586] [Medline: 34448727]
6. Lim H, Kakonge L, Hu Y, Turkstra L, Duff M, Toma C, et al. So, I can feel normal: participatory design for accessible social media sites for individuals with traumatic brain injury. 2023 Presented at: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; April 23-28, 2023; Hamburg, Germany p. 1-19. [doi: 10.1145/3544548.3581222]
7. Flynn MA, Rigon A, Kornfield R, Mutlu B, Duff MC, Turkstra LS. Characterizing computer-mediated communication, friendship, and social participation in adults with traumatic brain injury. Brain Inj 2019;33(8):1097-1104 [FREE Full text] [doi: 10.1080/02699052.2019.1616112] [Medline: 31100990]
8. Hu Y, Lim H, Kakonge L, Mitchell J, Johnson H, Turkstra L, et al. SMART-TBI: design and evaluation of the social media accessibility and rehabilitation toolkit for users with traumatic brain injury. 2024 Presented at: Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility; October 27-30, 2024; St John's, Newfoundland and Labrador, Canada p. 1-19. [doi: 10.1145/3663548.3675641]
9. Academy of Neurologic Communication Disorders Traumatic Brain Injury Writing Committee, Byom L, O'Neil-Pirozzi TM, Lemoncello R, MacDonald S, Meulenbroek P, et al. Social communication following adult traumatic brain injury: a scoping review of theoretical models. Am J Speech Lang Pathol 2020;29(3):1735-1748. [doi: 10.1044/2020_AJSLP-19-00020] [Medline: 32569483]
10. Tsai YC, Liu CJ, Huang HC, Lin JH, Chen PY, Su YK, et al. A meta-analysis of dynamic prevalence of cognitive deficits in the acute, subacute, and chronic phases after traumatic brain Injury. J Neurosci Nurs 2021;53(2):63-68. [doi: 10.1097/JNN.0000000000000570] [Medline: 33538456]
11. Diamond A. Executive functions. Annu Rev Psychol 2013;64:135-168 [FREE Full text] [doi: 10.1146/annurev-psych-113011-143750] [Medline: 23020641]
12. Lannoo E, Colardyn F, Vandekerckhove T, De Deyne C, De Soete G, Jannes C. Subjective complaints versus neuropsychological test performance after moderate to severe head injury. Acta Neurochir (Wien) 1998;140(3):245-253. [doi: 10.1007/s007010050091] [Medline: 9638261]
13. Rochat L, Beni C, Annoni J, Vuadens P, Van der Linden M. How inhibition relates to impulsivity after moderate to severe traumatic brain injury. J Int Neuropsychol Soc 2013;19(8):890-898 [FREE Full text] [doi: 10.1017/S1355617713000672] [Medline: 23816263]
14. Cassel A, McDonald S, Kelly M, Togher L. Learning from the minds of others: a review of social cognition treatments and their relevance to traumatic brain injury. Neuropsychol Rehabil 2019;29(1):22-55. [doi: 10.1080/09602011.2016.1257435] [Medline: 27899036]
15. Allain P, Hamon M, Saoût V, Verny C, Dinomais M, Besnard J. Theory of mind impairments highlighted with an ecological performance-based test indicate behavioral executive deficits in traumatic brain injury. Front Neurol 2019;10:1367 [FREE Full text] [doi: 10.3389/fneur.2019.01367] [Medline: 32038455]

XSL•FO
RenderX

16. Spikman JM, Timmerman ME, Milders MV, Veenstra WS, van der Naalt J. Social cognition impairments in relation to general cognitive deficits, injury severity, and prefrontal lesions in traumatic brain injury patients. J Neurotrauma 2012;29(1):101-111. [doi: 10.1089/neu.2011.2084] [Medline: 21933011]
17. Channon S, Crawford S. Mentalising and social problem-solving after brain injury. Neuropsychol Rehabil 2010;20(5):739-759. [doi: 10.1080/09602011003794583] [Medline: 20526955]
18. McDonald S. Impairments in social cognition following severe traumatic brain injury. J Int Neuropsychol Soc 2013;19(3):231-246. [doi: 10.1017/S1355617712001506] [Medline: 23351330]
19. Bibby H, McDonald S. Theory of mind after traumatic brain injury. Neuropsychologia 2005;43(1):99-114. [doi: 10.1016/j.neuropsychologia.2004.04.027] [Medline: 15488910]
20. Bosco FM, Angeleri R, Sacco K, Bara BG. Explaining pragmatic performance in traumatic brain injury: a process perspective on communicative errors. Int J Lang Commun Disord 2015;50(1):63-83 [FREE Full text] [doi: 10.1111/1460-6984.12114] [Medline: 25039503]
21. Turkstra LS, Norman RS, Mutlu B, Duff MC. Impaired theory of mind in adults with traumatic brain injury: a replication and extension of findings. Neuropsychologia 2018;111:117-122 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2018.01.016] [Medline: 29366949]
22. Vascello MGF, Pizzighello S, Spada MS, Martinuzzi A, Dalmaso M. Social face processing in chronic severe traumatic brain injury: altered decoding of emotions and mental states but preserved gaze cueing of attention. Neuropsychologia 2024;203:108975 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2024.108975] [Medline: 39179200]
23. Rodríguez-Rajo P, García-Rudolph A, Sánchez-Carrión R, Aparicio-López C, Enseñat-Cantallops A, García-Molina A. Social and nonsocial cognition: are they linked? A study on patients with moderate-to-severe traumatic brain injury. Appl Neuropsychol Adult 2022;29(5):1039-1048. [doi: 10.1080/23279095.2020.1845171] [Medline: 33174449]
24. Byom L, Turkstra LS. Cognitive task demands and discourse performance after traumatic brain injury. Int J Lang Commun Disord 2017;52(4):501-513 [FREE Full text] [doi: 10.1111/1460-6984.12289] [Medline: 27891725]
25. Filipčíková M, Quang H, Cassel A, Darke L, Wilson E, Wearne T, et al. Exploring neuropsychological underpinnings of poor communication after traumatic brain injury: the role of apathy, disinhibition and social cognition. Int J Lang Commun Disord 2024;59(2):433-448. [doi: 10.1111/1460-6984.12836] [Medline: 36541559]
26. Westerhof-Evers HJ, Fasotti L, van der Naalt J, Spikman JM. Participation after traumatic brain injury: the surplus value of social cognition tests beyond measures for executive functioning and dysexecutive behavior in a statistical prediction model. Brain Inj 2019;33(1):78-86 [FREE Full text] [doi: 10.1080/02699052.2018.1531303] [Medline: 30325200]
27. Togher L, Wiseman-Hakes C, Douglas J, Stergiou-Kita M, Ponsford J, Teasell R, INCOG Expert Panel. INCOG recommendations for management of cognition following traumatic brain injury, part IV: cognitive communication. J Head Trauma Rehabil 2014;29(4):353-368. [doi: 10.1097/HTR.0000000000000071] [Medline: 24984097]
28. VanSolkema M, McCann C, Barker-Collo S, Foster A. Attention and communication following TBI: making the connection through a meta-narrative systematic review. Neuropsychol Rev 2020;30(3):345-361. [doi: 10.1007/s11065-020-09445-5] [Medline: 32712759]
29. Wise EK, Mathews-Dalton C, Dikmen S, Temkin N, Machamer J, Bell K, et al. Impact of traumatic brain injury on participation in leisure activities. Arch Phys Med Rehabil 2010;91(9):1357-1362. [doi: 10.1016/j.apmr.2010.06.009] [Medline: 20801252]
30. Jourdan C, Bayen E, Pradat-Diehl P, Ghout I, Darnoux E, Azerad S, et al. A comprehensive picture of 4-year outcome of severe brain injuries. Results from the PariS-TBI study. Ann Phys Rehabil Med 2016;59(2):100-106 [FREE Full text] [doi: 10.1016/j.rehab.2015.10.009] [Medline: 26704071]
31. Hoofien D, Gilboa A, Vakil E, Donovick PJ. Traumatic brain injury (TBI) 10-20 years later: a comprehensive outcome study of psychiatric symptomatology, cognitive abilities and psychosocial functioning. Brain Inj 2001;15(3):189-209. [doi: 10.1080/026990501300005659] [Medline: 11260769]
32. Engberg AW, Teasdale TW. Psychosocial outcome following traumatic brain injury in adults: a long-term population-based follow-up. Brain Inj 2004;18(6):533-545. [doi: 10.1080/02699050310001645829] [Medline: 15204335]
33. Doig E, Fleming J, Tooth L. Patterns of community integration 2-5 years post-discharge from brain injury rehabilitation. Brain Inj 2001;15(9):747-762. [doi: 10.1080/02699050110034343] [Medline: 11516344]
34. McDonald S, Genova H. The effect of severe traumatic brain injury on social cognition, emotion regulation, and mood. Handb Clin Neurol 2021;183:235-260. [doi: 10.1016/B978-0-12-822290-4.00011-6] [Medline: 34389120]
35. Hu Y, Lim H, Johnson HL, O'Shaughnessy J, Kakonge L, Turkstra L, et al. Investigating day-to-day experiences with conversational agents by users with traumatic brain injury. 2023 Presented at: Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility; October 22-25, 2023; New York, United States p. 1-15. [doi: 10.1145/3597638.3608385]
36. Dahlberg C, Hawley L, Morey C, Newman J, Cusick CP, Harrison-Felix C. Social communication skills in persons with post-acute traumatic brain injury: three perspectives. Brain Inj 2006;20(4):425-435. [doi: 10.1080/02699050600664574] [Medline: 16716988]

XSL•FO
RenderX

37. MacDonald S. Introducing the model of cognitive-communication competence: a model to guide evidence-based communication interventions after brain injury. Brain Inj 2017;31(13-14):1760-1780 [FREE Full text] [doi: 10.1080/02699052.2017.1379613] [Medline: 29064304]
38. Brunner M, Palmer S, Togher L, Dann S, Hemsley B. "If I knew what I was doing on Twitter then I would use it more": Twitter experiences and networks of people with traumatic brain injury (TBI). Brain Impairment 2019;21(1):1-18. [doi: 10.1017/brimp.2019.12]
39. Toma CL, Hwang J, Kakonge L, Morrow EL, Turkstra LS, Mutlu B, et al. Does Facebook use provide social benefits to adults with traumatic brain injury? Cyberpsychol Behav Soc Netw 2024;27(3):214-220 [FREE Full text] [doi: 10.1089/cyber.2023.0211] [Medline: 38466929]
40. Brunner M, Hemsley B, Palmer S, Dann S, Togher L. Review of the literature on the use of social media by people with traumatic brain injury (TBI). Disabil Rehabil 2015;37(17):1511-1521. [doi: 10.3109/09638288.2015.1045992] [Medline: 25985717]
41. Tsaousides T, Matsuzawa Y, Lebowitz M. Familiarity and prevalence of facebook use for social networking among individuals with traumatic brain injury. Brain Inj 2011;25(12):1155-1162. [doi: 10.3109/02699052.2011.613086] [Medline: 21961574]
42. Feuston JL, Marshall-Fricker CG, Piper AM. The social lives of individuals with traumatic brain injury. 2017 Presented at: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, Colorado p. 182-194. [doi: 10.1145/3025453.3025784]
43. Brunner M, Palmer S, Togher L, Dann S, Hemsley B. Content analysis of tweets by people with traumatic brain injury (TBI): implications for rehabilitation and social media goals. 2019 Presented at: Proceedings of the 52nd Annual Hawaii International Conference on System Sciences (HICSS 2019); January 8-11, 2019; Wailea, Hawaii. [doi: 10.24251/hicss.2019.523]
44. Clough S, Morrow E, Mutlu B, Turkstra L, Duff MC. Emotion recognition of faces and emoji in individuals with moderate-severe traumatic brain injury. Brain Inj 2023;37(7):596-610 [FREE Full text] [doi: 10.1080/02699052.2023.2181401] [Medline: 36847497]
45. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of social media. Bus Horiz 2010;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
46. Brunner M, Rietdijk R, Avramovic P, Power E, Miao M, Rushworth N, et al. Developing social-ABI-lity: an online course to support safe use of social media for connection after acquired brain injury. Am J Speech Lang Pathol 2023;32(2S):924-940. [doi: 10.1044/2022_AJSLP-22-00099] [Medline: 36492293]
47. Ahmadi R, Lim H, Mutlu B, Duff M, Toma C, Turkstra L. Facebook experiences of users with traumatic brain injury: a think-aloud study. JMIR Rehabil Assist Technol 2022;9(4):e39984 [FREE Full text] [doi: 10.2196/39984] [Medline: 36525296]
48. Zhao F, Lim H, Morrow EL, Turkstra LS, Duff MC, Mutlu B. Designing evidence-based support aids for social media access for individuals with moderate-severe traumatic brain injury: a preliminary acceptability study. Front Digit Health 2022;4:991814 [FREE Full text] [doi: 10.3389/fdgth.2022.991814] [Medline: 36606124]
49. IBM Watson Natural Language Understanding. IBM. URL: https://www.ibm.com/products/natural-language-understanding [accessed 2023-09-13]
50. Check your texts online: grammar, spelling, and readability. TextGears. URL: https://textgears.com/ [accessed 2023-09-13]
51. Enabling online conversations. Perspective. URL: https://perspectiveapi.com/ [accessed 2023-09-13]
52. Duff MC, Morrow EL, Edwards M, McCurdy R, Clough S, Patel N, et al. The value of patient registries to advance basic and translational research in the area of traumatic brain injury. Front Behav Neurosci 2022;16:846919 [FREE Full text] [doi: 10.3389/fnbeh.2022.846919] [Medline: 35548696]
53. Malec JF, Brown AW, Leibson CL, Flaada JT, Mandrekar JN, Diehl NN, et al. The Mayo classification system for traumatic brain injury severity. J Neurotrauma 2007;24(9):1417-1424. [doi: 10.1089/neu.2006.0245] [Medline: 17892404]
54. Lund AM. Measuring usability with the use questionnaire12. Usability interface 2001;8(2):3-6 [FREE Full text]
55. Ledbetter AM. Measuring online communication attitude: Instrument development and validation. Commun Monogr 2009;76(4):463-486. [doi: 10.1080/03637750903300262]
56. Rosenberg M. Society and the Adolescent Self-Image. Princeton, New Jersey: Princeton University Press; 1965.
57. Diener E. Satisfaction with life scale. (J Pers Assess 1985;49(1):71-75. [doi: 10.1037/t01069-000]
58. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
59. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qual Psychol 2022;9(1):3-26. [doi: 10.1037/qup0000196]
60. Dedoose version 8.0.35: web application for managing, analyzing, and presenting qualitative and mixed method research data. Dedoose. 2018. URL: https://www.dedoose.com/ [accessed 2024-06-30]

XSL•FO
RenderX

Abbreviations

GCS: Glasgow Coma Scale
HIPAA: Health Insurance Portability and Accountability Act
IRB: institutional review board
LOC: loss of consciousness
PTA: posttraumatic amnesia
SMART-TBI: Social Media Accessibility and Rehabilitation Toolkit
TBI: traumatic brain injury
USE: Usefulness, Satisfaction, and Ease of Use

Edited by S Munce; submitted 11.Sep.2025; peer-reviewed by J Fielder, D Bethere, L Herdiman; comments to author 02.Jul.2026; accepted 13.Jul.2026; published 26.Aug.2026.

Please cite as:
Sullivan D, Patel NN, Kakonge L, Kim CY, Hu Y, Johnson HL, Toma CL, Duff M, Turkstra L, Mutlu B
Digital Aids for Social Media in Adults With Traumatic Brain Injury: Longitudinal Evaluation Study
JMIR Rehabil Assist Technol 2026;13:e83639
URL: https://rehab.jmir.org/2026/1/e83639
doi:10.2196/83639
PMID:

©Dakota Sullivan, Nirav N Patel, Lisa Kakonge, Callie Y Kim, Yaxin Hu, Hailey L Johnson, Catalina L Toma, Melissa Duff, Lyn Turkstra, Bilge Mutlu. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The SmartGlove BISINDO Wearable Sensor and Large Language Model–Based Sign Language–to-Speech System for Users With Communication Disorders: Design Science Research Study

Suhartono Suhartono[1*], Prof Dr; Bambang Yulianto[1*], Prof Dr; Burhanuddin Arafah[2*], PhD; Kusubakti Andajani[3*], Prof Dr; Abdul Kholiq[1*], MPd; Muhammad Ridha Anugrah Latief[2*], MHum; Bambang Prastio[3], MPd; Diding Wahyudin Rohaedi[1*], MHum; Deshinta Arrova Dewi[4], Prof Dr; Taqdir Taqdir[2], MHum

[1]Department of Indonesian Language and Literature Education, Universitas Negeri Surabaya, Jalan Lidah Wetan, Lakarsantri, Surabaya, East Java, Indonesia

[2]Hasanuddin University, Makassar, South Sulawesi, Indonesia

[3]State University of Malang, Malang, East Java, Indonesia

[4]INTI International University, Putra Nilai, Negeri Sembilan, Malaysia

[*]these authors contributed equally

Corresponding Author:
Suhartono Suhartono, Prof Dr
Department of Indonesian Language and Literature Education, Universitas Negeri Surabaya, Jalan Lidah Wetan, Lakarsantri, Surabaya, East Java, Indonesia

Abstract

Background: In Indonesia, people with communication disorders predominantly communicate using Bahasa Isyarat Indonesia (BISINDO). Nevertheless, all current systems for BISINDO recognition generate fragmented, word-by-word output and lack natural language generation (NLG) capabilities, thereby fundamentally constraining their communicative utility.

Objective: The aim of this study is to develop, validate, and field-evaluate SmartGlove BISINDO, an AI-assisted wearable sign language–to-speech system that embeds large language model (LLM)–powered NLG into a BISINDO translation pipeline for the first time.

Methods: This study followed the design science research (DSR) framework, covering phases 1 through 5 (problem identification, solution objectives, prototype design and development, demonstration, and evaluation). Data were collected from 12 purposively selected informants (6 domain experts and 6 deaf education companions/teachers) across 3 Indonesian cities (Malang, Surabaya, and Makassar), using a needs-analysis questionnaire and an expert validation sheet, and analyzed through statistical triangulation of agreement percentage, content validity ratio (CVR), and Aiken V. The validated prototype was field-tested in Makassar and Malang with 15 participants with communication disorders and 4 observers or companions; it was evaluated by a disability expert in Surabaya.

Results: The needs analysis yielded a grand mean of 4.77 (SD 0.51) across 4 clusters, with 89% mean agreement, indicating very high priority for all identified needs. The prototype's 5-layer architecture integrates wearable sensors, on-device gesture classification, and LLM-based NLG. Expert validation showed an overall Aiken V of 0.94, with all 14 items exceeding the 0.75 validity cutoff and with 11/14 items (79%) rated highly valid (V>0.90). The grand mean CVR was 0.76 (SD 0.16), exceeding the Lawshe minimum threshold of 0.56 (P<.05), with 3 items—gesture differentiation, AI contextual relevance, and settings configuration ease—falling below the threshold and indicating areas for refinement. During the testing phase, the system received a lower practicality score (mean 3.35, SD 0.16; 67%), particularly for clarity of speech output. However, observers and companions gave a significantly higher score of 88%, with a mean of 4.40 (SD 0.08). A disability expert also gave a score of 81% (with a mean value calculated across the 20 questionnaire items of 4.05, SD 0.60). Regarding language, the AI-generated output underwent three transformation processes: addition, deletion, and substitution. Of the 3, substitution is the most likely to alter the intended meaning.

Conclusions: SmartGlove BISINDO is a wearable sign language–to-speech system that introduces the BISINDO research landscape to LLM-based NLG connect-and-translate approaches and demonstrates strong content validity. This study presents preliminary evidence toward a replicable model for interdisciplinary DSR-based prototype development that combines augmentative and alternative communication (AAC) theory, systemic functional linguistics, and design science methodology, intended as a reference for other underrepresented sign language contexts in Southeast Asia and beyond. Field evaluation further indicated favorable practicality overall, alongside a gap between expert-perceived feasibility and user-rated experience that will guide subsequent design refinement.

XSL•FO
RenderX

(JMIR Rehabil Assist Technol 2026;13:e96258) doi:10.2196/96258

KEYWORDS

assistive communication technology; sign language recognition; natural language generation; large language model; wearable sensing system; BISINDO; communication disorders; augmentative and alternative communication; inclusive technology; process innovation; Bahasa Isyarat Indonesia

Introduction

Background

Communication is a fundamental human capacity that enables individuals to express thoughts and needs while building the social relationships necessary for daily life. Around 1.3 billion people worldwide have some form of disability [1]; despite being described as a "global crisis," communication disorders remain underrepresented in global health policy, even though their negative impacts on education, employment, and social participation, as well as their associated psychosocial effects such as stigma, low self-confidence, and reduced quality of life, are well documented [2-4]. Effective assistive systems for this population must therefore be natural, meaningful, and socially appropriate [2,3]. In Indonesia, approximately 1.02% of the population experiences severe communication difficulties [5], resulting in systematic marginalization for hundreds of thousands of people.

Most deaf people in Indonesia communicate using Indonesian Sign Language (Bahasa Isyarat Indonesia [BISINDO]), which emerged naturally within the Indonesian deaf community and has its own grammar and vocabulary distinct from spoken Indonesian [6]. Because very few people outside the deaf community understand BISINDO, individuals who are deaf face limited opportunities for social interaction [7]. This study responds to the growing need for pragmatic assistive technology (AT) research addressing unmet rehabilitation needs [8], building on recent evidence confirming the feasibility and value of novel AT solutions for this population [9].

Current Trends and Challenges in Sign Language Recognition Technology

Several previous studies on BISINDO recognition have shown a consistent relationship between accuracy, sentence processing capability, and practical application. For example, a convolutional neural network–long short-term memory (CNN-LSTM) architecture achieved individual gesture classification accuracy of up to 96%, but only for isolated word-level gestures, not for continuous sentences [10]. A mobile CNN-MobileNetV2-based app enabled real-time BISINDO-to-speech conversion via camera, but its classification accuracy was only 55% across 30 gesture categories. This significant decrease limits its reliability in the real world [11]. More recent research has attempted to extend recognition to the sentence level using 3D CNN-LSTM and 3D CNN–bidirectional LSTM (CNN-BiLSTM) models with connectionist temporal classification. These studies successfully overcome the limitations of isolated gestures, but these approaches remain camera-based and only go as far as mapping gestures to text, without natural language generation (NLG) capabilities [12].

Related research developments also come from the 2026 SIBI (Indonesian Sign System) recognition system, which integrates LSTM-based gesture recognition with Indo Bidirectional Encoder Representations from Transformers (BERT) contextual modeling. Word prediction accuracy reached 89%, but that was only for word-level correction, not full sentence generation. The system also did not specifically address BISINDO [13]. On the other hand, a 2025 transformer-based benchmark using the Sign Pose-based Transformer (SPOTER), Siformer, and Multiscale Vision Transformer version 2 (MViTv2) architectures achieved up to 97% accuracy in isolated gesture classification. Meanwhile, a 2026 comparison between MobileNetV3 and EfficientNetV2B0 successfully improved the device's efficiency. However, both were still limited to gesture classification, without addressing sentence-level meaning [14,15]. Across these studies, 3 limitations recurred: reliance on camera input (rather than wearable sensing), focus on isolated word or gesture recognition, and lack of generative NLG capabilities.

For resource-intensive sign languages, sign language recognition has advanced beyond sentence processing. Transformer-based approaches, for example, are capable of performing end-to-end recognition and translation simultaneously [16]. Sequential parallelism across continuous signage has also been modeled by CNN–hidden Markov model systems that use deep learning on sequential data [17,18]. However, these advances were made in American Sign Language and Deutsche Gebärdensprache, which have largely focused on annotated corpora. BISINDO lacks such resources. Therefore, existing architectures and training methods cannot be readily adapted to sign languages with lower resources and different structures like BISINDO.

Beyond BISINDO-specific work, the broader sign language translation literature has moved decisively toward large language model (LLM)–based, gloss-to-text pipelines conceptually similar to the present system. One approach couples a lightweight Sign Pose-based Transformer (SPOTER) with an off-the-shelf LLM prompted to convert sequences of spotted glosses into fluent spoken-language sentences, requiring no additional translation-specific training or paired corpora [19]. Others have extended this idea to explicitly low-resource sign languages through in-context prompting rather than fine-tuning, to compensate for scarce parallel data [20]. However, a 2025 systematic review of LLM- and multimodal-LLM–based sign language recognition and translation work covering 2020 - 2025 cautions that most reported gains remain concentrated in a small number of well-resourced sign languages, chiefly American Sign Language, and that evaluation continues to rely heavily on machine-translation metrics such as Bilingual Evaluation Understudy (BLEU), which are known to be weak proxies for communicative adequacy and are rarely validated with deaf or hard-of-hearing end users [21]. This pattern reinforces the gap

XSL•FO
RenderX

motivating the present study: existing gloss/SPOTER-to-LLM pipelines have so far been designed and evaluated for camera-based, high-resource sign languages, and none have been extended to a wearable-sensor pipeline or evaluated directly with the target BISINDO user population.

The current BISINDO system does not incorporate automatic NLG to transform detected gesture sequences into semantically coherent complete sentences. This is a critical gap. Sign languages convey meaning at the level of complete sentences, not just isolated gestures but also using gestures in context. In other words, systems not designed for AAC (augmentative and alternative communication) do not allow users to share the same language as their interlocutors [2,3]. Other researchers have overlooked some of the real potential of AAC from LLMs. For example, an LLM-based interface achieved a 29% text input rate in AAC users with amyotrophic lateral sclerosis [22]. In 2025, another study incorporated an LLM into a social support robot intended to interact with humans in real-time as part of geriatric care. The results showed a significant increase in interaction success and user engagement compared to a non-LLM dialogue system [23]. This demonstrates that LLMs as part of assistive devices for real-time interaction are not only capable but also beneficial in many areas. However, at the level of BISINDO translation, this potential is substantially underused.

Theoretical Framework

Three theoretical perspectives converge in this research, together informing the design science research (DSR) process's rigor cycle, each addressing a different layer of the same design problem. AAC theory sets the functional goal: assistive technology (AT) output must be natural and socially acceptable [2,3], a standard that isolated words, carrying no social or contextual meaning, cannot meet. Systemic functional linguistics (SFL) translates this goal into concrete linguistic properties that sign languages, like all languages, express through 3 simultaneous metafunctions [24]: the ideational metafunction conveys experiential content, the interpersonal metafunction encodes the social relationship between speaker and listener, and the textual metafunction structures the message as a coherent whole. Lexical translation alone cannot meet these functional criteria; LLM-based NLG, built on transformer architectures [25], provides the computational mechanism to derive grammatical, coherent, context- and intent-appropriate text from a queue of detected gestures [22,23]. In this chain, AAC theory defines the target, SFL specifies the linguistic properties required to reach it, and LLM-based NLG provides the technical means of fulfilling that demand. This chain directly informed the expert validation instrument, whose items evaluate grammaticality, naturalness, and contextual relevance, operationalizing the SFL metafunctions as validity metrics for the prototype.

Study Objectives

Within this framework, this study aims to design and evaluate the BISINDO SmartGlove, which will be recognized as a wearable device in the context of AI and the implementation of a sign language–to-speech system based on DSR [26,27]. In this study, DSR phases 1 to 5 were carried out: problem identification and motivation; definition of solution objectives; design and development of artifacts carried out through demonstration and evaluation [28]. This study focuses on indexing the theories of AAC, SFL, and DSR under one interdisciplinary lens. The study hopes that the DSR-based prototype development model can serve to develop inclusive AI AT. This can provide a template for other researchers that can be used in underrepresented sign language environments. Practically, this research meets the communication needs of Indonesian users with severe speech disorders. This is directly related to SDG (Sustainable Development Goal) 4 on quality education and SDG 10 on reducing inequalities. The way to do this is to allow people with hearing impairments and speech disorders to communicate independently in schools and social spaces. SmartGlove BISINDO supports SDG 4, which promotes equal access to education for persons with disabilities, and SDG 10, which calls for full social, economic, and political inclusion regardless of disability.

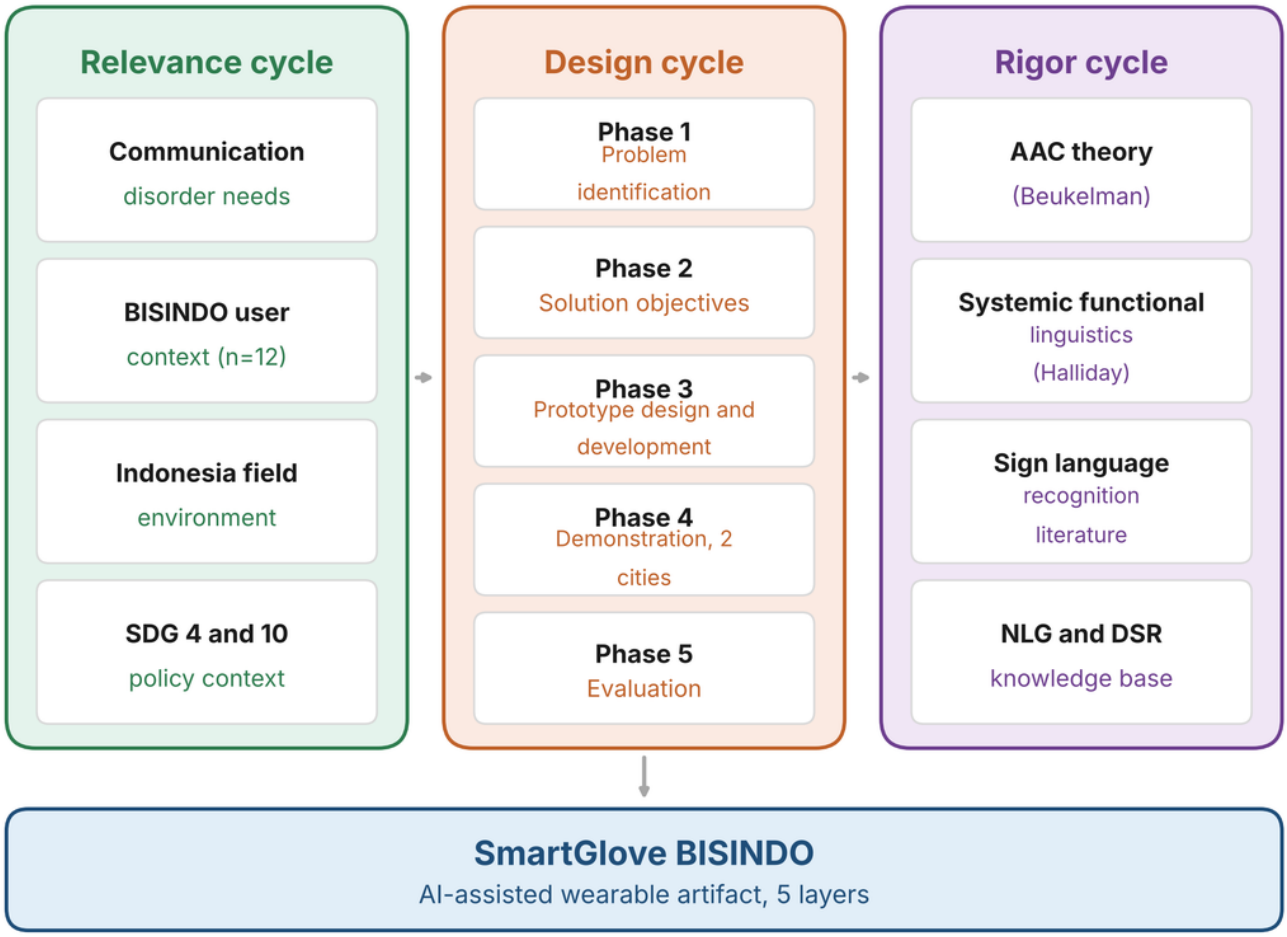
Methods

Research Design and Framework

This study follows 2 complementary DSR frameworks: Hevner [29] 3-cycle model and Peffers et al [27] 6-phase DSR, which covers problem identification, solution objectives, design and development, demonstration, evaluation, and communication. The DSR is one of the most widely adopted procedural frameworks in DSR [28]. DSR was chosen because it is specifically tailored to research that results in innovative technological artifacts as solutions to real-world problems. This iterative, user-centered approach reflects established AT design practices that incorporate stakeholder input throughout development [26-30]. The relevance cycle is operationalized via an analysis of needs with 12 participants, the design cycle through iterative design of a 5-layer prototype, and the rigor cycle through AAC theory and SFL as the guiding knowledge bases, as summarized in Figure 1.

XSL•FO
RenderX

Figure 1. Design science research (DSR) framework of SmartGlove Bahasa Isyarat Indonesia (BISINDO) study [26]. AAC: augmentative and alternative communication; NLG: natural language generation; SDG: Sustainable Development Goal.



Instrument Development

Both instruments were developed specifically for this study rather than adapted from preexisting validated scales, reflecting the absence of standardized measurement tools for BISINDO-specific, LLM-integrated AT. Item development followed a theory-driven operationalization process: constructs were first derived from the 3 theoretical perspectives described above (AAC theory, SFL, and DSR), then translated into observable, measurable indicators. For example, the SFL textual and interpersonal metafunctions were operationalized into needs-analysis items on sentence completeness and context-appropriate register and into validation items on grammaticality, naturalness, and contextual relevance. Both instruments operationalize the theoretical constructs associated with the DSR relevance and rigor cycles [27]. Beyond this theory-driven design process, construct validity was assessed through content validity review by 2 independent experts prior to field use and further evidenced empirically through the content validity ratio (CVR) and Aiken V results reported for each item.

Patient and Public Involvement

This study involved 6 deaf education companions and teachers from 3 Indonesian cities (Malang, Surabaya, and Makassar) as respondents in the needs analysis and as expert validators. Dataset curation was performed by a certified BISINDO dataset specialist. This informed the key design decisions. After validation, the prototype directly involves users with communication disorders, aligning with JMIR rehab AT's interest in research that includes relevant knowledge users [8].

Twelve participants were purposively sampled for their expertise and direct experience in communication contexts involving users with communication disorders in Indonesia. Although a sample of 12 is not statistically significant when using the proper computation of CVR, it is sufficient for CVR analysis, as per Warner's recommendations, since 12 participants meet the minimum required ceiling threshold (CVRthres) of 0.56 at $P<.05$. The participants were divided into 2 groups, namely, 6 domain experts (consisting of 2 sign language specialists, 2 AI technology experts, and 2 linguists) and 6 deaf education companions and teachers drawn from 3 Indonesian cities: Malang (n=2), Surabaya (n=2), and Makassar (n=2). All participants acted as both respondents for needs analysis and validators of prototype quality assessment.

Needs Analysis

The needs analysis instrument was a questionnaire covering 4 clusters and 10 items, rated on a 5-point Likert scale (full item list in Multimedia Appendix 1), developed to operationalize the theoretical requirements of AAC and SFL within the DSR relevance cycle. Content validity of the questionnaire was reviewed by 2 independent experts, who were not part of the 12-member study sample, before field use. After a brief introduction to the system concept, 12 participants completed

XSL•FO
RenderX

a requirements analysis questionnaire. Responses were then analyzed using descriptive statistics, including percentages, means, and SDs. Priority levels were divided into 2 categories: very high (mean ≥4.60) or high (mean 4.00 - 4.59).

Prototype Design

The prototype development was based on the DSR design cycle. Based on requirement analysis and literature review on existing hardware and software, the initial version was further refined based on observations from wearable use in practice and developments in NLG systems. The design resulted in an architecture specification with technical descriptive analysis and system architecture diagrams, which are 5-layered.

BISINDO Dataset

The BISINDO dataset has been constructed with the help of a real-life handwritten BISINDO gesture specialist. The expert acted as a demonstrator and authenticity validator during curation. Expert reviews were performed to refine all input gestures derived from the BISINDO standard vocabulary reference. To decrease the ambiguity of classification, some gestures were infrequently used or simply too similar to others (like most coordinate conjunctions, certain particular prepositions like “from” and “for,” or some physical verbs like “input” and “extract”). Additionally, the expert enriched the vocabulary with local references. This dataset comprises terms from the social and religious life in Indonesia such as greetings, emergency phrases, and words for extended families. An expert guiding such a process fits in well with user-centered AAC design methods [2,3] and aligns directly with the cultural and linguistic relevance criteria uncovered in needs analysis.

Expert Validation

The expert validation instrument was designed in a spreadsheet format. It consisted of 3 components and 14 items (full item list in Multimedia Appendix 2), all rated on a 5-point Likert-type scale, designed to operationalize the theoretical constructs underlying the DSR rigor cycle. Prior to the application of the instrument, 2 independent experts evaluated its content validity. Note that these experts are not part of the 12-person sample, who were followed by the research. The 12 participants undertook the role of validators after a complete demonstration of the system architecture. The team used statistical triangulation to validate the data. They examined 3 indices in parallel: percentage agreement, CVR, and Aiken V (where CVR is calculated using the Lawshe formula [31], representing a minimal threshold of 0.56, for n=12 at $P<.05$). In Aiken V [32], items are considered valid if V is greater than or equal to 0.75 and highly valid if V is greater than 0.90. The triangulation of multiple indices is more trustworthy than relying on a single index [30]. This method helps to solve the well-known limitation of using a single outcome between-participants evaluation framework in previous studies reviewing rehabilitation technologies [32]. A sample of 12 participants was purposively selected. The sample consisted of 2 sign language specialists, 2 AI technology experts, 2 linguists, and 6 deaf education companions and teachers drawn from 3 Indonesian cities: Malang, Surabaya, and Makassar. This composition was intended to be interdisciplinary, consistent with the relevance and rigor cycle of the DSR framework. Use of percentage agreement, CVR, and Aiken V concurrently was a check on interrater consistency across this heterogeneous panel. For the sample size, formal interrater reliability coefficients were not calculated.

Field Demonstration and Evaluation

Following expert validation, the validated prototype was field-tested with user participants and observers/companions in 2 Indonesian cities, Makassar and Malang, with an independent disability-expert assessment conducted in Surabaya, corresponding to DSR phases 4 and 5 (demonstration and evaluation). The field evaluation was measured using 3 instruments from different perspectives, consistent with recommended approaches for evaluating the usability of rehabilitation technologies across multiple stakeholder groups [33]. First, a practicality questionnaire was completed by users with communication impairments. Second, a parallel questionnaire was administered to an observer or companion present during use. Third, a validation instrument was administered by a disability expert. Because each instrument measures something different and results are expected to vary, analyses of these 3 sets were conducted separately. Furthermore, the study analyzed linguistic performance by taking each interaction and comparing the raw word sequences of gestures (system input) with the final sentences generated by the AI (system output). This was done qualitatively. The observed changes were linguistic shifts generated through the AI-assisted process.

Ethical Considerations

The study was carried out in accordance with the Declaration of Helsinki and applicable national guidelines for research involving human participants. All participants provided written informed consent, including details on the study objective, procedures, voluntary participation, data confidentiality, and withdrawal rights. Twelve participants signed informed consent before data collection. Field demonstration and evaluation participants, comprising 15 user participants with communication disorders, 4 observers or companions, and 1 disability expert, separately provided informed consent directly, obtained with the assistance of a sign language interpreter for user participants with communication disorders. There was no collection of personal health information; data from the 12 expert and stakeholder participants related to their professional knowledge and attitudes toward system design and validation, while data from field demonstration and evaluation participants related solely to their experience using the prototype and their perceptions of its practicality. Data collection for all stages of the research (needs analysis, expert validation, and field demonstration) was conducted in accordance with research ethics. Ethical approval was formally granted as an Ethical Exemption (reference number 183/UN38.16/KEP/EC/2026) by the Research Ethics Committee of the Lembaga Penelitian dan Pengabdian kepada Masyarakat (Institute for Research and Community Service) at Universitas Negeri Surabaya, confirming that the study met the 7 World Health Organization ethical standards of 2011, based on the 2016 Council for International Organizations of Medical Sciences guidelines.

XSL•FO
RenderX

Results

Needs Analysis

The needs analysis used a 5-point Likert-scale questionnaire completed by 12 participants, covering 4 clusters and 10 indicators as presented in Table 1.

Table . SmartGlove Bahasa Isyarat Indonesia (BISINDO) user needs analysis results.

Cluster and indicator	Mean agreement, %	Mean (SD)
Response speed and accuracy		
Speech output <1 second after gesture completion	100	4.92 (0.29)
Real-time gesture recognition without delay	100	4.83 (0.39)
Minimization of unintended gesture detection errors	83	4.67 (0.65)
Naturalness of language output		
Output as complete grammatical sentences	92	4.83 (0.39)
Sentences comprehensible to lay conversation partners	83	4.75 (0.62)
Context adjustment for formal/informal situations	75	4.58 (0.79)
Portability and ease of use		
Lightweight portable device across contexts	75	4.67 (0.78)
Independent operation without assistance	92	4.75 (0.45)
Cultural and linguistic relevance		
Based on authentic BISINDO vocabulary	100	4.92 (0.29)
Coverage of at least 150 daily-life gestures	92	4.75 (0.45)
Total mean score	89	4.77 (0.51)

Across all 4 clusters, the needs analysis showed very high and homogeneous agreement (grand mean 4.77, SD 0.51; mean agreement 89%). Cultural and linguistic relevance had the highest cluster mean (4.84, SD 0.37; mean agreement 96%) and lowest SD, indicating strong consensus among experts; its top items, response speed and accuracy (item 1.1) and authentic vocabulary basis (item 4.1), each achieved unanimous agreement (mean 4.92, SD 0.29; 12/12, 100%). The response speed and accuracy cluster averaged 4.81 (SD 0.44; mean agreement 94%), naturalness of language output averaged 4.72 (SD 0.60; mean agreement 83%) with one item (context adjustment for formal/informal situations) falling into the High rather than very high category (mean 4.58, SD 0.79), and portability and ease of use averaged 4.71 (SD 0.62; mean agreement 83%).

Each cluster translated directly into a design principle for the prototype. The response speed and accuracy requirement set a sub-1-second end-to-end latency budget across the architecture. The naturalness of language output requirement made the NLG module a required rather than optional architectural element, driving the layer 4 design. Portability and independent-operation requirements pushed the hardware toward wearable, microcontroller-based systems with wireless communication that do not depend on external infrastructure. Finally, the requirement for authentic BISINDO vocabulary coverage of at least 150 gestures set the minimum dataset size.

Prototype Design

Informed by the 4 need clusters, SmartGlove BISINDO was designed as a 5-layer integrated system connecting sensor data acquisition to speech output. The prototype design was based on the DSR design cycle.

The workflow is illustrated in Figure 2. Once set up, both gloves (left as slave and right as master) sample sensor data at 100 Hz at matching times. The left glove itself acquires 5 flex values and 6 inertial measurement unit (IMU) values, which are transmitted to the right glove through ESP-NOW. The right glove uses a data aggregator to aggregate the data with sync timestamping, sending an array of 22 Float32 features per frame over BLE 5.0 to the Android app. The data then goes through a 3-step preprocessing pipeline before being classified by the TFLite (Google LLC) model. Gestures with confidence levels above the determined threshold of 85% or higher were piped into the AI language wrapper to convert NLG-generated text from the LLM into normal sentences in Bahasa Indonesia, and then vocalized using the Google text-to-speech (TTS) Engine detailed in Table 2.

XSL•FO
RenderX

Figure 2. Flowchart of SmartGlove Bahasa Isyarat Indonesia (BISINDO) integrated system architecture. BLE: Bluetooth Low Energy; LLM: large language model.

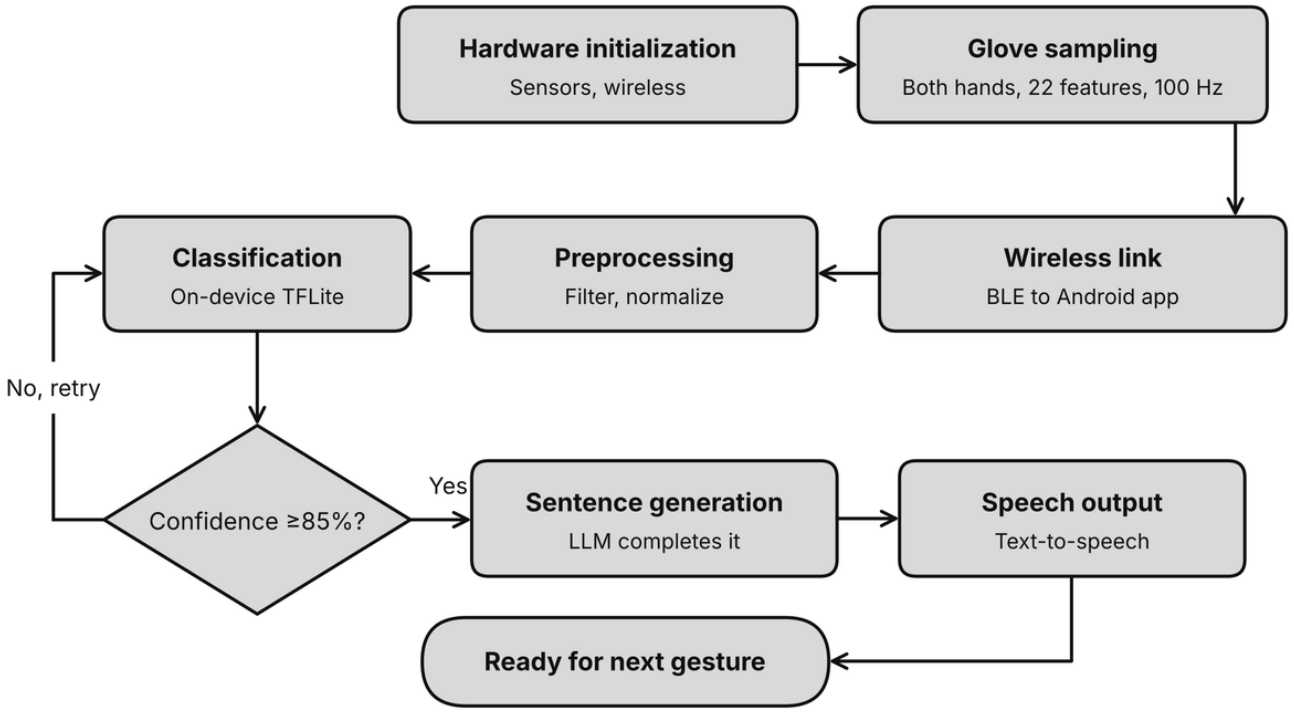


Table . Technical specifications: 5-layer SmartGlove Bahasa Isyarat Indonesia (BISINDO) architecture.

Layer	Key components	Key specifications	Primary function
Layer 1: sensor acquisition	ESP32-S3 ZERO (×2), Flex Sensor Spectra 4.5” (×10), MPU[a]-6050 IMU[b] (×2), and Custom PCB[c]	Dual-core 240 MHz; 12-bit ADC[d]; 100 Hz sampling; 22 features/frame (Float32)	Raw data acquisition from both hands simultaneously
Layer 2: wireless communication	ESP-NOW, BLE[e] 5.0, and data aggregator	ESP-NOW latency <5 ms; BLE MTU[f] 512 bytes; throughput 2 Mbps	Sensor data transmission to Android
Layer 3: processing and inference	Preprocessing pipeline, TFLite LSTM[g]+dense, and confidence threshold	Tensor (1,30,22); 150 classes; target accuracy ≥85%; latency <50 ms	On-device BISINDO gesture classification
Layer 4: AI language wrapper	Sentence buffer, LLM[h] REST[i] API, and postprocessing UI[j]	Pause-based segmentation; API response max 500 ms; RecyclerView history	LLM-based NLG[k] for natural Bahasa Indonesia sentence generation
Layer 5: Speech output	Google TTS[l] engine, Android speaker	Locale id-ID; speed 0.9; real-time playback	Conversion of sentences to Bahasa Indonesia speech

[a]MPU: motion processing unit.
[b]IMU: inertial measurement unit.
[c]PCB: printed circuit board.
[d]ADC: analog-to-digital converter.
[e]BLE: Bluetooth Low Energy.
[f]MTU: maximum transmission unit.
[g]LSTM: long short-term memory.
[h]LLM: large language model.
[i]REST: Representational State Transfer.
[j]UI: user interface.
[k]NLG: natural language generation.
[l]TTS: text-to-speech.

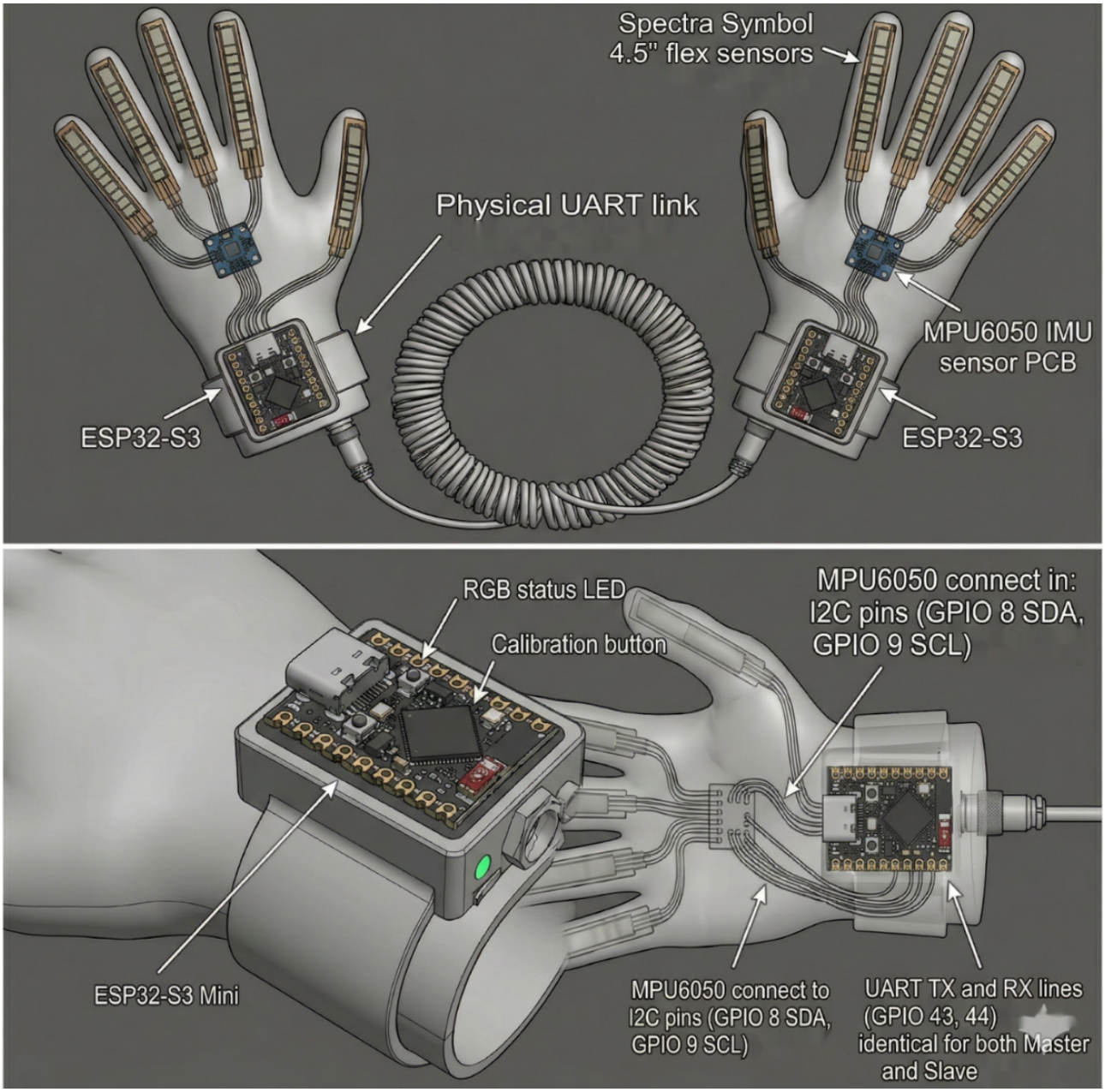
Layer 1 is the physical layer. Five Spectra Symbol 4.5 flex sensors (Spectra Symbol) in each glove measure finger bending, while a single MPU-6050 6-axis IMU (InvenSense/TDK InvenSense) tracks hand orientation, acceleration, and rotation.

XSL•FO
RenderX

Combining 2 sensor types yields 22 features per frame. This provides a richer representation than camera-based systems and performance comparable to the validated multi-IMU glove [34]. Dual-core 240-MHz ESP32-S3 (Espressif Systems) ZERO microcontrollers control both gloves at a sampling rate of 100 Hz.

Wireless protocols perform 2 distinct functions at layer 2. The ESP-NOW communicates between the gloves via an inter-ping in less than 5 milliseconds, using peer-to-peer communication, and transmits a Float32 array via BLE 5.0 to the Android application (Google LLC) using a data aggregator, where timestamping is performed as shown in Figure 3.

Figure 3. Glove design and specifications. GPIO: general-purpose input/output; I2C: interintegrated circuit; IMU: inertial measurement unit; PCB: printed circuit board; RGB: red-green-blue; RX: receive; SCL: serial clock line; SDA: serial data line; TX: transmit; UART: universal asynchronous receiver-transmitter.



Layer 3 is performed with an input tensor (1, 30, and 22), where the data is passed through a Kalman filter to min-max normalization in a sliding window (30 frames with a stride of 5). The TFLite LSTM+Dense model classifies movements into 150 gestures; only those with 85% or higher accuracy are declared valid. At the same time, on-device inference does not hinder time performance and eliminates cloud reliance, adhering to the portable AT design principles [35].

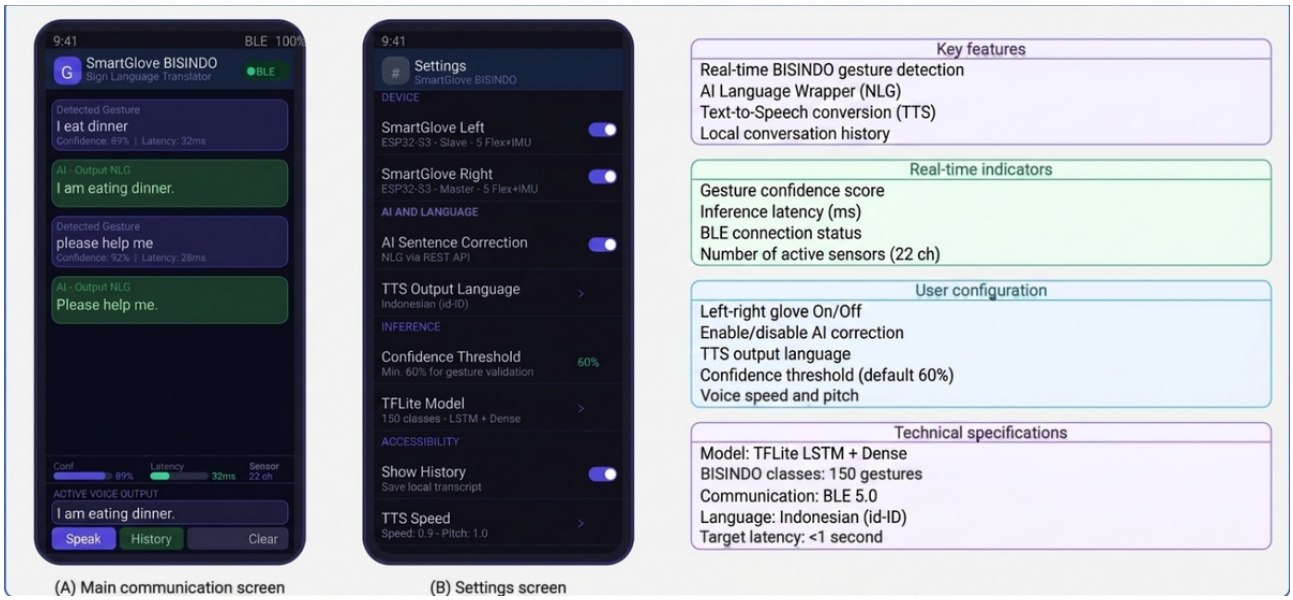
Layer 4 is the key differentiating component of the BISINDO SmartGlove, integrating LLM-based NLG, which is not present in any previous BISINDO system [36]. Valid gesture labels are collected by the sentence buffer through temporal lag detection. The raw word sequence is sent to the language model via the Representational State Transfer (REST) API (maximum 500 ms) for grammatical and contextual correction, generating sentences like (1) "*Saya sedang makan malam*" from the inputs

XSL•FO
RenderX

"*saya*" + "*makan*" + "*malam*"; and (2) "I am having dinner" from the inputs "I" + "eat" + "night."

The LLM-AAC framework is directly applied to sign language translation. Android's RecyclerView displays the final sentence in real time and draws data from the local conversation history. At layer 5, the final sentence is converted into Indonesian. The Google TTS Engine is used via locale id-ID with a rate of 0.9. Voice is replayed in real time, and end-to-end latency remains below 1 second throughout the process as shown in Figure 4.

Figure 4. SmartGlove Bahasa Isyarat Indonesia (BISINDO) app interface. BLE: Bluetooth Low Energy; CH: channel; NLG: natural language; LSTM: long short-term memory.



The app interface (Figure 4) is based on 2 primary screens. The communication screen displays the translation step. The raw gestures that are detected, and below them, the sentence corrected by the AI. Quick access to confidence scores, system latency, and BLE connection status indicators, too! Second, there's a settings screen. You may set up device connections that connect to the inference threshold and voice output update. Accessibility and ease of use (without any specific training) are the whole design around it. This is exactly what the needs analysis showed.

BISINDO Dataset

The BISINDO gesture dataset comprised 150 classes across 6 categories: 26 letters, 42 numbers, 21 greetings and social expressions, 18 family terms, 2 core essential words, and 41 additional vocabulary items, with 80% (120/150) classes completeness at the time of the field demonstration (full breakdown in Multimedia Appendix 3). This directly addresses the cultural and linguistic relevance criteria identified in the needs analysis and reflects the participatory curation process described in the Methods section.

Expert Validation

Performance and validity of the prototype were checked using an expert validation sheet. It has 3 areas of interest: greeting the data to calculate how well BISINDO gestures were displayed, accuracy and naturalness should be calculated by NLG-generated sentences, and an app should have a user-friendly interface. Twelve validators completed the sheet. Responses were analyzed using CVR, Aiken V, and triangulated percentage agreement. The summary findings of the study are shown in Table 3.

XSL•FO
RenderX

Table . Expert validation results: SmartGlove Bahasa Isyarat Indonesia (BISINDO) prototype (N=12).

Code	Assessment item	Values, n (%)	Values, mean (SD)	CVR[a]	Aiken V	Category
K1	BISINDO gesture vocabulary is authentic and meets standards	12 (100)	4.92 (0.29)	1.00	0.98	Highly valid
K2	Coverage of 150 classes represents daily communication needs	11 (92)	4.75 (0.45)	0.83	0.94	Highly valid
K3	The system can differentiate gestures with similar finger configurations	10 (83)	4.67 (0.65)	0.67	0.83	Valid
K4	A flex and IMU[b] sensor combination is appropriate for capturing BISINDO gesture characteristics	11 (92)	4.83 (0.39)	0.83	0.96	Highly valid
N1	AI-generated sentences are grammatical and conform to Bahasa Indonesia rules	11 (92)	4.83 (0.39)	0.83	0.96	Highly valid
N2	Sentence output sounds natural and communicative to lay conversation partners	10 (83)	4.75 (0.62)	0.67	0.94	Highly valid
N3	AI can correct gesture translation word order into complete sentences	11 (92)	4.83 (0.39)	0.83	0.96	Highly valid
N4	Sentence output is contextually relevant to user gesture intent	9 (75)	4.58 (0.79)	0.50	0.90	Valid
N5	AI response time does not disrupt communication fluency	10 (83)	4.67 (0.65)	0.67	0.92	Highly valid
U1	The interface is easy to understand and operate without specialized training	11 (92)	4.75 (0.45)	0.83	0.94	Highly valid
U2	BLE[c] connection status and confidence score indicators are clearly presented	10 (83)	4.67 (0.65)	0.67	0.92	Highly valid
U3	Conversation history display helps users monitor communication flow	11 (92)	4.83 (0.39)	0.83	0.96	Highly valid
U4	Settings features (threshold, TTS[d] speed, language) can be configured easily	9 (75)	4.58 (0.79)	0.50	0.90	Valid

XSL•FO
RenderX

Code	Assessment item	Values, n (%)	Values, mean (SD)	CVR[a]	Aiken V	Category
U5	Interface design is inclusive and meets the needs of users with communication disorders	12 (100)	4.92 (0.29)	1.00	0.98	Highly valid
_[e]	Overall mean across all items, n/N (%)	148/168 (88%)	4.76 (0.51)	0.76	0.94	Highly valid

[a]CVR: content validity ratio.

[b]IMU: inertial measurement unit.

[c]BLE: Bluetooth Low Energy.

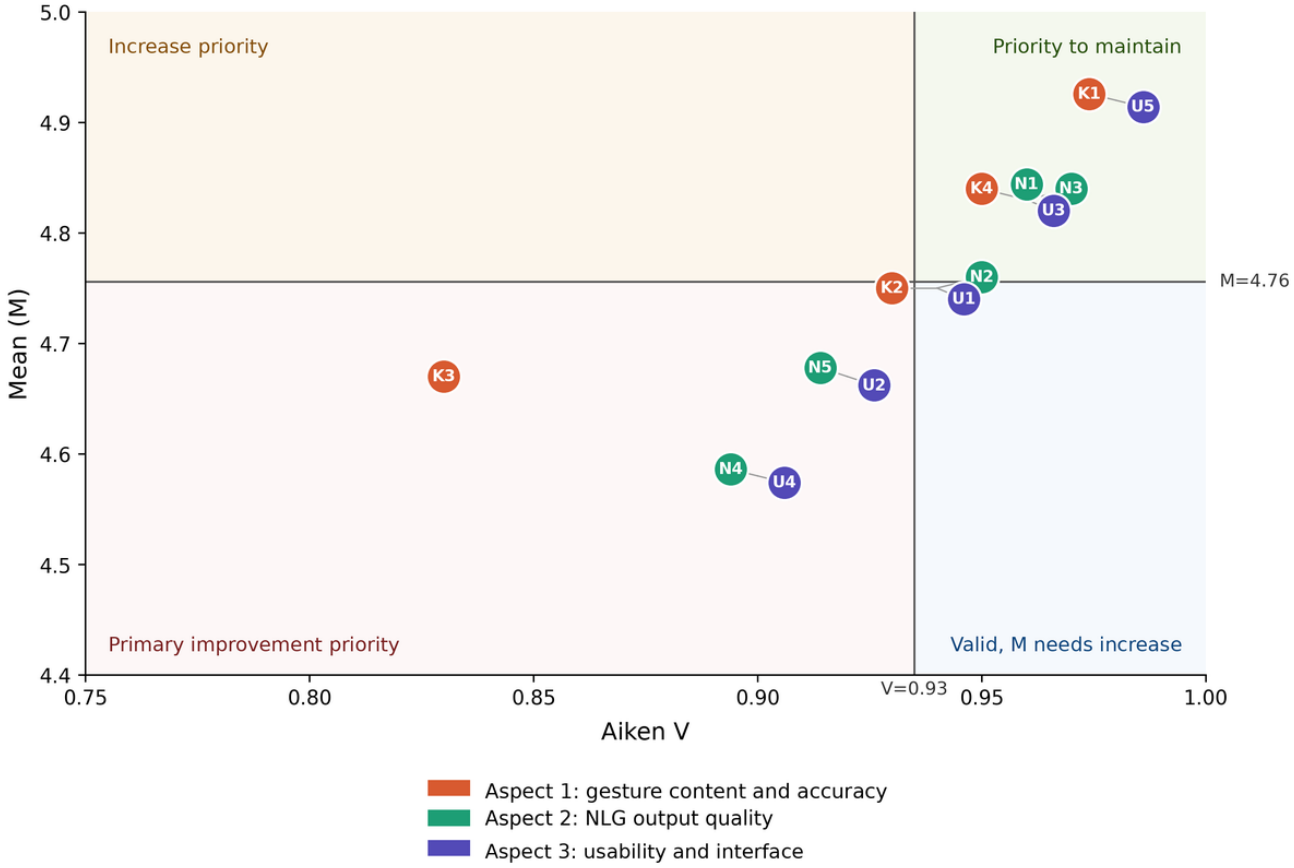
[d]TTS: text-to-speech.

[e]Not applicable.

It performed relatively well on content validity. The average Aiken V was 0.94 (SD 0.04), the mean Likert scale was 4.76 (SD 0.51), and the agreement rate ranged from 88% (148/168) to 100% (12/12) among all technical items tested. No item fell below the threshold of 0.75 for Aiken V, and in fact, 11 out of 14 (79%) items were classified as highly valid (V≥0.90). With an average CVR of 0.76 (SD 0.16) across 12 participants, you can see that this value was well above the Lawshe cutoff number of 0.56. For the separated aspects, aspect 1 received an Aiken V of 0.93, and aspects 2 and 3 had a value of 0.94 each.

However, K3 (overlapping movement), N4 (app context of AI in question), and U4 (low-configurable by hand) were only partly CVRs. The next iteration of the DSR design cycle should have improvements on all three. Figure 5: The Cartesian diagram will clarify this even more. The X-axis is Aiken V; the Y-axis is the mean Likert score.

Figure 5. Cartesian diagram of expert validation results. NLG: natural language generation.



The Cartesian diagram of Figure 5 shows the 2 indices plotted simultaneously against the 14 validation items. The line is drawn at this average Aiken V of 0.94 (SD 0.04) with the average of 4.76 (SD 0.11). In the first quadrant in the top right (priority to maintain) are K1, K4, N1, N3, U3, and U5. The base strengths of the prototype are indicated by a V value greater than 0.94 and a mean value greater than 4.76 (SD 0.11). K3, N4, and U4 are located in the lower left quadrant. The 3 are interrelated systematically, and these 3 items would constitute the targeted area of improvement in the next round.

XSL•FO
RenderX

Field Demonstration and Evaluation

Field testing was conducted using 3 complementary instruments across 2 Indonesian cities (Table 4).

Table . Field evaluation results (Makassar and Malang)[a].

Respondent	Makassar, mean (SD); n/N (%)	Malang, mean (SD); n/N (%)	Combined, mean (SD); n/N (%)
User participants	3.34 (0.16); 8/15 (53)	3.36 (0.17); 7/15 (47)	3.35 (0.16); 15/15 (100)
Observers/companions	4.35 (0.07); 2/4 (50)	4.45 (0.07); 2/4 (50)	4.40 (0.08); 4/4 (100)

[a]Mean scores are on a 5-point Likert scale. SD is the sample SD (n−1) of respondent-level mean scores. Mean score as a percentage of the maximum score (5): user participants, 67% (Makassar), 67% (Malang), 67% (combined); observers/companions, 87% (Makassar), 89% (Malang), 88% (combined).

In both Makassar and Malang, ratings from observers and companions were consistently higher than ratings from user participants. The reported practicality percentages represent mean Likert-scale ratings converted to percentages of the maximum possible score, rather than proportions of participants. For overall practicality, user participants with communication disorders gave an average score of 3.35 (SD 0.16) out of 5, or 67% (n=15). Conversely, observers and companions gave an average score of 4.40 (SD 0.08), or 88% (n=4). The gap in scores was as high as about 24 percentage points in Makassar and 18 percentage points in Malang. In Surabaya, an independent disability specialist rated the utility of the prototype at 81% (with a mean value calculated across the 20 questionnaire items of 4.05, SD 0.60; n=1), which backs up the higher scores from observers and companions. The Makassar (only user) responses revealed the lowest-scoring aspects of voice output suitability and clarity. Both sites repeatedly mentioned the gloves' sensitivity, that they are easy to use but not intuitive, the regional variation in BISINDO terminology, and that user guides would need to be produced.

Our comparison involves linguistic analysis of the unprocessed raw input created by gestures (the system input) and the sentence generated by AI (the system output) for each interactive simulation. Three transformation processes were identified. Addition was used to combine elements for a fuller answer, that is, "I" + "eat" became "I am eating." The process of elision removed redundant elements, for example, "I" + "already" + "drink" became "I drink water." Finally, "air" was introduced and "already" omitted, which shows how an output can go through multiple operations at the same time. Substitution occurred when the elements were replaced, causing the meaning to change, for example, "mother" + "running" became mother walking, not mother running. Among the 3 processes, substitution was the worst, because it not only introduced grammar changes but also changed what the user intended to say. Thus, this is consistent with the relatively lower user ratings of the speech output with respect to appropriateness and clarity. Any replacement effectively changes the meaning of the original sentence, causing it to become no longer what the speaker intended. Regional differences in BISINDO vocabulary were also found across sites, indicating that additional training data of the same target words from different regions is necessary.

Discussion

Principal Results

Four principal findings emerged from this study. First, cross-disciplinary stakeholder consensus, assessed quantitatively through the needs analysis, confirmed a very high and consistent priority for BISINDO-based assistive communication technology capable of generating natural, real-time speech (grand mean 4.77, SD 0.51; mean agreement 89%). Second, the SmartGlove BISINDO prototype directly addresses this need through a 5-layer architecture that, to our knowledge, is the first BISINDO system to integrate LLM-based NLG into the translation pipeline. Third, expert validation indicated strong content validity (grand mean Aiken V 0.94, SD 0.04; CVR 0.76, SD 0.16; 11 of 14 items highly valid), providing preliminary evidence that the prototype design is ready to proceed to the next DSR cycle. Fourth, field demonstration with actual users with communication disorders showed favorable practicality ratings from observers, companions, and a disability expert (81% - 88%), but lower ratings from user participants themselves (67%), revealing a gap between expert-perceived feasibility and lived user experience that motivates further design refinement.

From a linguistic perspective, items assessing the grammaticality and naturalness of AI-generated sentences (N1, V=0.96) received highly valid ratings, suggesting that the AI Language Wrapper output meets the ideational and textual metafunctions defined by SFL. Perfect agreement on interface inclusivity (U5, V=0.98) similarly supports the AAC principle of independent operability. Field-based linguistic analysis supports the expert's validity ratings. This analysis identified substitution, rather than addition or deletion, as the transformation process most likely to disrupt ideational metafunctions. This is because substitution alters the intended meaning (field demonstration and evaluation, and results).

In this study, the design of the BISINDO dataset was conducted in collaboration with certified BISINDO experts. This meant prioritizing vocabulary that people actually use in conversation while eliminating classes that are rarely encountered or difficult to visually distinguish. This stakeholder-driven refinement aligns with user-centered design principles, which are crucial for achieving adoption of rehabilitation technology in real-world practice settings [37]. Co-design with experts is a key implementation driver that requires ongoing effort.

These findings carry implications at 3 levels. At the practice level, SmartGlove BISINDO has the potential to reduce communication barriers for individuals with speech-impairment in Indonesian inclusive education classrooms, public service interactions, and daily social contexts, settings in which no deployable BISINDO-to-speech solution currently exists. Given that an estimated 1.02% of Indonesia's population experiences severe communication difficulties [4], even modest adoption

XSL•FO
RenderX

of this prototype could yield meaningful benefits for public health and educational equity, supporting SDG 4 and SDG 10. At the methodological level, the DSR-grounded, co-design-intensive development process demonstrated here offers a preliminary, potentially transferable model for AT research in other underrepresented sign language communities, particularly those facing similar constraints in dataset availability and NLG integration; this transferability remains to be empirically tested rather than assumed. At the policy level, the strong stakeholder consensus on authentic BISINDO vocabulary and independent operability provides preliminary evidence to inform advocacy for context-sensitive, locally grounded assistive communication solutions, consistent with calls to address structural barriers to technology adoption in low- and middle-income countries [38].

Comparison With Prior Work

Table 5 situates SmartGlove BISINDO against previous BISINDO recognition systems, and the most recently published related work is shown in Table 5.

Table . Comparison of SmartGlove Bahasa Isyarat Indonesia (BISINDO) with prior BISINDO recognition systems.

Study	Input modality	Recognition level	NLG[a]/sentence generation	Reported accuracy, %
Aljabar and Suharjito [10], 2020	Camera (CNN-LSTM)[b]	Isolated gesture	No	96
Candra et al [11], 2024	Camera (CNN-MobileNetV2, mobile)	Isolated gesture (30 classes)	No	55
Wiguna and Rojali [12], 2025	Camera (3D CNN-LSTM/BiLSTM[c]+CTC[d])	Sentence-level (gesture-to-text only)	No	Not directly comparable
Kindy et al [14], 2024	Camera (Transformer: SPOTER[e], Siformer, MViTv2[f])	Isolated sign, word-level	No	97
Fadzli and Rahardi [15], 2026	Camera (MobileNetV3, EfficientNetV2-B0)	Isolated hand sign	No	90 - 96
Maharjan et al [39], 2025	Wearable glove (hybrid CNN-Transformer)	Isolated gesture	No	98
SmartGlove BISINDO (this study)	Wearable dual-glove (flex+IMU[g] sensors)	Isolated gesture, LLM[h]-corrected to full sentence	Yes (LLM-based NLG)	Target ≥85%; empirical validation pending

[a]NLG: natural language generation.

[b]CNN-LSTM: convolutional neural network–long short-term memory.

[c]CNN-BiLSTM: convolutional neural network–bidirectional long short-term memory.

[d]CTC: connectionist temporal classification.

[e]SPOTER: Sign Pose-based Transformer.

[f]MViTv2: Multiscale Vision Transformer, version 2.

[g]IMU: inertial measurement unit.

[h]LLM: large language model.

A closely related system for Indonesian Sign Language, though developed for SIBI rather than BISINDO, integrated an LSTM gesture classifier with an IndoBERT language model and reported 89% gesture accuracy alongside a 3-fold improvement in word-prediction accuracy when the language model was applied [13]. This confirms that combining gesture recognition with language modeling improves output quality even in closely related Indonesian sign systems, though this approach remained limited to word-level correction rather than full sentence generation. Among wearable glove-based systems, a category of approaches reviewed extensively in the broader wearable sensor-based sign language recognition literature [40], Maharjan et al [39] reported 98% gesture classification accuracy using a hybrid CNN-Transformer model, though this system, like other prior work, does not incorporate NLG for sentence-level output.

Of all systems compared, only SmartGlove BISINDO stands apart. It's the only one that combines wearable sensors with LLM-based NLG to produce full sentences based not on words/gesture-to-text, but on other gestures. Remember, this comparison used a narrative synthesis of previous results rather than a direct head-to-head trial. This is not a fair apples-to-apples comparison since the systems are different in data sets, hardware, and evaluation protocols. This limitation is expanded upon in the Limitations section below. Additionally, it was not used to measure the gesture-recognition accuracy of SmartGlove BISINDO itself in this study. Instead, evaluation looked at content validity through an expert panel and field practicality with users, observers, and a disability expert. Technical classification performance wasn't the focus, and that's discussed under Future Directions.

Limitations

Consistent with common methodological challenges in evaluating rehabilitation technologies [30], field evaluation was limited to practicality ratings from a small sample (15 user participants, 4 observers/companions, 1 disability expert) across 2 to 3 cities, rather than a controlled effectiveness trial with a

XSL•FO
RenderX

comparison condition or pre-post outcome measures; results should therefore be interpreted as an early-stage field demonstration (DSR phases 4‐5) rather than a definitive effectiveness evaluation. Similarly, the 12-participant expert panel, while meeting the minimum threshold for CVR and Aiken V analyses [31,32], limits the generalizability of the needs-analysis and validation findings, and a formal interrater reliability coefficient was not computed given the sample size. Because both instruments were developed specifically for this study, empirical construct validation beyond expert content review, such as confirmatory factor analysis with a larger sample, was not conducted.

Several sources of uncertainty in real-time gesture recognition were not directly quantified, including sensor noise, classification ambiguity near the 85% or higher confidence threshold, and intersigner variability in gesture speed and amplitude; because the AI Language Wrapper generates sentences from already-classified gestures, such uncertainty can propagate into the output, for example, a misclassified gesture with falsely high confidence could yield a sentence that is grammatically fluent but semantically incorrect. Quantifying this uncertainty, for example through confidence-aware generation or user-facing indicators, is an important direction. Additional hardware limitations include the need for more ergonomic glove materials and sensor straps to support longer sessions, and the fact that the target TFLite model accuracy of ≥85% across diverse user cohorts and all 150 gesture classes has not yet been empirically validated.

Future Directions

Building on the preliminary evidence for content validity established here, 4 directions define the research agenda going forward. First, the gap between favorable observer/expert ratings and lower user-participant ratings, particularly for speech-output clarity, warrants deeper investigation; we recommend incorporating the technology acceptance model [41] alongside AAC and SFL as an interpretive framework, since the technology acceptance model's core constructs of perceived usefulness and ease of use are well suited to explaining why expert-perceived feasibility diverges from lived user experience. Second, empirical performance testing is needed to verify the TFLite gesture-classification model's accuracy across diverse user cohorts and hand morphologies and to quantify the classification and language-generation uncertainties identified in the Limitations section under real-world conditions. Third, the priority-improvement areas identified through expert validation point to specific refinements for the next DSR design cycle: similar-gesture differentiation (K3) toward sensor augmentation or finer feature engineering, contextual relevance of AI-generated output (N4) toward more sophisticated, BISINDO-specific prompt engineering, and settings configuration ease (U4) toward interface simplification. Fourth, scalability testing to other underrepresented sign language systems in Southeast Asia is planned, building on the language-agnostic design of Layers 1 to 3; given that over 50% of individuals in some low- and middle-income countries lack access to needed rehabilitation services [42], portable, low-infrastructure solutions such as SmartGlove BISINDO carry urgency beyond Indonesia, though this broader applicability remains to be empirically demonstrated rather than assumed.

Conclusions

This study developed and validated SmartGlove BISINDO, which is, to our knowledge, the first BISINDO sign language–to-speech system to integrate LLM-based NLG, addressing a systemic gap in prior BISINDO recognition research. A needs analysis of 12 cross-disciplinary participants confirmed a very high and consistent priority (grand mean 4.77, SD 0.51, mean agreement 89%) across all 4 need clusters: response speed and accuracy, naturalness of language output, portability and ease of use, and cultural and linguistic relevance.

The resulting prototype comprises a 5-layer architecture spanning dual-sensor glove hardware, wireless communication, on-device gesture classification, LLM-based sentence generation, and speech output. Expert validation by the same 12 participants indicated strong content validity (grand mean Aiken V 0.94 (SD 0.04), CVR 0.76 (SD 0.16), both exceeding their respective thresholds), with 11 of 14 items rated highly valid. Three items, concerning similar-gesture differentiation, contextual relevance of generated output, and settings configuration ease, were identified as priorities for refinement in the next DSR design cycle iteration. Field demonstration with actual users, observers, and a disability expert further indicated favorable practicality overall, alongside a gap between expert-perceived feasibility and user-rated experience that will guide the next design iteration.

Acknowledgments

During the preparation of this manuscript, the authors used the generative AI tools Gemini (Google) to generate initial drafts of figures and Claude (Anthropic) to edit figures and to assist in drafting responses to editorial queries during revision. The authors reviewed and verified all AI-assisted output and take full responsibility for the accuracy, originality, and integrity of the content, including all references and citations.

Funding

This research is funded by the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology, and managed under the EQUITY Program (contract number 4312/B3/DT.03.08/2025).

XSL•FO
RenderX

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: SS, BY, BA
Formal analysis: MRAL, BP, DWR, DAD, TT
Investigation: MRAL, DAD
Methodology: SS, BY, BA, TT
Supervision: BA, KA
Validation: BP, DWR
Writing – original draft: SS, AK, MRAL, TT
Writing – review & editing: SS, BY, BA, KA, AK, MRAL, BP, DWR, DAD, TT
SS had full access to all data and takes responsibility for the integrity of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Clusters and items of the needs analysis questionnaire.
[DOCX File, 23 KB - rehab_v13i1e96258_app1.docx]

Multimedia Appendix 2
Aspects and items of the expert validation sheet.
[DOCX File, 24 KB - rehab_v13i1e96258_app2.docx]

Multimedia Appendix 3
SmartGlove BISINDO dataset composition by category.
[DOCX File, 24 KB - rehab_v13i1e96258_app3.docx]

References

1. Global report on health equity for persons with disabilities. : World Health Organization; 2022 Dec 2 URL: https://www.who.int/publications/i/item/9789240063600 [accessed 2026-09-29]
2. Beukelman DR, Light JC. Augmentative and Alternative Communication: Supporting Children and Adults with Complex Communication Needs, 5th edition: Paul H Brookes Publishing Co; 2020.
3. Light J, Mcnaughton D. Designing AAC research and intervention to improve outcomes for individuals with complex communication needs. Augment Altern Commun 2015 Jun;31(2):85-96. [doi: 10.3109/07434618.2015.1036458] [Medline: 25904008]
4. Cummings L. Pragmatic Disorders: Springer; 2014. [doi: 10.1007/978-94-007-7954-9]
5. Alabshar N, Pujiwati LA, Munawaroh T, Fatoni Z. Disabilitas dan kemiskinan ekstrem di Indonesia: analisis data Survei Sosial Ekonomi Nasional tahun 2020 [Article in Indonesian]. J Kawistara 2024;14(1):86-102. [doi: 10.22146/kawistara.83519]
6. Handhika T, Zen RIM, Lestari DP, Sari I. Gesture recognition for Indonesian sign language (BISINDO). J Phys Conf Ser 2018 Jun 1;1028:012173. [doi: 10.1088/1742-6596/1028/1/012173]
7. McAllister L, McCormack J, McLeod S, Harrison LJ. Expectations and experiences of accessing and participating in services for childhood speech impairment. Int J Speech Lang Pathol 2011 Jun;13(3):251-267. [doi: 10.3109/17549507.2011.535565] [Medline: 21563899]
8. Munce SEP. Introducing JMIR Rehabilitation and Assistive Technologies: a venue for publishing interdisciplinary research on the development, implementation, and evaluation of health innovations and emerging technologies in the field of rehabilitation. JMIR Rehabil Assist Technol 2024 Apr 22;11(1):e56348. [doi: 10.2196/56348] [Medline: 38648632]
9. Vaezipour A, Aldridge D, Koenig S, et al. Rehabilitation supported by immersive virtual reality for adults with communication disorders: semistructured interviews and usability survey study. JMIR Rehabil Assist Technol 2023 Oct 31;10(1):e46959. [doi: 10.2196/46959] [Medline: 37906228]
10. Aljabar A, Suharjito S. BISINDO (Bahasa Isyarat Indonesia) sign language recognition using CNN and LSTM. Adv Sci Technol Eng Syst J 2020;5(5):282-287. [doi: 10.25046/aj050535]
11. Candra A, Intan TK, Rosmalinda, et al. Development of machine learning-based sign language translator for Bahasa Isyarat Indonesia (BISINDO). AIP Conf Proc 2024;2987(1). [doi: 10.1063/5.0199747]

XSL•FO
RenderX

12. Wiguna KL, Rojali. Sentence-level Indonesian Sign Language (BISINDO) recognition using 3D CNN-LSTM and 3D CNN-BiLSTM models. Int J Adv Comput Sci Appl 2025;16(9):433. [doi: 10.14569/IJACSA.2025.0160941]
13. Maritza AB, Ardiansyah R. Real-time Indonesian sign language recognition using long short-term memory and IndoBERT contextual modeling. MALCOM: Indones J Mach Learn Comput Sci 2026;6(3):1425-1436. [doi: 10.22146/jnteti.v15i3.27789]
14. Kindy GO, Leonali G, Lucky H. Word-level BISINDO: a novel video Indonesian Sign Language dataset and baseline methods. Procedia Comput Sci 2025;269:249-258. [doi: 10.1016/j.procs.2025.08.277]
15. Fadzli AK, Rahardi M. Comparative analysis of MobileNetV3 and EfficientNetV2B0 in BISINDO hand sign recognition using MediaPipe landmarks. J Appl Inform Comput 2026;10(1):737-746. [doi: 10.30871/jaic.v10i1.11878]
16. Cihan Camgöz N, Koller O, Hadfield S, Bowden R. Sign language transformers: joint end-to-end sign language recognition and translation. Proc IEEE/CVF Conf Comput Vis Pattern Recognit 2020:10023-10033. [doi: 10.1109/CVPR42600.2020.01004]
17. Koller O, Camgoz NC, Ney H, Bowden R. Weakly supervised learning with multi-stream CNN-LSTM-HMMs to discover sequential parallelism in sign language videos. IEEE Trans Pattern Anal Mach Intell 2020 Sep;42(9):2306-2320. [doi: 10.1109/TPAMI.2019.2911077] [Medline: 30990421]
18. Adaloglou N, Chatzis T, Papastratis I, et al. A comprehensive study on deep learning-based methods for sign language recognition. IEEE Trans Multimedia 2022;24:1750-1762. [doi: 10.1109/TMM.2021.3070438]
19. Sincan OM, Camgoz NC, Bowden R. Using an LLM to turn sign spottings into spoken language sentences. arXiv. Preprint posted online on Mar 15, 2024 URL: https://arxiv.org/html/2403.10434v1 [accessed 2026-09-29]
20. Bulla L, Tuccio G, Mongiovì M, Gangemi A. Leveraging large language models for accurate sign language translation in low-resource scenarios. In: Volume 413: ECAI 2025: IOS Press; 2025:4217-4224. [doi: 10.3233/FAIA251315]
21. González-Rodríguez JR, Córdova-Esparza DM, Terven J, Romero-González JA. Sign language recognition and translation with large multimodal models: a systematic review. In: Martínez-Villaseñor L, Vázquez RA, Ochoa-Ruiz G, editors. Advances in Soft Computing: MICAI 2025: Springer; 2025:272-290. [doi: 10.1007/978-3-032-09044-7_20]
22. Cai S, Venugopalan S, Seaver K, et al. Using large language models to accelerate communication for eye gaze typing users with ALS. Nat Commun 2024 Nov 1;15(1):9449. [doi: 10.1038/s41467-024-53873-3] [Medline: 39487163]
23. Blavette L, Dacunha S, Alameda-Pineda X, et al. Acceptability and usability of a socially assistive robot integrated with a large language model for enhanced human-robot interaction in a geriatric care institution: mixed methods evaluation. JMIR Hum Factors 2025 Aug 1;12:e76496. [doi: 10.2196/76496] [Medline: 40750072]
24. Halliday MAK. An Introduction to Functional Grammar, 1st edition: Edward Arnold; 1985.
25. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: Guyon I, von Luxburg U, Bengio S, Wallach HM, Fergus R, Vishwanathan SVN, et al, editors. Advances in Neural Information Processing Systems 30 (NeurIPS 2017) 2017:5998-6008 URL: https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf [accessed 2026-09-29]
26. Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Q 2004 Mar 1;28(1):75-106. [doi: 10.2307/25148625]
27. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inf Syst 2007 Dec;24(3):45-77. [doi: 10.2753/MIS0742-1222240302]
28. De Sordi JO, Azevedo M, Meireles M, Pinochet L, Jorge C. Design science research in practice: what can we learn from a longitudinal analysis of the development of published artifacts? Inform Sci 2020 Aug;23:1-23. [doi: 10.28945/4493]
29. Hevner AR. A three-cycle view of design science research. Scand J Inf Syst 2007;19(2):87-92 [FREE Full text]
30. Venable J, Pries-Heje J, Baskerville R. FEDS: a framework for evaluation in design science research. Eur J Inf Syst 2016 Jan;25(1):77-89. [doi: 10.1057/ejis.2014.36]
31. Lawshe CH. A quantitative approach to content validity. Pers Psychol 1975;28(4):563-575. [doi: 10.1111/j.1744-6570.1975.tb01393.x]
32. Aiken LR. Three coefficients for analyzing the reliability and validity of ratings. Educ Psychol Meas 1985 Mar;45(1):131-142. [doi: 10.1177/0013164485451012]
33. Alhasani R, George N, Radman D, Auger C, Ahmed S. Methodologies for evaluating the usability of rehabilitation technologies aimed at supporting shared decision-making: scoping review. JMIR Rehabil Assist Technol 2023 Aug 15;10:e41359. [doi: 10.2196/41359] [Medline: 37581911]
34. Ji A, Wang Y, Miao X, et al. Dataglove for sign language recognition of people with hearing and speech impairment via wearable inertial sensors. Sensors (Basel) 2023 Jul 26;23(15):6693. [doi: 10.3390/s23156693] [Medline: 37571476]
35. Lancioni GE, Singh NN, O'Reilly MF, et al. Low-cost technology-aided programs for supporting people with motor, visual, and intellectual disabilities in functional forms of occupation and communication: proof-of-concept study. JMIR Rehabil Assist Technol 2023 Mar 24;10:e44239. [doi: 10.2196/44239] [Medline: 36961483]
36. Reiter E, Dale R. Building Natural Language Generation Systems: Cambridge University Press; 2000.
37. Alder G, Taylor D, Rashid U, et al. A brain computer interface neuromodulatory device for stroke rehabilitation: iterative user-centered design approach. JMIR Rehabil Assist Technol 2023 Dec 11;10:e49702. [doi: 10.2196/49702] [Medline: 38079202]

XSL•FO
RenderX

38. Marshall J, Wylie K, McLeod S, et al. Communication disability in low and middle-income countries: a call to action. BMJ Glob Health 2024 Jul 10;9(7):e015289. [doi: 10.1136/bmjgh-2024-015289] [Medline: 38991579]
39. Maharjan S, Fernando S, Shrestha S, Mariyanayagam D, Virdee B. Sign language detection and translation using smart glove 2.0 using hybrid CNN-transformer model. In: Swaroop A, Virdee B, Correia SD, Polkowski Z, editors. Proceedings of Data Analytics and Management ICDAM 2025, Volume 3: Springer; 2025, Vol. 1597:38-51. [doi: 10.1007/978-3-032-02831-0_4]
40. Kudrinko K, Flavin E, Zhu X, Li Q. Wearable sensor-based sign language recognition: a comprehensive review. IEEE Rev Biomed Eng 2021;14:82-97. [doi: 10.1109/RBME.2020.3019769] [Medline: 32845843]
41. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
42. Rehabilitation. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/rehabilitation [accessed 2026-09-29]

Abbreviations

AAC: augmentative and alternative communication
AT: assistive technology
BERT: Bidirectional Encoder Representations from Transformers
BISINDO: Bahasa Isyarat Indonesia
BLE: Bluetooth Low Energy
BLEU: Bilingual Evaluation Understudy
CNN: convolutional neural network
CVR: content validity ratio
DSR: design science research
IMU: inertial measurement unit
LLM: large language model
LSTM: long short-term memory
MViTv2: Multiscale Vision Transformer version 2
NLG: natural language generation
REST: Representational State Transfer
SDG: Sustainable Development Goal
SFL: systemic functional linguistics
SIBI: Indonesian Sign System
SPOTER: Sign Pose-based Transformer
TTS: text-to-speech

Edited by S Munce; submitted 27.Mar.2026; peer-reviewed by B Pereira, O Tutsoy; revised version received 21.Aug.2026; accepted 22.Sep.2026; published 22.Oct.2026.

Please cite as:
Suhartono S, Yulianto B, Arafah B, Andajani K, Kholiq A, Latief MRA, Prastio B, Rohaedi DW, Dewi DA, Taqdir T
The SmartGlove BISINDO Wearable Sensor and Large Language Model–Based Sign Language–to-Speech System for Users With Communication Disorders: Design Science Research Study
JMIR Rehabil Assist Technol 2026;13:e96258
URL: https://rehab.jmir.org/2026/1/e96258
doi:10.2196/96258

© Suhartono Suhartono, Bambang Yulianto, Burhanuddin Arafah, Kusubakti Andajani, Abdul Kholiq, Muhammad Ridha Anugrah Latief, Bambang Prastio, Diding Wahyudin Rohaedi, Deshinta Arrova Dewi, Taqdir Taqdir. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 22.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Toward a Digitally Informed Knitted Prosthetic Interface With Graded Stiffness to Enhance Comfort in Transtibial Amputees: Proof-of-Concept Case Study

Trevor Binedell[1], BAppSc, MSc, PhD; Pei Zhi Chia[2], BSocSci; Ying Yi Tan[2], BA, MArch, PhD; Hong Yee Low[2], BSc, MSc, PhD

[1]Foot Care & Limb Design Centre, Tan Tock Seng Hospital, Singapore, Singapore

[2]Engineering Product Design Pillar, Singapore University of Technology and Design, 8 Somapah Road, Singapore, Singapore

Corresponding Author:
Hong Yee Low, BSc, MSc, PhD
Engineering Product Design Pillar, Singapore University of Technology and Design, 8 Somapah Road, Singapore, Singapore

Abstract

Background: Despite considerable advancements in prosthetic technology, a substantial proportion of lower limb amputees reduce or discontinue prosthesis use, with reported nonuse rates ranging from 12% to 53%. This reflects the multifactorial challenges associated with long-term prosthetic use, among which comfort and skin health are consistently identified as key determinants. More specifically, studies point toward nonbreathable silicone liners trapping heat and sweat, leading to skin and hygiene problems. These persistent limitations underscore the need for alternative interface materials that offer improved breathability, moisture management, and tunable mechanical properties.

Objective: This study aimed to introduce Flexoknit, a transtibial prosthetic liner that integrates user-specific digital skin strain analysis with computer numerical control multimaterial knitting to create a mechanically tuned, breathable, and anatomically customized interface. Using digital biomechanical data as the primary design driver—rather than clinician heuristics alone—Flexoknit aims to determine the feasibility and performance of a skin strain–guided, computer numerical control–knitted prosthetic interface in terms of material function, clinical performance, and user experience.

Methods: Flexoknit uses programmable multimaterial knitting, incorporating thermal-reactive yarns that stiffen when heated to create structural support zones, alongside spandex yarns that provide elastic compression and breathable zones. Uniaxial tensile tests showed that yarn and stitch combinations can generate distinct stiffness grades, with nearly order-of-magnitude differences. The spatial layout of these graded zones aligns high-stiffness regions with the lines of nonextension, and low-stiffness regions with areas of greater skin strain. With the new prosthetic interface, a series of controlled tests was conducted to compare performance against the participant's existing prosthesis with a conventional silicone liner. User testing was organized into 3 domains (ie, mobility, suspension, and comfort) using standardized quantitative assessments and structured qualitative data collection.

Results: User testing demonstrated a 22.5% improvement in total range of motion, a 37.5% reduction in interface mass, and improved thermal regulation in hot, humid environments compared to that of a conventional silicone liner. The user walked unaided and performed sit-to-stand movements, reporting positive comfort and usability feedback.

Conclusions: This work establishes Flexoknit as a promising direction for future prosthetic development—one that integrates principles of biomechanics, textile engineering, and digital fabrication to create user-centered interface solutions. The findings suggest that digitally engineered knitted interfaces can provide a highly customizable, breathable, and compliant alternative to conventional silicone liners, particularly for lower-activity amputees or individuals prioritizing comfort and ease of use.

(JMIR Rehabil Assist Technol 2026;13:e91396) doi:10.2196/91396

KEYWORDS

prosthetic; health care; comfort; lines of nonextension; graded stiffness; multimaterial; knitting; textile; thermal reactive

Introduction

Despite considerable advancements in prosthetic technology, a substantial proportion of lower limb amputees reduce or discontinue prosthesis use, with reported nonuse rates ranging from 12% to 53% [1-3]. This reflects the multifactorial challenges associated with long-term prosthetic use, among which comfort and skin health are consistently identified as key determinants. Prosthetic comfort is central to user satisfaction and long-term adherence and strongly influences mobility and quality of life [4-8]. Interface-related discomfort—particularly heat, perspiration, and skin complications—has been repeatedly identified as a major contributor to reduced wear time and

XSL•FO
RenderX

dissatisfaction among prosthesis users. Comfort is shaped by several factors, including socket design, interface materials, and suspension approaches [9], and each prosthetic socket is typically crafted through a manual, iterative process that aims to optimize fit [10]. For individuals with transtibial amputation, this precision is especially important because soft tissues must accommodate substantial loads and shear forces during everyday movement [11].

As the primary point of contact between the residual limb and the prosthetic socket, the interface plays a decisive role in user comfort. Dissatisfaction rates with this interface remain high: in a study of 78 lower limb amputees, only 43% reported being satisfied with prosthetic comfort [12]. Excessive sweating is one of the most frequently reported complaints among prosthesis users, affecting up to 70% of individuals with lower limb amputation and significantly impacting daily activities. Approximately half of users report discomfort related to heat and perspiration, highlighting the clinical significance of thermal burden within the socket-limb interface [13,14].

Silicone liners, while widely used to enhance suspension and load distribution, are inherently nonbreathable and act as thermal insulators, trapping sweat against the skin and exacerbating these issues. Among 83 amputees using silicone liners, 31% to 60% reported hygiene-related problems, including itching, perspiration, odors, and scarring [15]. Silicone's insulating properties contribute to dermatological issues, such as heat rash, folliculitis, soreness, and odor [16], and users wearing liners have been shown to experience increased residual limb temperatures during exercise, whereas those without liners did not [17]. These issues are further intensified in hot and humid climates, where thermal load and moisture accumulation exacerbate discomfort [18]. Traditional strategies such as wearing socks under or over liners may improve fit but further increase insulation, adding to the thermal burden or reducing friction, which leads to a decreased coupling grip. These findings suggest that limitations in existing liner materials, particularly their inability to effectively regulate heat and moisture, represent a key modifiable contributor to prosthetic discomfort and reduced use.

These persistent limitations underscore the need for alternative interface materials that offer improved breathability, moisture management, and tunable mechanical properties. Textile-based solutions, particularly knitted structures, present a promising approach due to their inherent porosity, efficient moisture transport, and capacity for precise mechanical tuning. The intermeshing loop architecture of knitted fabrics can produce wide-ranging stiffness, stretch, and ventilation characteristics. Recent advances in textile engineering demonstrate that multimaterial knits can achieve spatially graded mechanical behavior through stitch topology [19], thermal-reactive yarns that form rigid 3D structures [20], and hybrid yarn systems that regulate functional responses [21]. However, existing prosthetic liners primarily rely on homogeneous elastomeric materials and do not exploit these textile capabilities. Early attempts to incorporate textile and knitted interfaces into prosthetic design have demonstrated several functional advantages but remain limited in scope. For example, fabric-covered or differentiated-stretch liner concepts—such as knitted tubes layered with silicone—have aimed to reduce pistoning by restricting distal elongation while preserving proximal compliance or modifying longitudinal elongation over pressure-sensitive areas, thereby improving suspension using the textile architecture [22]. Similarly, breathable textile liner designs using spacer fabrics and partial silicone coatings have shown improved ventilation, moisture transport, and cushioning, although these were largely reported in upper-limb applications and were individualized through manual trimming rather than digitally generated geometry [23]. In parallel, knitted compression systems for amputated limbs have demonstrated that stitch density, loop structure, and yarn selection can be used to tune compression and mechanical response; however, these designs were primarily intended for compression therapy or limb shaping rather than as load-bearing prosthetic interfaces [24,25]. More recently, digital workflows linking 3D scanning to knitting have demonstrated the feasibility of producing geometrically customized liners, but these approaches remain focused on geometric shape conformity and automated manufacturing [26].

Taken together, these studies establish that textile-based interfaces can improve breathability, cushioning, compression control, and even geometric customization. However, they do not yet demonstrate a prosthetic interface in which the regional mechanical properties of the material are prescribed based on subject-specific biomechanical data. In particular, prior work has not integrated digital skin strain mapping with multimaterial textile design to spatially align stiffness, compliance, and reinforcement with the mechanical demands of the residual limb.

A critical missing link in the pursuit of personalized prosthetic interface design is the integration of digital skin strain analysis into material architecture. Skin strain reflects the deformation of superficial tissues during movement, offering direct insight into where support, compliance, or protection is most needed. The concept of lines of nonextension (LoNEs)—regions of the skin that undergo minimal stretch during motion—first described by Iberall [27], has become increasingly adopted in the design of wearable devices. Aligning stiff materials with LoNEs and compliant materials in regions of high strain can minimize shear, reduce friction-related injuries, and enhance comfort.

Our previous work introduced a clinically accessible, low-cost digital workflow for mapping the surface strain field and identifying LoNEs on transtibial residual limbs using marker-based photogrammetry and open-source 3D reconstruction software [28]. This method quantifies maximum, minimum, and directional skin strains and reveals consistent patterns across individuals, including peak tensile strain over the patella, compression in the popliteal region, and anatomical yet individualized LoNE trajectories. These results extend prior studies using more complex imaging approaches [29], demonstrating that skin strain mapping provides actionable biomechanical data for prosthetic interface design.

The convergence of digital biomechanics and computer numerical control (CNC) multimaterial knitting presents an opportunity for medical device development. CNC knitting machines can precisely position yarns of different elasticity,

XSL•FO
RenderX

stiffness, and thermal and moisture management properties, enabling the creation of functionally graded textiles that respond to the unique strain environment of each user. By integrating skin strain data with the knitting design process, stiff knitted structures can be strategically aligned along LoNE pathways to enhance suspension and reduce pistoning, while compliant, breathable knit zones can be placed in areas of high deformation to reduce friction, pressure, and heat accumulation.

Building on these innovations, this study introduces *Flexoknit*, a transtibial prosthetic liner that integrates user-specific digital skin strain analysis with CNC multimaterial knitting to create a mechanically tuned, breathable, and anatomically customized interface. Using digital biomechanical data as the primary design driver—rather than clinician heuristics alone—Flexoknit aims to determine the feasibility and performance of a skin strain–guided, CNC-knitted prosthetic interface in terms of material function, clinical performance, and user experience.

Methods

Study Design

This study used an experimental, mixed methods design to develop, characterize, and evaluate Flexoknit, a transtibial prosthetic interface created through CNC multimaterial knitting and informed by residual limb skin strain patterns. The study comprised three major phases: (1) *digital design and fabrication* of multimaterial knitted structures, including yarn selection, stitch programming, and interface construction; (2) *material and fabric characterization* involving thermal analysis, mechanical testing, and moisture management assessment of the knitted samples; and (3) *user testing* to evaluate mobility, suspension, thermal and moisture comfort, and overall user satisfaction compared with a conventional silicone liner.

This design enabled a comprehensive evaluation of both the engineered textile properties and the user-level functional response to a digitally informed prosthetic interface.

Recruitment

Participants were recruited from a prosthetics clinic in Singapore using convenience sampling. Inclusion criteria included participants aged ≥21 years, unilateral transtibial amputation, regular prosthesis users, and the ability to ambulate independently with or without assistive devices. Exclusion criteria included active dermatological issues on the residual limb, acute musculoskeletal conditions, or medical conditions affecting ambulation.

One long-term transtibial prosthesis user (76-y-old man) completed the full user testing protocol, including mobility, suspension, thermal, moisture, and satisfaction assessments. An additional healthy participant (35-y-old man) performed the outdoor thermal test in hot and humid conditions to simulate typical Singapore climate exposure. The difference in participants between the indoor and outdoor tests was due to the age of the amputee participant. As the intention of the outdoor sweat and thermal test was to generate adequate sweat within a relatively short duration (15 min) by having a user who could exert themselves under the hot sun, the team opted for a healthy younger participant to perform this outdoor test instead.

Materials and Fabrication

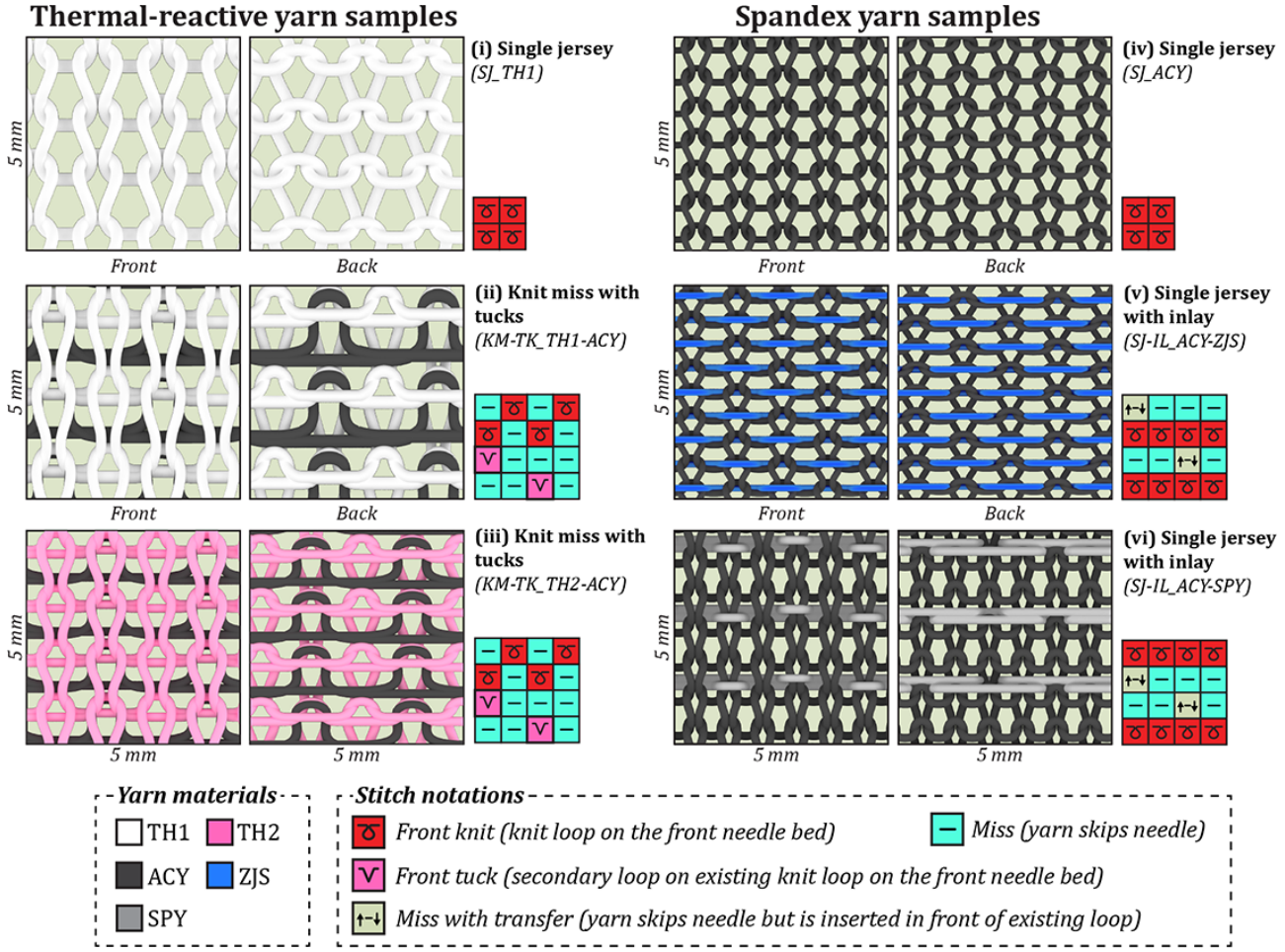
For the main body of the Flexoknit interface, 2 types of thermal-reactive yarns were used. TH1 consists of a 10/1 bicomponent polyester binder spun yarn supplied by Unitika, branded under the model ESPORAN. This yarn is constructed with a polyester core that has a melting point of 256 °C and a co-polyester sheath with a melting point of 110 °C. TH2 is made from a blend of polyester, polyurethane, and nylon, featuring a fusing point of 100 °C, supplied by Asahi Kasei. Three types of spandex yarns were used: ACY (300 D polyester DTY (96 F) and 100 D spandex); ZJS (210 D spandex and 2 ends of 75 D polyester, under the model number 2107575) from Zhejiang Kangjiesi New Material Technology Co., Ltd; and SPY, a low-stiffness elastic yarn from Shima Seiki called Shima Set Up Yarn.

The yarns were knitted into fabric on a Shima Seiki MACH2XS153-15L WHOLEGARMENT 15-gauge flat-bed weft knitting machine programmed using the SDS-ONE KnitPaint software. Three stitch patterns were used in the knitted prosthetic interface (see Figure 1 for diagrams of the stitch patterns). Single jersey (SJ), also known as plain jersey or stockinette stitch, is formed with only front knit stitches. SJ with inlay (SJ-IL) consists of 1 yarn to knit a SJ fabric, while a second yarn is woven in and out of the knitted loops of the first yarn. The second yarn forms approximately straight floats of yarn along the course of the fabric. Knit-miss (cross-miss) 1×1 with tucks (KM-TK) uses 1 yarn to knit the knit-miss 1×1 stitch pattern, where alternate needles knit and miss stitches, with the order of the knit and miss stitches reversing in every course. A second yarn tucks on the knit-miss 1×1 fabric at regular intervals.

More information on the fabrication of the prosthesis can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 1. Stitch patterns and yarn combinations used in the knitted prosthetic interface. For each stitch pattern, a diagram of the front and back of a 5 mm square of fabric is shown, with a knit notation of the stitch pattern at the bottom right. There are three combinations that contain thermal-reactive yarn: (i) single jersey (yarn: TH1), (ii) knit-miss (TH1) with tucks (ACY), and (iii) knit-miss (TH2) with tucks (ACY). There are 3 combinations that do not contain thermal-reactive yarn: (iv) single jersey (ACY), (v) single jersey (ACY) with inlay (ZJS), and (vi) single jersey (ACY) with inlay (SPY).



Material and Fabric Characterization

Yarn Materials

The 3 yarn materials that were used most extensively in the knitted prosthetic interface were characterized using a PerkinElmer DSC 6000 differential scanning calorimetry (DSC) machine. A sample of yarn weighing 7.5 to 8.6 mg was placed in a 50 µL aluminum pan and heated and cooled from 15 °C to 200 °C to 15 °C for 2 cycles at a rate of 10.00 °C per minute in a nitrogen atmosphere (20 mL per minute; see Table S1 in Multimedia Appendix 1 for more details).

Knitted Fabric

The uniaxial tensile properties of the knitted fabrics were characterized according to a modified standard referencing ISO (International Organization for Standardization) 13934 - 1: 2014 (strip test) using an Instron 5943 with a 1.0 kN load cell (see Uniaxial tensile test sample preparation in Multimedia Appendix 1 for more details). The samples were first prepared by heating tubes of fabric on plywood blocks. As the knitted prosthetic interface is stretched in the course-wise (CW) direction when the interface is worn on the residual limb, the samples were made in 2 sizes: *regular* (tube circumference: 390 mm) and *narrow* (tube circumference: 286 mm or 26.7% narrower than regular tubes). The narrow tubes had to stretch to fit onto the plywood blocks. After heating, flat samples were cut out from the tubes to fit onto the tensile testing machine. Each sample size was 50 mm in width and 50 mm in gauge length, and each knit configuration was tested separately in 2 orientations: *wale wise (WW)* or 0° and *CW* or 90°. The strain rate used was 50 mm per minute across all samples. For each knit configuration, the force-strain curves were plotted, and the average breaking force, tensile strain at break, and elastic modulus were calculated.

For thermal and moisture characterization, 8 stitch patterns were tested using ACY yarn: interlock, mesh, pique lacoste (cross-tuck 1×1), purl (garter stitch), rib 1×1, SJ (plain jersey), single pique (cross-tuck interlock), and tuck 1×1. For each test, 3 samples were tested for each stitch pattern. Thermal resistance (K/W) and relative water vapor permeability (%) were measured using a fast skin model Permetest instrument (Sensora Instruments) in a manner similar to ISO 11092 [30]. Wicking rate was measured using a vertical wicking rate test [31,32], with the WW axis of the fabric aligned to the vertical axis. More information on the vertical wicking rate test can be found in Multimedia Appendix 1.

Prosthetic Interface Fabrication

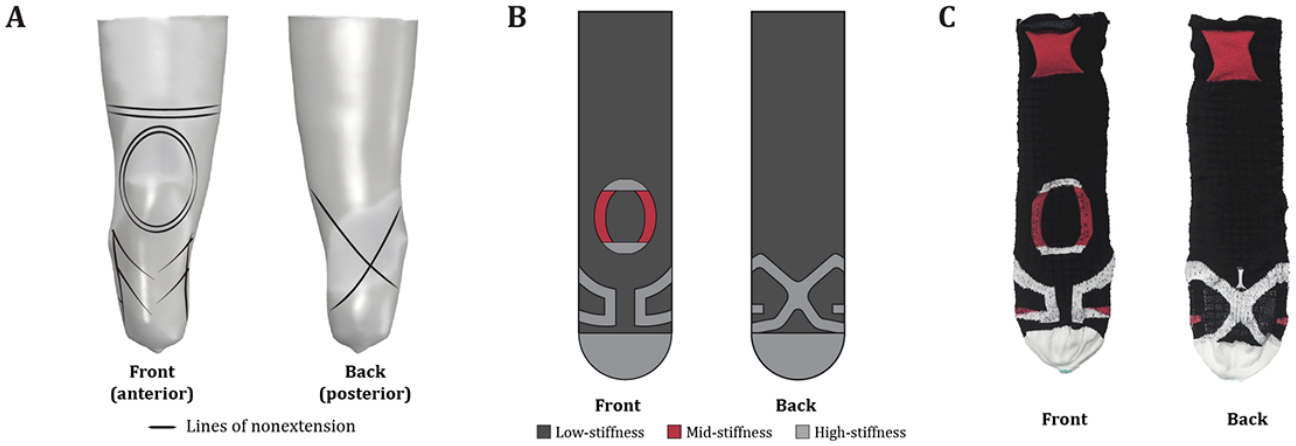
LoNEs are regions of the body that undergo negligible skin strain during movement [27]. Prior studies have suggested using

XSL•FO
RenderX

LoNEs to guide material placement in wearable systems—including prosthetic sockets [33] and mechanical counterpressure suits [34]—by positioning stiff materials along low-strain areas and reserving compliant materials for regions of higher deformation. This strategy allows structural reinforcement without restricting natural skin mobility. Similar approaches have been applied in orthotic design using selectively reinforced 3D-printed elements [35].

In this study, LoNEs informed the spatial layout of stiffness zones within the knitted prosthetic interface. The user's residual limb was photographed in multiple poses, reconstructed as a 3D model, and analyzed to generate directional strain fields (Figure 2A; see [28] for full methodology). High-, medium-, and low-stiffness zones were then mapped based on these strain patterns (Figure 2B), forming the structural template for the knitted design. In this study, mapping was conducted under controlled, non–weight-bearing conditions, with the participant supported and not loading the residual limb. Skin deformation was captured between 2 standardized knee positions: full extension (0°) and flexion at approximately 60°, measured using a goniometer to ensure consistency.

Figure 2. Design of the graded stiffness knitted prosthetic interface. (A) Diagram of lines of nonextension based on directional strain field from photographs of the user's residual limb. (B) Diagram of high-, medium-, and low-stiffness zones on the prosthetic interface. (C) Photograph of the knitted prosthetic interface.



For each posture, a circumferential video of the residual limb was recorded over approximately 90 seconds, from which a sequence of images (~300 frames) was extracted for 3D reconstruction. The strain field was computed by comparing the reconstructed limb geometry between the extended and flexed states, and the resulting directional strain data were used to identify LoNEs for interface design. A single mapping sequence was performed for the participant, and the derived strain field was used directly for design implementation.

This mapping approach captures kinematic skin strain behavior under controlled conditions and does not account for additional deformation under load-bearing scenarios. Accordingly, LoNEs are interpreted as functionally low-strain design pathways rather than fixed anatomical lines under dynamic loading.

High-stiffness zones were positioned directly over LoNE pathways, providing mechanical reinforcement without restricting mobility. These areas resist elongation under load, supporting suspension, minimizing pistoning, and redistributing pressure toward more tolerant regions of the residuum, such as the medial and lateral tibial surfaces and calf musculature [36]. A stiffened distal zone was added to improve load transfer at the socket-interface junction while reducing overall weight by confining rigid materials only to targeted regions.

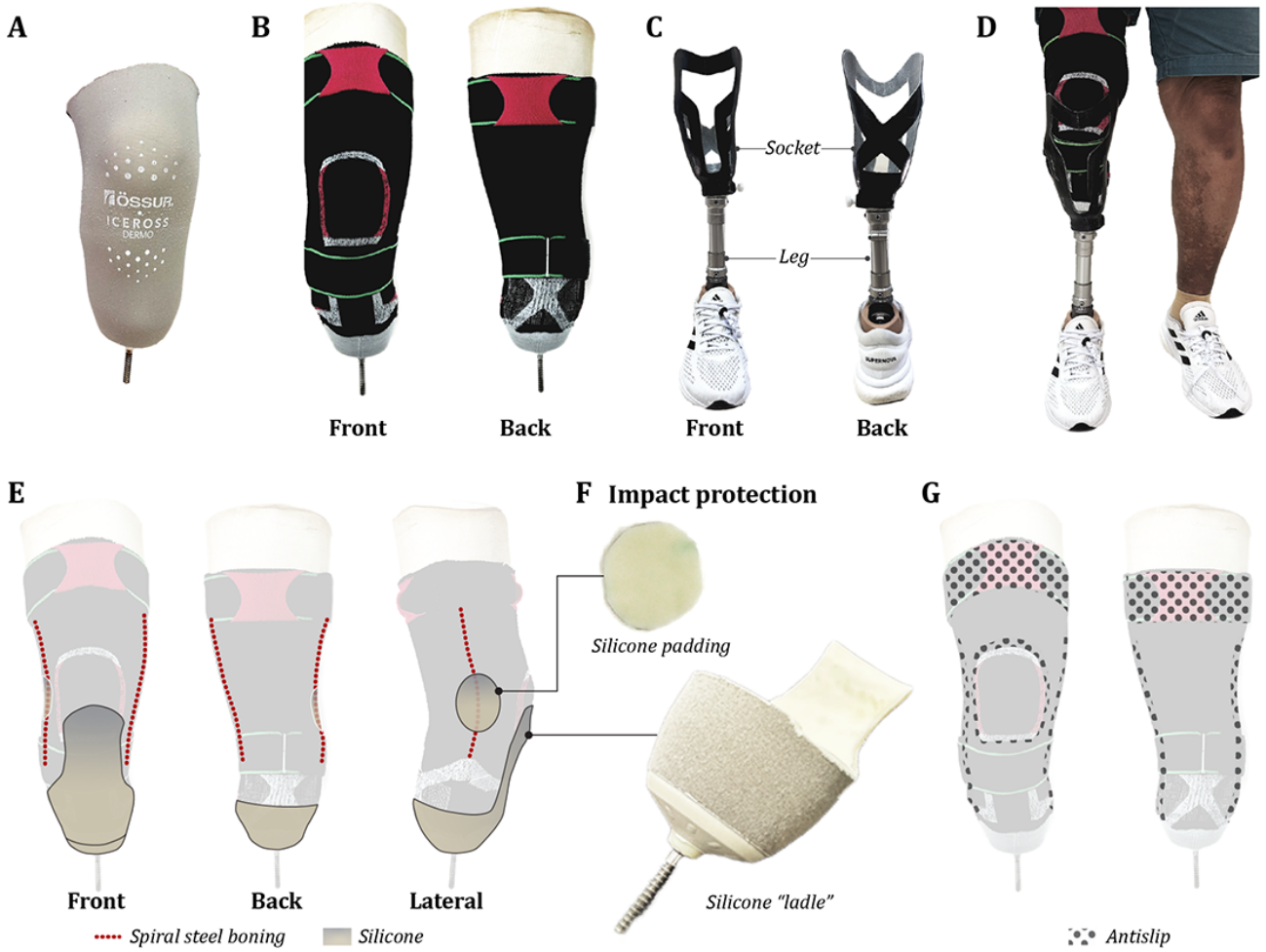
Medium-stiffness zones were placed adjacent to the patella, where controlled flexibility is required to accommodate knee movement. This configuration provided directional stiffness in the vertical axis to prevent interface deformation and limit migration while preserving transverse flexibility for knee flexion.

Low-stiffness zones corresponded to regions experiencing high skin strain and were designed to stretch freely, supporting soft tissue deformation during movement. As these zones form most of the interface surface, using breathable, compliant yarns in these areas substantially improved comfort and thermal and moisture management. Integrated compression structures in both layers further improved mechanical coupling between limb and interface and helped manage shear.

The final Flexoknit interface was constructed as a double-layered knitted tube (Figure 3). The outer layer incorporated spatially graded stiffness elements through specific yarn and stitch combinations. High-stiffness zones were primarily knitted using the KM-TK_TH1-ACY configuration, while SJ_TH1 was used at the distal end to allow 3D shaping without compromising rigidity. Medium-stiffness zones used KM-TK_TH2-ACY, and lateral and medial regions included integrally knitted channels containing spiral steel boning. The boning provided vertical stiffness to reduce pistoning while preserving transverse flexibility for knee movement.

XSL•FO
RenderX

Figure 3. Prosthetic interface components. (A) Photograph of current Iceross silicone liner used by the user. (B) Photograph of experimental Flexoknit prosthetic interface. (C) Photograph of hard socket with artificial leg. (D) Photograph of complete Flexoknit prosthetic system. (E) Diagram showing locations of spiral steel boning and silicone padding for impact protection. (F) Photographs of silicone padding for impact protection. (G) Diagram showing the location of antislip.



Low-stiffness regions were knitted with compliant spandex yarns (eg, SJ_ACY) to enable close limb conformity and accommodate local strain. Compression was strategically integrated: SJ-IL_ACY-SPY provided CW compression in the outer layer, while SJ-IL_ACY-ZJS applied surface-level compression in the inner layer, enhancing mechanical coupling and reducing shear.

Additional design features addressed impact protection, suspension, and adjustability. Cushioning pads were placed over vulnerable anatomical areas, including the fibular head, tibial crest, and distal stump. Suspension was facilitated by the thermoformed fit of the knitted tube, augmented by hook-and-loop straps, a patellar silicone ring, and antislip regions aligned with LoNEs. These elements improved security during ambulation without compromising donning ease or adaptability to limb volume fluctuations.

The knitted prosthetic interface was fabricated on a Shima Seiki MACH2XS153-15L WHOLEGARMENT 15-gauge flat-bed weft knitting machine, with all stitch programming executed in SDS-ONE APEX/KnitPaint. Thermal-reactive yarns (TH1 and TH2) were used to enable structural fixation after heat setting. During the user tests, the knitted prosthetic interface was inserted into a rigid laminated transtibial socket and aligned to ensure correct placement of stiffness zones.

User Testing

Following fabrication of the Flexoknit prosthetic interface, a series of controlled tests was conducted to compare performance against the participant's existing prosthesis with a conventional silicone liner. User testing was organized into 3 domains (ie, mobility, suspension, and comfort) using standardized quantitative assessments and structured qualitative data collection. A summary of the test protocol is provided in Table S2 in Multimedia Appendix 1.

Mobility Testing

Range of Motion

Range of motion (ROM) measurements were obtained using a standard goniometer. Joint angles were recorded for the hip, knee, and ankle with and without the prosthesis in place. ROM measurements followed the procedures described in Range of Motion in Multimedia Appendix 1.

Ten-Minute Walk Test

The participant was instructed to walk continuously for 10 minutes at a self-selected pace within a designated indoor gym

XSL•FO
RenderX

space. Distance covered was recorded. Temperature readings were collected concurrently as described in Skin temperature measurements in Multimedia Appendix 1.

Ten-Meter Walk Test

Walking speed was assessed on a straight 10 meter walkway. To reduce participant burden, timings were obtained during the first 3 laps of the 10-Minute Walk test (10MinWT). The test was performed at a comfortable and safe pace.

Timed Up and Go

The Timed Up and Go (TUG) test followed a standardized procedure: starting from a seated position, the participant stood on cue, walked 3 meters, turned, returned to the chair, and sat down. Total completion time was recorded.

Suspension Testing

A static elongation test was carried out. Two reference marks were applied along the proximal brim of the rigid socket—one at the highest lateral point and the other aligned with the patella. With the participant standing, the prosthetic limb was lifted while maintaining full knee extension. Vertical displacement between the original and displaced marks was measured to determine socket migration relative to the knitted interface. Displacement was measured manually using a ruler aligned to the reference markings, with a measurement resolution of 1 mm, and measurements were repeated across 3 trials for each prosthetic condition, with the mean value reported. The measurement error for vertical displacement using manual reference markings was estimated to be +1 or −1 mm based on repeated pilot measurements. Any measured variability across the 3 trials would also be reported as well.

Testing was performed with the knee maintained in full extension throughout to minimize variability due to joint position. Limb loading was standardized using the participant's body weight during standing prior to limb lifting, followed by manual unloading of the prosthesis; however, no instrumented load measurement was used, and the applied force during limb lifting was not quantified.

This method represents a static assessment of suspension and does not account for dynamic loading conditions encountered during gait, including cyclic loading, shear forces, and stance-phase transitions. Given the manual measurement resolution and estimated error margin, small between-condition differences should be interpreted with caution. Supplementary feedback on suspension was collected through brief unstructured interviews.

Comfort Testing

Mass Measurement

The total mass of each prosthetic configuration was recorded using a digital scale (Model 103272; Redman). Socket mass was measured separately, and interface mass was calculated as the difference between total system mass and socket mass. The same procedure was used for both the silicone liner system and the Flexoknit system.

Thermal Comfort

Thermal measurements were collected during the indoor 10MinWT. The indoor environment was measured at 23 °C, and the participant was dressed in a T-shirt and shorts on the day of the test. A ProtoCentral MAX30205 wearable body thermometer was secured to the center of the calf muscle belly. This location was selected due to the muscle's role in generating heat during physical activity [37,38]. Continuous skin temperature data were recorded throughout the 10-minute walking period for both liner conditions.

As the 10MinWT was conducted indoors in an air-conditioned clinic, a secondary experiment was designed to evaluate thermal performance in a warm, humid outdoor environment, reflective of typical daytime conditions in Singapore (31 °C - 33 °C) [39] and reflective of a more accurate simulation of prolonged heat exposure. A healthy participant (35-y-old man), who was dressed in a T-shirt and shorts, wore customized versions of the silicone and knitted interfaces with dual openings for ease of donning. Each test included a resting period (5 - 15 min or until temperature stabilization) followed by 15 minutes of continuous walking on a partially sun-exposed track. Temperature was recorded using the same protocol as the indoor testing.

Moisture Comfort

Moisture comfort was assessed through verbal feedback and postactivity inspection of the prosthetic interface. Examinations focused on the presence of sweat accumulation and tactile dampness. Observations were documented immediately following each walking trial for both interface types.

User Experience and Satisfaction

Standardized Questionnaires

Two validated user-reported outcome measures were administered: (1) Socket Comfort Score (SCS) [40] and (2) Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0).

The participant provided independent ratings for the conventional liner and Flexoknit conditions.

Qualitative Interviews

Unstructured interviews were conducted to gather additional feedback related to perceived comfort, usability, adjustability, and overall experience with each interface. Real-time observations of interface behavior during walking and donning were also recorded.

Data Analysis

Material properties (DSC peaks, breaking force, strain at break, modulus, thermal resistance, vapor permeability, and wicking rate) were analyzed. Tensile results were plotted as force-strain curves. Mobility and suspension outcomes were compared between the Flexoknit interface and the participant's silicone liner. Temperature data were analyzed as continuous time series profiles. Moisture comfort and qualitative feedback were thematically summarized. Questionnaire scores (SCS and QUEST 2.0) were tabulated and compared across configurations.

XSL•FO
RenderX

This study comprised a material characterization phase and a single-user proof-of-concept evaluation. Owing to the single participant design of the user testing component (n=1), no inferential statistical analysis was performed for the clinical, functional, thermal, or satisfaction outcomes. Normality testing and hypothesis testing were not appropriate for single participant data. User testing results are therefore presented descriptively as raw values and percentage differences between the experimental and conventional prosthetic conditions.

For specific outcome measures, ROM, gait, mass, thermal comfort, and SCS data are reported as observed values, while QUEST 2.0 results are summarized using median scores. Suspension testing using the static elongation method is reported descriptively, with mean displacement values presented for each condition.

Material and fabric characterization data were also analyzed descriptively. Thermal and moisture management results are presented as mean values, as presented in Table 1, and mechanical behavior is presented using observed ranges and representative force-strain responses across yarn-stitch combinations. As this study was designed as a proof-of-concept investigation, no inferential comparisons were performed.

Table . Thermal and moisture comfort results.

Stitch pattern	Thermal resistance		Water vapor permeability		Vertical wicking rate	
	Mean (m^2·K/W)	Rank	Mean (%)	Rank	Mean (mm/s)	Rank
Interlock	40.20	3	35.97	6	0.1437	7
Mesh	40.67	4	45.13	2	0.1971	3
Pique lacoste	41.63	5	41.87	5	0.2115	2
Purl	35.68	2	42.33	4	0.1439	6
Rib 1×1	58.97	8	34.27	7	0.0843	8
Single jersey	31.95	1	47.47	1	0.1964	4
Single pique	48.23	7	30.60	8	0.1755	5
Tuck 1×1	42.48	6	42.60	3	0.2778	1

Ethical Considerations

This study was approved by the Institutional Review Board of the Singapore University of Technology and Design (HBR-22 - 00507). All participants provided written informed consent prior to participation. Participants were informed of the study objectives, procedures, potential risks, and benefits, and were free to withdraw from the study at any time without penalty. All data were deidentified prior to analysis and reporting to protect participant privacy and confidentiality. Only the study investigators had access to identifiable participant information, which was stored securely in accordance with institutional data protection requirements. Participants did not receive financial compensation for participation in this study.

Results

Results are presented according to the 4 methodological domains: material and fabric characterization, user testing results (mobility, suspension, and comfort), user experience, and user satisfaction.

Participant Characteristics

One transtibial prosthesis user (76-y-old man, height 173 cm, weight 78 kg, and 15 y post amputation due to diabetes) completed the full mobility, suspension, comfort, and user experience testing. The participant's conventional prosthesis consisted of a 6 mm silicone locking liner, laminated composite socket, pylon components with a pin lock suspension system, and a 1C30 Trias foot, worn for an average of 9 hours daily.

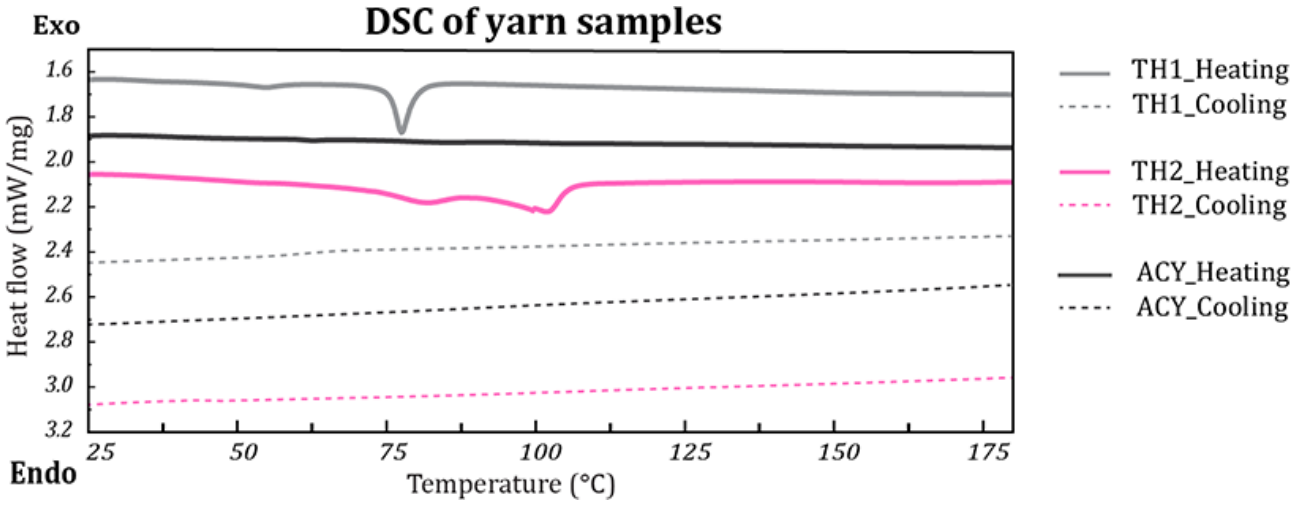
Material and Fabric Performance

Thermal Characteristics (DSC Results)

DSC revealed distinct thermal behaviors among the yarns (Figure 4). TH1 exhibited a single melting peak at approximately 75 °C, while TH2 showed 2 broad melting transitions at approximately 75 °C and 110 °C. The ACY spandex yarn demonstrated no significant thermal transitions within the tested range.

XSL•FO
RenderX

Figure 4. Differential scanning calorimetry (DSC) thermograms for the 3 types of yarns: TH1, TH2, and ACY. The solid lines represent heating, while the dashed lines represent cooling.

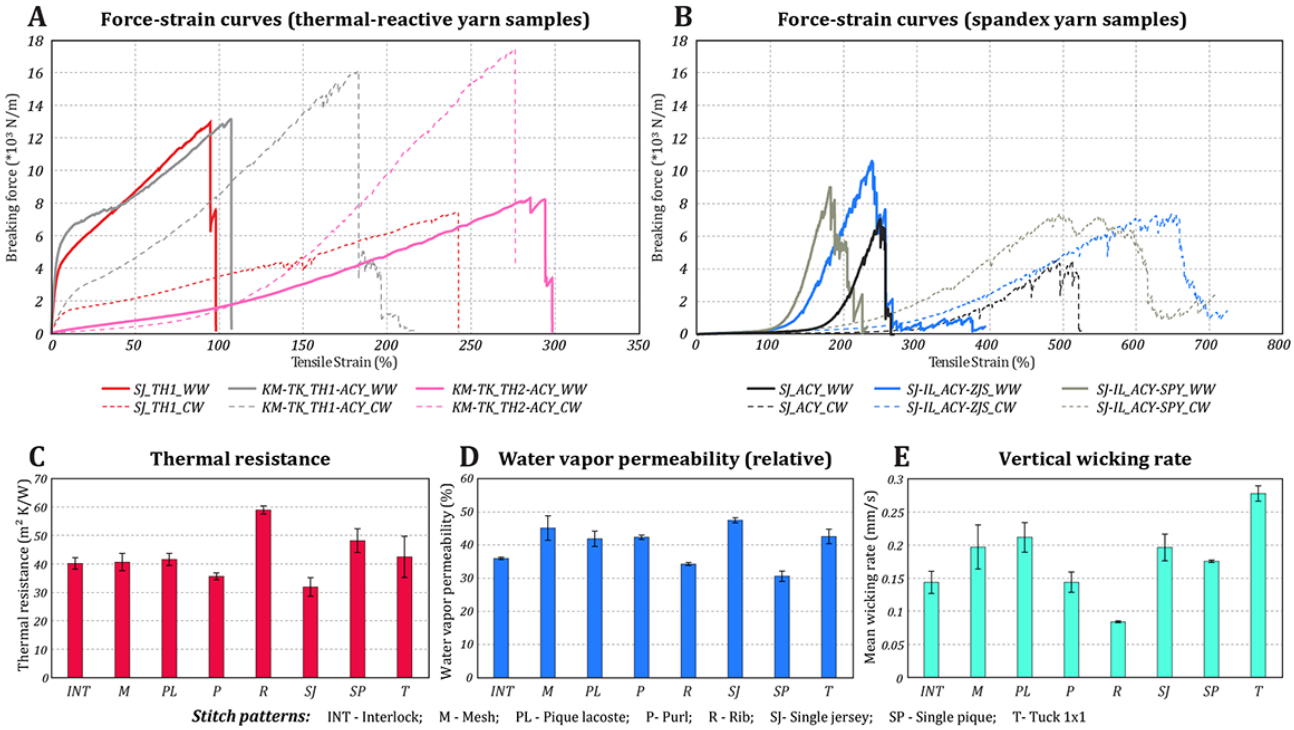


Mechanical Behavior of Knitted Fabrics

Uniaxial tensile testing demonstrated clear differentiation in mechanical behavior across the yarn-stitch combinations (Figure 5). Fabrics incorporating thermal-reactive yarns exhibited higher breaking forces and lower tensile strains at break compared with those knitted solely from spandex-based yarns. Analysis of elastic modulus values showed that the knitted samples consistently fell into 3 stiffness categories:

- Low stiffness (38 - 288 N/m): SJ_ACY, SJ-IL_ACY-SPY, SJ-IL_ACY-ZJS
- Medium stiffness (1224 - 2978 N/m): KM-TK_TH2-ACY
- High stiffness (12,886 - 110,337 N/m): SJ_TH1, KM-TK_TH1-ACY

Figure 5. Material characterization results. (A, B) Uniaxial tensile tests for 6 stitch pattern and yarn combinations (regular-width samples only). (A) Force-strain curves for samples containing thermal-reactive yarn and (B) force-strain curves for samples that do not contain thermal-reactive yarn. (C-E) Thermal and moisture comfort results for 8 stitch patterns knitted with ACY yarn. (C) Thermal resistance, (D) relative water vapor permeability, and (E) vertical wicking rate.



The inclusion of thermal-reactive yarns resulted in substantially higher stiffness and strength at break. For example, KM-TK_TH1-ACY exhibited an elastic modulus nearly an order of magnitude greater than KM-TK_TH2-ACY, despite both samples using the same stitch architecture. KM-TK_TH1-ACY showed a more brittle stress-strain response, whereas KM-TK_TH2-ACY demonstrated greater elasticity, with pronounced anisotropy characterized by higher stiffness in the CW direction than in the WW direction.

To evaluate the effect of prestretching during heat setting, both regular and narrow samples were tested. Narrow samples, which replicated the stretch applied when forming the knitted interface

XSL•FO
RenderX

over a mold, generally exhibited increased stiffness and breaking force in the CW direction, along with reduced tensile strain at break. In the WW direction, narrow samples typically showed decreased breaking force, increased strain at break, and reduced stiffness. An exception was SJ_TH1, which demonstrated increased stiffness in both directions when tested in the narrow state. Across all combinations, narrow samples showed more homogeneous mechanical behavior between axes, indicating reduced anisotropy after prestretching. No systematic differences were observed between front and back sample orientations.

In addition, the force-strain curves in Figures 5A and 5B revealed the distinct failure modes of the different knit samples. Upon testing the samples to their breaking point, the thermal-reactive yarn samples exhibited brittle failure with a sudden force drop, possibly due to the presence of hardened thermoplastic (TH1 and TH2) in the knit. The failure mode is akin to a cleavage in the middle of the sample's testing area. On the other hand, the spandex yarn samples exhibited ductile tensile failure with significant "noise" upon breaking. The tested samples showed necking before failure, followed by several consecutive instances of columns or rows of elastic yarn loops breaking and unraveling upon failure. This behavior was reflected as several step-down peaks after the sample reached its breaking force.

Thermal and Moisture Management Performance

Eight stitch patterns were evaluated for thermal resistance, water vapor permeability, and wicking rate (Table 1). SJ demonstrated the lowest thermal resistance and highest vapor permeability, with a mid-range wicking rate, ranking highest overall for thermal and moisture management performance.

User Testing

Mobility Outcomes

Range of Motion

ROM testing showed measurable differences across prosthetic configurations. Using the knitted interface alone, maximum knee extension remained unchanged, while maximum flexion decreased by 10°. When used with the LoNE-aligned rigid socket, total knee ROM was 98°, compared with 80° when using the participant's current prosthesis (silicone liner and hard socket). In both conditions, terminal knee extension was limited to 20° of flexion (ie, a 20° extension deficit). Total ROM was calculated as the difference between peak flexion and terminal extension, and the 18° increase in ROM for the knitted interface and LoNE-aligned rigid socket is therefore attributed to increased flexion capability rather than a change in extension (Table 2).

Table . Range of motion (ROM) measurements of the various prosthetic configurations.

Prosthesis type	Extension (°)	Flexion (°)	Total ROM (°)
None (bare skin)	0	130	130
Knitted prosthetic interface	0	120	120
Current prosthesis	20	100	80
Experimental prosthesis	20	118	98

Gait Performance

During the 10MinWT, the participant covered 10.5% less distance with the experimental prosthesis compared with the conventional system. In the 10-Meter Walk Test (10mWT), walking time increased by 8.3% when using the knitted interface. The TUG test duration increased by 1.7 seconds. The participant ambulated independently during all tests.

Suspension Outcomes

Socket migration measured using the static elongation test showed a mean displacement of 4 mm (+1 or −1 mm) with the experimental system and 3 mm (+1 or −1 mm) with the conventional system. Given that the observed between-condition difference (1 mm) lies within the estimated measurement error of the manual marking method (+1 or −1 mm), this finding should be interpreted with caution. During indoor walking, the participant reported a gradual loosening sensation and sinking at the distal tibia when using the experimental prosthesis, despite the small measured difference in static displacement. These observations were recorded during the structured suspension assessments.

Comfort Outcomes

Mass

The experimental prosthesis weighed 26.1% less than the participant's current prosthesis. The knitted interface alone was 37.5% lighter than the silicone liner.

Thermal Comfort

In the indoor air-conditioned setting (23 °C), the user's resting posterior limb skin temperature increased by 0.89 °C with the knitted interface and 0.18 °C with the silicone liner. During the 10MinWT, the skin temperature increased by 0.82 °C (knitted) and 0.62 °C (silicone).

However, in the warmer outdoor setting (34 °C), the nonamputee's resting posterior limb skin temperature increased by 0.21 °C (knitted) and 0.71 °C (silicone). During the 10MWT, the skin temperatures rose by 1.95 °C with the knitted interface and 2.88 °C with the silicone liner.

Moisture Comfort

No sweat accumulation was observed in indoor trials. In outdoor trials, visible perspiration accumulated beneath the silicone liner, and the participant reported a sweaty sensation. The knitted interface left the skin dry on doffing, although the fabric itself was damp. A sharp drop in skin temperature after removing the

XSL•FO
RenderX

silicone liner was observed; no similar drop occurred with the knitted interface.

User Experience and Satisfaction

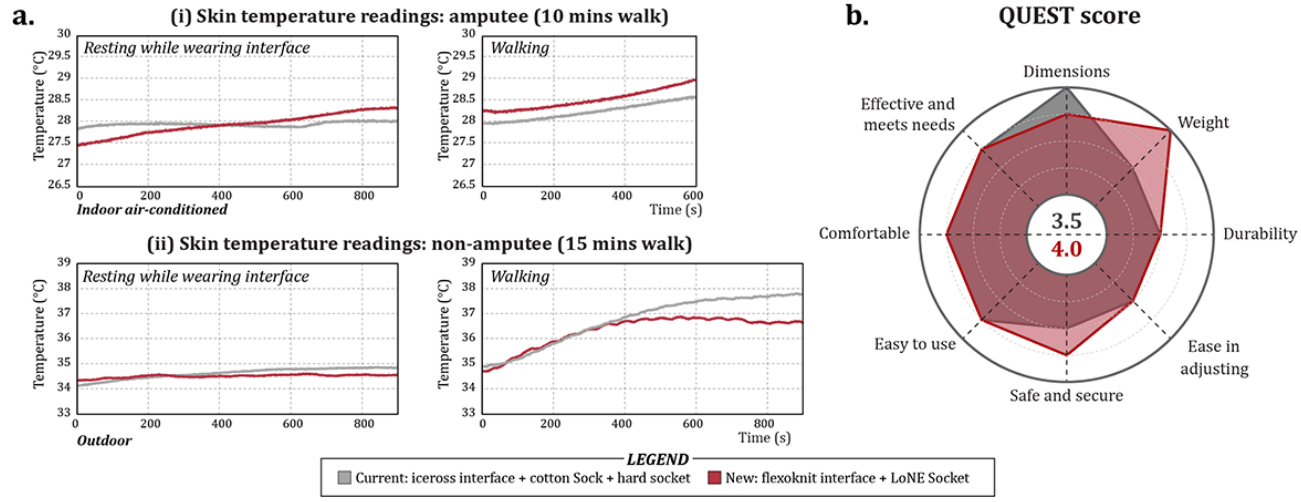
Socket Comfort Score

The participant rated the knitted interface 7/10 and the current prosthesis 9/10.

QUEST 2.0 Satisfaction

The knitted prosthesis received a higher median satisfaction score (4.0) than the conventional prosthesis (3.5; Figure 6B). The IQR of the knitted prosthesis's score was 0.5, while the conventional prosthesis was 1.0, demonstrating that there was less variation in the participant's responses for the knitted prosthesis than the conventional one.

Figure 6. (A) Skin temperature measurements when resting while wearing the prosthetic interface and walking while wearing the prosthetic interface and socket for (i) the amputee and (ii) the nonamputee users. (B) Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST) score showing the user-reported outcomes for the current and experimental prostheses. The numbers in the center of the spider diagram show the median scores for the current (black) and experimental (red) prostheses. A score of 1 represents "not satisfied at all," and a score of 5 represents "very satisfied." LoNE: line of nonextension.



Qualitative Feedback

Qualitative feedback from interviews and real-time observation provided additional insight into the participant's experience using both prosthetic systems. When using the Flexoknit interface, the participant consistently remarked on the noticeably reduced weight of the device, describing it as feeling lighter and more compliant during walking. The knitted structure was perceived as "springy," and the participant noted a softer underfoot feel compared with his usual prosthesis. In indoor trials, the participant reported that the knitted interface felt warmer around the limb, whereas in outdoor trials, the nonamputee participant described the knitted interface as feeling cooler and less sweaty than the silicone liner. The textile appearance of the knitted interface was preferred, with the participant expressing positive reactions to its aesthetic qualities.

During donning, the participant reported that the hook-and-loop suspension straps added steps to the usual routine and occasionally caught on the surface of the knitted fabric, although no structural damage was observed. While walking with the experimental prosthesis, the participant described a gradual sinking sensation at the distal tibia, corresponding to the mild static elongation observed during suspension testing. After returning to the conventional prosthesis following the experimental trials, the participant described the device as feeling comparatively "hard" during stance and walking, indicating a perceptible difference in interface compliance between the 2 systems.

Discussion

Principal Findings

This proof-of-concept study demonstrates the feasibility of a digitally informed, multimaterial knitted prosthetic interface for transtibial users while also highlighting the biomechanical trade-offs introduced by increased interface compliance. By combining residual limb skin strain analysis with CNC multimaterial knitting, we created Flexoknit—a double-layered, graded stiffness interface integrated with a LoNE-informed rigid socket. Material characterization showed that targeted yarn-stitch combinations can generate 3 distinct stiffness tiers spanning nearly an order of magnitude in elastic modulus. When implemented in a transtibial system, this graded design reduced interface mass relative to a conventional silicone liner, preserved or improved knee ROM, and showed context-dependent thermal and moisture advantages, particularly in hot outdoor conditions. However, suspension-related challenges, including the participant's reported distal "sinking" sensation, and the reduced distance walked during the 10MinWT indicate that improvements in ROM and comfort-related properties do not necessarily translate into better gait performance or interface stability. Despite the suspension-related challenges and donning complexity, user-reported satisfaction scores and qualitative feedback indicate that the concept is acceptable and potentially attractive for lower-activity community amputees. Taken together, these findings suggest that increased interface flexibility may have facilitated greater knee excursion during isolated ROM testing, while simultaneously reducing

XSL•FO
RenderX

stance-phase stability during walking through less consistent limb-socket coupling. These findings therefore support feasibility but also highlight the need for further refinement of load transfer and suspension security before broader clinical application.

Design of Multimaterial Graded Stiffness Knitted Fabric

The material characterization phase confirms that both yarn composition and stitch architecture can be systematically tuned to control mechanical behavior. Thermal-reactive yarns (TH1 and TH2) exhibited clear melting transitions and, once heat-set, produced knitted structures with substantially higher stiffness and strength than fabrics knitted solely from spandex yarns, consistent with prior work on thermal-reactive knits [20]. Uniaxial tensile testing showed that KM-TK_TH1-ACY and SJ_TH1 formed the high-stiffness group, KM-TK_TH2-ACY formed the medium-stiffness group, and SJ spandex configurations formed the low-stiffness group. The separation between these categories—38 - 288 N/m; 1224 - 2978 N/m and 12,886 - 110,337 N/m, respectively—provided a clear design palette for allocating support and compliance within the interface.

Prestretching the knitted tubes prior to heat setting, a process that mimics forming the fabric over a limb mold, further modified the mechanical response. Narrow (stretched) samples typically showed increased stiffness and breaking force in the CW direction and more homogeneous behavior between WW and CW axes, suggesting reduced anisotropy once the fabric is tensioned to its in-use configuration. This behavior is advantageous for prosthetic applications, where a predictable, direction-independent response can improve suspension and load transfer.

Thermal and moisture management testing indicated that stitch patterns can be used to optimize comfort-related properties independently of stiffness. Among the patterns assessed, SJ achieved the best overall balance of low thermal resistance, high water vapor permeability, and adequate vertical wicking, supporting its use as the dominant structure in low-stiffness, high-contact zones. Collectively, these findings show that multimaterial CNC knitting can deliver a controlled spectrum of mechanical, thermal, and moisture management properties within a single seamless interface.

LoNE-Guided Reinforcement Design

The LoNE-guided, digitally manufactured construct illustrates a shift from uniform material liners toward spatially tailored interfaces in which stiffness, compression, and breathability are all driven by patient-specific biomechanical data. LoNE-guided reinforcement may improve suspension by placing stiffer load-bearing elements along relatively low-strain pathways, where they can resist migration with less disruption to natural tissue motion. However, LoNEs are not assumed to remain perfectly fixed under load-bearing conditions and should instead be interpreted as functionally stable design zones. Nevertheless, the present findings suggest that anatomical alignment of stiffness zones alone is insufficient; effective suspension depends not only on alignment with LoNE pathways but also on the absolute magnitude and spatial distribution of interface stiffness required to maintain stable load transfer during dynamic gait.

Interpretation of User Testing and Clinical Implications

In functional testing with the participant, the Flexoknit system preserved maximum knee extension and increased overall knee flexion when compared with the conventional prosthesis, yielding a 22.5% improvement in total ROM. This suggests that aligning stiff zones with LoNEs and reserving compliant structures for high-strain regions may permit greater sagittal-plane excursion at the knee without immediately compromising basic structural support. However, the increase in ROM should not be interpreted as a direct proxy for improved functional mobility, as gait performance during the 10MinWT, 10mWT, and TUG was modestly worse with the experimental system. Despite the 26.1% reduction in prosthesis mass, the participant walked more slowly with the experimental system. This suggests that prosthetic weight alone was not the dominant determinant of performance in this case. From a biomechanical perspective, walking efficiency depends not only on limb mass but also on the quality of load transfer, proprioceptive consistency, and the user's confidence during the stance phase. It is therefore plausible that the benefits of reduced mass were offset by less stable mechanical coupling at the limb-socket interface. Importantly, the participant was able to ambulate independently and perform functional tasks such as sit-to-stand, indicating that the experimental system met minimum requirements for community mobility.

The participant's report of a gradual distal tibial "sinking" sensation is biomechanically significant, suggesting incomplete load transfer coupling between the residual limb, knitted interface, and rigid socket. In an optimally coupled system, axial and shear forces are distributed proximally and circumferentially to maintain stable limb positioning. In this case, the increased compliance of the knitted interface likely permitted localized deformation—particularly distally—reducing the efficiency of axial load transfer. This may explain the perceived distal settling despite only a 1 mm difference in static elongation, highlighting that small displacements can be clinically meaningful when repeatedly experienced at sensitive anatomical sites.

Although the study was not designed to isolate the precise structural cause, this behavior likely reflects a combination of compliant material response and insufficient longitudinal reinforcement in low-stiffness regions. While these compliant zones were intended to accommodate soft tissue and enhance comfort, they may have allowed excessive vertical deformation under load. Future iterations should therefore refine not only overall stiffness but also the spatial distribution of reinforcement to improve suspension without compromising comfort.

This phenomenon also has implications for gait stability. While compliance can enhance pressure distribution, excessive deformation may introduce micromovements that reduce proprioceptive consistency and perceived security during stance. This may contribute to a more cautious gait pattern, reduced efficiency, and slower functional performance, even in the

XSL•FO
RenderX

absence of pain. The participant's slower walking speed and longer TUG time are therefore likely influenced not only by device unfamiliarity but also by altered mechanical coupling at the interface, potentially reducing confidence during weight acceptance and forward progression.

The graded stiffness design also achieved a substantial reduction in mass, with the experimental prosthesis approximately one-quarter lighter overall and the knitted interface over one-third lighter than the silicone liner. The participant's perception of the prosthesis as "lighter" and "springy" aligns with these findings and may explain the increased knee flexion observed during ROM testing. However, this reduction in mass did not translate to faster gait. The perceived "springiness" likely reflects increased compliance, whereby some mechanical energy during stance was absorbed within the interface rather than efficiently transferred to forward progression. While this may enhance comfort, it may also reduce the stability required for confident movement. These findings suggest that reductions in prosthetic weight should be considered alongside interface stability, as improvements in one may be offset by compromises in the other.

Thermal and moisture outcomes were environment dependent. In the air-conditioned clinic, the knitted interface produced slightly greater temperature rises than the silicone liner, suggesting increased insulation when ambient temperatures are low. In contrast, in a hot outdoor setting representative of tropical climates, the knitted interface limited skin temperature rise and left the skin dry after walking, while the silicone liner trapped sweat and was perceived as hot and sticky. These findings imply that an open, wicking knitted structure may offer particular advantages in warm, humid environments, where dermatological complications and heat intolerance are common.

Despite generally adequate suspension, the static elongation test and user reports indicated minor distal "sinking" with the experimental system. This finding likely reflects the combined mechanical behavior of the compliant knitted interface, the stiffness zoning strategy, and the structural response of the LoNE-informed rigid socket under dynamic load. Rather than representing gross pistoning alone, the sensation may indicate localized compression and delayed load sharing across the limb-liner-socket complex, particularly during repeated stance loading. Clinically, this is relevant because even subtle distal displacement can alter pressure concentration, disturb the user's perception of stability, and reduce confidence during walking. This observation may also help explain the slower walking speed and longer TUG duration despite the reduced prosthetic mass. Even subtle distal settling or micromovement at the interface can alter how ground reaction forces are transmitted through the prosthesis and may reduce the user's confidence during loading response and mid-stance. In response, the user may adopt a more cautious gait strategy, with slower progression and more conservative transfer during transitional movements such as turning and sit-to-stand. This likely reflects the combined behavior of the compliant interface and the LoNE-informed rigid socket, including possible flexing of the latter under load.

The lower SCS for Flexoknit compared with the user's own device appears to be driven by this sensation of gradual loosening rather than by gross discomfort or pain. This interpretation is also consistent with the higher overall QUEST score and positive qualitative feedback regarding appearance, weight, and perceived compliance, suggesting that user acceptance of the concept remained favorable despite the mechanical shortcoming. Taken together, these findings suggest that interface compliance must be carefully balanced: sufficient to accommodate tissue motion and improve comfort, but not so great that it compromises distal support or mechanical coupling.

Future iterations should therefore investigate increasing distal and longitudinal stiffness, refining the suspension mechanism, and optimizing outer socket rigidity to reduce perceptible settling during gait. Nonetheless, the higher overall QUEST score and the participant's preference for the device's appearance and perceived security suggest that, with refinements to the outer socket stiffness and suspension strategy, user comfort could be further improved.

Conclusions

This study presents a proof-of-concept transtibial prosthetic interface—Flexoknit—engineered through a multimaterial, graded stiffness knitting approach informed by LoNEs. By combining thermal-reactive yarns, spandex-based elastic yarns, and spatially targeted stitch architectures, the knitted interface achieved distinct zones of mechanical behavior that were aligned with the user's residual limb strain patterns and functional requirements. The fabrication process demonstrated that CNC knitting can reliably integrate structural reinforcement, compliance, compression, and breathability within a single seamless textile construct tailored to an amputee's limb geometry.

Material testing confirmed that the yarn-stitch combinations produced the intended stiffness gradients, while functional testing showed that the interface supported safe ambulation and preserved or enhanced joint mobility when used with a LoNE-aligned rigid socket. Improvements in thermal and moisture regulation, particularly in warm outdoor conditions, and a substantial reduction in overall prosthesis weight further demonstrated the potential benefits of integrating knitted structures into prosthetic design. User feedback highlighted the comfort, lightness, and aesthetic appeal of the device, reinforcing the feasibility of this approach for improving user experience.

As an early prototype, some usability challenges remain, including potentially inadequate prosthesis stability under load and the need for more intuitive suspension and donning mechanisms. Nonetheless, the findings suggest that digitally engineered knitted interfaces can provide a highly customizable, breathable, and compliant alternative to conventional silicone liners, particularly for lower-activity amputees or individuals prioritizing comfort and ease of use.

Overall, this work establishes Flexoknit as a promising direction for future prosthetic development—one that integrates principles of biomechanics, textile engineering, and digital fabrication to create user-centered interface solutions. Further research is warranted to refine structural performance, evaluate long-term

XSL•FO
RenderX

durability, and validate the design across diverse user groups and real-world conditions.

Limitations and Future Work

This study has several limitations that should be considered when interpreting the findings. First, the prosthetic evaluation was conducted with a single transtibial amputee user, and thermal testing in a hot outdoor environment involved only one additional healthy participant. The results therefore reflect an in-depth case study rather than population-level evidence and may not generalize to users with different limb geometries, activity levels, or comorbidities. Second, the familiarization period with the Flexoknit system was short, so the gait and comfort outcomes likely include early adaptation effects. Longer-term use may alter both objective performance and subjective preferences. Third, the prototype system combined a novel knitted interface with a bespoke LoNE-informed rigid socket that has not yet been fully optimized. The biomechanical interpretation of distal "sinking" is based on user reports, static elongation data, and observational findings rather than instrumented assessment of in-socket motion or pressure distribution. In addition, while material and bench-top testing were extensive, real-world durability, laundering behavior, and the long-term skin response to continuous fabric–skin contact were not assessed. These aspects are critical for translation to everyday clinical practice.

Future work should therefore focus on multiparticipant studies that include a broader range of ages, activity classifications, and residual limb shapes, with follow-up over weeks to months. Such studies should incorporate more detailed biomechanical outcomes (eg, 3D gait analysis, interface pressure, and shear mapping) and dermatological assessments to evaluate skin integrity over time. On the design side, iterative optimization of the rigid socket stiffness, strap layout, and donning sequence is needed to reduce distal sinking and simplify daily use. Finally, further development of the digital workflow—from automated LoNE extraction to direct translation into knit patterns—along with systematic durability and wash-cycle testing, will be essential to move Flexoknit from proof-of-concept to a clinically deployable, manufacturable system.

Acknowledgments

The authors would like to thank the Singapore Institute of Manufacturing Technology (SIMTech) at Nanyang Technological University for allowing the use of their humidity chamber for material characterization experiments, and ST Engineering Land Systems for allowing the use of their Permetest instrument. The authors would like to thank the National University of Singapore's Department of Material Science and Engineering, Pornsruang Wiwattanakul (Nadia) (SUTD), and Taylor Jade Self (SUTD/NUS) for their assistance with conducting differential scanning calorimetry experiments. The authors would like to thank Ng Ming Liang (SUTD) and Wong Yun Han Madeline (NUS) for their assistance with conducting material characterization experiments, Benjamin Lim En How (SUTD) for assistance with photography, and Christyasto Priyonggo Pambudi (SUTD) for assistance with illustrations. During the preparation of this work, the authors used Claude.AI to improve the readability and language consistency of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This work was supported by the National Healthcare Group Centre for Medical Technologies & Innovations and National Health Innovation Centre Singapore Joint MedTech Grant (CMTiNHIC3-21-01-04).

Authors' Contributions

TB contributed to conceptualization, methodology, investigation, resources, writing of the original draft, supervision, and funding acquisition. PZC contributed to methodology, investigation, writing of the original draft, reviewing and editing, and project administration. YYT contributed to methodology, formal analysis, investigation, writing of the original draft, reviewing and editing, and visualization. HYL contributed to conceptualization, reviewing and editing, supervision, and funding acquisition.

Conflicts of Interest

TB reports a relationship with Tan Tock Seng Hospital that includes funding and grants. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Multimedia Appendix 1
Supplementary methods and materials used in this paper.
[DOCX File, 31 KB - rehab_v13i1e91396_app1.docx]

References

XSL•FO
RenderX

1. Roffman CE, Buchanan J, Allison GT. Predictors of non-use of prostheses by people with lower limb amputation after discharge from rehabilitation: development and validation of clinical prediction rules. J Physiother 2014 Dec;60(4):224-231. [doi: 10.1016/j.jphys.2014.09.003] [Medline: 25450484]
2. Chamlian TR. Use of prostheses in lower limb amputee patients due to peripheral arterial disease. Einstein (Sao Paulo) 2014;12(4):440-446. [doi: 10.1590/S1679-45082014AO3132] [Medline: 25628194]
3. Hurwitz M, Czerniecki J, Morgenroth D, et al. Racial disparities in prosthesis abandonment and mobility outcomes after lower limb amputation from a dysvascular etiology in a veteran population. PM R 2025 Feb;17(2):137-146. [doi: 10.1002/pmrj.13240]
4. Agrawal VR, Skrabek RQ, Embil JM, Gross P, Trepman E. Effect of socioeconomic and health factors on prosthetic use after lower-limb amputation. JPO Journal of Prosthetics and Orthotics 2014;26(2):79-86. [doi: 10.1097/JPO.0000000000000027]
5. Ebrahimzadeh MH, Moradi A, Bozorgnia S, Hallaj-Moghaddam M. Evaluation of disabilities and activities of daily living of war-related bilateral lower extremity amputees. Prosthetics & Orthotics International 2016;40(1):51-57. [doi: 10.1177/0309364614547410]
6. Gailey R, McFarland LV, Cooper RA, et al. Unilateral lower-limb loss: prosthetic device use and functional outcomes in servicemembers from Vietnam war and OIF/OEF conflicts. J Rehabil Res Dev 2010;47(4):317-331. [doi: 10.1682/jrrd.2009.04.0039] [Medline: 20803401]
7. Gholizadeh H, Abu Osman NA, Eshraghi A, Ali S, Yahyavi ES. Satisfaction and problems experienced with transfemoral suspension systems: a comparison between common suction socket and seal-in liner. Arch Phys Med Rehabil 2013 Aug;94(8):1584-1589. [doi: 10.1016/j.apmr.2012.12.007] [Medline: 23262380]
8. Schaffalitzky E, Gallagher P, Maclachlan M, Wegener ST. Developing consensus on important factors associated with lower limb prosthetic prescription and use. Disabil Rehabil 2012;34(24):2085-2094. [doi: 10.3109/09638288.2012.671885] [Medline: 22494367]
9. Stevens PM, DePalma RR, Wurdeman SR. Transtibial socket design, interface, and suspension: a clinical practice guideline. J Prosthet Orthot 2019;31(3):172-178. [doi: 10.1097/JPO.0000000000000219]
10. Steer JW, Worsley PR, Browne M, Dickinson AS. Predictive prosthetic socket design: part 1-population-based evaluation of transtibial prosthetic sockets by FEA-driven surrogate modelling. Biomech Model Mechanobiol 2020 Aug;19(4):1331-1346. [doi: 10.1007/s10237-019-01195-5] [Medline: 31256276]
11. Silver-Thorn MB, Steege JW, Childress DS. A review of prosthetic interface stress investigations. J Rehabil Res Dev 1996 Jul;33(3):253-266. [Medline: 8823673]
12. Dillingham TR, Pezzin LE, MacKenzie EJ, Burgess AR. Use and satisfaction with prosthetic devices among persons with trauma-related amputations. Am J Phys Med Rehabil 2001 Aug;80(8):563-571. [doi: 10.1097/00002060-200108000-00003]
13. Berke GM, Fergason J, Milani JR, et al. Comparison of satisfaction with current prosthetic care in veterans and servicemembers from Vietnam and OIF/OEF conflicts with major traumatic limb loss. JRRD 2010;47(4):361. [doi: 10.1682/JRRD.2009.12.0193]
14. Meulenbelt HE, Geertzen JH, Jonkman MF, Dijkstra PU. Determinants of skin problems of the stump in lower-limb amputees. Arch Phys Med Rehabil 2009 Jan;90(1):74-81. [doi: 10.1016/j.apmr.2008.07.015]
15. Hachisuka K, Nakamura T, Ohmine S, Shitama H, Shinkoda K. Hygiene problems of residual limb and silicone liners in transtibial amputees wearing the total surface bearing socket. Arch Phys Med Rehabil 2001 Sep;82(9):1286-1290. [doi: 10.1053/apmr.2001.25154] [Medline: 11552206]
16. Lake C, Supan TJ. The incidence of dermatological problems in the silicone suspension sleeve user. JPO Journal of Prosthetics and Orthotics 1997 ;9(3):97-106. [doi: 10.1097/00008526-199700930-00003]
17. Williams RJ, Takashima A, Ogata T, Holloway C. A pilot study towards long-term thermal comfort research for lower-limb prosthesis wearers. Prosthetics & Orthotics International 2019;43(1):47-54. [doi: 10.1177/0309364618791604]
18. Rezvanifar SC, Conklin S, Davis BL. Experimental thermal analysis of a novel prosthetic socket along with silicone and PCM liners. J Biomech 2020 May 7;104:109788. [doi: 10.1016/j.jbiomech.2020.109788] [Medline: 32312555]
19. Singal K, Dimitriyev MS, Gonzalez SE, Cachine AP, Quinn S, Matsumoto EA. Programming mechanics in knitted materials, stitch by stitch. Nat Commun 2024 Mar 23;15(1):2622. [doi: 10.1038/s41467-024-46498-z] [Medline: 38521784]
20. Tan YY, Quek YH, Chia PZ, Gupta U. Geometric control of thermoformable knitted textiles using raster images. Smart Mater Struct 2023 Jul 1;32(7):075001. [doi: 10.1088/1361-665X/acd66e]
21. Gupta U, Lau JL, Chia PZ, et al. All knitted and integrated soft wearable of high stretchability and sensitivity for continuous monitoring of human joint motion. Adv Healthc Mater 2023 Aug;12(21):1-16. [doi: 10.1002/adhm.202202987] [Medline: 36977464]
22. Vint N, Laghi A. Knit prosthetic liner textile with differentiated knit fabric exterior. 2019 Mar inventors. [Medline: 31984365]
23. Bertels T, Kettwig T. Breathable liner for transradial prostheses. 2011 Presented at: Proceedings of the 2011 MyoElectric Controls/Powered Prosthetics Symposium; Aug 14-19, 2011 URL: https://tinyurl.com/479fjxj8 [accessed 2026-06-13]
24. Mikucioniene D, Halavska L, Melnyk L, Milašius R, Laureckiene G, Arabuli S. Classification, structure and construction of functional orthopaedic compression knits for medical application: a review. Appl Sci (Basel) 2024;14(11):4486. [doi: 10.3390/app14114486]

XSL•FO
RenderX

25. Mikucioniene D, Halavska L, Laureckiene G, Melnyk L, Arabuli S, Milašius R. Development of knitted compression covers for amputated limbs. Fibers (Basel) 2024;12(10):80. [doi: 10.3390/fib12100080]
26. Naake A, Peraza I, Kyosev Y, Wieczorek F. 3D scan to knit - workflow and challenges of automated data processing and knit program generation for prosthesis liners. Presented at: 3DBODYTECH 2023 - 14th International Conference and Exhibition on 3D Body Scanning and Processing Technologies, Lugano, Switzerland, 17-18 October 2023 p. 2023. [doi: 10.15221/23.56]
27. Iberall AS. The Use of Lines of Nonextension to Improve Mobility in Full-Pressure Suits.: Rand Development Corporation; 1964. [Medline: 14262984]
28. Binedell T, Gupta U, Sithanathan B, Subburaj K, Blessing L. Mapping lines of non-extension in persons with lower limb amputation to aid comfort-driven prosthetic socket design. Med Eng Phys 2023 Aug;118(1):104018. [doi: 10.1016/j.medengphy.2023.104018] [Medline: 37536839]
29. Solav D, Moerman KM, Jaeger AM, Herr HM. A framework for measuring the time-varying shape and full-field deformation of residual limbs using 3D digital image correlation. IEEE Trans Biomed Eng 2019 Oct;66(10):2740-2752. [doi: 10.1109/TBME.2019.2895283] [Medline: 30676943]
30. Hes L. Non-destructive determination of comfort parameters during marketing of functional garments and clothing. Indian J Fibre Text Res 2008;33:239-245 [FREE Full text]
31. Test method 197-2011 vertical wicking of textiles. In: AATCC Technical Manual: AATCC:373-376 URL: https://members.aatcc.org/store/tm197/617/ [accessed 2026-06-13]
32. Chen Q, Mei M, Zheng R, Fan J, Fu B, Ma P. Wicking and drying behavior of knitted cotton fabrics under modified water drop test. Journal of Natural Fibers 2022 Jan 2;19(1):232-247. [doi: 10.1080/15440478.2020.1738308]
33. Lin B, Moerman KM, McMahan CG, Pasch KA, Herr HM. Low-cost methodology for skin strain measurement of a flexed biological limb. IEEE Trans Biomed Eng 2017 Dec;64(12):2750-2759. [doi: 10.1109/TBME.2016.2626442] [Medline: 27849521]
34. Wessendorf AM, Newman DJ. Dynamic understanding of human-skin movement and strain-field analysis. IEEE Trans Biomed Eng 2012 Dec;59(12):3432-3438. [doi: 10.1109/TBME.2012.2215859] [Medline: 22961262]
35. Barrios-Muriel J, Romero Sánchez F, Alonso FJ, Salgado DR. Design of semirigid wearable devices based on skin strain analysis. J Biomech Eng 2019 Feb 1;141(2):021008. [doi: 10.1115/1.4040250]
36. Smith DG, Michael JW, Bowker JH. Atlas of Amputations and Limb Deficiencies, 3rd edition: American Academy of Orthopaedic Surgeons; 2004.
37. Silbernagl S, Despopoulos A. Color Atlas of Physiology, 7th edition: Thieme Publishing Group; 2015.
38. Diment LE, Thompson MS, Bergmann JHM. Comparing thermal discomfort with skin temperature response of lower-limb prosthesis users during exercise. Clin Biomech (Bristol, Avon) 2019 Oct;69:148-155. [doi: 10.1016/j.clinbiomech.2019.07.020]
39. Fong M. The Weather and Climate of Singapore, 1st edition: Meteorological Service Singapore; 2012.
40. Hanspal RS, Fisher K, Nieveen R. Prosthetic socket fit comfort score. Disabil Rehabil 2003 Nov 18;25(22):1278-1280. [doi: 10.1080/0963828031000160398 3] [Medline: 14617445]

Abbreviations

10MinWT: 10-Minute Walk Test
10mWT: 10-Meter Walk Test
CNC: computer numerical control
CW: course-wise
DSC: differential scanning calorimetry
ISO: International Organization for Standardization
KM-TK: knit-miss (cross-miss) 1×1 with tucks
LoNE: lines of nonextension
QUEST: Quebec User Evaluation of Satisfaction with Assistive Technology
ROM: range of motion
SCS: Socket Comfort Score
SI-IL: Single jersey with inlay
SJ: single jersey
TUG: Timed Up and Go
WW: wale-wise

XSL•FO
RenderX

Edited by S Munce; submitted 15.Jan.2026; peer-reviewed by D Menychtas, G Harih, L Lou; revised version received 16.Apr.2026; accepted 29.Apr.2026; published 26.Jun.2026.

Please cite as:
Binedell T, Chia PZ, Tan YY, Low HY
Toward a Digitally Informed Knitted Prosthetic Interface With Graded Stiffness to Enhance Comfort in Transtibial Amputees: Proof-of-Concept Case Study
JMIR Rehabil Assist Technol 2026;13:e91396
URL: https://rehab.jmir.org/2026/1/e91396
doi:10.2196/91396

© Trevor Binedell, Pei Zhi Chia, Ying Yi Tan, Hong Yee Low. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Multigesture Electromyographic Control Complexity in Upper Limb Prostheses Actuated via Single Sensor Input Contraction Magnitude: Qualitative Study for Evaluating Performance and Cognitive Load

Abrianna Lalle[1]; Samantha Migliore[1], BSHS; Jeffrey Stevenson[1], BA; Ethan Bell[1]; Maanya Pradeep[1], BSBS; Delaney Gunnell[1], BSBS; Sophie Bennett[1]; Viviana Rivera[1], BSBS, MSIMS; Peter Smith[1], BSCPE, MSMS, PhD; Matt Dombrowski[1], BFA, MFA; John Sparkman[1], BSME, MSME; Albert Manero II[1], BSAE, MAeroE, PhD

Limbitless Solutions, University of Central Florida, Orlando, FL, United States

Corresponding Author:
Albert Manero II, BSAE, MAeroE, PhD
Limbitless Solutions
University of Central Florida
12703 Research Parkway
STE 100
Orlando, FL, 32826
United States
Phone: 1 4078236757
Email: albert@limbitless-solutions.org

Abstract

Background: Lack of functionality is one factor that contributes to prosthetic rejection rates. Electromyographic upper limb prostheses are controlled through muscle contractions in the user's residual limb. The incorporation of multigesture controls into a novel, in-house developed upper limb prosthesis requires users to differentiate between the strengths of muscle contractions to trigger programmed gestures. Little research exists on the limitations of expanding device capabilities. This expansion may lead to a decline in accuracy and perceived usability or an increase in training time and cognitive workload.

Objective: This study aimed to determine the feasibility of implementing multiple gestures when learning electromyographic controls during a single training session.

Methods: Participants with full upper extremity control were fitted with a Flex Controller, a surface electromyography device that measures muscle contraction. Contractions were visualized as peaks and calibrated through an adjustable scale on a tablet. A training app was developed in-house to test novice users on an electromyography control system. Users interacted with 1, 3, or 5 zones on the screen. Each horizontal zone represented a threshold required to trigger a distinct gesture on the prosthesis. The cohorts were labeled A1 (n=9), A2 (n=10), A3 (n=9), and B1 (n=26). Every participant completed 3 trials per arm, and each trial consisted of 15 randomized cues. Each cue was represented by a green color change, with 1 point earned after a successful peak. Collected outcomes included performance, the System Usability Scale, and the National Aeronautics and Space Administration Task Load Index.

Results: Scores decreased significantly as zones increased (Kruskal-Wallis H_3=24.9, P<.001). The mean scores were 15.0 (SD 0.0) for 1 zone, 9.1 (SD 1.1) for 3 zones, and 5.5 (SD 1.1) for 5 zones. Perceived usability, measured by System Usability Scale, showed modest omnibus difference across cohorts (Kruskal-Wallis H_3=5.22, P=.16); however, a pairwise comparison showed the 5-gesture cohort rated usability lower than the progressive cohort (2-tailed Welch t_{11}=–2.19, P=.05). The 5-gesture cohort rated the system lowest (mean 63.3, SD 16.2). Cognitive workload, assessed through the National Aeronautics and Space Administration Task Load Index, increased with the number of gestures. The performance subscale showed a significant omnibus difference across cohorts (Kruskal-Wallis H_3=21.4, P<.001). Mean performance subscale scores were 84.4 (SD 14.7) for the single-gesture condition, 30.6 (SD 21.7) for the 5-gesture condition, and 44.2 (SD 21.2) for the progressive cohort, reflecting increasing perceived difficulty with more gestures. The sample size for quantitative analysis was 54.

XSL•FO
RenderX

Conclusions: These findings support the implementation of progressive training for 3 gestures. Usability perceptions were the highest among the more complicated progressive cohort, which is likely related to perceived improvement. Progressively learning 3 gestures enables a balance between device capability, user intention, perceived usability, and cognitive workload.

(JMIR Rehabil Assist Technol 2026;13:e88528) doi:10.2196/88528

KEYWORDS

prostheses; upper limb prosthesis; gamified training; usability; biomedical technology; artificial limbs; biomedical engineering; user centered design

Introduction

Upper limb prostheses must meet the functional demands of daily life [1,2]. Varied factors cited as contributors to prosthesis rejection demonstrate the need to further understand how to optimize training and device usage to meet these demands [3]. Electromyographic prostheses have fewer reported cases of device abandonment, attributed to increased capabilities [4]. Despite this, the lack of device functionality that contributes to prosthetic rejection rates [5] underscores the need to further expand device capabilities while maintaining user accuracy and perceived usability for novice users. The integration of advanced grips, finger configurations, and gesture recognition enables greater adaptability and reliability in a system [6]. Similarly, incorporating multiple gestures may improve prostheses' functional capabilities, but they must be feasible for users to adopt the control scheme in daily life.

Electromyographic prostheses are controlled by muscle activation, which triggers the device to perform a preprogrammed gesture. Noninvasive surface electromyography (sEMG) detects microvoltages from the skeletal muscle, which are detectable from the skin's surface [7,8]. To receive the electrical signals, sEMG can be embedded anywhere the device makes contact with the user's residual limb. Multigesture electromyographic prostheses similarly require muscle activation but can perform multiple distinct movements. There are a variety of ways to process data collected by sEMG sensors to trigger a specific motion. Control schemes can use machine learning models in conjunction with several placed sEMG sensors [9], pattern recognition sequences [10], or magnitude-based discretization using a single sensor [11]. While a single-gesture prosthesis typically performs an open-close grasp with each muscle contraction, a multigesture prosthesis offers a variety of hand motions, such as pinch-point gestures, task-specific grasps, or individualistic expressions. Multigesture prostheses can be approached from a direct, single-sensor control perspective or from a pattern recognition standpoint [10]. Pattern classifiers have been found to potentially identify a user's intention with up to 90% accuracy under ideal conditions, but present significant issues, including electrode shift, position variation, and force variation [11]. However, using a single electromyography (EMG) electrode with multigesture capabilities through force grading has been shown to be effective for pediatric populations [12].

This research aims to evaluate multigesture training with single sensor control. The research team has been developing a novel multigesture prosthesis that leverages a single sEMG sensor. Patients using the device can intentionally contract their muscles at different strength magnitudes to perform various gestures with the prosthetic limb. For example, a "light" contraction would prompt a different command than a "hard" contraction. Thresholds can be set to respond to a user's individual strength range through calibration. This EMG signal processing includes amplification, rectification, band-pass filtration, and smoothing [13]. This control scheme is thought to perform better for people with congenital limb differences, who often have limited muscle volume and discretization compared with acquired amputations later in life.

All recorded EMG voltages from the sensors are processed and reported as values from 0 to 1023, with higher values indicating greater muscle contraction intensity. The input is filtered and rectified using bandpass filtering to reduce the reference scale to a comfortable range of contraction. When reading and processing muscle contraction for multigesture control, a single "flex" is quantified as the highest recorded EMG value within a series of recorded values above a set resting baseline. Each user can adjust this resting baseline to account for natural variance in base noise sensor readings. Any values below this resting threshold are effectively ignored by the prosthesis, allowing for intentional contractions to be easily differentiated from resting noise. When the continuous EMG signal from the prosthesis reaches a value above the resting threshold, all values are recorded within the current "flex" until the signal dips below the resting threshold again. Once the EMG signal falls below the resting threshold, the highest recorded value for the previous contraction is identified as the peak. This peak value is then compared with the preset multigesture thresholds to determine which threshold was hit. Each gesture threshold has a minimum contraction value, ranging from 0 to 1024, that can be manually set by each user. A threshold's range spans from its minimum value to the next highest minimum value of the following threshold, or 1024 if there are no higher thresholds. When a discretized threshold is met, the corresponding hand gesture is triggered.

Implementation of multigesture EMG control using a single sensor is limited [11,14,15], but has leveraged pattern recognition and magnitude control. One example of this control implementation used to perform complex gestures is an assistive device that incorporates a hands-free wheelchair interface [14]. In this interface, engagement of the user's temporalis muscle controls a wheelchair's movement. Different levels of contraction correspond to wheelchair commands, including turning left, right, and moving forward similar to previous prosthesis controls [16]. This may be applicable in populations without the manual dexterity required to operate a traditional wheelchair joystick, such as those affected by a stroke or

XSL•FO
RenderX

amyotrophic lateral sclerosis. This interface includes 3 levels of discretization, and research demonstrated that users could successfully gain control of the system [17], with three user control methods that allow independent use of the wheelchair and increase human-machine interaction and user autonomy. This successful schema demonstrates the potential for using a single-channel EMG system to control multiple outcomes.

Gamified training applies game design elements to nongame situations [18] and offers an engaging option to familiarize a patient with a system or motion. For electromyographic prostheses, gamified training can help teach novice users how to engage the targeted muscle and learn novel controls. There is a need to explore the impacts of serious gaming further, as discrepancies in the literature suggest that evaluating training platforms should be done on a case-by-case basis [19]. In a pediatric population, compliance with a training regimen is a multifactorial struggle that often falls on the caregiver [20]. Gamified training can encourage consistent preparation for an electromyographic prosthetic arm. The purpose of such training games is to expose and strengthen the children's residual muscles before receiving their prostheses, which is intended to lower the time they need to become proficient with the device [21]. Game-based intervention shows an overall improvement in sEMG control and fine muscle activation [22].

The gamification in this study is built upon a preexisting gamified training strategy. The intention of the system is to prepare pediatric patients before using an EMG-powered prosthetic arm, with past research demonstrating gamified training leading to positive outcomes [23]. Training that starts at the time of initial device or before receiving a prosthesis may be beneficial [24,25]. Rehabilitation protocols are often progressive, requiring the establishment of baseline skills [26]. Some muscular rehabilitation settings have observed a benefit from using progressive training over static training [27]. In other settings, such as pain rehabilitation, no significant difference was found when comparing static and progressive training [28]. Various factors affect pediatric rehabilitation compliance, including engagement, motivation, and cognitive workload [29]. The cumulation of these factors has led to the need to examine progressive training in regard to pediatric electromyographic prosthesis use.

A gap in the current literature exists regarding the usability of electromyographic prosthesis systems, specifically those using a single-sensor system as opposed to the more common practice of including multiple sensors on the residual limb to gather EMG input [30]. The magnitude of contraction is less often the activation input for EMG devices; instead, multiple sensors working in conjunction using further signal analysis are more widespread. However, the incorporation of multiple sensors can add additional complexity to the system and make it more prone to error [15]. Although there is a range of electromyographic prostheses within the field of assistive technology, it has become apparent that standardization is necessary in the evaluation of EMG controls for single-sensor device models [31]. Additionally, opting to prioritize a single sensor may increase ease of learning in the pediatric population. While these multichannel systems using a field of muscle responses may be more beneficial toward amputees, who are more likely to have numerous feasible muscle groups for targeting with sEMG, this may differ in those with congenital limb differences. Amputees have been shown to learn new motor control schemes and muscle activation patterns via sEMG training [32].

As the etiology of congenital limb difference may vary, it is important to consider that the muscular architecture of each individual may not allow for multiple muscle targets. Methods of control for multigesture vary greatly, with each potentially benefiting from its own method of quantification. The use of a single-channel EMG system is uncommon, necessitating the development of a specific evaluation process for user acclimatization to the device. Additionally, an evaluation to determine the limit of the number of possible discretization thresholds should be developed. This testing software, in conjunction with the survey instruments, is intended to effectively measure the perceived difficulty level of the varying modes (1, 3, or 5 discretization zones within a single sensor) that are analogous to the multigesture electromyographic prosthetic device.

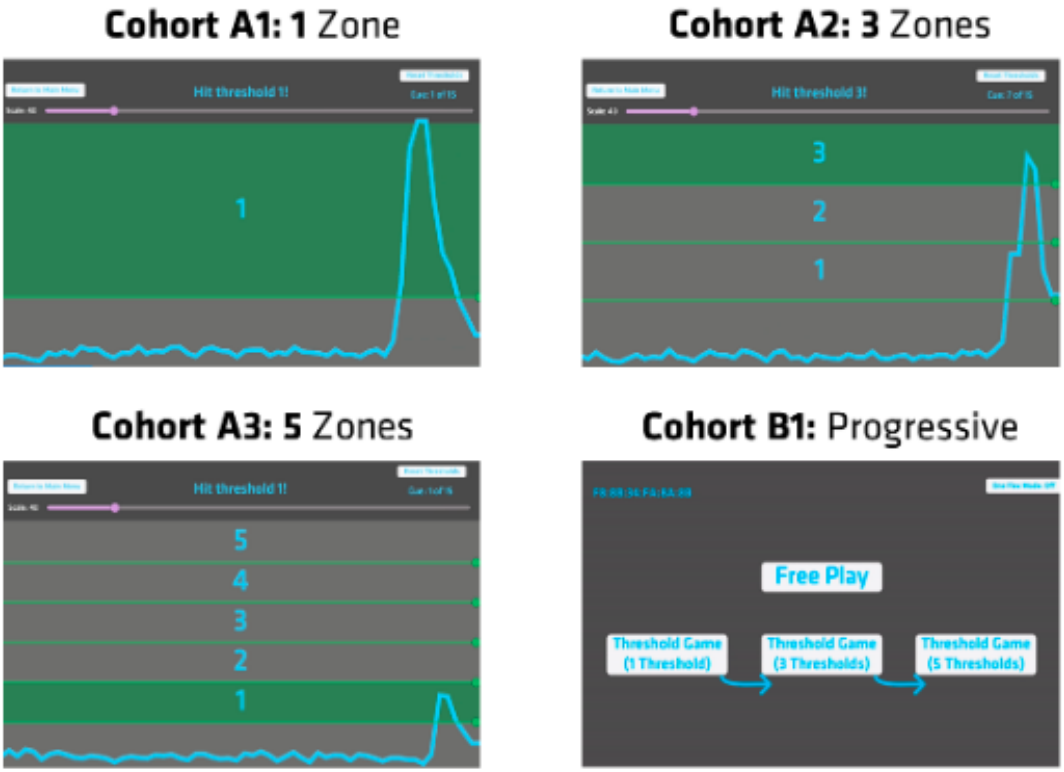
Methods

Study Design

To observe the results of novice users attempting to target various levels of contractions, a gamified training app, the Muscular Discretization Assessment Tool (MDAT), was developed in-house for user testing. The research team designed the MDAT to resemble the calibration system used by participants in a clinical trial testing the Limbitless Solutions prosthetic arm. Ensuring the effectiveness of the gamified MDAT controls before usage by prosthesis users was intended to optimize gameplay before distribution. When designing the gamified training app, complete parity with the functionality of the multigesture prosthesis was one of the main design aims. Each participant's muscular contractions were displayed in a graph showing real-time voltage changes, as illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Representation of the progression scheme of each of the four cohorts. Cohort A1 included 1 section for participants to flex within. Cohort A2 consisted of the screen divided into three sections for participants to flex within. Cohort A3 consisted of the screen divided into five sections for participants to flex within. Cohort B1 had participants complete the first round with the screen divided into 1 section, the second round with 3 sections, and the third round with 5 sections.



The app recorded and processed EMG input through the same mechanisms as the electromyographic prosthesis; once the user's EMG signal passed the base resting threshold, all EMG values were recorded and saved until the signal fell below the resting threshold again. The EMG processing provides filtered data that directly corresponds to a muscular contraction, with a sampling rate of 1000 Hz. These steps include amplification, rectification, band-pass filtration ranging from 2 to 1000 Hz, and smoothing [13]. The highest recorded value between these 2 points was then identified and was considered the peak value of this contraction. This peak was used to determine which multigesture zone the contraction fell within. The app also allowed for the manual adjustment of each threshold, as well as a global scaling value, which amplified the measured EMG signal to allow users with weaker readings to still hit all thresholds easily. The prosthesis similarly allows for these levels to be adjusted and calibrated for a personalized experience.

The EMG Flex Controller sends data to the multigesture app approximately 60 times per second. These data are validated and processed by the app every time a new visual frame is calculated, which occurs at a capped rate of 30 times per second on the tablets available for user gameplay. This translates to a rate of processing and data reading from the controller once every 33.3 milliseconds. The app reads in muscle strength from the user directly as captured by the EMG sensors on the Flex Controller, and no additional adjustments are made to the readings by the game. The app also reflects the threshold system for triggering different gestures on the prosthetic, which provides one-to-one parity. Before app implementation in the study, internal testing was conducted. Since the software of the app was built on the framework of the calibration app used for the existing EMG-controlled prosthetic, the data are as reliable and seamless as they would be in real-world applications. The accuracy of this software has been tested before this study and has been implemented in ongoing clinical trials.

The calibration period consisted of participants completing a maximum muscular contraction. Based on where this signal ended, the sensitivity was adjusted. For example, if the participant was not able to easily reach the top of the screen via their muscular contraction, the signal sensitivity would be increased. This scale effectively "zooms in" the reading captured by the controller to bring the maximum captured value closer to the user's readings without losing the accuracy and fidelity of data. This allowed for a standardized normalization process that was based on the user's maximum contraction. The zone areas are initially segmented equally, but are customized to each participant such that they can feasibly hit all thresholds, with their baseline signal remaining below the lowest threshold. For additional support, the participant was prompted to hit each of the zones before actual gameplay.

The MDAT had a freeplay mode and various settings to replicate the multigesture thresholds analogous to the prosthesis control. The participant was fitted with an in-house developed Flex Controller around the forearm that contained surface electrodes to detect microvoltages from muscle contractions. The freeplay mode was used before data collection to calibrate the user's Flex Controller to their personal muscle capabilities. The gameplay modes provided an environment for users to contract

XSL•FO
RenderX

at various thresholds, recreating the muscle engagement required to operate the corresponding multigesture prosthetic device. In these gameplay modes, the screen was divided into 1, 3, or 5 sections, as presented in Figure 1.

The user interface (UI) for the gamified MDAT was designed to communicate the EMG signal and thresholds to the user as clearly as possible. Other elements, such as score, game instructions, and options, were placed at the top of the screen space to maximize the area available for EMG-related elements. Clarity was provided through explicit instructions for which randomized muscle strength grading thresholds are hit, delivered in a similar fashion as a game of *Simon Says* or *Bop It* (Hasbro Inc). A limited yet bright color palette was chosen for the UI to ensure that important elements stand out to users. A thick, bright blue line represented the user's EMG signal from the Flex Controller, allowing users to understand how their muscle activity was translated in-game. Horizontal green lines across the screen indicated gesture thresholds. These were the only green elements in the UI and were intentionally chosen to make thresholds easy to identify at any moment during playtime. To further enhance threshold identification, each zone was visibly labeled as a numerical range. This aided users in identifying which threshold they were being prompted to hit during gameplay, as the game presented messages in the form of "Hit threshold X!." X, in this instance, would correspond to the assigned threshold number that was cued to be hit. The cued threshold zone was highlighted in green, intuitively demonstrating to users which threshold they must hit without having to read the instruction prompt.

The participant had 1 attempt to reach within the threshold, and there was no time limit on completing an attempt. An attempt was considered successful if the participant's peak contraction was within the cued highlighted threshold. Successful attempts were scored out of 15 cues per round to record overall performance.

A total of 54 participants completed data collection, with ages ranging from 18 to 30 years old. Anyone was invited to participate if they were between the ages of 18 and 64 years, able to consent, had full control of their upper extremities, had no injuries to their hands or arms that would impact their upper extremity control during data collection, and if they had never used a Limbitless Solutions myoelectric prosthesis or Flex Controller. Participants were assigned a cohort (A1, A2, A3, or B1) based on participation order. Cohort A1 included 1 section for the users to contract within, while sections A2 and A3 corresponded to 3 and 5 threshold sections, respectively. Cohort A1 had an analytical population of n=9. Cohort A2 had an analytical sample of n=10. Cohort A3 had an analytical sample of n=9. Cohort B1 had an analytical sample of n=26. Each of these cohorts completed 3 rounds with their respective number of zones. B1 was the progressive cohort, in which participants progressed through the 1, 3, and 5 zone modes in sequence. Users were first fitted with the Flex Controller on the right forearm. The controller was calibrated in freeplay mode based on the ease of comfortably reaching each threshold. This was guided by the researcher and manipulated by changing the sensitivity of the surface electrodes. Before gameplay, each user was informed of their respective cohort and shown each of the thresholded screens. For participants in the progressive cohort, researchers further explained the thresholds as fixed increments that changed progressively after each of the 15 cues was presented. Each participant played through the 15-cue round 3 times, taking 1-minute breaks in between each round to avoid fatigue. This procedure was then repeated with the Flex Controller on the left forearm.

After the testing period concluded, postassessment surveys were delivered to participants. These surveys included the Systems Usability Scale (SUS) and the National Aeronautics and Space Administration Task Load Index (NASA-TLX) [33,34]. The SUS is a 10-item scale commonly used as an objective measure of a digital interface's usability [33]. The NASA-TLX measures cognitive workload through 6 categories [34]. The Edinburgh Handedness Inventory (EHI) was also completed to determine the participants' dominant hand [35]. Finally, the user's grip strength was measured using a handheld dynamometer. Grip strength was measured on both the right and left hands, and each was tested twice. The highest score for each hand was recorded. Participants also reported confidence with video games, confidence with exercise, confidence with visual cues, and activities they have participated in during the current week.

Ethical Considerations

These procedures were approved by the University of Central Florida Institutional Review Board on October 1, 2024 (STUDY00007067). The inclusion criteria for participation consisted of people aged 18-64 years with full control of their upper extremities. Full informed consent was presented in both a visual and an auditory manner. Participants were encouraged to ask questions throughout the informed consent process and to request extra rest periods as needed. Participant confidentiality was preserved wherever possible by assigning a randomly generated 8-digit user ID based on participation order. These IDs were generated and stored on a premade linking spreadsheet. Once an ID was used, it was marked as "claimed" on the spreadsheet to ensure each was used only once. The Qualtrics forms containing all of the study data were linked to this ID, and no identifiable participant information was retained or used for analysis. The only potential participant compensation was extra credit from an instructor, which was offered at the discretion of the instructor. Each instructor who offered extra credit from this study was required to provide an alternative assignment of equal time and effort for the same amount of credit in the case that a participant could not complete the study.

Statistical Analysis

Analysis was conducted in Python (version 3.12.9) using the libraries *pandas*, *NumPy*, and *scipy.stats*. The analytic sample was 54 participants (N=54). Tests were 2-tailed with α=.05; results were not adjusted for multiple comparisons. Normality was assessed for within-subject differences for paired comparisons using the Shapiro-Wilk test. When a paired comparison showed evidence of nonnormal differences (Shapiro P<.05), a Wilcoxon signed-rank test was run as a nonparametric sensitivity analysis and both results were reported. Nonparametric checks were used for ordinal or multilevel factors. SDs use $t_{(n-1)}$=1 for effect size testing.

XSL•FO
RenderX

Score construction was in accordance with SUS standard guidelines (odd items: x–1; even items: 5–x; sum×2.5 to yield 0-100 scores) [33]. NASA-TLX raw inputs (1-21) were rescaled to 0-100 scores via (x–1)×5 [34]. The EHI laterality quotient was calculated via (R–L)/(R+L)×100, with R representing the responses associated with right-handedness, and L representing left-hand responses.

Within cohorts, trial scores were compared using paired, 2-sided *t* tests. The mean difference (95% CI), $t_{(df)}$, *P*, and paired effect size Cohen *d* with Hedges *g* small-sample correction were calculated.

To assess the difference between a progressive and static mode of learning, in-between cohort comparisons were made between cohort B1's trials and each of the respective A cohorts. For example, B1, B2, and B3 were compared with trial 1 of cohort A1, trial 2 of cohort A2, and trial 3 of cohort A3, respectively. This was done to make analogous comparisons based on the number of zones and trial numbers. Differing trials were isolated to account for users' exposure time to the game and control system. Comparisons between cohorts were done using Welch 2-sided *t* tests due to unequal variances. The mean difference (95% CI), Welch $t_{(df)}$ and *P*, and effect sizes (Cohen *d*, Hedges *g*, and average d_{av}), were determined using pooled Cohen *d*, Hedges *g*, and average *d* (d_{av}).

Associations between secondary variables (eg, confidence with video games, exercise, and cue use) and outcomes (trial scores, SUS, NASA-TLX, etc) were assessed using Spearman rank correlation computed within cohorts; ρ and *P* were reported. For ordinal factors with multiple levels, Kruskal-Wallis omnibus tests were conducted, with *H* and *P* reported where relevant. For multiselect variables (eg, videogame experience categories and activities), binary flags were created and compared groups via Welch *t* tests, reporting d_{av}. Odds ratios were not calculated in this study, as they were not appropriate considering the data. Gender comparisons used Mann-Whitney *U* tests with rank-biserial r as effect size, as normality could not be assumed in the smaller subsample of men (n=12).

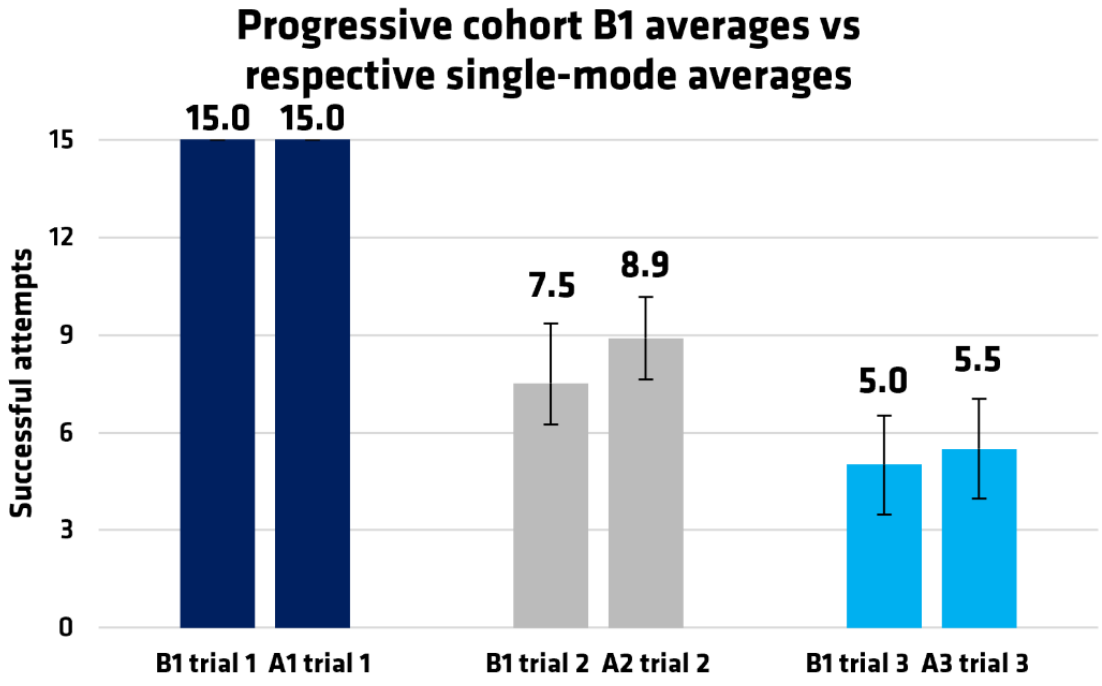
Results

Performance Scores

Overview

Comparisons were made between cohort B1's trial 1, trial 2, and trial 3, and the Al, A2, and A3 respective cohorts. These respective trials are matched for comparison based on the number of thresholds, as well as the time spent exposed to the game, which was measured through sequential trials. The outcomes demonstrate that user accuracy decreased as the number of zones increased. When playing with 1 threshold, as illustrated in Figure 2, as B1 trial 1 and A1 trial 1, the average score was 15 (SD 0) successful attempts out of 15 possible points per round. Scores from playing with 3 thresholds are displayed in the middle column, labeled B1 trial 2 and A2 trial 2. Scores from playing with 5 thresholds are displayed on the right side of the graph, labeled as B1 trial 3 and A3 trial 3. This trend is displayed in Figure 2.

Figure 2. A comparison was made between the nonprogressive and progressive cohorts to determine performance differences. These performances are scored out of 15 potential successful attempts and visualized to include pairwise comparisons between Cohort B1's Trial 1, Trial 2, and Trial 3, and the Al, A2, and A3 respective trials. Cohort A1 averaged a score of 15, A2 averaged 9.1, and A3 averaged 5.5. The progressive B1 cohort also followed most of these averages with average scores of 15, 7.5, and 5.0, respectively. Error bars correspond to plus/minus one standard deviation. A general trend observed was that performance decreased as the number of zones increased in both cohorts; however, there were no statistically significant differences between the cohorts.



XSL•FO
RenderX

Cohort A1 averaged a score of 15 (SD 0.0), A2 averaged 9.1 (SD 1.1), and A3 averaged 5.5 (SD 1.1). The progressive B1 cohort also followed most of these averages with average scores of 15 (SD 0.1), 7.5 (SD 1.8), and 5.0 (SD 1.5) for each of the 3 trials, respectively. SD was included in parentheses, as seen in Table 1. Table 2 compares the B1 progressive gameplay mode to each of the respective A cohorts, paired by the number of zones. Cohort A1 averaged a score of 15 (SD 0.0), A2 averaged 8.9 (SD 1.5), and A3 averaged 5.5 (SD 1.1); the progressive B1 cohort also followed most of these averages. Overall, no statistically significant differences were found between the B1 cohorts and their corresponding A cohorts, indicating that the groups are comparable.

Table 1. Performance scoring by successful attempts.

Cohort	Trial 1, mean (SD)	Trial 2, mean (SD)	Trial 3, mean (SD)	Average score, mean (SD)
A1	15.0 (0.0)	15.0 (0.0)	15.0 (0.0)	15.0 (0.0)
A2	8.9 (1.5)	8.9 (1.3)	9.6 (1.8)	9.1 (1.1)
A3	5.6 (1.9)	5.3 (0.7)	5.5 (1.5)	5.5 (1.1)
B1	15 (0.1)	7.5 (1.8)	5.0 (1.5)	9.2 .9)

Table 2. Between-cohort comparisons.

Cohort and trial	Mean difference (95% CI)	Welch *t* (*df*)	Hedges *g*	Cohen $d_{(av)}$	*P* value
B1, A1, Trial 1	–0.04 (–0.09 to 0.02)	–1.44 (25.0)	–0.32	–0.40	.16
B1, A2, Trial 2	–1.38 (–2.49 to –0.27)	–2.57 (23.8)	–0.79	–0.88	.02
B1, A3, Trial 3	–0.48 (–1.75 to 0.79)	–0.81 (13.7)	–0.31	–0.32	.43

Performance Scoring by Successful Attempts

Performance was determined by the number of correct randomized cues a participant hits with a peak contraction out of 15 total cues. Performance scores were compared across cohorts and separated by trial, with SDs included (Table 1). Cohort A1 averaged a score of 15 (SD 0.0), A2 averaged 9.1 (SD 1.1), and A3 averaged 5.5 (SD 1.1). The progressive B1 cohort also followed most of these averages, with an overall average of 9.2 (SD 0.9).

Within-cohort paired comparisons across trials did not show statistically significant changes in performance. In cohort A2, trial 1 vs trial 3 differences deviated from normality (Shapiro *P*=.01); therefore, a Wilcoxon signed-rank test was conducted as a sensitivity analysis. Results were consistent with the paired *t* test (t_9=−1.07, *P*=.32; Wilcoxon W=13.5, *P*=.59).

Comparison Between Cohorts

Between-cohort comparisons were also assessed to determine statistical significance in user performance. Cohort A1 averaged a score of 15 (SD 0.0), A2 averaged 8.9 (SD 1.5), and A3 averaged 5.5 (SD 1.1). The progressive B1 cohort also followed most of these averages. There were no statistically significant differences between the B1 and the respective counterparts for A1 (*P*=.16) and A3 (*P*=.43). Cohort A2 performed significantly higher in Trial 2 than Cohort B1 (*P*=.02, ρ=0.8). A negative mean difference favored the A Cohorts. Hedges *g* was also listed in the table (Table 2). The average (SD) standardized mean difference was $d_{(av)}$.

SUS

Cohorts A1, A2, and B1 had the highest perceived usability as measured through the SUS, as seen in Table 3 [33]. The SUS is measured using a 100-point scale, with 100 portraying the best imaginable product, and 0 indicating an unusable product [33]. Cohorts A1 and A2 both had an average SUS score of 75 (cohort A1: mean 75.0, SD 18.9 and cohort A2: mean 75.0, SD 16.8). Cohort A3 had the lowest SUS score with an average of 63.3 (SD 16.2). Cohort B1 had the highest average SUS score with 76.2 (SD 11.7). The overall group SUS score was 73.6. A Welch *t* test showed a significant difference between cohorts A3 and B1 (*P*=.05).

Table 3. Overall System Usability Scale scores by cohort.

Cohort	Overall SUS[a], mean (SD)
A1	75.0 (18.9)
A2	75.0 (16.8)
A3	63.3 (16.2)
B1	76.2 (11.7)
Overall	73.6 (15.1)

[a]SUS: System Usability Scores.

XSL•FO
RenderX

SUS is a measure of perceived usability and was taken as a postassessment survey [33]. Average overall SUS scores and the respective SDs are separated by cohort. Cohorts A1, A2, A3, and B1 had overall average scores of 75.0 (SD 18.9), 75.0 (SD 16.8), 63.3 (SD 16.2), and 76.2 (SD 11.7), respectively. The overall average score between all the cohorts was 73.6 (SD 15.1), which falls in the "Good" category on the corresponding subjective scoring of the SUS [33].

NASA-TLX and Subscales

Inversely, NASA-TLX scores are on a 100-point scale, with 0 indicating the desired lowest cognitive workload, and 100 indicating the highest cognitive workload [34]. The NASA-TLX consists of 6 different subscales that are the main contributors to cognitive workload—mental demands, physical demands, temporal demands, frustrations, effort, and performance [34]. These subscales were selected due to underlying theories of workload that correspond to demands that a user perceives during a task [34]. These subscales are depicted visually per cohort in Figure 3. The performance subscale has an inverse relationship, where higher self-reported performance scores indicate a lower cognitive workload and are considered more favorable [34]. Average NASA-TLX scores were highest among cohort A3 in 5 out of 6 subcategories. From a Kruskal-Wallis test, each subscale had a *P* value. Only the performance subscale passed the significance threshold. These results are presented in Table 4. This follows the expected trend of cognitive workload increasing as the number of zones expands. Differences in perceived effort and perceived frustration were not statistically significant between cohorts. To clarify how these workload patterns varied between specific cohorts, pairwise comparisons of performance and temporal demand subscales are presented in Tables 5 and 6. The largest and most significant differences emerged in the performance subscale, where the single-gesture cohort A1 outperformed all other cohorts with large mean differences. The temporal demand subscale had significant comparisons involving cohort B1, with higher perceived time pressure than cohort A1 and cohort A3. Within the mental demand subscale, cohort A1 tested significantly lower than cohort B1 (P=.01, $d_{(av)}$ =–0.80). In the physical demand subscale, no tests achieved statistical significance, but cohort A2 had a lower score than A3 (P=.15, $d_{(av)}$=0.70).

Figure 3. The National Aeronautics and Space Administration Task Load Index (NASA-TLX) measures perceived cognitive workload [34]. It is divided into 6 subscales, each on a 100 point scale where 0 is the desired low cognitive workload [34]. Performance is on an inverse scale, where 100 would indicate a more desirable performance [34]. This subscale is depicted on the right in its own graph to account for this difference. The trends overall confirmed the hypothesis that demand would increase, while performance would decrease, due to increased zones of control. The progressive training Cohort B1 countered the expectations, having less perceived difficulty than Cohort A3. Note: *:P<.05; **:P<.01; and ***:P<.001.

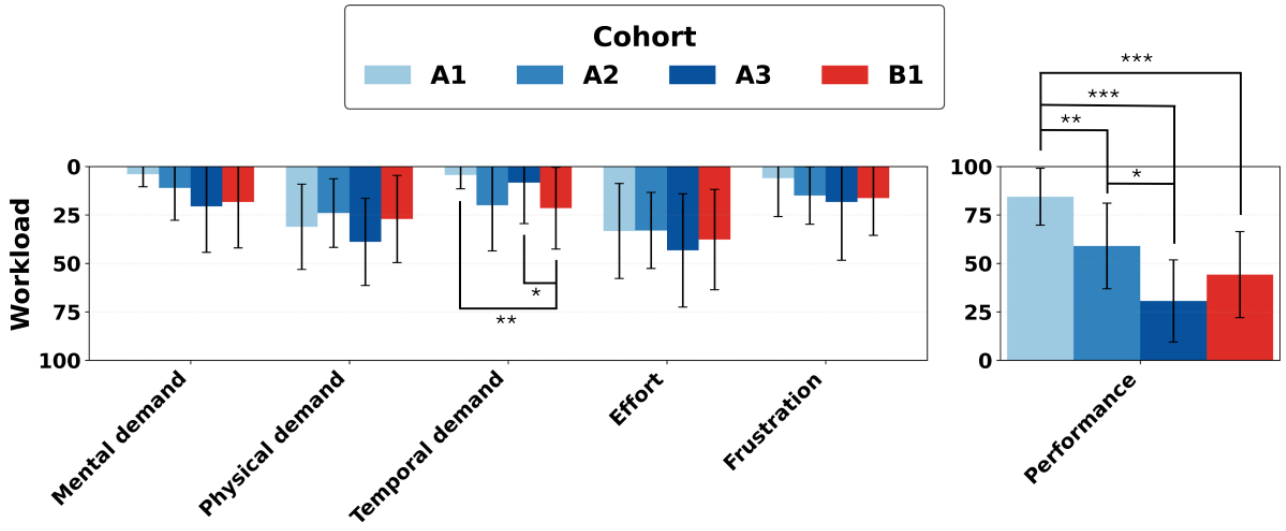


Table 4. Average NASA-Task Load Index subscores (separated by cohorts, with corresponding SDs included for each of the scores).

Cohort[a]	Mental demand	Physical demand	Temporal demand	Performance	Effort perceived	Frustration perceived	Overall
A1, mean (SD)	3.9 (6.6)	31.1 (22.0)	4.4 (7.1)	84.4 (14.7)	33.3 (24.5)	6.1 (19.7)	16.4 (10.9)
A2, mean (SD)	11.0 (16.7)	24.0 (17.7)	20.0 (23.6)	59.0 (22.1)	33.0 (19.6)	15.0 (14.8)	24.5 (14.1)
A3, mean (SD)	20.6 (29.7)	38.9 (22.5)	8.3 (8.9)	30.6 (21.2)	43.3 (29.2)	18.3 (30.1)	33.3 (19.6)
B1, mean (SD)	18.3 (23.7)	27.1 (22.5)	21.5 (21.1)	44.2 (21.2)	37.7 (25.9)	16.3 (19.2)	29.7 (14.5)
P value	.49	.38	.09	<.001	.86	.29	.12

[a]Scores for each subsection are on a 100-point scale, with 0 indicating a desired low cognitive workload [34]. The performance subscale is the only exception to this, where 100 would indicate a more desirable score [34]. The NASA-Task Load Index measures perceived cognitive workload through a postassessment survey [34]. A general trend of cognitive workload increasing with a higher number of zones was observed.

XSL•FO
RenderX

Table 5. Pairwise comparisons for the NASA-TLX performance subscale. This subscale is scored inversely, with higher values indicating better perceived performance and lower cognitive workload [34]. This subscale had the largest and most consistent differences between cohorts. Cohort A1 rated its performance significantly higher than Cohorts A2, A3, and B1 with large standardized mean differences. Rows are organized with the most significant results at the top.

Comparison	Cohen $d_{(av)}$	Cohen $d_{(av)}$ (SE)	*P* value
A1, A3	2.91	0.70	<.001
A1, B1	2.20	0.47	<.001
A1, A2	1.36	0.52	.009
A2, A3	1.23	0.51	.01
A2, B1	0.68	0.38	.09
A3, B1	–0.64	0.39	.12

Table 6. Pairwise comparisons for the NASA-TLX temporal demand subscale [34]. Temporal demand showed fewer significant results than the performance subscale, but still reflected meaningful differences between cohorts. Cohort B1 reported significantly higher temporal demand than cohort A1 and cohort A3. Other pairwise contrasts within this subscale did not reach significance, although cohort A1 and cohort A2 showed a medium effect size trend. Rows are organized with the most significant results at the top.

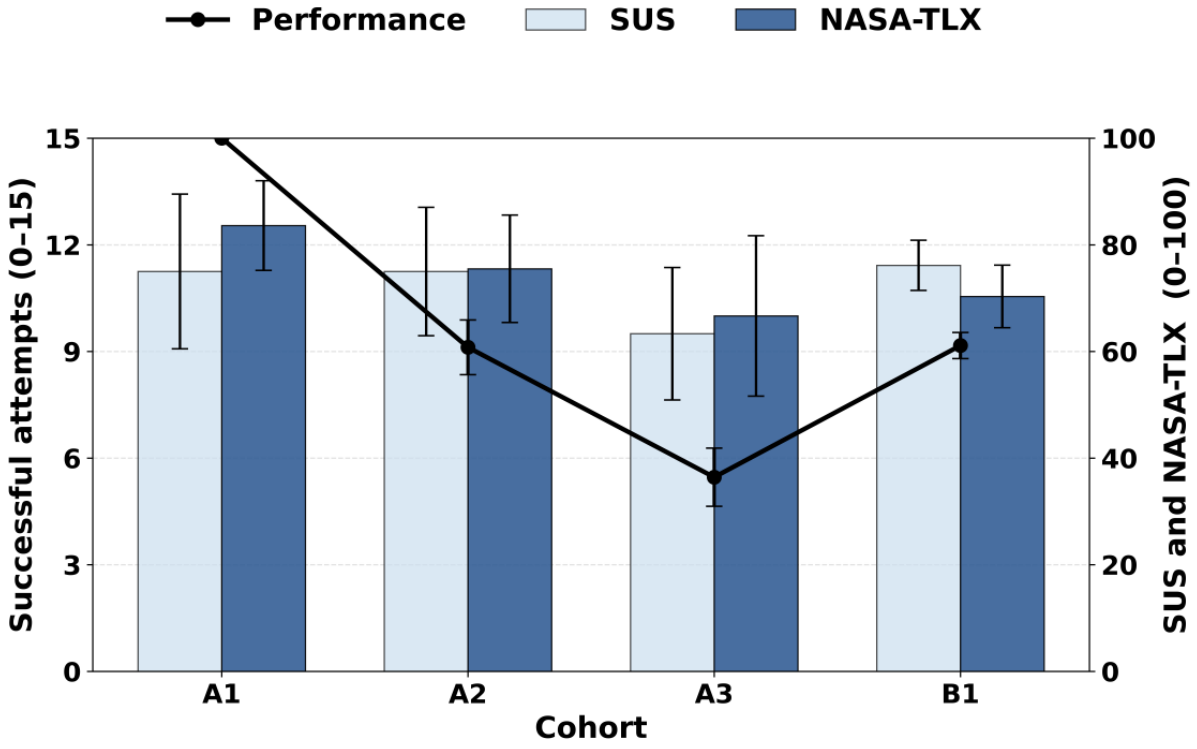
Comparison	Cohen $d_{(av)}$	Cohen $d_{(av)}$ (SE)	*P* value
A1, B1	–1.06	0.41	.001
A3, B1	–0.79	0.40	.02
A1, A2	–0.89	0.48	.07
A2, A3	0.65	0.47	.17
A1, A3	–0.48	0.48	.32
A2, B1	–0.05	0.37	.89

Performance and Composite SUS and NASA-TLX Scores Compared Between Cohorts

Successful attempts at scoring, SUS scores, and NASA-TLX scores are compared across cohorts in Figure 4. This allows for the visualization of trends and tradeoffs between the number of zones and accuracy, usability, and cognitive workload.

XSL•FO
RenderX

Figure 4. NASA-TLX (National Aeronautics and Space Administration Task Load Index) and SUS (System Usability Scale) composite scores are displayed in a combined bar chart, and successful attempts are overlaid as a line graph for Cohorts A1, A2, A3, and B1 [33-34]. This allows for the visualization of trends and tradeoffs between number of zones and cognitive load, usability, and accuracy. The number of successful attempts decreased as the number of zones increased from Cohort A1 to A3. No meaningful conclusions could be drawn from composite NASA-TLX scores between cohorts. The SUS composite scores did not differ significantly between Cohorts A1, A2, and A3, but a significant difference was found between Cohorts A3 and B1. NASA-TLX composite scores are inverted so that a higher score is favorable for the sake of graph continuity and tradeoff visualization.



Dominant Hand Grip Strength

Grip strength was measured in kilograms using a dynamometer after the completion of the postassessment surveys. The average grip strength in kilograms is segmented by gender. The frequency of each reported gender and the SDs are included in Table 7.

Table 7. Dominant hand grip strength by self-reported gender.

Gender	Frequency, n	Grip strength (kg), mean (SD)
Men	12	44.0 (11.2)
Women	40	27.3 (4.6)
Other	2	21.3 (0.6)

Secondary Assessments

Secondary measurements included the EHI and participants' confidence with video games, exercise, and following visual cues. Participants were also asked to share activities they completed within the last 7 days that required the use of fine or gross motor skills. Based on EHI calculations, all participants tested into pure right-handedness, with a laterality total of >40. Across correlation tests per cohort, most associations were small and nonsignificant (median |ρ|=0.12, IQR 0.05-0.34). Only 2 showed significance after inspection, that is, exercise frequency vs grip strength (ρ=0.53-0.75, *P*=.03-.002) and cue-use confidence vs workload (ρ=0.45, *P*=.01).

Preassessment surveying also included demographics, video game experience, exercise, crafts, instruments, and other fine-motor tasks. Handedness inventory, self-reported dominant hand, and grip strength were assessed. There was no correlation between gender and performance (ρ=–0.11, *P*=.43), despite large grip strength differences between male and female participants (ρ=–0.57, *P*<.001).

Discussion

Principal Findings

This study aimed to evaluate potential trade-offs associated with increasing gesture capability in an electromyographic prosthetic arm actuated by a single sensor. Increased training time, increased perception of cognitive workload, decreased accuracy, and decreased usability may all be factors discouraging the implementation of multiple gestures. This study explored 1, 3, and 5 distinct gestures triggered by magnitude control and their training feasibility in a single-session rapid learning environment. In addition, differences between static and progressive learning modes were also evaluated. The progressive

XSL•FO
RenderX

training mode (cohort B1) demonstrated the feasibility of reducing demand on the user.

Extrapolating from past work and the results of this study, a long-term training regime using a gamified interface that closely corresponds to the prosthetic device's interface may allow users to improve multigesture use. Developing control of electromyographic inputs in a low-stress, low-consequence manner is critical to unlocking the prosthesis's advanced functionality without patient rejection. Devices with more features and more advanced controls may result in frustration for patients and lead to rejection. Providing a "challenge mode" to test controls that can unlock additional device functionality may be a solution to this issue. The results of this work indicate that a 3-tiered gesture control scheme is capable with a limited training timeline. However, it is important to note that able-bodied users may experience different fatigue patterns or motivation levels compared with individuals with limb differences. In the prosthesis clinical trial protocol, the training games are delivered before the prosthetic arm in an effort to train basic muscle control and discretization. Once the prosthesis is delivered, it has a single threshold control for open and close functionality. Gamified training continues until a 3-threshold differentiation is easily achievable by the user, and only then is multigesture support unlocked in the prosthetic device. The goal is to have the majority of failures occur in the training games and increase initial success with the prosthetic device. The outcomes of this work show that a tiered approach is appropriate for training 3 thresholds. Future work will help determine if long-term training can support a 5-threshold model.

Cohort Performance Scoring

To determine if including multiple gestures is feasible, it is necessary to consider performance, usability, and cognitive workload. The 3-gesture mode demonstrated similar usability scores as the single-gesture cohort, with better cognitive workload and performance scores than the 5-gesture mode. This suggests that incorporating 3 gestures may strike a functional balance between difficulty and usability for first-time interactions with the system. Additionally, a progressive mode of learning demonstrated the highest perceived usability, despite incorporating up to 5 gestures, which had the lowest scores and the highest cognitive workload in the static mode of learning. Capping initial progressive training at 3 gestures has the potential to increase perceived usability further while retaining the higher accuracy and lower cognitive workload of cohort A2.

Previous studies have shown that increased time on task is associated with higher user performance outcomes with the electromyographic controls [23,36]. Additionally, close correspondence of a training period with posttest tasks resulted in greater performance scores, demonstrating the importance of task-specific training.

SUS

The overall SUS scoring of 73.6 falls into the "ok" designation, demonstrating that this system was likely deemed usable by the participants [33]. Individual cohort scoring indicated that cohort A3 had the lowest perceived usability. Since this cohort had also reported the greatest perceived mental workload and the lowest performance among participants, this was expected to have the lowest usability scoring. Cohort B1 has the highest perceived usability, which is supported by lower NASA-TLX scores and similar performance scores to the nonprogressive cohorts [34]. The 12.9-point SUS difference between cohort A3 and cohort B1 is both statistically significant (P=.05) and meaningful with a Cohen d of −0.91. This indicates a substantial difference between cohort averages relative to within-group variability. Clinically, a significant difference with a large effect size implies that users in a progressive learning environment experience are likely related to higher levels of usability than users in a static 5-zone learning environment.

NASA-TLX

The NASA-TLX, measuring perceived cognitive workload, demonstrated that cognitive workload increased as the number of gestures increased [34]. Similar to the trend in scoring, difficulty tended to increase with the inclusion of more gestures, so it was expected that cognitive workload would follow the same pattern. NASA-TLX scores were aligned with scoring deficits instead of perceived usability as measured through SUS [33,34]. Cohort A3 demonstrated the highest cognitive workload in 5 of the 6 NASA-TLX subcategories [34].

Limitations

One limitation of this study was the unintentional inclusion of exclusively right-handed participants. Although a small proportion of participants self-reported as left-handed (n=6), the EHI score indicated right-hand bias [35]. This may not be representative of true population proportions, however, and instead may be due to the relatively small sample size of this study.

Similarly, 40 of the 54 participants in the study were female. This can be explained by smaller sample sizes and convenient population sampling of college students. Throughout the data collection process, minor collection discrepancies were accounted for. One participant withdrew from the study due to a personal situation during data collection. The participant revoked their consent to participate, and the research session was terminated immediately. Their data were removed from the dataset. During another session, one user ID was claimed by 2 distinct participants during the data collection process. Since both participants completed the study independently of each other, the datasets were differentiated with an (a) and (b) at the end of their identifier, and both datasets were used in the final analysis.

The aim of this study was to explore trends in novice users as they were exposed to different controls. This study used a proxy population of college-aged students who had full control of their upper extremities, which is not fully representative of our target population of prosthesis users. Although these findings cannot be completely generalized to prosthesis users, the intention of this initial testing is to ensure positive outcomes among users before implementation in limb-different population groups. Additionally, the target population is primarily pediatric patients, which is seen as a vulnerable population group for testing purposes. In order to limit confounding variables, such as

XSL•FO
RenderX

different levels of limb difference and the effect of differing anatomical muscle structure, a proxy, an able-bodied population, was then used. Past work in prosthesis development has shown a significant difference in muscular activation between amputees and nonamputees [32]. As such, the use of a proxy population then allows for standardization that may prove more difficult in the intended population, as well as identification of potential errors, streamlining eventual delivery. Serious games have also been tested in a proxy population of college-aged students and have been further implemented with patients with a limb difference [23,36]. The observed performance measures, including accuracy scoring, SUS, and NASA-TLX, have been used independently among studies that include different limb users [33,34]. Additionally, future work may strengthen findings through the use of a pre-post comparison. Past work in the facility has evaluated the efficacy of motor learning through serious games and evaluated participant scoring before and after intervention [23]. Future work may evaluate long-term changes in scoring using the training app. This study demonstrated the efficacy of the MDAT within a proxy population, with the intention that future work will include testing in the intended clinical population of upper limb prosthesis users.

The unequal and relatively small sample sizes in cohorts A1, A2, and A3 limit statistical power, particularly for between-cohort comparisons with the progressive cohort B1. Although effect sizes were calculated and in several cases were small to moderate, the absence of statistically significant differences between cohorts may reflect limited power rather than true equivalence. Therefore, findings comparing progressive and static modes should be interpreted cautiously and considered hypothesis-generating. Further testing with the application will be ongoing with the current proxy population and a representative limb-different population.

Another limitation of this study is the single-gesture scoring metric for cohort A1. Cohort A1 had a mean score of 15.0 (SD 0.0). This is because all contractions that are above threshold are considered a success, as there is only 1 zone. This creates a ceiling effect that may have affected variance and statistical analysis. This also highlights the functionality of single-gesture mode testing for novice users. There is a possibility of a fatigue effect due to 3 repetitive trials, which the researchers aimed to minimize by implementing mandatory 1-minute rest periods per round and offering continued rest if the participant wanted more recovery time.

Conclusions

The most clinically relevant and statistically significant results were found from the SUS scoring [33]. Specifically, the significant increase in perceived usability for the progressive cohort seen in this study may influence user adherence to training, addressing one of the primary concerns surrounding prosthetic device rejection. There were no statistically significant differences between cohorts in scoring for progressive mode and static mode, and there were only 2 statistically significant subscales of the NASA-TLX [34]. Both the SUS scores and the NASA-TLX scores displayed a trend that difficulty increases with the inclusion of more gestures as anticipated [33,34]. This is supported by the results from cohort A3, which had the lowest accuracy scoring and the highest cognitive workload.

Especially among young prosthesis users, maintaining adherence to a training regime is particularly difficult [20]. The usability of training systems is necessary to encourage motivation in pediatric training regimes [37]. Overall, these results suggest that implementing up to 3 gestures with a progressive training mode may have the potential to balance the increasing complexities required for additional gestures. This work aims to increase the functionality while not introducing an undue burden. Since reducing the complexity of a training regimen may discourage device rejection, this is an essential consideration. This may enable expanding device capabilities while maintaining accuracy, perceived usability, and a lower cognitive workload.

This control schema will be implemented in a gamified training format to teach users the musculature engagement and control required to operate an upper limb prosthesis with multiple gesture states. This progression will be introduced to the intended population through long-term clinical trials of individuals with a limb difference who are training to use a prosthetic device.

Implementing these training controls into a gamified regimen may increase time on task [23]. Prolonging the time engaged in a training measure can increase its effectiveness; this engagement may decrease the time needed to learn controls, as reinforcement learning can begin with the first introduction to controls. Reinforcement learning in gamified training may significantly improve performance scores [38-41]. Applying this principle to future training applications has the potential to encourage pediatric compliance with a training program, as time on task and reinforcement learning lead to improved performance and engagement.

Future work aims to evaluate these controls in a longer-term research trial incorporating a 2-gesture step toward training up to 3 gestures progressively. This can be implemented into the pre-existing gamified training protocol used in current pediatric clinical trials. Increasing the in-game difficulty may prepare users for the increased difficulty associated with additional gestures on the prosthetic device. One eventual aim of future work is to discern what form of gamified training would be most effective to establish device use on the first day of delivery.

Acknowledgments

The manuscript benefited from minor edits for grammar and syntax via writing software that may leverage artificial intelligence (Grammarly). The authors are accountable for the content of the manuscript.

XSL•FO
RenderX

Data Availability

The datasets analyzed during this study are available from the corresponding author upon reasonable request.

Funding

The authors would like to thank the many members and alumni of the research team who have supported the project throughout its development. The Limbitless Solutions program has received financial and in-kind support from the 3D printer manufacturer Stratasys LTD and 3D design software manufacturer Autodesk. The authors want to disclose that the funders had no role in the design of the overarching research study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. The researchers appreciate the philanthropic support for research from the Paul B. Hunter & Constance D. Hunter Charitable Foundation, The Don and Lorraine Freeberg Foundation, the Albert E. & Birdie W. Einstein Foundation, the Unity for Humanities Grant, and the Disney Grant.

Authors' Contributions

Formal analysis: AL, SM, EB, MP, SB
Visualization: AL, SM, EB, SB, VR
Software: J Stevenson, MD, J Sparkman
Conceptualization, PS, J Sparkman, AM
Methodology: PS, J Sparkman, AM
Funding acquisition: PS, MD, AM
Resources: MD, J Sparkman
Supervision: AM
Investigation: AL, SM, EB, MP, DG, SB
Writing – original draft: AL, SM, J Stevenson, EB, MP, DG, SB. VR, PS, MD, AM, J Sparkman
Writing – review & editing: AL, SM, J Stevenson, EB, MP, DG, SB. VR, PS, MD, AM, J Sparkman

Conflicts of Interest

None declared.

References

1. Yamamoto M, Chung KC, Sterbenz J, Shauver MJ, Tanaka H, Nakamura T, et al. Cross-sectional international multicenter study on quality of life and reasons for abandonment of upper limb prostheses. Plastic and Reconstructive Surgery Global Open 2019;7(5):e2205. [doi: 10.1097/gox.0000000000002205]
2. Benz HL, Yao J, Rose L, Olgac O, Kreutz K, Saha A, et al. Upper extremity prosthesis user perspectives on unmet needs and innovative technology. 2016 Presented at: In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 16 Aug 2016; Florida USA p. 287-290. [doi: 10.1109/embc.2016.7590696]
3. Biddiss EA, Chau TT. Upper limb prosthesis use and abandonment: A survey of the last 25 years. Prosthetics and Orthotics International 2007;31(3):236-257. [doi: 10.1080/03093640600994581] [Medline: 17979010]
4. Park J, Zahabi M. Cognitive workload assessment of prosthetic devices: a review of literature and meta-analysis. IEEE Transactions on Human-Machine Systems 2022;52(2):181-195. [doi: 10.1109/thms.2022.3143998]
5. Smail LC, Neal C, Wilkins C, Packham TL. Comfort and function remain key factors in upper limb prosthetic abandonment: findings of a scoping review. Disability and Rehabilitation: Assistive Technology 2021 Jan;16(8):821-830. [doi: 10.1080/17483107.2020.1738567] [Medline: 32189537]
6. Shinde A, Shete V. Advancements in myoelectric robotic hands and prosthetic limbs. 2025 Presented at: International Conference in Advances in Power, Signal, and Information Technology (APSIT); 2025 May 23; Bhubaneswar, India p. 1-7. [doi: 10.1109/apsit63993.2025.11086319]
7. William Marras. Selected Topics in Surface Electromyography for Use in The Occupational Setting: Expert Perspectives. Morgantown, West Virginia: U.S. Department of Health and Human Services, National Institute for Occupational Safety and Health; Mar 01, 1992:13-15.
8. Farina D, Jiang N, Rehbaum H, Holobar A, Graimann B, Dietl H, et al. The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2014;22(4):797-809. [doi: 10.1109/tnsre.2014.2305111]
9. Li W, Shi P, Yu H. Gesture recognition using surface electromyography and deep learning for prostheses hand: state-of-the-art, challenges, and future. Frontiers in Neuroscience 2021;15:621885 [FREE Full text] [doi: 10.3389/fnins.2021.621885] [Medline: 33981195]
10. Resnik L, Huang H, Winslow A, Crouch DL, Zhang F, Wolk N. Evaluation of EMG pattern recognition for upper limb prosthesis control: A case study in comparison with direct myoelectric control. Journal of Neuroengineering and Rehabilitation 2018;15(1):23 [FREE Full text] [doi: 10.1186/s12984-018-0361-3] [Medline: 29544501]

XSL•FO
RenderX

11. Manero A, Smith P, Sparkman J, Dombrowski M, Courbin D, Barclay P, et al. Utilizing additive manufacturing and gamified virtual simulation in the design of neuroprosthetics to improve pediatric outcomes. MRS Communications 2019;9(3):941-947. [doi: 10.1557/mrc.2019.99]
12. Scheme E, Englehart K. Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. Journal of Rehabilitation Research & Development 2011;48(6):643-659. [doi: 10.1682/jrrd.2010.09.0177] [Medline: 21938652]
13. Smith P, Dombrowski M, Rivera V, Pradeep M, Gunnell D, Sparkman J, et al. Analysis of a novel training game with eye tracking and electromyography for autonomous wheelchair control. Applied Sciences 2025;15(10):5268. [doi: 10.3390/app15105268]
14. Manero AC, McLinden SL, Sparkman J, Oskarsson B. Evaluating surface EMG control of motorized wheelchairs for amyotrophic lateral sclerosis patients. Journal of NeuroEngineering and Rehabilitation 2022;19(1):88 [FREE Full text] [doi: 10.1186/s12984-022-01066-8] [Medline: 35965311]
15. Tavakoli M, Benussi C, Lourenco JL. Single channel surface EMG control of advanced prosthetic hands: a simple, low cost and efficient approach. Expert Systems with Applications 2017;79:322-332. [doi: 10.1016/j.eswa.2017.03.012]
16. Manero A, Sparkman J, Dombrowski M, Buyssens R, Smith PA. Developing and training multi-gestural prosthetic arms. In: Virtual, Augmented and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation. Cham, Switzerland: Springer International Publishing; 2018:427-437.
17. Smith P, Dombrowski M, McLinden S, MacDonald C, Lynn D, Sparkman J, et al. Advancing dignity for adaptive wheelchair users via a hybrid eye tracking and electromyography training game. In 2022 Symposium on Eye Tracking Research and Applications 2022:1-7. [doi: 10.1145/3517031.3529612]
18. Khoshnoodifar M, Ashouri A, Taheri M. Effectiveness of gamification in enhancing learning and attitudes: a study of statistics education for Health School Students. Journal of Advances in Medical Education & Professionalism 2023;11(4):230-239 [FREE Full text] [doi: 10.30476/JAMP.2023.98953.1817] [Medline: 37901759]
19. Almeida J, Nunes F. The practical work of ensuring effective use of serious games in a rehabilitation clinic: a qualitative study. JMIR Rehabilitation and Assistive Technologies 2020;7(1):e15428 [FREE Full text] [doi: 10.2196/15428] [Medline: 32130177]
20. Santer M, Ring N, Yardley L, Geraghty AWA, Wyke S. Treatment non-adherence in pediatric long-term medical conditions: systematic review and synthesis of qualitative studies of caregivers' views. BMC Pediatrics 2014;14:63 [FREE Full text] [doi: 10.1186/1471-2431-14-63] [Medline: 24593304]
21. Dombrowski M, Smith PA, Buyssens R. Utilizing digital game environments for training prosthetic use. In: Virtual, Augmented and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation. Cham: Springer; 2016:481-489.
22. Prahm C, Kayali F, Sturma A, Aszmann O. PlayBionic: Game‐based interventions to encourage patient engagement and performance in prosthetic motor rehabilitation. PM&R 2018;10(11):1252-1260. [doi: 10.1016/j.pmrj.2018.09.027]
23. McLinden S, Smith P, Dombrowski M, MacDonald C, Lynn D, Tran K, et al. Utilizing electromyographic video games controllers to improve outcomes for prosthesis users. Appl Psychophysiol Biofeedback 2024;49(1):63-69 [FREE Full text] [doi: 10.1007/s10484-023-09598-y] [Medline: 37526785]
24. Davids JR, Wagner LV, Meyer LC, Blackhurst DW. Prosthetic management of children with unilateral congenital below-elbow deficiency. The Journal of Bone & Joint Surgery 2006;88(6):1294-1300. [doi: 10.2106/jbjs.e.00982]
25. Anderson F, Bischo WF. Augmented reality improves myoelectric prosthesis training. In: International Conference on Disability, Virtual Reality & Associated Technologies. Aalborg University Copenhagen in Denmark: AAU Open Publishing; 2012.
26. Andrews JR, Harrelson G, Wilk K. Physical Rehabilitation of the Injured Athlete: Expert Consult Online and Print. New York, Philadelphia: Elsevier Health Sciences; 2012.
27. Sherry MA, Best TM. A comparison of 2 rehabilitation programs in the treatment of acute hamstring strains. Journal of Orthopaedic & Sports Physical Therapy 2004;34(3):116-125. [doi: 10.2519/jospt.2004.34.3.116] [Medline: 15089024]
28. Mohamad Chan EW, Nadzalan AM, Othman Z, Hafiz E, Hamid MS. The Short-Term effects of progressive vs. Conventional core stability exercise in rehabilitation of nonspecific chronic low back pain. Sains Malaysiana 2020;49(10):2527-2537. [doi: 10.17576/jsm-2020-4910-18]
29. Levac DE. Individual and contextual factors influencing children's effort in pediatric rehabilitation interventions. Developmental Medicine & Child Neurology 2024;66(1):23-31 [FREE Full text] [doi: 10.1111/dmcn.15609] [Medline: 37082901]
30. Campbell E, Phinyomark A, Scheme E. Current trends and confounding factors in myoelectric control: Limb position and contraction intensity. Sensors (Basel) 2020;20(6):1613 [FREE Full text] [doi: 10.3390/s20061613] [Medline: 32183215]
31. Prakash A, Sharma S. Single-channel surface electromyography (sEMG) based control of a multi-functional prosthetic hand. Instrumentation Science & Technology 2021;49(4):428-444. [doi: 10.1080/10739149.2021.1880933]
32. Wang X, Wang J, Fei N, Duanmu D, Feng B, Li X, et al. Alternative muscle synergy patterns of upper limb amputees. Cognitive Neurodynamics 2024;18(3):1119-1133. [doi: 10.1007/s11571-023-09969-5] [Medline: 38826662]

XSL•FO
RenderX

33. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. International Journal of Human-Computer Interaction 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
34. Hart SG. NASA-Task load index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2006;50(9):904-908 [FREE Full text] [doi: 10.1177/154193120605000909]
35. Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 1971;9(1):97-113. [doi: 10.1016/0028-3932(71)90067-4] [Medline: 5146491]
36. MacDonald C, Barnum E, Pradeep M, Rivera V, Smith P, Dombrowski M, et al. Serious game-based training for improved control of a temporalis electromyography interface for controlling powered wheelchairs. Games for Health Journal 2025 [FREE Full text] [doi: 10.1177/2161783x251397932] [Medline: 41253416]
37. Levac D, Dumas HM, Meleis W. A tablet-based interactive movement tool for pediatric rehabilitation: development and preliminary usability evaluation. Journal of NeuroEngineering and Rehabilitation 2018;5(2):e10307 [FREE Full text] [doi: 10.2196/10307] [Medline: 30478025]
38. Freitag K, Karayiannidis Y, Zbinden J, Laezza R. Fine-Tuning myoelectric control through reinforcement learning in a game environment. IEEE Transactions on Biomedical Engineering 2026;73(1):168-179. [doi: 10.1109/TBME.2025.3578855] [Medline: 40498601]
39. Beltran JE, Castro YP, Mantilla MI. Video games as an influence in meaningful learning. 2017 Presented at: European Conference on Electrical Engineering and Computer Science (EECS); November 17-19, 2017; Bern, Switzerland p. 347-350. [doi: 10.1109/eecs.2017.71]
40. Van Twillert S, Postema K, Geertzen JH, Hemminga T, Lettinga AT. Improving rehabilitation treatment in a local setting: a case study of prosthetic rehabilitation. Clinical Rehabilitation 2009;23(10):938-947 [FREE Full text] [doi: 10.1177/0269215509338125] [Medline: 19717507]
41. Simon AM, Lock BA, Stubblefield KA. Patient training for functional use of pattern recognition-controlled prostheses. Journal of Prosthetics and Orthotics 2012;24(2):56-64 [FREE Full text] [doi: 10.1097/JPO.0b013e3182515437] [Medline: 22563231]

Abbreviations

EHI: Edinburgh Handedness Inventory
EMG: electromyography
MDAT: Muscular Discretization Assessment Tool
NASA-TLX: National Aeronautics and Space Administration Task Load Index
sEMG: surface electromyography
SUS: System Usability Scale
UI: user interface

Edited by S Munce; submitted 01.Dec.2025; peer-reviewed by P Wang, TH Hsieh; comments to author 19.Feb.2026; revised version received 28.Mar.2026; accepted 09.Apr.2026; published 20.May.2026.

Please cite as:
Lalle A, Migliore S, Stevenson J, Bell E, Pradeep M, Gunnell D, Bennett S, Rivera V, Smith P, Dombrowski M, Sparkman J, Manero II A
Multigesture Electromyographic Control Complexity in Upper Limb Prostheses Actuated via Single Sensor Input Contraction Magnitude: Qualitative Study for Evaluating Performance and Cognitive Load
JMIR Rehabil Assist Technol 2026;13:e88528
URL: https://rehab.jmir.org/2026/1/e88528
doi:10.2196/88528
PMID:

©Abrianna Lalle, Samantha Migliore, Jeffrey Stevenson, Ethan Bell, Maanya Pradeep, Delaney Gunnell, Sophie Bennett, Viviana Rivera, Peter Smith, Matt Dombrowski, John Sparkman, Albert Manero II. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Specialist Rehabilitation Providers' Experiences With an Online Self-Compassion Course: Reflexive Thematic Analysis

Eva Cohen[1], BSc, MSc; Kristina M Kokorelias[2,3,4], PhD; Dorothy Luong[5,6], MSc; McKyla McIntyre[7], MD; Marina Bastawrous Wasilewski[1,3,4], PhD; Sander L Hitzig[1,3,4], PhD; Sarah Munce[4,5,6,8], PhD; Rosalie Steinberg[1], MD; Anthony Feinstein[1], MD, PhD; Mark T Bayley[4,7], MD; Lawrence R Robinson[9], MD; Carolyn Steele Gray[8,10], PhD; Robert Simpson[1,4,7,11], MD, PhD

[1]Evaluative Clinical Sciences, Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Toronto, ON, Canada
[10]Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, ON, Canada
[11]Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom
[2]Department of Geriatric Medicine, Sinai Health System and University Health Network, Toronto, ON, Canada
[3]Department of Occupational Science & Occupational Therapy, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[4]Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, 550 University AveToronto, Toronto, ON, Canada
[5]KITE Research Institute, University Health Network, Toronto, ON, Canada
[6]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada
[7]Toronto Rehabilitation Institute, University Health Network, University of Toronto, Toronto, ON, Canada
[8]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[9]Department of Medicine, University of British Columbia, Victoria, BC, Canada

Corresponding Author:
Robert Simpson, MD, PhD
Evaluative Clinical Sciences, Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Abstract

Background: Rehabilitation health care providers (HCPs) report high levels of burnout. Self-compassion interventions have shown beneficial effects on HCP burnout, but they have never been explored in specialist rehabilitation settings where challenges may differ.

Objective: This study aimed to explore the experiences of specialist rehabilitation HCPs with an online Self-Compassion for Healthcare Communities (SCHC) course aimed at providing tools to regulate emotional well-being, including burnout.

Methods: Semistructured qualitative interviews (n=20) with specialist rehabilitation HCPs were used to explore experiences with the SCHC course. A reflexive thematic analysis study design was chosen to highlight participants' insights, and inductive coding was undertaken to analyze data and organize findings into themes.

Results: Six themes that reflected HCPs' experiences with the course were constructed: (1) the nature of working in rehabilitation; (2) different perspectives on burnout in specialist rehabilitation; (3) a new perspective, less self-criticism, more self-compassion; (4) growing recognition of the importance of compassion for oneself and others; (5) challenges engaging with the SCHC course; and (6) challenges to sustaining self-compassion in specialist rehabilitation.

Conclusions: Specialist rehabilitation HCPs who participated in the SCHC course developed a new understanding of the importance of fostering self-compassion and compassion for others. This shift may have supported HCPs in navigating workplace challenges, and HCPs described experiencing changes that could help with burnout, improve their care of patients and relationships with colleagues.

Trial Registration: ClinicalTrials.gov NCT05997082; https://clinicaltrials.gov/study/NCT05997082?cond=self%20compassion%20rehab&viewType=Card&rank=1

(JMIR Rehabil Assist Technol 2026;13:e81706) doi:10.2196/81706

KEYWORDS

self-compassion; burnout (Professional); rehabilitation; qualitative research; health personnel

XSL•FO
RenderX

Introduction

The ability to reach person-centered rehabilitation treatment goals relies on effective longitudinal communication and relational collaboration among patients, families, and health care providers (HCPs), and in this context HCP compassion is particularly important [1-3]. Compassion from HCPs enhances patient satisfaction, treatment adherence, and goal achievement [1]. Compassion describes sharing another's feelings coupled with a desire to relieve distress by taking action [2]. Compassion is often overlooked as a component of care in acute settings, but may be more prominent in rehabilitation settings due to the focus on each individual's recovery of independent function [2]. Indeed, a review of therapeutic interventions revealed that compassionate care is valued by people undergoing rehabilitation and their healing is promoted by a comforting environment [4].

Ensuring that compassion features routinely in specialist rehabilitation is challenged by HCP compassion fatigue [1], burnout [5], and secondary traumatic stress (STS) [6-11]. HCPs working in rehabilitation often endure an emotional burden throughout prolonged patient stays, especially as they manage patient and family expectations [12,13]. Compassion fatigue, a state where HCPs experience the reduced ability or motivation to care for their patients due to the sustained stressors of their duties [14,15], negatively impacts HCP quality of life and well-being [16]. Burnout is a linked construct characterized by emotional exhaustion, objectification of patients, and a reduction of effectiveness at work [17]. STS refers to a scenario where HCPs may develop symptoms typical of posttraumatic stress disorder, which mirror those of their patients, and arises when HCPs work with people who have endured traumatic events [6]. These constructs are highly correlated [18], but most research has focused on the effects of burnout, with well-developed literature on drivers and consequences [19,20]. Unfortunately, the COVID-19 pandemic and rise of digital health care in Canada have increased HCPs' risk of experiencing compassion fatigue, burnout, and STS [21,22].

To mitigate negative effects on HCPs, interventions focused on stress reduction and increasing resilience have shown effectiveness for reducing burnout and STS [23]. Mindfulness-based interventions (MBIs), which aim to systematically teach participants to "pay attention, in a particular way, on purpose and nonjudgmentally" have been widely used and shown to improve stress levels and reduce anxiety and depression among HCPs. While MBIs are helpful in mitigating the effects of burnout in-person [24] and online [25], experiences have not been assessed in rehabilitation settings. MBIs may not be feasible for busy HCPs as they are usually conducted over 8 weeks. Evidence suggests that MBIs may be the most beneficial (eg, accounting for reductions in stress) and the best attended in-person or synchronously online [3,26,27] and that brief, informal sessions (eg, less than 4 hours) improve accessibility [28]. Although the optimal MBI delivery method for HCPs is unknown, the combination of compassion cultivation training and mindfulness techniques may help mitigate the effects of burnout, compassion fatigue, and STS [29,30].

Like MBIs, compassion-based interventions effectively reduce anxiety, depression, and psychological distress, but they differ by placing greater emphasis on fostering supportive relationships and common humanity [31]. Similarly, self-compassion interventions also reduce anxiety and depression [32], and were found to do so to a greater extent than MBIs in a 2018 review [33]. In fact, an analysis of a large cross-sectional survey of nurses revealed that self-compassion was protective against burnout-related reductions in compassion [34]. Compassion-based interventions may be better than MBIs at improving HCP well-being and compassion [32], but they are complex interventions and the optimal delivery method is also unknown.

It is crucial to consider the digital landscape in which Western health care is situated with the rise of telemedicine and frequent use of digital technology since COVID-19 [35]. A mix of "asynchronous" (prerecorded online) [36] and "synchronous" (live online) components may best fit the needs of busy HCPs whose personal and professional duties require increasing levels of engagement with online platforms. An intervention which can be accessed "whenever and wherever" is likely to increase accessibility and thus improve engagement [25]. Indeed, the rise of digital health care may facilitate the success of MBIs delivered online rather than in-person [37]. A 2024 narrative review of randomized controlled trials assessing the efficacy of digital mental health interventions for HCPs revealed positive effects for targeting burnout, and acceptability was reported [38]. Factors such as flexibility, ease of use, the availability of support and feedback, and the likelihood of achieving positive outcomes were identified as facilitators of intervention effectiveness [38]. Furthermore, a 2024 scoping review on the assessment of audio-guided online MBIs suggested the need for future research to assess the diverse methods of online delivery (eg, videos and podcasts) [25].

Therefore, the purpose of this study is to explore the experiences of specialist rehabilitation HCPs with an online Self-Compassion for Healthcare Communities (SCHC) course [32].

Methods

Study Design

Overview

This study presents the qualitative process evaluation of a linked feasibility trial testing the SCHC course in specialist rehabilitation providers. No quantitative trial data are presented in this manuscript. Specialist rehabilitation providers are defined in this context as regulated health professionals working in "specially trained interprofessional teams with a focused skill set," [39]. The SCHC course on which participants were interviewed was delivered over Zoom (Zoom Video Communications, Inc), with weekly 60-minute sessions on Fridays between 12 PM and 1 PM, by a certified SCHC instructor (a specialist neuro-occupational therapist) and assistant (trainee clinical psychologist), with participants attending whole group sessions and regular smaller group break-out sessions. Participants were assigned home practice activities that varied weekly. An overview of intervention

XSL•FO
RenderX

content and week-by-week session guide is available in Multimedia Appendix 1.

Individual postcourse interviews were conducted by the study coordinator on Zoom. The Zoom platform was chosen due to widespread use across the institutional settings involved in recruitment, and to stay consistent with the course delivery. In addition, research participants generally view online interviewing methods for qualitative studies favorably, even over in-person interviews [40,41]. In-depth semistructured interviews were conducted following the final SCHC session. The interview topic guide (see Multimedia Appendix 2) explored participants' experiences with the SCHC course, its perceived effects, usage, barriers and facilitators to participation, and suggestions for improving future iterations.

Theoretical Framework

A qualitative interpretive methodology with a reflexive thematic analysis design and an inductive approach to theme generation was used without relying on preexisting theories or frameworks [42,43] to explore participants' experiences with the SCHC course. This methodology was selected to learn from qualitative data directly and to highlight participants' experiences. An inductive, rather than a theoretical analysis, relies on researchers' interpretations of the raw data, which may stray from the specific questions posed to participants [44].

As a reflexive thematic analysis, the aim was to achieve "information power" with an emphasis on data quality rather than data saturation, which prioritizes data quantity [45,46]. Whereas data saturation occurs when no new themes arise from coding new information gathered from the sample, information power refers to "the richness and relevance of data collected from a sample" [47]. Aiming to achieve a level of meaning in the data (ie, information power), and not solely redundance, is consistent with the method of a reflexive thematic analysis [45,46].

The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was followed to ensure ethical data collection and rigorous reporting methods (see Checklist 1) [48].

Participant Selection

Eligibility criteria for the feasibility trial in which this study's data were drawn from were being an HCP, 18 years of age or older, working substantially (ie, not agency staff) in a specialist rehabilitation setting (in-patient and out-patient) across existing specialist rehabilitation sites in a large metropolitan area (ie, the Greater Toronto Area), having the ability to communicate in spoken and written English, and not having undertaken a mindful-self-compassion course or other MBI in the past 12 months. All participants in the SCHC course (n=26) were eligible to interview. Participants were informed they would be invited for an interview following the course, but that participation would be voluntary. Following the course, all enrolled participants received up to 3 successive email invitations to schedule an interview with the study coordinator.

Data Collection and Setting

Demographic data (eg, age, gender, and professional role), available in Multimedia Appendix 3, were collected prior to course commencement via an online survey. Interviews were held and recorded on Zoom by the study coordinator (EC), who was present throughout the course only to manage administrative work and technology issues. On average, interviews lasted 50 minutes. Virtual interviews were hosted from the study coordinators' place of work. Participants joined their one-on-one interview from their workplaces or their homes. The interview audio files were transcribed verbatim by a secure third-party service and not returned to participants for review.

Research Team and Reflexivity

Personal Characteristics and Relationship With Participants

The research team situates themselves within a pragmatic paradigm, centering their beliefs and motives on the real-world significance and usefulness of research findings (ie, what is practical and achievable). In this case, this applies to the SCHC course as a tool for helping specialist rehabilitation providers address the well-documented challenges of busy and stressful frontline clinical care [49].

The lead author (EC) is a younger Canadian woman and an early career researcher. Her undergraduate and postundergraduate experience with qualitative research made her well-suited for involvement in this project while bringing a fresh lens to the analysis as a student who studied psychology in undergrad and studied aging and health in graduate school. Without firsthand experience as an HCP, previous knowledge of the SCHC course, and prior relationships with participants, she was well-placed to immerse herself in the data while maintaining a broader perspective of HCP experiences with burnout. The lead author (EC) conducted the interviews to limit influence on participants as she is not an HCP herself. The lead author (EC) was known to the participants as the study coordinator, but she did not provide any of her own personal insights throughout the course or during the interviews.

The senior author is a middle-aged Scots Canadian man who works as a physiatrist with a research background in developing and evaluating complex interventions, including those based in mindfulness and compassion. His holistic understanding of working in chronically stressful health care environments and his firsthand experience with burnout among fellow HCPs in diverse settings provided crucial insider perspectives, shaping how burnout was conceptualized and interpreted. To minimize potential influence, the senior author did not attend the sessions or conduct the interviews.

The third researcher (KMK) involved with data analysis (female) is a research associate experienced in qualitative research and thematic analysis. Her background in developing and evaluating complex interventions, including those mindfulness-based, provided a valuable lens to data interpretation. Additional authors are researchers and clinicians in rehabilitation science providing guidance and expertise in the design, conduction, and review of this project.

XSL•FO
RenderX

Reflexivity

The research team included a diverse range of perspectives and backgrounds, and this richness of experience and diversity was leveraged during the interpretive process with reflexivity practiced throughout. For example, the lead author (EC) kept a reflective set of notes of her thoughts and experiences of sitting-in on the SCHC course, during the semistructured interviews (which she conducted), when working through the development of a theme journal (ie, a flexible and evolving analytic tool used to keep track of theme development), and shared these reflections with the team during regular collaborative meetings where a rich discussion and a deeper exploration of the meaning of data was strongly encouraged. Meeting notes were made and continually updated as data were qualitatively coded and themes were developed. The primary investigator of this study is a physician and how this perspective, and associated power dynamics, might influence views were transparently acknowledged and discussed throughout the entire reflexive thematic analysis process. For example, team members (KMK, DL) with clinical expertise noted their inclination to prioritize certain health-related aspects of what participants described, while others from research-focused roles highlighted broader social and cultural contexts within the story being developed.

Data Analysis

Data were analyzed using a reflexive thematic analysis design with an inductive approach [44,50]. The steps followed were:

1. Familiarization: a total of 3 researchers (EC, KMK, and RobertSimpson) situated themselves with the deidentified interview data by reading transcripts several times and noting patterns or interesting points. The involvement of multiple researchers enhanced analytic depth and fostered reflexive dialogue, which was apparent throughout the entire thematic analysis process.
2. Initial code generation: initial codes were synthesized from each interview transcript in NVivo (version 12; QSR International Pty Ltd). Through an iterative coding process, 2 researchers labeled meaningful chunks of data with phrases that aligned with the research objective (eg, “course format was accessible”). A team meeting was held to facilitate collaborative discussion of initial codes, groupings, and patterns within the data. This discussion led to the re-organization of some groupings to maintain relevance to the research objective while assumptions were questioned. Initial code generation continued until all transcripts were fully labeled and organized, which itself was an iterative and recursive process.
3. Initial theme generation: subsequent team meetings facilitated theme development through discussion, during which initial themes (eg, “patterns of shared meaning” [51]) were constructed by grouping initial code labels together based on similarity and naming these groupings in NVivo. Preliminary theme definitions were also synthesized. Initial theme names and definitions were noted in a shared theme journal document to keep track of meeting progress. At this stage, mind maps were also created and continuously revised during and between meetings on an online design platform to assist with visualizing the meaning of initial groupings and the connections between them, while being able to zoom out and grasp a holistic view of the data.
4. Reviewing and developing themes: several weekly meetings were held to continue developing themes through collaborative discussion of the patterns within the data, ensuring that meaningful insights were representative of the groupings. Continuous reflexive engagement supported the development of cohesive storylines for each theme.
5. Refining, redefining, and naming themes: the research team iteratively refined the theme journal during regular conversational meetings by revisiting themes to ensure relevance to the objective. This process involved ongoing iterative refinement of themes and definitions by continuing reflexive engagement with the data. Researchers continued to question their initial assumptions and interpretations of the data and always attempted to view the data from several stances. Theme names were finalized based on theme definitions. Updated mind maps helped researchers visualize the story told by the data.
6. Producing the report: the analysis was written up to present a coherent story of the data, with continued reflexive engagement guiding development. Direct quotes from participant interviews were extracted, supplemented with researchers’ careful interpretations of each theme and how they connect to one another. Findings were situated among current knowledge in the field.

Ethical Considerations

Ethical approval was obtained from Sunnybrook Health Science Centre Research Ethics Board (SUN-5492). All study participants provided prospective written informed consent.

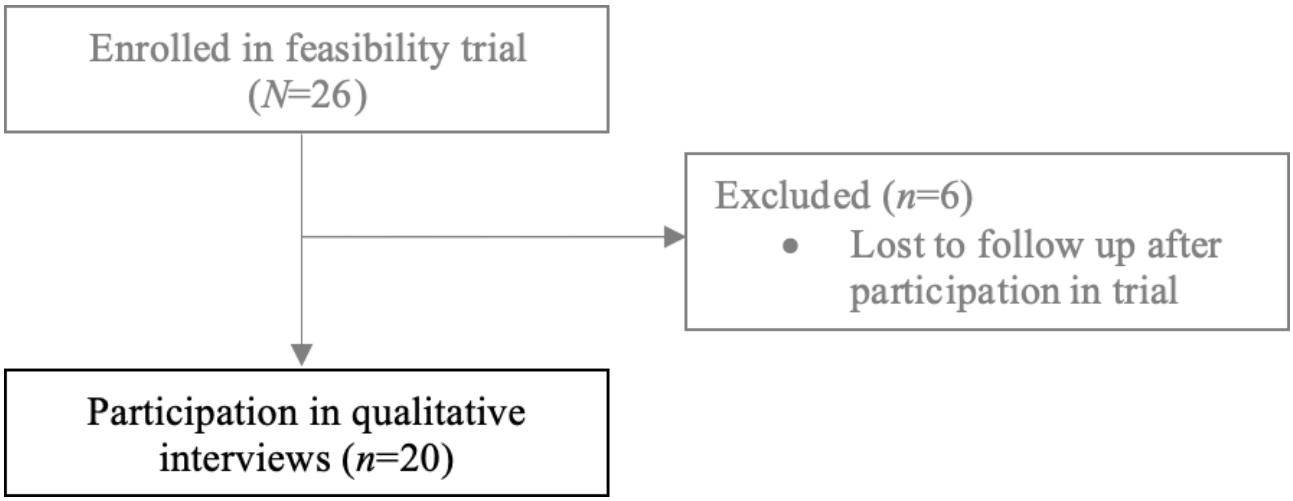
Results

Overview

Rehabilitation providers (n=20; see Multimedia Appendix 3 for participant demographics) who engaged in the SCHC course shared their experiences working in rehabilitation as well as their experiences with the course, thus resulting in 6 interconnected themes that portray HCPs’ intricate experiences. Only a subset of participants was interviewed, and we do not have information on the reasons why some individuals chose not to participate despite efforts to follow up after participation in the course (see Figure 1). This may have introduced self-selection bias, as those with more positive experiences may have been more likely to take part in interviews. The mean (SD) age of participants was 43.5 (10.5), the majority identified as women, almost one-third identified as White North American, and 96% worked in urban environments whereas one participant worked in both urban and rural areas. For an average of 13.6 years, HCPs worked primarily as occupational therapists and social workers, but there was indeed a vast spread of professions within specialist rehabilitation represented.

XSL•FO
RenderX

Figure 1. Participant flow diagram. Gray represents the linked feasibility trial in which this qualitative study draws data from. Detailed trial feasibility metrics are reported separately (under review).



Participants spoke of (1) the nature of working in specialist rehabilitation, an interplay between longitudinal relationships and focus on adaptation, which heavily impacted (2) different perspectives on burnout in specialist rehabilitation and experiences with the course. Throughout the course, (3) a new perspective, less self-criticism, more self-compassion, was gained by many participants as they learned how to alter their self-deprecating attitudes. Once self-compassion was fostered using various techniques learned throughout course sessions, (4) growing recognition of the importance of compassion for oneself and others arose as participants learned how to extend their newfound sense of self-compassion to their patients, family members, and colleagues. Unfortunately, (5) challenges engaging with the SCHC course were inevitable as HCPs endured busy schedules and individuals varied in their comfortability being vulnerable in an online, group setting alongside their peers. Although it was reported by participants that the course helped cultivate self-compassion and enhanced their compassion for others, several factors limited the ability of HCPs to extend this compassion to their patients. These factors are described in the last theme: (6) challenges to sustaining self-compassion in specialist rehabilitation. Supplemental quotes categorized by theme are provided in Multimedia Appendix 4.

Theme 1: The Nature of Working in Rehabilitation

The first theme, the nature of working in rehabilitation, provides context into the aspects of working in rehabilitation that can render HCPs at risk of compassion fatigue, burnout, and STS and how these factors mirror and differ from those faced by HCPs that work in other settings. Reflecting on their experiences prior to taking the course, HCPs described how they are challenged with distinct stressors due to the nature of the specialty. Rather than curing, HCPs in this field are tasked with helping their patients adapt to new circumstances, often following life-changing illness or injury. Often, patients in rehabilitation are perceived by HCPs as initially resistant to adaptation as their expectations had not been addressed in acute care. Due to the nature of rehabilitation, HCPs build long-term relationships that can last from weeks to decades. HCPs may also connect with their patients on an emotional level, which can have positive and negative implications for well-being. Some HCPs endure long-term maltreatment from patients, and others are affected by the implications of difficult cases outside of working hours.

> *I only recognized in the last couple years … a lot of abuse and dumping … negative energy put onto you. People … or my other colleagues don't realize after four to six weeks of someone being so negative, there is an impact on you.* [Woman, aged 45 years]

> *I would be left thinking about the traumas that they had experienced…And worrying if they'd be alive tomorrow…It was harder to disconnect…I wouldn't be working, but the thoughts would still be there, or I'd have nightmares.* [Woman, aged 30 years]

Like other HCPs, rehabilitation providers experience occupational stress from a lack of agency, large caseloads, staff shortages, time constraints, and excessive administrative work. Rehabilitation HCPs described how they often experienced moral distress when discharging patients to what were perceived to be unfit home environments, simply to meet organizational pressures around bed occupancy. Stress was accentuated for specialist rehabilitation HCPs when required to perform their duties, and sometimes colleagues' duties, within an overwhelmed system.

> *Some days are stressful because I'm asked to see four patients at a time. And there's no helping around. Or…there's three patients in the gym that I have to leave unattended… the nonsense that goes on around here is frustrating beyond.* [Woman, aged 34 years]

These stressors, some of which may be unique to rehabilitation, some which are common throughout health care, made the SCHC course relevant for participants as they sought out strategies to manage burnout and build resilience.

Theme 2: Different Perspectives on Burnout in Specialist Rehabilitation

Participants from different backgrounds and professional roles described the concept of burnout in various ways, whether they had learned about the concept from the course or elsewhere.

XSL•FO
RenderX

Participants described experiences with burnout affecting numerous aspects of their lives, especially in relation to their work in rehabilitation.

Participant feedback suggested the SCHC course might help improve participants' abilities to recognize and describe burnout and compassion fatigue, phenomena that were described as common. Indeed, burnout was described by participants as poorly functioning or performing at work. Accordingly, burnout was described as affecting more than one dimension of a person's life, including emotionally and physically, which was seen to impact participants' personal lives in an "…exponential downward spiral…" (Man, age 35 years).

> *Burnout is being both emotionally, physically…spiritually, like just out of steam, and not able to get through the day without interventions. So, like there was a time when I was burned out and I needed to take naps during the day. I just was emotionally exhausted and in the afternoon I would take two hours out of work, and then I would work later, because I needed that energy to keep going.* [Woman, aged 50 years]

Participants discussed burnout and compassion fatigue as distinct but related concepts. Specifically, participants described compassion fatigue as a concept constituting a precursor to burnout, which encompasses emotional exhaustion and limits the quality of compassionate care provided to patients.

> *It starts to weigh you down, especially when you hear some of the stories. Because the patients, they open up to you … they share so much of their life experience. And you listen with compassion. But it does wear you down after a while.* [Woman, aged 34 years]

Through attending the course, participants explained gaining a clearer vocabulary to describe burnout and compassion fatigue, enhancing their self-awareness and recognition of these phenomena in their daily practice.

Theme 3: A New Perspective, Less Self-Criticism, More Self-Compassion

The third theme speaks to gaining a new perspective when it came to teaching participants how to cultivate more self-compassion and less self-criticism. Participants recognized the importance of incorporating this new perspective to shape their attitudes about themselves and others in a positive way.

Following the completion of the course, most participants shared that the concept of self-compassion had been new to them. Most described being conscious of compassion for their patients, including the circumstances which would activate this way of being, but acknowledged not recognizing ever being compassionate to themselves. A common description across interviews was that self-compassion is more difficult to inculcate than compassion for others,

> *It's easy just to blame yourself. Self-compassion I find is harder. I find it much easier to be kinder to others than to myself.* [Woman, aged 49 years]

Frequently, providers described putting patient needs and work tasks before their own. Participants often recognized a tendency to self-criticize and be hard on themselves.

> *But I do think I've had difficulties with self-compassion because I grew up in a very domineering system. You're either wrong or you're wrong. You're never right. And with that, I've internalized a lot of that.* [Man, aged 35 years]

Despite this, participants described curiosity about the concept of self-compassion, seeming to recognize a clear need for self-care and lacking a resource to do so.

> *I didn't know what to expect, because I think when I hear the word self-compassion, my mind can immediately go to, oh, get a massage, pamper myself…So, I was definitely curious as to what would come out of this. I was pleasantly surprised. I just went in with an open mind. But I think I took away something more than I thought I was going to take away.* [Woman, aged 44 years]

Participants reflected on gradually learning to recognize self-critical tendencies throughout the course and beginning to adopt new approaches over time, with deliberate attempts to incorporate compassionate responding to themselves. Recognizing the adverse effects of self-criticism, participants began to see self-compassion as humanizing how they responded to themselves.

> *Self-compassion means having compassion for yourself as an individual…This is something I really resonated with in the course and thought about it in a new way. It's like just feeling. Just remembering that you are a person too…Recognizing it's something I have to continue to work on from the course.* [Woman, aged 60 years]

Participants also noted that this had beneficial feed-forward effects in relation to others through common-humanity, with a renewed desire to alleviate the distress that all humans experience.

> *My peace can be restored when I remind myself that I'm not the only person navigating these feelings and thoughts. That common humanity piece is really important to me because it makes me feel less isolated.* [Woman, aged 49 years]

Theme 4: Growing Recognition of the Importance of Compassion for Oneself and Others

Participants described that recognizing the importance of compassion for oneself and others became more apparent throughout the course as they reflected on learning how to let go of their self-critical attitudes over time and focused on being kind to themselves, using the learned self-compassion practices, and listening to peers' similar experiences.

Participants described that a fundamental premise to fostering self-compassion is recognizing it is okay be kind to oneself, even as an HCP "it's okay sometimes to just step back and take care of yourself first, before other people," (Woman, aged 53 years). Participants reflected on using self-compassion practices

XSL•FO
RenderX

more naturalistically in their day-to-day lives as the course progressed while others described using what they had learned to incorporate in their clinical care of patients, or in their professional lives with students, or in their personal lives with family members.

> *I use the [hand on heart strategy]. I've been doing that with my family too. Like my wife when we talk, just to remind myself to be grounded. And just the practice of doing the exercises I think just made me more attuned to my patients and just helped me remember to hold space for people and have more patience. Immediately, I just felt like my practice was shifting a little bit.* [Man, aged 43 years]

Participants iterated a perceived importance of regularity of the self-compassion practice as a means of developing the skill and driving lasting change. In this regard, the course environment appeared to motivate continued engagement with the self-compassion practices, where participants described being inspired by others' reflections on personal insights and use.

> *I really did appreciate listening to others. And I found some of the other participants were so thoughtful and open and vulnerable and really shared both their struggles and how the practices were really meaningful to them. And that was inspiring in a way, to know other people are also struggling similarly.* [Woman, aged 42 years]

Indeed, the online nature allowing providers from otherwise separate rehabilitation sites reinforced a sense of camaraderie and solidarity.

> *A year or so ago, a patient died… We don't get a lot of deaths in rehab. And in rehab we're not trained for that. Being able to sit down as a group after that happened, no management, just us, and talk about how we were feeling from everybody's perspectives… was the most important thing for us to prevent burnout… and all help each other get past it.* [Woman, aged 40 years]

Setting work-life boundaries, "not taking any work home" or "not looking at emails," (Woman, aged 26 years) was also deemed a crucial element to self-compassion by participants, who described this as necessary to maintain their own well-being and ability to provide compassionate care to patients. According to HCPs, developing an understanding of cultivating inward compassion, through learning and applying specific strategies, played an integral role in fostering a grounded and clear mindset needed to extend compassion toward others.

Theme 5: Challenges Engaging With the SCHC Course

Participants described several barriers to course participation from the method of course delivery to challenges with course content, as well as individual variation in level of readiness to engage in sensitive conversations. A striking insight shared was that to make time for course engagement, participants had to give up time spent charting during their lunch-hour (ie, investing time in themselves came at a cost, causing anxiety surrounding completing work later in the day.)

> *I didn't attend all the courses. I think I only attended two and then I really had to back out because it was causing a lot more stress in terms of my workload, and I sat there more anxious than anything, so it was setting my afternoons off.* [Woman, aged 45 years]

Other challenges noted included a sense of discomfort in connecting emotionally to others online and participants worried about how exposing oneself in this way might undermine their professional identity,

> *When you're not surrounded by people, physically, you can't pick up on their body cues. It's just not a very comfortable environment to talk about your emotions or talk about your feelings.* [Woman, aged 34 years]

While some HCPs found aspects of the course to be conducive to engagement (eg, the diversity of attendees and the online format), there were many suggestions for improvement which primarily focused on timing, mode, and content. Suggestions did not follow a specific pattern and varied from implementing an asynchronous course to hosting the meetings before working hours to limiting the group to a smaller size. A persistent suggestion was tailoring content to typical challenging rehabilitation scenarios and implementing a follow-up course to reinforce new skills.

Theme 6: Challenges to Sustaining Self-Compassion in Specialist Rehabilitation

Several barriers to being compassionate in rehabilitation beyond course completion were perceived, ranging from individual (eg, changing habits) to systemic (eg, heavy workload) in nature. However, following course completion, participants explained that by necessity they learned how to implement self-compassion practices from the SCHC spontaneously in their daily lives despite obstacles, such as stepping aside for a brief period to "ground" oneself during unbidden stress.

Perceived role stipulations were cited to challenge practices from the SCHC in specialist rehabilitation, where an explicit emphasis is placed on HCPs prioritizing physical function, and less attention paid to emotional distress which often underlies impairment and impedes functioning.

> *I would say most people who are struggling physically are also struggling emotionally. I found often with [stroke patients], it was harder to get a read on really how much someone was struggling, especially if their speech was impacted. It was a little bit harder to connect with people because you couldn't quite tell all the time how much emotionally this has impacted.* [Woman, aged 40 years]

Participants also noted that implementing practices from the SCHC course could not address prevalent systemic issues directly, such as when patient and family expectations for rehabilitation were perceived as unrealistic due to not being addressed in the acute setting,

> *So that, I find, has always probably there been a consistent stress factor, dealing with families with high expectations and then managing their*

XSL•FO
RenderX

> *expectations. So, when you have those conversations and if they're very anxious or stressed and they come at you and they use you as a scapegoat sometimes.* [Woman, aged 44 years]

Systemic-level realities such as large caseloads, time constraints, and the fast pace of inpatient units were all seen as impediments to maintaining compassion, despite learnings from the course. Hinting at moral distress, there were clear depictions of the struggle to balance personal capacity (eg, energy) for delivering exceptional, compassionate care with the relentless demands of the work environment. This finding underscores the necessity of learning and implementing tools to foster self-compassion in an environment that may not inherently promote it.

> *I'm struggling with taking care of a lot of patients. I often-times feel like I don't give enough attention to them. It's more about I'm feeling guilty if I don't give enough attention to them, and I know, if I have the time, I can make a lot of difference. But I just cannot. And after all these years I do make sure I don't spend way too much time here. I know I can't sustain that.* [Woman, aged 59 years]

Discussion

Principal Findings

In this exploratory qualitative study, specialist rehabilitation providers were provided with the opportunity to engage in the SCHC course online. Reflexive thematic analysis generated six prevalent and interconnected themes: (1) the nature of working in rehabilitation, (2) different perspectives on burnout in specialist rehabilitation, (3) a new perspective, less self-criticism, more self-compassion, (4) growing recognition of the importance of compassion for oneself and others, (5) challenges engaging with the SCHC course, and (6) challenges to sustaining compassion in specialist rehabilitation.

The nature of working in rehabilitation highlights the complexities of specialist rehabilitation that can render HCPs at risk of compassion fatigue, burnout, and STS. Not only does the focus of specialist rehabilitation—adaptation—contrast with the focus of acute care—treatment or cure, but also patient and families' expectations of the goals of rehabilitation are rarely managed beforehand. Thus, managing expectations following a disabling life event becomes a daunting and burdensome task for HCPs to fulfill that extends beyond dealing with the emotional distress that may arise from completing their regular tasks. Additionally, the longitudinal nature of specialist rehabilitation affects HCPs in 2 capacities. On one hand, HCPs describe the value in getting to know their patients on a deeper level and the satisfaction of helping patients improve their functioning throughout time. On the other hand, HCPs may endure long-term maltreatment from their patients and their families if expectations are not managed properly and if functional improvements are not apparent. Further, these effects may persist outside of working hours, affecting HCPs in profound ways. Although familiar for specialist rehabilitation HCPs, these nuances may not be well understood among the general public [52,53], and effects on HCPs' well-being are not always incorporated into organizational guidelines or well-documented in the literature [54,55].

The nuances that a specialist rehabilitation setting offers may expose HCPs to highly traumatizing experiences [8] which may contribute to a significant risk of burnout [56]. Most participants in this study were unclear about how to define or differentiate between compassion fatigue, burnout, and STS. Many had heard of the concept of burnout before the course and could speculate at its meaning based on personal or colleagues' experiences. Most conveyed experiencing stress through providing patient care, some sounding personally traumatized. Participants discussed the course helping to grow their understanding of the concept of burnout and that it also provided a safe forum for discussing difficult experiences among peers who could understand, to the extent that this fed into suggestions for making the course even more relevant. Elsewhere, other self-compassion intervention studies have quantitatively reported reductions in symptoms of burnout among HCPs following participation or in comparison to a control group [57-60], and have also reported that HCPs developed an increased awareness of their own well-being [58]. In this exploratory qualitative reflexive thematic analysis, we did not assess quantitative changes in burnout pre- to postcourse, but participants did report that through participation in the SCHC course they became aware they may be experiencing burnout and also gained useful tools to combat these challenges. A better understanding among HCPs of how work in specialist rehabilitation can affect one's occupational well-being may facilitate engagement with evidence-based self-care strategies and may feed into concerted efforts to change a culture in health care that raises the risk of burnout for many HCPs [61].

HCPs have ethical duties to do no harm and to act in patients' best interests (eg, [62,63]), even if this means doing so at the expense of their own well-being, which is a prevailing and problematic cultural norm in health care [32]. Therefore, it is not surprising that participants described having an easier time being compassionate to their patients than to themselves. Most HCPs recognized that they are regularly compassionate to their patients yet had never labeled their actions as such. Even fewer participants acknowledged intentionally being self-compassionate while reflecting on their lives prior to the course. While the course attempted to emphasize the opposite, for many participants, to identify professionally as a specialist rehabilitation HCP, one cannot also show humanistic qualities. Whether this belief was perpetuated by personal factors or systemic norms and the complex association between necessitates future research. Being a "human being" (eg, being exhausted or not performing your best at work) lends well to practicing self-compassion, but for HCPs, having self-compassion was not possible without compromising professional identity. Accordingly, it appeared to take time for HCPs to gain a new perspective, less self-criticism, and more self-compassion to challenge their beliefs. In line with the findings of Neff et al [32], the SCHC course may be more beneficial for HCPs who report having low self-compassion to begin with, although in this study, changes over time were described retrospectively and not formally assessed longitudinally.

XSL•FO
RenderX

Following the new perspective participants described, many discussed the course being helpful in growing recognition for the importance of compassion for oneself and others. Participants found being intentional about one's thoughts, feelings, and actions important when seeking to cultivate compassion and debriefing about difficult situations as a health care team an important mechanism to bring compassion to bear on collective adverse experiences. Descriptions of growing compassion may align with current quantitative evidence that mindful self-compassion interventions are successful in minimizing HCPs' self-criticism [59] and improving self-compassion [32,57-60]. This study is the first study to report participant experiences of growing self- and other-focused compassion in specialist rehabilitation HCPs, despite these constructs not being assessed quantitatively.

The "new" concepts covered in the SCHC course were well-received by HCPs who were open to learning about how to improve their well-being by increasing their self-awareness and engaging with self-compassion practices. Similarly, existing literature affirms the importance of practicing the techniques from the SCHC course in a health care setting for oneself and for one's patients [58,59]. Accordingly, promoting common humanity [64] and fostering communities of support [58] continue to be useful tools to promote compassion in HCPs.

HCPs provided both positive feedback and constructive criticism based on the delivery of the SCHC course. The online format provided participants with the flexibility to join from several sites across a large metropolitan area, and the diversity of professional roles was appreciated and contributed to a sense of common humanity when peers discussed similar experiences in their duties even across sites. Additional advantages that supplement previous research [57] include saving travel time and facilitating joining from home. One of the most notable suggestions was to tailor the content to include rehab-specific examples. This suggestion was also reported in recent reviews. Emphasis on the importance of contextual adaptation underscores the need for tailoring digital interventions to the specific group of HCPs they intend to be delivered to [38] (eg, rehabilitation providers). A 2024 scoping review specific to the assessment of online MBIs primarily using audio-guided programs also suggested tailoring interventions for groups of HCPs [25]. This change is likely to enhance acceptability but would necessitate further feasibility testing.

Despite participants' acknowledgment of the benefit of self-compassion practices they learned, many noted challenges engaging with the course such as scheduling challenges, expected with busy HCPs [32]. Interestingly, not being ready to share personal struggles in a group setting was a less common challenge. It is possible that those who were unwilling to share did not participate in interviews entirely. The fact that some HCPs disclosed these feelings still emphasizes participants' curiosity about the self-compassion intervention and the importance of having trained facilitators to handle challenges that may arise. While the group format (eg, the presence of unfamiliar peers) hindered some participants' willingness to share and reduced their ability to build rapport with other course participants and facilitators, the added benefit of accessibility must be weighed with the impracticalities of in-person sessions. However, most participants appeared to value the format of the course. Also noted was the impracticability of engaging in self-compassion practices and actions at work, in line with Crandall et al [58] and in contrast to Delaney's [59] findings that most nurses found their self-compassion practices useful on the job. The lack of suitable work environments to practice self-compassion (eg, a quiet location to engage in the "hand on heart" practice) speaks to the need for leaders and organizations to understand the need to focus on employee well-being and appropriate infrastructure change (eg, staff "recovery" rooms; [4]).

To provide compassionate care to promote better patient satisfaction [1] and healing [4], both individual and systemic factors should be targeted. Participants spoke of their complex patient populations and difficulties managing patient expectations to limit the compassion they were able or willing to give. Patients undergoing specialist rehabilitation often exhibit distress in more than one area of their lives (eg, physical and emotional) as they typically experience life-altering diseases or catastrophic injuries [65,66]). HCPs should be aware that patients are frustrated with their new realities or the health care system more generally, which can be perceived as dehumanizing both by HCPs and patients alike [67]. Speculatively, increased awareness of shared distress and common humanity through participation in the SCHC course may limit HCPs' internalization of patient frustration, to instead respond with compassion, and this in turn may limit their risk of feeling burnt out, although this remains speculative and would have to be tested quantitatively to map out processes and determine proponent causality. Regardless, HCP training in compassionate care should arguably include strategies for helping identify the source of patient frustration and provide tools for HCP emotional awareness and regulation. Although not only rehab-specific [68], systemic challenges that feature in HCPs' day-to-day include being overwhelmed with case numbers and not having enough time to complete their duties. From participant accounts, the guilt that comes with wanting to be compassionate, but not being able to, may perpetuate specialist rehabilitation HCP moral distress and feelings related to burnout, and the cycle is likely to continue if changes are not made. Policy changes facilitated by leadership that foster a sense of shared humanity and a healthy workplace environment are likely to tackle barriers to compassion, such as implementing protected time with management to reflect on challenges as well as taking short breaks throughout the day [69,70].

Limitations

Although rich data were gathered to portray HCPs' diverse experiences with workplace-related stress and the SCHC course, this study comes with limitations. First, generalizability to the population of specialist rehabilitation providers is restricted. The sample all worked in the Greater Toronto Area, which limits applicability of findings to other populations across Canada or countries outside of Canada, especially those from rural areas. Fewer HCPs are employed in rural compared to urban environments, which can intensify workplace stress and increase the prevalence of role sharing among rehabilitation disciplines [71]. Such factors may intensify HCPs' experiences with burnout and thus, the course may need tailored in relation to the needs

XSL•FO
RenderX

of the population served. For instance, rural HCPs may find themselves experiencing more debilitating workplace stress that may or may not classify as burnout, as they may take on additional roles outside their scope of practice or work more hours than urban HCPs. Despite this limitation, the sample was heterogeneous, containing people of various ages and professional disciplines in rehabilitation who attended most sessions. This study recruited participants from a specific geographical area due to the nature of the research protocol. Future iterations should consider expanding the sample to increase generalizability and assess feasibility on a wider scale, something which offering an online SCHC, as opposed to in-person, can help facilitate.

Further, it is entirely possible that participants who did not interview had different experiences with the course than those who did interview, reflecting potential self-selection bias. Common barriers for participation in qualitative research include the time requirement (eg, more time intensive than survey-based research) and lack of incentives [72]. Noninterviewees could have been too busy at work to participate in an interview and thus, they may have described more severe workplace-related stresses than the group who interviewed. All participants were offered gift cards to interview. In addition to time and incentives factoring into one's decision, those who did not interview may have had poorer experiences with the course than counterparts who interviewed. Alternatively, nonrespondents may have felt as though the course was not relevant to them, which may explain the lack of desire to provide feedback. These characteristics may or may not have altered the themes that arose from this analysis if the nonrespondents provided interview data. Ultimately, it was difficult to conclude why the 6 participants did not choose to interview as they did not respond to our interview invites even after following up. Positive findings, including the perceived growth of compassion and course accessibility, should be interpreted while considering the potential reasons why a subset of participants chose not to interview. Nevertheless, the number of participants who interviewed was high and motivation for participation was captured, important pieces of the puzzle when analyzing acceptability of a program in a specific setting.

Social desirability bias, the tendency to respond in accordance with perceived socially acceptable characteristics or favorable outcomes [73], may have influenced interviewee responses. The more engagement with researchers students in a learning setting experience and the more stress in one's life (eg, being overworked or experiencing burnout), the more likely participants are to respond to questions favorably, without responses reflecting their true experiences [74,75]. Social desirability bias may explain the tendency for participants to report that the SCHC course was helpful in growing their understanding of the concept of self-compassion or limiting feelings associated with burnout or workplace stress. Additionally, HCPs may fear disclosing burnout-related feelings due to stigma surrounding mental health issues, worries about it impacting their employment [76], or privacy concerns [77]. While semistructured interviews were used to gather responses, which may increase the effects of social desirability bias, the study coordinator limited her involvement during the trial to limit potential influence, being on camera only when necessary to solve technical issues or answer logistical questions. Furthermore, the purpose of the study was communicated to participants as not to solely provide tools to improve emotional well-being, but rather to gain insights into the experiences of engagement with the course. Questions were phrased from both sides (eg, in which ways was the course helpful or unhelpful?) or as neutral as possible, and participants were assured their responses would remain anonymous. Social desirability bias has been shown to be minimized in these ways (eg, [73]) and the mix of positive and negative responses, including suggestions for course improvement, help demonstrate this. Aside from the risk of socially desirable responding, online interviewing is liked by participants, easy to use by researchers and facilitators, and conveys perceptions of empathy and rapport well when used for qualitative research [40,41,78].

Implications and Future Directions

Participation in the online SCHC course appeared to help improve HCPs' knowledge of the concept of burnout, validate common struggles with their peers, and teach them tools to foster self-compassion and compassion for others. Some participants experienced challenges feeling connected to the course and fellow participants because of the online synchronous format, course timing, and other barriers described. Accordingly, participants described program strengths and areas for improvement. Both positive feedback and constructive criticism provided by HCPs can inform the development of knowledge translation tools (eg, a website), educational sessions, organizational strategies to better support staff well-being.

This is the first study exploring HCPs' experiences with the SCHC course in a specialist rehabilitation context. Variations of the course should be tested iteratively for acceptability, accessibility, and efficacy by adjusting course timing (eg, time of day), format (eg, synchronous), and class size in line with suggestions provided by our participants. The most prominent suggestion was to make content rehabilitation specific. Future trials should include concrete examples of stressors that may arise in rehabilitation when discussing the need for self-compassion so that participants can relate. Course delivery conditions also matter; protected learning time and continuing education credits may be required.

Accordingly, to better understand potential mechanisms of action and effectiveness (eg, by measuring pre- and postburnout, compassion fatigue, and STS scores), the next phase of research will compare the SCHC course and a matched comparator, such as a Health Enhancement Program. The Health Enhancement Program is an intervention specifically designed to match MBIs in terms of group context, instructor time and attention, education on health behaviors, but not on any content on mindfulness or self-compassion [79]. This would constitute a phase 2 randomized controlled trial design [80]. Additional future studies, such as one that gathers perspectives from health leaders (eg, Division heads) can provide valuable insights into organizational structures that may facilitate participation (eg, the availability of quiet spaces) and clarify responsibility for participation.

XSL•FO
RenderX

Conclusion

An online SCHC course appears relatively acceptable and accessible to some specialist rehabilitation providers, who noted gaining a deeper understanding of the concepts of self-compassion and burnout and discussed a desire to extend more compassion to themselves, their patients, and colleagues. Barriers to participation, such as difficulties making the time to join sessions and challenges with being vulnerable online were also discussed by participants, which may limit relevance for some busy HCPs. Self-compassion interventions like this one may be a useful strategy for improving HCPs' well-being in a postpandemic context; however, optimization changes appear necessary to maximize relevance and use for all specialist rehabilitation HCPs, and this should be the focus of future research.

Acknowledgments

Thank you to the St. John's Research Team and Jillian Scandiffio for your ongoing support with this project. We are grateful for all study participants for taking time out of your busy schedules to participate in this important project.

Artificial intelligence was not used in any stages of the hypothesis, data collection, data evaluation, or manuscript preparation.

Funding

This work was funded entirely by an Associated Medical Services Fellowship in Compassion and AI.

Data Availability

Data may be made available upon reasonable request.

Authors' Contributions

Conceptualization: R Simpson (lead), MM (supporting), MBW (supporting), SLH (supporting), R Steinberg (supporting), SM (supporting), AF (supporting), MB (supporting), LRR (supporting), CSG (supporting)
Formal Analysis: EC (lead), KMK (equal), R Simpson (equal), DL (supporting)
Writing – Original Draft: EC (lead)
Writing – Review & Editing: KMK (lead), R Simpson (equal), DL (supporting), MM (supporting), MBW (supporting), SLH (supporting), SM (supporting), R Steinberg (supporting), AF (supporting), MB (supporting), LRR (supporting)

Conflicts of Interest

SM is the current editor in chief at *JMIR Rehabilitation and Assistive Technologies*. KMK is a current member of the editorial board at *JMIR Rehabilitation and Assistive Technologies*.

Multimedia Appendix 1
Week-by-week session guide.
[DOCX File, 16 KB - rehab_v13i1e81706_app1.docx]

Multimedia Appendix 2
Semistructured interview guide.
[DOCX File, 16 KB - rehab_v13i1e81706_app2.docx]

Multimedia Appendix 3
Sociodemographic characteristics of participants.
[DOCX File, 4455 KB - rehab_v13i1e81706_app3.docx]

Multimedia Appendix 4
Supplemental quotes.
[DOCX File, 35 KB - rehab_v13i1e81706_app4.docx]

Checklist 1
COREQ checklist.
[DOCX File, 4455 KB - rehab_v13i1e81706_app5.docx]

References

XSL•FO
RenderX

1. Posa S, Wasilewski MB, Mercer SW, Simpson S, Robinson LR, Simpson R. Conceptualization, use, and outcomes associated with empathy and compassion in physical medicine and rehabilitation: a scoping review. Int J Rehabil Res 2022 Dec 1;45(4):291-301. [doi: 10.1097/MRR.0000000000000542] [Medline: 35837691]
2. Simpson R, Szigeti Z, Sheppard CL, et al. The experiences of patients, family caregivers, healthcare providers, and health service leaders with compassionate care following hospitalization with COVID-19: a qualitative study. Disabil Rehabil 2023 Sep;45(18):2896-2905. [doi: 10.1080/09638288.2022.2113564] [Medline: 36354063]
3. Wasilewski MB, Vijayakumar A, Szigeti Z, et al. Patient and provider experiences with compassionate care in virtual physiatry: qualitative study. J Med Internet Res 2024 Aug 6;26:e51878. [doi: 10.2196/51878] [Medline: 39106094]
4. Halstead LS. The power of compassion and caring in rehabilitation healing. Arch Phys Med Rehabil 2001 Feb;82(2):149-154. [doi: 10.1053/apmr.2001.22207]
5. Swensen SJ, Shanafelt TD. Mayo Clinic Strategies to Reduce Burnout: 12 Actions to Create the Ideal Workplace: Oxford University Press; 2020.
6. Stamm BH. The Concise proQOL Manual. 2010 URL: https://proqol.org/ [accessed 2026-06-25]
7. Sliwa JA, Clark GS, Chiodo A, et al. Burnout in diplomates of the American Board of Physical Medicine and Rehabilitation—prevalence and potential drivers: a prospective cross‐sectional survey. PM R 2019 Jan;11(1):83-89. [doi: 10.1016/j.pmrj.2018.07.013]
8. Teel J, Reynolds M, Bennett M, et al. Secondary traumatic stress among physiatrists treating trauma patients. Proc (Bayl Univ Med Cent) 2019 Apr;32(2):209-214. [doi: 10.1080/08998280.2018.1559694] [Medline: 31191130]
9. Silver JK, Bhatnagar S. Physician Burnout in Physical Medicine and Rehabilitation (PM&R): should we focus more on physiatrists' mission? Am J Phys Med Rehabil 2017 Aug;96(8):e159-e161. [doi: 10.1097/PHM.0000000000000674] [Medline: 28081027]
10. Li Calzi S, Farinelli M, Ercolani M, Alianti M, Manigrasso V, Taroni AM. Physical rehabilitation and burnout: different aspects of the syndrome and comparison between healthcare professionals involved. Eura Medicophys 2006 Mar;42(1):27-36. [Medline: 16565683]
11. Bruschini M, Carli A, Burla F. Burnout and work-related stress in Italian rehabilitation professionals: a comparison of physiotherapists, speech therapists and occupational therapists. Work 2018;59(1):121-129. [doi: 10.3233/WOR-172657] [Medline: 29439375]
12. Peel R, Jones S, Miller C, Gibson J. Great expectations? A qualitative study of health professionals' perspectives on breaking bad news about rehabilitation potential after traumatic brain injury or spinal injury. Disabil Rehabil 2020 Nov;42(22):3209-3214. [doi: 10.1080/09638288.2019.1589586] [Medline: 31006346]
13. Anestis E, Eccles FJR, Fletcher I, Triliva S, Simpson J. Healthcare professionals' involvement in breaking bad news to newly diagnosed patients with motor neurodegenerative conditions: a qualitative study. Disabil Rehabil 2022 Dec;44(25):7877-7890. [doi: 10.1080/09638288.2021.2002436] [Medline: 34783624]
14. Nolte AG, Downing C, Temane A, Hastings-Tolsma M. Compassion fatigue in nurses: a metasynthesis. J Clin Nurs 2017 Dec;26(23-24):4364-4378. [doi: 10.1111/jocn.13766] [Medline: 28231623]
15. Sinclair S, Raffin-Bouchal S, Venturato L, Mijovic-Kondejewski J, Smith-MacDonald L. Compassion fatigue: a meta-narrative review of the healthcare literature. Int J Nurs Stud 2017 Apr;69:9-24. [doi: 10.1016/j.ijnurstu.2017.01.003] [Medline: 28119163]
16. Yelvington ML, Corson T, Tarver T, Teague T, Oguntola-Blount A, Reynolds S. A scoping review of compassion fatigue and compassion satisfaction in burn therapists. J Burn Care Res 2024 May 6;45(3):608-613. [doi: 10.1093/jbcr/irae029] [Medline: 38375782]
17. Shanafelt TD, Noseworthy JH. Executive leadership and physician well-being: nine organizational strategies to promote engagement and reduce burnout. Mayo Clin Proc 2017 Jan;92(1):129-146. [doi: 10.1016/j.mayocp.2016.10.004] [Medline: 27871627]
18. Rauvola RS, Vega DM, Lavigne KN. Compassion fatigue, secondary traumatic stress, and vicarious traumatization: a qualitative review and research agenda. Occup Health Sci 2019 Sep;3(3):297-336. [doi: 10.1007/s41542-019-00045-1]
19. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med 2018 Jun;283(6):516-529. [doi: 10.1111/joim.12752] [Medline: 29505159]
20. West CP, Dyrbye LN, Erwin PJ, Shanafelt TD. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. Lancet 2016 Nov 5;388(10057):2272-2281. [doi: 10.1016/S0140-6736(16)31279-X] [Medline: 27692469]
21. Bennett J, Paech G, Hodges BD. Without Compassion, There Is No Healthcare: Compassionate Care in a Technological World: McGill-Queen's University Press; 2020.
22. Maunder RG, Heeney ND, Strudwick G, et al. Burnout in hospital-based healthcare workers during COVID-19. : Ontario COVID-19 Science Advisory Table (OCSAT); 2021. [doi: 10.47326/ocsat.2021.02.46.1.0]
23. Cocker F, Joss N. Compassion fatigue among healthcare, emergency and community service workers: a systematic review. Int J Environ Res Public Health 2016 Jun 22;13(6):618. [doi: 10.3390/ijerph13060618] [Medline: 27338436]
24. Goodman MJ, Schorling JB. A mindfulness course decreases burnout and improves well-being among healthcare providers. Int J Psychiatry Med 2012;43(2):119-128. [doi: 10.2190/PM.43.2.b] [Medline: 22849035]

XSL•FO
RenderX

25. Baek G, Cha C, Lee M, Cho A. Online mindfulness-based interventions for healthcare professionals: a scoping review. Workplace Health Saf 2024 Nov;72(11):460-482. [doi: 10.1177/21650799241254554] [Medline: 39215553]
26. Wolever RQ, Finn MTM, Shields D. The relative contributions of live and recorded online mindfulness training programs to lower stress in the workplace: longitudinal observational study. J Med Internet Res 2022 Jan 21;24(1):e31935. [doi: 10.2196/31935] [Medline: 35060911]
27. Wolever RQ, Bobinet KJ, McCabe K, et al. Effective and viable mind-body stress reduction in the workplace: a randomized controlled trial. J Occup Health Psychol 2012 Apr;17(2):246-258. [doi: 10.1037/a0027278] [Medline: 22352291]
28. Gilmartin H, Goyal A, Hamati MC, Mann J, Saint S, Chopra V. Brief mindfulness practices for healthcare providers - a systematic literature review. Am J Med 2017 Oct;130(10):1219. [doi: 10.1016/j.amjmed.2017.05.041] [Medline: 28687263]
29. Conversano C, Ciacchini R, Orrù G, Di Giuseppe M, Gemignani A, Poli A. Mindfulness, compassion, and self-compassion among health care professionals: what's new? A systematic review. Front Psychol 2020;11:1683. [doi: 10.3389/fpsyg.2020.01683] [Medline: 32849021]
30. Roca P, Vazquez C, Diez G, Brito-Pons G, McNally RJ. Not all types of meditation are the same: mediators of change in mindfulness and compassion meditation interventions. J Affect Disord 2021 Mar 15;283:354-362. [doi: 10.1016/j.jad.2021.01.070] [Medline: 33578349]
31. Kirby JN, Tellegen CL, Steindl SR. A meta-analysis of compassion-based interventions: current state of knowledge and future directions. Behav Ther 2017 Nov;48(6):778-792. [doi: 10.1016/j.beth.2017.06.003] [Medline: 29029675]
32. Neff KD, Knox MC, Long P, Gregory K. Caring for others without losing yourself: an adaptation of the mindful self-compassion program for healthcare communities. J Clin Psychol 2020 Sep;76(9):1543-1562. [doi: 10.1002/jclp.23007] [Medline: 32627192]
33. Wilson AC, Mackintosh K, Power K, Chan SWY. Effectiveness of self-compassion related therapies: a systematic review and meta-analysis. Mindfulness (N Y) 2019 Jun;10(6):979-995. [doi: 10.1007/s12671-018-1037-6] [Medline: 31354877]
34. Dev V, Fernando AT III, Lim AG, Consedine NS. Does self-compassion mitigate the relationship between burnout and barriers to compassion? A cross-sectional quantitative study of 799 nurses. Int J Nurs Stud 2018 May;81:81-88. [doi: 10.1016/j.ijnurstu.2018.02.003] [Medline: 29518624]
35. Lawrence K, Levine DL. The digital determinants of health: a guide for competency development in digital care delivery for health professions trainees. JMIR Med Educ 2024 Aug 29;10:e54173. [doi: 10.2196/54173] [Medline: 39207389]
36. Finlay-Jones A, Kane R, Rees C. Self-compassion online: a pilot study of an internet-based self-compassion cultivation program for psychology trainees. J Clin Psychol 2017 Jul;73(7):797-816. [doi: 10.1002/jclp.22375] [Medline: 27787877]
37. Zhang A, Borhneimer LA, Weaver A, et al. Cognitive behavioral therapy for primary care depression and anxiety: a secondary meta-analytic review using robust variance estimation in meta-regression. J Behav Med 2019 Dec;42(6):1117-1141. [doi: 10.1007/s10865-019-00046-z] [Medline: 31004323]
38. Aye LM, Tan MM, Schaefer A, et al. Self-help digital mental health intervention in improving burnout and mental health outcomes among healthcare workers: a narrative review. Digit Health 2024;10:20552076241278313. [doi: 10.1177/20552076241278313] [Medline: 39257871]
39. Specialized rehabilitation: guidance for Ontario Health Teams. : GTA Rehab Network; 2021 URL: https://gtarehabnetwork.ca/wp-content/uploads/2022/08/Specialized_Rehab_-_All_Populations_Nov_2021.pdf [accessed 2026-06-25]
40. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. Int J Qual Methods 2019 Jan 1;18:1609406919874596. [doi: 10.1177/1609406919874596]
41. Oliffe JL, Kelly MT, Gonzalez Montaner G, Yu Ko WF. Zoom interviews: benefits and concessions. Int J Qual Methods 2021 Jan;20:16094069211053522. [doi: 10.1177/16094069211053522]
42. Vears DF, Gillam L. Inductive content analysis: a guide for beginning qualitative researchers. Focus Health Prof Educ 2022;23(1):111-127. [doi: 10.11157/fohpe.v23i1.544]
43. Clarke V, Braun V. Thematic analysis. In: Teo T, editor. Encyclopedia of Critical Psychology: Springer; 2014:1947-1952. [doi: 10.1007/978-1-4614-5583-7_311]
44. Campbell K, Orr E, Durepos P, et al. Reflexive thematic analysis for applied qualitative health research. Qual Rep 2021;26(6):2011-2028. [doi: 10.46743/2160-3715/2021.5010]
45. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
46. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
47. Sureshkumar A, Scandiffio J, Luong D, et al. Implementing a new clinical service - what's your elevator pitch? BMC Health Serv Res 2025 Mar 28;25(1):462. [doi: 10.1186/s12913-025-12633-9] [Medline: 40156017]
48. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

XSL•FO
RenderX

49. Long KM, McDermott F, Meadows GN. Being pragmatic about healthcare complexity: our experiences applying complexity theory and pragmatism to health services research. BMC Med 2018 Jun 20;16(1):94. [doi: 10.1186/s12916-018-1087-6] [Medline: 29921277]
50. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
51. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
52. Bialosky JE, Bishop MD, Cleland JA. Individual expectation: an overlooked, but pertinent, factor in the treatment of individuals experiencing musculoskeletal pain. Phys Ther 2010 Sep;90(9):1345-1355. [doi: 10.2522/ptj.20090306] [Medline: 20592270]
53. Wade S. Physical therapy is 'the best-kept secret in health care'. The Associated Press. 2024. URL: https://apnews.com/article/be-well-physical-therapy-ffd7de3680fd4e63322cd63085cfbeda [accessed 2026-06-10]
54. Bushuven S, Trifunovic-Koenig M, Bunz M, et al. Applicability and validity of second victim assessment instruments among general practitioners and healthcare assistants (SEVID-IX Study). Healthcare (Basel) 2024 Jan 30;12(3):351. [doi: 10.3390/healthcare12030351] [Medline: 38338236]
55. Ogińska-Bulik N, Gurowiec PJ, Michalska P, Kędra E. Prevalence and predictors of secondary traumatic stress symptoms in health care professionals working with trauma victims: a cross-sectional study. PLoS ONE 2021;16(2):e0247596. [doi: 10.1371/journal.pone.0247596] [Medline: 33621248]
56. Simpson R, Cohen E, Kiss A, et al. Burnout in Canadian physiatrists: a national cross-sectional survey. Am J Phys Med Rehabil 2025 Nov 1;104(11):967-975. [doi: 10.1097/PHM.0000000000002748] [Medline: 40444938]
57. Eriksson T, Germundsjö L, Åström E, Rönnlund M. Mindful self-compassion training reduces stress and burnout symptoms among practicing psychologists: a randomized controlled trial of a brief web-based intervention. Front Psychol 2018;9:2340. [doi: 10.3389/fpsyg.2018.02340] [Medline: 30538656]
58. Crandall J, Harwood L, Wilson B, Morano C. Mindful self-compassion training and nephrology nurses' self-reported levels of self-compassion, burnout, and resilience: a mixed methods study. Nephrol Nurs J 2022;49(5):405-417. [Medline: 36332121]
59. Delaney MC. Caring for the caregivers: evaluation of the effect of an eight-week pilot mindful self-compassion (MSC) training program on nurses' compassion fatigue and resilience. PLoS ONE 2018;13(11):e0207261. [doi: 10.1371/journal.pone.0207261] [Medline: 30462717]
60. Super A, Yarker J, Lewis R, Keightley S, Summers D, Munir F. Developing self-compassion in healthcare professionals utilising a brief online intervention: a randomised waitlist control trial. Int J Environ Res Public Health 2024 Oct 11;21(10):1346. [doi: 10.3390/ijerph21101346] [Medline: 39457319]
61. Cloninger CR, Stoyanov D, Stoyanova KK, Stutzman KK. Empowerment of health professionals: promoting well-being and overcoming burn-out. In: Mezzich JE, Appleyard WJ, Glare P, et al, editors. Person Centered Medicine: Springer International Publishing; 2023:703-723. [doi: 10.1007/978-3-031-17650-0_42]
62. Code of ethics regulations and standards. Ontario College of Pharmacists. 2022. URL: https://www.ocpinfo.com/regulations-standards/code-of-ethics/ [accessed 2026-06-10]
63. CMA code of ethics and professionalism. Canadian Medical Association. 2018. URL: https://policybase.cma.ca/link/policy13937 [accessed 2026-06-10]
64. Ling D, Petrakis M, Olver JH. The use of common humanity scenarios to promote compassion in healthcare workers. Aust Soc Work 2021 Jan 2;74(1):110-121. [doi: 10.1080/0312407X.2020.1808031]
65. Yang Q, Young IF, Wan J, Sullivan D. Culturally grounded scapegoating in response to illness and the COVID-19 pandemic. Front Psychol 2021;12:632641. [doi: 10.3389/fpsyg.2021.632641] [Medline: 33889110]
66. Daker-White G, Hays R, McSharry J, et al. Blame the patient, blame the doctor or bBlame the system? A meta-synthesis of qualitative studies of patient safety in primary care. PLoS ONE 2015;10(8):e0128329. [doi: 10.1371/journal.pone.0128329] [Medline: 26244494]
67. Roupa A, Patelarou A, Giakoumidakis K, et al. Measurement of dehumanization, self-dehumanization, and empathy as mediating factors among healthcare professionals. Healthcare (Basel) 2025 Jan 3;13(1):75. [doi: 10.3390/healthcare13010075] [Medline: 39791682]
68. Malenfant S, Jaggi P, Hayden KA, Sinclair S. Compassion in healthcare: an updated scoping review of the literature. BMC Palliat Care 2022 May 18;21(1):80. [doi: 10.1186/s12904-022-00942-3] [Medline: 35585622]
69. Nielsen CL, Lindhardt CL, Näslund-Koch L, Frandsen TF, Clemensen J, Timmermann C. What is the state of organisational compassion-based interventions targeting to improve health professionals' well-being? Results of a systematic review. J Adv Nurs 2025 May;81(5):2246-2276. [doi: 10.1111/jan.16484] [Medline: 39373033]
70. Paakkanen M, Martela F, Pessi AB. Battling the barriers to compassion in organizations. Aust J Manag 2025 Aug;50(3):836-875. [doi: 10.1177/03128962241286071]
71. Knox K, Rohatinsky N, Rogers M, Goodridge D, Linassi G. Access to traumatic spinal cord injury care in Saskatchewan, Canada: a qualitative study on community healthcare provider perspectives. Can J Disabil Stud 2014;3(3):83 [FREE Full text] [doi: 10.15353/cjds.v3i3.174]

XSL•FO
RenderX

72. Gürsoy E, Vatansever M. Perceptions, challenges, and opportunities of qualitative research among inexperienced medical academics. BMC Med Educ 2025 Aug 29;25(1):1220. [doi: 10.1186/s12909-025-07824-1] [Medline: 40883726]
73. Bispo Júnior JP. Social desirability bias in qualitative health research. Rev Saude Publica 2022;56:101. [doi: 10.11606/s1518-8787.2022056004164] [Medline: 36515303]
74. Lavidas K, Papadakis S, Manesis D, Grigoriadou AS, Gialamas V. The effects of social desirability on students' self-reports in two social contexts: lectures vs. lectures and lab classes. Information 2022;13(10):491. [doi: 10.3390/info13100491]
75. Papadakis S, Semerikov SO, Striuk A, et al. Embracing digital innovation and cloud technologies for transformative learning experiences. 2024 Presented at: Proceedings of the 11th Workshop on Cloud Technologies in Education (CTE 2023) URL: https://ceur-ws.org/Vol-3679/paper00.pdf [accessed 2026-06-25]
76. Hudson E, Arnaert A, Lavoie-Tremblay M. Healthcare professional disclosure of mental illness in the workplace: a rapid scoping review. J Ment Health 2025 Feb;34(1):78-90. [doi: 10.1080/09638237.2021.1979485] [Medline: 34582294]
77. Arnhart K, Privitera MR, Fish E, et al. Physician burnout and barriers to care on professional applications. J Leg Med 2019;39(3):235-246. [doi: 10.1080/01947648.2019.1629364] [Medline: 31626578]
78. Anthony K, Miller-Day M, Dupuy M, et al. Is there really a difference? A comparison of in-person and online qualitative interviews. Int J Qual Methods 2025 Apr;24:16094069251349580. [doi: 10.1177/16094069251349580]
79. Mahaffey BL, Mackin DM, Vranceanu AM, et al. The Stony Brook Health Enhancement Program: the development of an active control condition for mind-body interventions. J Health Psychol 2020;25(13-14):2129-2140. [doi: 10.1177/1359105318787024] [Medline: 30010422]
80. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
HCP: health care provider
MBI: mindfulness-based intervention
SCHC: Self-compassion for healthcare communities
STS: secondary traumatic stress

Edited by J Sarvestan; submitted 04.Aug.2025; peer-reviewed by R Martin, S Papadakis; revised version received 19.May.2026; accepted 25.May.2026; published 15.Jul.2026.

Please cite as:
Cohen E, Kokorelias KM, Luong D, McIntyre M, Wasilewski MB, Hitzig SL, Munce S, Steinberg R, Feinstein A, Bayley MT, Robinson LR, Steele Gray C, Simpson R
Specialist Rehabilitation Providers' Experiences With an Online Self-Compassion Course: Reflexive Thematic Analysis
JMIR Rehabil Assist Technol 2026;13:e81706
URL: https://rehab.jmir.org/2026/1/e81706
doi:10.2196/81706

© Eva Cohen, Kristina M Kokorelias, Dorothy Luong, McKyla McIntyre, Marina Bastawrous Wasilewski, Sander L Hitzig, Sarah Munce, Rosalie Steinberg, Anthony Feinstein, Mark T Bayley, Lawrence R Robinson, Carolyn Steele Gray, Robert Simpson. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 15.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding the Origins and Factors of Burnout in Physical Medicine and Rehabilitation: Grounded Theory Analysis

Robert Simpson[1,2,3], MBChB, PhD; Eva Cohen[4], BSc; Stephanie Posa[1], MSc; Marina Wasilewski[1,5], PhD; Anthony Feinstein[1,5], MD, PhD; Mark Bayley[1,2], MD; Larry Robinson[1,5], MD; Sarah Munce[1,6], PhD; Carolyn Steele Gray[1,7], PhD; Kristina Kokorelias[1,7], PhD

[1]University of Toronto, Toronto, ON, Canada
[2]Toronto Rehabilitation Institute, Toronto, ON, Canada
[3]University of Glasgow, Glasgow, Glasgow City, United Kingdom
[4]Queen's University, Kingston, ON, Canada
[5]Sunnybrook Research Institute, Toronto, ON, Canada
[6]Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada
[7]Lunenfeld-Tanenbaum Research Institute, Toronto, ON, Canada

Corresponding Author:
Robert Simpson, MBChB, PhD
University of Toronto
1 King's College Circle
Toronto, ON, M5S 3K3
Canada
Phone: 1 (416) 864 5377
Email: robert.simpson@uhn.ca

Abstract

Background: Physician burnout is highly prevalent in Physical Medicine and Rehabilitation (PM&R), but its origins and drivers remain poorly understood.

Objective: This study aims to explore the factors contributing to burnout among Canadian physiatrists.

Methods: Using Charmaz's Constructed Grounded Theory within a qualitative interpretivist paradigm, we interviewed 30 Canadian physiatrists about their experiences with burnout. Analysis was informed by Cooley's looking-glass self theory.

Results: Burnout in PM&R in Canada stems from a medical culture prioritizing academic excellence over compassionate care. Canadian physiatrists report shame and self-criticism when unable to meet these high standards. Retrospective accounts from Canadian physiatrists suggest that burnout peaks during residency, where autonomy is low and demands are high. Participants also described feeling unprepared to handle patients' emotional needs and experiencing moral distress when necessary care cannot be delivered due to systemic barriers. Health care bureaucracy further compounds burnout.

Conclusions: Addressing burnout in PM&R in Canada requires upstream systemic and contemporary cultural change.

(JMIR Rehabil Assist Technol 2026;13:e80499) doi:10.2196/80499

KEYWORDS

burnout; physical medicine and rehabilitation; physiatrists; grounded theory; medical culture; residency training; moral distress; health care bureaucracy; emotional labor; stigma

Introduction

Background

Burnout was originally described as an occupational stress syndrome associated with chronically stressful and emotionally intense work conditions [1]. In health care, burnout has been further classified as having 3 overlapping domains, including emotional exhaustion, diminished sense of achievement, and depersonalization, the latter of which refers to objectification of patients, or loss of compassion [2]. Of concern, burnout in physicians around the world is widespread [3], representing a crisis in health care globally. Physician burnout has significant individual, organizational, and societal costs through worse job

XSL•FO
RenderX

performance, workforce attrition, and lower patient satisfaction [4,5].

Among physician specialties in the United States, Physical Medicine and Rehabilitation (PM&R) consistently ranks among the most burned out [6] (odds ratio [OR] 1.30), with only Emergency Medicine (OR 2.53, 95% CI 2.18-2.93), Family Medicine (OR 1.46, 95% CI 1.29-1.65), and Radiology (OR 1.31, 95% CI 1.12-1.54) reporting higher odds of burnout. Burnout affects a significant proportion of PM&R specialists and trainees—ranging from 48% to 62% for specialists and up to 83.3% for trainees [7]. Putative factors contributing to burnout in PM&R physicians come mostly from US survey data, mirror those seen in physicians overall, and do not clarify the specialty-specific causes that drive or sustain burnout [6]. PM&R uniquely focuses on the longitudinal care of people with disabling conditions, aiming to treat impairments, support independence, enhance social participation, and improve quality of life [8]. These philosophical underpinnings are based on a biopsychosocial model of health [9]. This is the core work of PM&R specialists ("physiatrists") [10].

Existing literature has highlighted several potential contributors to burnout among physiatrists, including high patient load, administrative burden, limited resources [10,11], and may also include the emotional toll of managing patients with chronic conditions [6,11,12], especially those subject to trauma [13]. However, these studies have largely not explored the personal experiences of physiatrists that lead to burnout in PM&R. To our knowledge, only one qualitative study has examined burnout in PM&R, focusing on personal strategies for occupational well-being during the COVID-19 pandemic; however, it used descriptive methods and did not develop a theory explaining burnout in the specialty.

Physiatrists' work with people affected by disabling long-term conditions, coordination of multidisciplinary care, and prioritizing of functional recovery entails unique stressors not fully captured by burnout theories from other medical contexts [14,15]. This gap highlights the need for a PM&R-specific theoretical framework to clarify the unique stressors driving burnout and guide targeted interventions aimed at improving physiatrists' well-being and patient care.

Aims

By capturing the lived experiences and perspectives of Canadian physiatrists, this study aims to provide a comprehensive understanding of the origins and factors contributing to burnout in PM&R through a grounded theory approach.

Methods

Design

A qualitative exploratory approach using grounded theory was chosen for its suitability in exploring underresearched areas and generating deeper insights into attitudes, cultures, and social processes [16], specifically Charmaz's constructed grounded theory [17-19]. Grounded theory as a qualitative methodology is used when the focus is on the process and the creation of theory or a conceptual framework [17,18,20]. In this research, grounded theory was used to construct a conceptual understanding of how physiatrists comprehend and experience burnout. Charmaz's [17-19] approach to grounded theory emphasizes the coconstruction of meaning between researchers and participants, aligning well with our theoretical grounding in Cooley's [21] looking-glass self midrange theory. The looking-glass self theory posits that individuals form their self-concepts based on how they perceive others to view them. This idea was instrumental in shaping our analysis of how physiatrists internalize and respond to burnout—through both personal experiences and external reflections from colleagues, patients, health care institutions, and social systems overall [21]. Using the looking-glass self theory, this study explored how Canadian physiatrists' burnout perceptions are shaped by interactions with and perceived judgments of others. This perspective enriched our grounded theory approach by highlighting relational dynamics such as how social comparison and external validation influence feelings of adequacy or support. It also guided the creation of visual and narrative materials, which complemented narrative data by capturing the interplay between internal experiences and external perceptions central to both the sensitizing theory and our subsequent analysis [22]. Visual materials were developed to reflect scenarios where physiatrists interacted with their environments, colleagues, and patients, illustrating how external judgments and expectations contribute to the emotional and psychological responses to burnout. In this context, the visuals served as a creative tool to depict how social comparisons, external validation, and perceived judgments play a role in shaping physiatrists' experiences of burnout.

The study received approval from the Sunnybrook Research Ethics Board, ensuring voluntary participation, confidentiality, and the right to withdraw.

Data Collection

Recruitment and Sample

Between September 2023 and March 2024, 30 physiatrists from across Canada were recruited for the study. There are an estimated 501 physiatrists in Canada [23]. Eligibility required English proficiency, current employment as a physiatrist, and willingness to join an online interview. Recruitment used advertisements distributed by the Canadian Association of Physical Medicine and Rehabilitation, outreach by PM&R Division Directors, and via public licensing registries for remote areas. Potential participants received an initial email plus up to 2 reminders. The team monitored recruitment to ensure diversity in age, gender, race, geography, subspecialty, and experience with burnout, expanding outreach to PM&R leaders when needed. Participants provided consent before interviews and received a CAD $50 (US $36.05) Amazon gift card afterward.

Methods of Data Collection

Data collection occurred from September 2023 to March 2024 through in-depth qualitative interviews conducted by RS and a trained research assistant EC. Questions were designed to be open-ended, to explore participants' experiences with burnout in PM&R, with the interviewer paying particular attention to views and actions, as well as emergent areas of potential theoretical interest. Audio and video recordings were made via

XSL•FO
RenderX

Zoom (Zoom Communications). The semistructured interview guide, focused on burnout experiences among Canadian physiatrists, is available in Multimedia Appendix 1.

Data Analysis

In this study, data analysis began alongside data collection, allowing for constant comparison between new data and emerging categories. This method ensured that each round of data collection was informed by the preceding analysis, allowing the research team to refine interview questions and explore areas that required deeper understanding. We used an inductive and iterative process to derive final insights [18]. Coding was done in multiple stages—open coding to identify initial concepts, axial coding to link these concepts together, and selective coding to integrate them into a coherent framework. First, two researchers open-coded the data in NVivo (Lumivero) using descriptive codes on participant responses [24]. The reviewers then met with the principal investigator to discuss the preliminary coding and developed an inductive codebook that was applied to the transcripts. The looking-glass self theory informed our coding process by encouraging us to attend to moments in the interviews where participants referenced feedback from their work environment—whether from peers, supervisors, or patients—as central to their feelings of burnout or professional fulfillment. This theoretical approach underscored the importance of external perceptions in shaping burnout, thus enriching our analysis of the interplay between personal identity, professional role expectations, health institutions, and systemic pressures. From this, the same researchers engaged in a constant comparison analysis where meetings were held to discuss the common higher-order themes implied by clusters of open codes, leading to the development of axial codes reflecting participants' insights on the conceptualization and understanding of burnout and experiences with this phenomenon [25,26]. The looking-glass self theory served as a sensitizing concept, framing how burnout is shaped through social interactions in PM&R and guiding the interpretation of nuanced reflections on acknowledgment by self, professional peers, and institutions. Two researchers applied axial coding to transcripts, discussing the process weekly [26]. Once team discussions suggested a preliminary understanding of a theoretical framework, the researchers engaged in a final reflexive process of selective (focused) coding [27]. Selective coding involved the deliberate examination of specific aspects of the preliminary framework within the transcripts to deepen our understanding of the framework, explore relationships between the data, and extract meaningful insights [19]. Categories were identified from comparative analysis, culminating in a core category that encapsulated the study's main theme [18]. Through team discussions, the researchers were able to extrapolate prominent themes within the data to inform an overarching conceptual understanding [19]. In between the weekly meetings, the researchers went independently back to the data to further refine their ideas against the emerging conceptual framework.

We systematically searched for disconfirming cases that challenged emerging themes, documenting and analyzing them to refine our findings. Analysis and team discussions continued until a conceptual framework achieving theoretical saturation was developed [28], where themes around how physiatrists understood and experienced burnout were consistently repeating, with no new dimensions or variations emerging, and the conceptual framework appeared sufficiently dense to explain the origins and factors sustaining burnout in PM&R [27].

Positionality

Our positionality as a team influenced this process, as we were mindful of how our own experiences and professional backgrounds shaped the way we coded and interpreted the data [29]. The lead author, a middle-aged Scots Canadian physiatrist with expertise in mindfulness and compassion, contributed insider insight on burnout in high-stress health care settings, framing it through the looking-glass self theory and emphasizing emotional and relational aspects. An interdisciplinary team, including a PhD-trained rehabilitation scientist and early-career researchers, incorporated structural and institutional factors and reflected on their roles, producing a nuanced, reflexive framework capturing the complexity of burnout in PM&R.

Ethical Considerations

This study was conducted in accordance with the World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Participants and prospectively approved by the Sunnybrook Research Ethics Board (project ID: SUN-5835). Informed consent was obtained electronically from all study participants after the nature and possible consequences of the study were explained. Respect for individual autonomy was emphasized, including that even following the granting of informed consent, participants could choose to withdraw from the study for any reason, without needing to give a reason and without adverse consequence. All participants who provided a semistructured interview received a CAD $50 (US $36.05) Amazon gift card in compensation. All reasonable efforts were made throughout to protect participant safety, dignity, privacy, and confidentiality. All personal identifying information was deidentified, and any interview quotes with potentially identifying information were not included in the study findings.

Results

Participant Characteristics

Most participants were women (19/30, 63%), with a mean age of 45.5 years. Most participants were recruited from Ontario (18/30, 60%). The most common racial identity was White European (12/30, 40%), followed by White North American (9/30, 30%), East Asian (6/30, 20%), South Asian (2/30, 7%), and Latin European (1/30, 3%). Participants worked in a wide range of PM&R settings, but most worked in an urban setting (28/30, 93%), full time (22/30, 73%), with 19 (63%) participants working 50 or more hours per week, with mixed inpatient and outpatient populations (24/30, 80%), and in most cases in neurological rehabilitation (13/30, 43%). No PM&R residents were interviewed (Table 1).

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of participants.

Characteristic	Values (N=30)
Sex/gender, n (%)	
Female/woman	19 (63.3)
Male/man	11 (36.7)
Age (years), mean (SD)	45.5 (10.5)
Number of years practicing, mean (SD)	13.6 (10.7)
Province of practice, n (%)	
Ontario	18 (60)
British Columbia	4 (13)
Quebec	3 (10)
Saskatchewan	2 (6.7)
Alberta	1 (3.3)
New Brunswick	1 (3.3)
Nova Scotia	1 (3.3)
Race/ethnicity[a] , n (%)	
White European	12 (40)
White North American	9 (30)
East Asian	6 (20)
South Asian	2 (6.7)
Mixed	2 (6.7)
Latin European	1 (3.3)
Specialty[a] , n (%)	
Neurorehabilitation	13 (43.4)
Electromyography	9 (30)
Other	8 (26.7)
Spinal rehabilitation	5 (16.7)
Musculoskeletal	4 (13.3)
General rehabilitation	3 (10)
Pain	3 (10)
Primary professional role, n (%)	
Clinician, full-time	22 (73.3)
Clinician, part-time	8 (11.8)
Secondary professional role[a] , n (%)	
Clinician-teacher	14 (46.7)
Clinician-scientist	13 (43.3)
Administrative roles	10 (33.3)
Clinician-quality improvement	1 (3.3)
Hours worked per week, n (%)	
50+	19 (63.3)
30-40	7 (23.3)
40-50	4 (13.3)
Time spent providing virtual care, n (%)	

XSL•FO
RenderX

Characteristic	Values (N=30)
0%-25%	27 (90)
25%-50%	3 (10)
Provides out-of-hours[b], n (%)	18 (60)
Clinic setting, n (%)	
Mixed inpatient and outpatient	24 (80)
Outpatient	6 (20)
Rurality, n (%)	
Urban	28 (93.3)
Urban and rural	2 (6.7)

[a]Participants could choose more than one option. Percentages may not add up to 100.

[b]Reflects the number and percentage of participants answering "yes" to this question.

Main Findings

The study findings present a conceptually grounded theory titled "Burnout in Physical Medicine and Rehabilitation is the Hierarchical Performance Paradox." The theory comprises 3 interconnected concepts: hierarchical structure, perception of performance, and their interaction in PM&R. In this theory, which is based on retrospective accounts of the lived experience of practicing physiatrists, the origins of burnout in PM&R are sown early in the training of budding physiatrists. Specifically, the hierarchical structure of medical training and practice imposes escalating demands amidst constrained autonomy as Canadian physiatrists progress through their careers, whilst simultaneously the perceived necessity of perfection in performance—shaped by internalized appraisals by peers, patients, and regulatory authorities—foments feelings of imposter syndrome. This hidden curriculum prioritizes performance proficiency above all else, leading to a mindset of "Doctor first, human second" (Woman, aged 33 years) (the hidden curriculum here refers to the implicit lessons, values, and attitudes conveyed through the behaviors and expectations of mentors, colleagues, and the health care system, which are not formally part of the medical educational syllabus [30]). Burnout risk appears highest when autonomy is low and stress is high, fostering self-criticism and shame. Social comparison and perfectionism are reported by Canadian physiatrists to begin in medical school and intensify during PM&R residency, when autonomy and social support decline, and burnout peaks. Later, increased autonomy does not eliminate risk, as expanded responsibilities, emotional demands, health care system inadequacies, and bureaucratic pressures perpetuate moral distress and burnout throughout a Canadian physiatrist's career. Illustrative quotes from participants are woven throughout the text. Figure 1 depicts the grounded theory process from coding to theory for burnout in Canadian physiatrists. Figure 2 depicts the hierarchical performance paradox.

XSL•FO
RenderX

Figure 1. Grounded theory process.

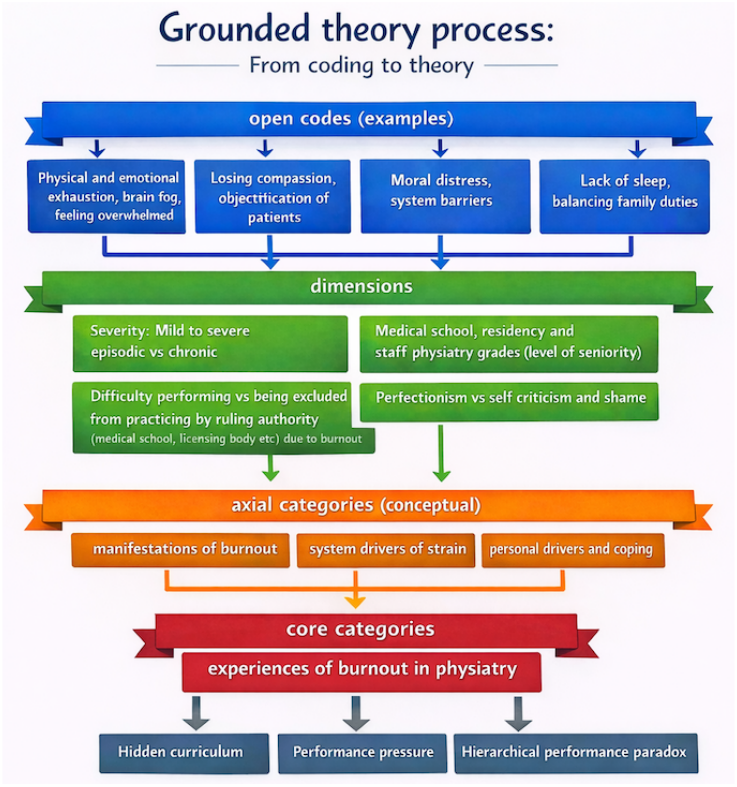
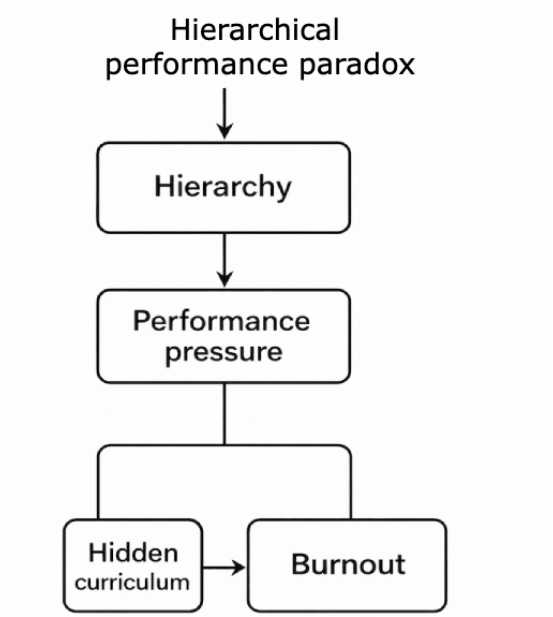


Figure 2. The hierarchical performance paradox and burnout in physiatry.



The Role of Hierarchy in Burnout in Physical Medicine and Rehabilitation

Retrospective accounts from practicing Canadian physiatrists indicate that burnout in PM&R originates from a hierarchical and performance-driven medical education system that inadvertently de-emphasizes compassionate care, leading to stress and emotional exhaustion:

> *Medicine is the stodgiest, most conformist profession that there is, and it has not adapted or evolved to the 20th century, and I think that creates a burnout environment.* [Man, aged 41 years]

Hierarchy in this context refers to the structured levels of authority and responsibility within medical education and practice, where individuals at lower levels, such as medical students and residents, are subject to the oversight, evaluation, and directives of those higher in the system, such as attending physicians and administrators. This hierarchical framework prioritizes academic performance, efficiency, and adherence to institutional norms over compassionate care, creating an environment where Canadian physiatrists face intense pressure,

XSL•FO
RenderX

reduced autonomy, and a heightened risk of burnout. A hidden curriculum in this context perpetuates these values, fostering emotional detachment and prioritizing efficiency and exceptional levels of academic performance. These factors, together with systemic constraints and relentless vertical oversight from within the hierarchy, and horizontal peer appraisal, contribute to burnout outcomes such as emotional exhaustion, depersonalization, and a diminished sense of achievement among many Canadian physiatrists.

Participant accounts indicated that the motivations for becoming a physiatrist are largely altruistic, reflecting core values of improving quality of life for people with disabling long-term conditions:

> *I try my hardest to be compassionate to patients because that's why I wanted to get into what I wanted to do in physiatry. I feel that you have a longer relationship with people, so you need more of a compassionate approach versus a one-off here and there if you're doing surgery or something like that.* [Woman, aged 35 years]

However, participants also stressed that, in spite of altruistic motivations, the reality is that performance stratification determines who is eligible to apply for admission to medical school, emphasizing academic performance over other qualities:

> *I think being in medicine and needing the grades you need to get into medical school, things like that require a certain degree of perfectionism, so it's something that's been valued throughout your life.* [Woman, aged 35 years]

Participants iterated that once admitted to medical school, proficiency in academic performance is strongly emphasized. Failure to meet the bar is associated with adverse social comparisons, shame, and self-criticism:

> *We're moving away from that shame-based teaching environment, but we still have a lot of that within our clinical training. Shame-based learning is by no means gone and a lot of the times it's just under the surface rather than directly somebody yelling at you for not being good enough.* [Woman, aged 33 years]

Compassion was described as not being a primary or explicit focus in this milieu; rather, the emphasis was on knowledge of disease and learning how the system really works:

> *Yes, I think part of it that I never learned that [self-compassion], it's not part of training and medicine training magnifies mistakes. We have whole rounds about how you learn from how people, quote unquote, screwed up. So, I think that that's part of it, is that the culture of medicine doesn't necessarily lend itself to that.* [Man, aged 39 years]

Participants recalled how, upon graduating from medical school, as residents in PM&R, they often faced excessive workloads fulfilling the demands stipulated by their superiors, sacrificing their autonomy and sense of self:

> *I was a resident, so I had no control over. I was just told how much call I needed to take. It was different back then, there was payroll, which is the governing body to the residents, but I don't think it was as strict as it is now. The residents were fairly powerless* [Woman, aged 48 years]

Based on participant recollections, autonomy appeared inversely related to hierarchy within the medical education and practice they experienced. In a hierarchical structure such as this, individuals at lower levels, that is, residents, have limited control over their workload, schedules, and decision-making, as these are dictated by superiors. This lack of autonomy was described as resulting in residents prioritizing institutional demands over personal needs or values, leading to a diminished sense of self.

Only upon becoming fully qualified did a sense of autonomy return, at which point physiatrists in Canada often assume a leadership role in the multidisciplinary team:

> *How I really got better was I stopped being a resident and then I became a staff, and I had more autonomy. I think that is a factor to relieving burnout, being able to control your work schedules. Being able to say yes to some things and no to other things. I think that's actually how it got helped* [Woman, aged 37 years]

However, participants recounted that the reality is that the job continues to be dictated by hierarchical health systems, regulatory body oversight, and peer appraisal of performance:

> *But I'll say that even at [rehab hospital] there have certainly been instances where I feel like the response from the organisation in response to something that happened with an error that then caused harm was not, I think the kindest and most compassionate response* [Woman, aged 37 years]

Indeed, as Canadian physiatrists advance in their careers and move up the hierarchy, although in one sense autonomy is greater, on the other hand, their responsibilities and expectations also increase, creating a persistent tension between autonomy and normative performance standards.

Regardless, Canadian physiatrists of all stages struggle to provide the care they believe is necessary for their patients due to health system constraints and fear authoritarian reprisal for failing to meet expected high performance standards:

> *I know that in Ontario, people are always like, don't mess with Ontario College of Physicians because it's a big problem.* [Woman, aged 32 years]

Where burnout is concerned, regulatory bodies were seen as threatening due to their ability to "take away your license" to practice medicine:

> *The College of Physicians was not chill about it though. Because that was PGY2, right, and then when I graduated, you had to answer some questions about did you ever get time off because of this?...the College of Quebec they were just like, have you or are you in a situation where this will be a problem? There was no differentiating it and then they were like, we need a letter from your doctor saying that you are up to practise medicine* [Woman, aged 32 years]

XSL•FO
RenderX

This highly stressful situation may partly explain the difficulties Canadian physiatrists face in providing care in this hierarchical practice climate.

> *And now they're looking at even things like being rude in your private life as being a reportable offense. I mean, really? 24/7, we have to be models of perfection, or we could be reported and lose our livelihood, which would impact whether we can pay our bills. It's from all sides. To the point where I struggle with whether or not to encourage my own children to enter the profession. Yes.* [Woman, aged 49 years]

> *But the burden of care, I think more and more what happens earlier on is your lifeline is that licence to practise, and if that gets taken away, which seems always to be this looming threat, that if I get caught doing something that's not according to the College's mandate or whatever then basically I can't earn a living because my licence is off.* [Man, aged 60 years]

Participants recalled that, as medical students, despite altruistic motivations for entering PM&R, low hierarchical status, limited autonomy, and strict academic expectations fostered vulnerability and feelings of inadequacy, contributing to early burnout experiences:

> *Again, obviously I'm a woman so I have a bit of a lens there, but imposter syndrome, which I think is more common in women, but I still feel…* [Woman, aged 35 years]

Participants described that medical schools exert authority over those who were underperforming, dictating eligibility for module completion or exclusion from training. Self-care was not explicitly emphasized and seen as antithetical to the culture of medicine:

> *And rather than being harsh towards yourself when you make mistakes or whatever, it's being non-judgmental and kind to yourself pretty much…Well that's definitely not a part of medicine.* [Man, aged 39 years]

> *Well, I'm saying that I don't think that there is a curriculum in medicine that teaches people self-compassion and work-life balance.* [Woman, aged 58 years]

Participants described how this scenario continued into their PM&R residencies, where, as trainees, they had to meet performance standards based on long working hours and antisocial working patterns. These demands prevented them from spending time with friends and family, exercising, and attending to diet and sleep. Autonomy became restricted due to strict shift patterns, and as residents, they remained dependent on guidance from their superiors. During this stage, participants described how burnout loomed large as an ever-present threat:

> *I did burn out in PGY2 and that was the reason I left, is when I had no more compassion for my patients, like I had no empathy for them. I just didn't care. I think it was more me being exhausted and other things and then when my patients started pissing me off to be honest. It was just whenever they would talk to me about things, I was like, why are you complaining about this.* [Woman, aged 32 years]

Participants described that as they transitioned from being a trainee to becoming a fully qualified physiatrist, they took on a greater responsibility and became exposed to greater distress in their patients, "Being a physician, especially for a long time, especially dealing with nasty stuff, I think, if you do let that stuff get to you, it's going to be really hard to succeed in the long run" (Man, aged 41 years), and increasingly aware of health care system limitations, which was a source of moral distress:

> *I want to be compassionate for the people in front of me, whatever role I'm having, whether it's for students and residents or whether it's for the patients and families. But I'm coming up against either structural or system or contextual issues that don't let me do the thing that I think is the right thing to do, and that causes distress* [Woman, aged 43 years]

Risk of burnout was described by participants as ever-present as they struggle to balance growing personal and professional responsibilities. Constant oversight and stringent reporting requirements were seen to create a culture prioritizing protocols and metrics over patient-centered care, participants perceiving institutions to prioritize efficiency over quality in interpersonal interactions:

> *And when you're dealing with human beings, that's extremely difficult to individualize a program for somebody in a system where you're actually a datapoint.. success is measured by the number of patients sometimes just seen, not even healed, not even heard, not even spoken to.* [Woman, aged 49 years]

The chronic lack of resources, including inadequate staffing and time constraints, was described by participants as leaving them feeling pressured to prioritize logistical demands over holistic patient needs. Heavy administrative duties and systemic constraints were seen to undermine emotional openness and sensitivity to patient distress. In this context, participants described how they may come to view compassionate care as ancillary rather than a core element:

> *I think it's like going over and beyond for patients. And also, being there for them emotionally, even though that's not really part of your job description, but sometimes they just need a Kleenex and need somebody to give them a hug.* [Woman, aged 37 years]

Consequently, as Canadian physiatrists' advance in their careers within such a framework, there may be a gradual erosion of emphasis on compassionate care in favor of meeting normative peer appraisal, health care institution, and regulatory demands. Institutional wellness initiatives were perceived as misdirected, emphasizing personal responsibility for burnout and failing to address core cultural and systemic issues:

> *We can exercise, we can mediate. We know those things. Those are not the things that we need changes on. We've all made those changes already. We need*

XSL•FO
RenderX

> *the help in the change in the system. They're the system factors that get us down* [Woman, aged 49 years]

The Hidden Curriculum and Its Damaging Effects in Physical Medicine and Rehabilitation

A hidden curriculum, rooted in medical school and reinforced through professional role modeling, was described by participants as shaping behaviors and attitudes that can drive burnout and erode well-being across a physiatrist's career. Such a hidden curriculum was seen as subtly prioritizing professional achievement over compassion, system efficiency over patient-centered care, and physical treatments over addressing emotional distress.

> *I think compassion is really a hidden curriculum kind of thing, and depending on who you're exposed to, you might or might not get any direct instruction on how it is to be compassionate to people. [Compassion] is not really emphasized a lot in medicine compared to clinical expertise. Our recruitment process and training is directed around shutting down self-compassion.* [Woman, aged 33 years]

These implicit lessons were described as conflicting with the altruistic motivations that draw many physiatrists into medicine:

> *Yes, of course. If you're a human being, you're going to be emotionally affected by other people's struggles, your patients' struggles. And that's why you end up in medicine, right, because you really want to help people.* [Man, aged 45 years]

Participants described that in the hidden curriculum, experiencing burnout is seen as a sign of weakness. To be successful in this arena, participants described how they must meet peer and professional standards at all costs:

> *And then, some of that worry about what other clinicians are going to think about my decision making and choices, and how that's going to affect long-term relationships, both with the patient, but also with the other clinicians in the workplace.* [Woman, aged 33 years]

The stakes were described as high, margins for error small, and risk of shame apparent:

> *…If we mess up, it's a big deal…the culture of we expect a lot from our students and learners in, I find it's sometimes not realistic.* [Woman, aged 32 years]

Indeed, participants recounted how to meet the demands of their work, they often had to sacrifice parts of themselves:

> *I definitely have at different times throughout my career. Having relatively recently finished residency, trying to balance work as well as studying for the [regulatory board] exam. Trying to have any semblance of social life and I mean that very loosely, seeing your parents every now and then. It's not like I see my friends or do anything really for myself all that often.* [Woman, aged 35 years]

Maintaining this level of professional performance was described as exhausting, physically and emotionally, and to deal with this, physiatrists take steps to minimize the stress they are exposed to in their work.

Participants also described that, to be a successful physiatrist in Canada, this often involved maintaining emotional distance from their patients to avoid burnout and manage the high demands of the profession:

> *I find those very challenging because I'm just trying to sling through and get through my day treating physical problems. But then when you have these mental problems that make it more challenging to treat the physical problem, I think that takes a lot of time and effort and just makes your job more difficult and it requires more empathy which is draining.* [Man, aged 32 years]

And, although participants recognized that distress is both physical and emotional, many did not feel they had been well prepared through their medical education to deal with emotional distress prevalent among their patients:

> *I was downright angry at my medical education early on.. we never even talked about stuff like this, let alone how feelings and emotions get in the way of healthcare.* [Man, aged 60 years]

> *I am generally treating physical problems and often times peoples' psychological distress or mental health problems can worsen their physical symptoms and I just, that's not a population that I do well with.* [Man, aged 32 years]

Thus, physiatrists may view emotional distress as not within their purview, leading to compartmentalization of care:

> *And I think a lot of my colleagues, especially on the outpatient side in Canada, end up seeing a lot of run-of-the mill chronic pain, which ends up being people complaining a lot. I don't have to see a lot of that and so that helps me, because I actually see real suffering. Not real suffering, psychological suffering is still real, but I'm not trained to deal with the psychological suffering of things like chronic pain.* [Man, aged 41 years]

As such, Canadian physiatrists may become desensitized to the emotional needs of themselves and their patients, focusing primarily on efficiency and protocolized medicine. However, burnout can still result from the forces of this hidden curriculum, and results can be stark. For some, there is frustration and anger towards the system, as they feel there is no time to be compassionate due to the overwhelming demands of the job,

> *So, the things that get me burnt out would be like when I have a huge amount of work in a short period of time and I don't know how I'm going to be able to get it done. A lot of work and no time, and uncertainty about how I'm going to be able to get through.* [Man, aged 48 years]

XSL•FO
RenderX

Others might become disenfranchised completely, adopting the belief that fixing the broken health care system is not their responsibility:

> *No, I have no interest in doing that. It's too much of an uphill battle.. I would have to put all of my other interests and hobbies in life aside and have to focus on mobilising myself to participate in the process of improving Canadian healthcare, which I am completely unwilling to do.* [Man, aged 41 years]

Physiatrists Lead Multidisciplinary Teams and Focus on Adaptation Rather Than Cure

Participants in this study described the work that they do in PM&R as about leading teams, with a focus on adaptation rather than cure, a source of optimism, enthusiasm, and reward for physiatrists, with many iterating that this was where they derived most joy in their work, and felt most able to help:

> *I think that's why rehab is so rewarding. I know that [short-term wins] make a big difference… It's often those little extra moments that you take.* [Woman, aged 34 years]

Participants emphasized helping patients adapt to disabling conditions and improve quality of life, rather than "fix" a diagnosis. Participants described that involving patients in their care fosters control and confidence but requires physiatrists to have a deep understanding of individual circumstances, often eliciting profound and painful empathy in the process.

> *Really severe impairments. Patients that are so injured or have terrible conditions that are very debilitating. It's very easy to find compassion and empathy for those people. When you also see people who lose a limb, or have a, are paralysed or something. It's very easy to build your compassion up and feel for that person because it's terrible. It's subjectively an awful thing.* [Man, aged 32 years]

> *I also try not to absorb the terrible situations all the time and say it's a terrible situation, but I think subconsciously, because I see it all the time and they're all terrible situations. Usually people that are coming to our services are not coming because of a good situation.* [Woman, aged 47 years]

However, using the knowledge derived from a deep understanding of patient distress, participants described how, as physiatrists, they sought to create tailored care plans that optimize functional independence in their patients. This holistic approach was seen as fundamental to the conceptualization of PM&R, but also exposes physiatrists to difficult emotions:

> *I think the patients often express frustration in different ways… It's not as glamorous or as immediately effective as they had anticipated. And so, that often comes out to us as, 'why aren't you fixing this and why is that not working?* [Woman, aged 47 years]

Participants spoke about the need to boundary themselves in relation to the emotions of others, and to see the limitations of the health care system in which they work with realism and detachment, otherwise they might be quickly overcome by cynicism and burnout. Indeed, participants perceived that forbearance was expected from them in the face of abuse and that, in the predominant culture, their emotional well-being was not a priority:

> *We take a lot of verbal, emotional, sometimes even physical abuse. The expectation is often to just keep providing care and keep being compassionate. But that's not so simple in practice. I do find that can be a real challenge.* [Woman, aged 34 years]

> *We should be, somebody, at the end of the day has said, you know what, I heard about that patient, it was rough. How are you doing? That must be a really tough thing that you had to deal with, how are you feeling? I think that everybody just presumes that they're… Everybody else is stressed out, how is that any different than yours.* [Woman, aged 47 years]

Breaking bad news was another area of PM&R, which participants described as perennially difficult in an emotional sense, especially when patient and family expectations were unrealistic, or were not met, and this experience undermined physiatrists' sense of proficiency:

> *I think sometimes [patients] want a magic pill to make all their problems go away without any side effects of any work on their part… That is that part that I find the hardest. When you strive to go above and beyond and give people ten things they can do but they have excuses… they don't ever get better.* [Woman, aged 35 years]

As physiatrists coordinate and guide multidisciplinary specialist rehabilitation teams, including deliberating over access to care, participants described how they could often be a focal point for venting from other health care providers:

> *Eventually my team pushed me to act on it, like could you do something to potentially protect her, but me knowing that there really isn't an ethical or legal way to go about making this decision on behalf of the patient, I felt very uncompassionate towards my team in that situation.* [Man, aged 35 years]

Participants also spoke about a sense of not being appreciated broadly, by patients, institutions, system leaders, and even peers. They described how feeling unappreciated by others left them feeling isolated, disconnected, and disillusioned, "Nobody, from a physiatry perspective, has said 'we appreciate what you do.' For that reason, I've tended to work more collaboratively with my colleagues in other specialties rather than my own colleagues here, which is maybe unhealthy" (Woman, aged 47 years), and disillusioned with the health care system at large. When feeling this way, participants were less inclined to listen to the nuances in patient stories and more inclined to objectify patients:

> *seeing them as their diagnosis. Like, oh that's the cervical myelopathy, and seeing it as their list, the tick list of things that you've got to do on rounds, to follow up and you're just doing it to make sure that you've done your job.* [Woman, aged 47 years]

XSL•FO
RenderX

Bureaucracy Alongside Hierarchy

Bureaucracy and hierarchy were described by participants as contributing to burnout in PM&R by shifting focus from patient care to rules, paperwork, and complex approval processes, particularly for disability income support. This pervasive administrative burden was felt to foster powerlessness, frustration, and demoralization, subtly eroding autonomy and engagement more insidiously than direct authoritarian control.

> *Stupid rules that make no sense and that take a lot of time and effort to work around. So, dealing with the bureaucracy of stupid rules... And that causes a lot of frustration and wasted energy that could be used on other things. Stupid rules... we need to have some ways to be able to say, some direct lines to be able to call somebody and say, this is why this one shouldn't follow the rules and have someone at the other end being able to say okay. I get it. Yes.* [Woman, aged 55 years]

Participants perceived rigid rule enforcement as nonsensical, adding unnecessary distress to patients and physiatrists alike. In this context, participants described feeling disconnected from their initial sense of purpose, their professional identity overshadowed by the sheer volume of administrative work, which they felt obliged to do in their derived social and professional roles:

> *Oh, I think the paperwork. So, usually, rehab forms or disability forms are at least six to ten, maybe 20 pages long.* [Woman, aged 49 years]

This chronic exposure to perceived impersonal and inflexible bureaucratic hurdles, rather than direct authoritarian mandates, appeared to truly wear down physiatrists, contributing significantly to burnout:

> *Certainly, the burden of documentation, trying to get down all the documents and organize, sometimes some people, they'll need forms filled out, and those can be very time-consuming and at times frustrating. Because you feel like you're writing down things that are already in the record but you have to write them down for the completion of a form for a referral or completion of a form for a disability tax credit, I think documentation definitely.* [Man, aged 61 years]

Discussion

Summary of Main Findings

Using a constructed grounded theory approach, based on retrospective accounts of lived experience among Canadian physiatrists, this study developed a conceptual model of burnout in PM&R, highlighting how medical hierarchy—both explicit and via the hidden curriculum—insidiously shapes professional identity from medical school throughout the practice of PM&R. Unrealistic performance standards, emphasis on academic achievement over compassion, and lack of self-care training erode well-being, foster shame, and limit Canadian physiatrists' ability to respond empathetically. Coupled to this, excessive paperwork, complex administrative processes, and strict regulatory compliance wear Canadian physiatrists down over time.

While PM&R's focus on adaptation and holistic care is fulfilling, burnout undermines Canadian physiatrists' capacity to lead multidisciplinary teams effectively, negatively impacting both colleagues and patients.

Comparison With Existing Literature

No previous studies have used grounded theory to explore burnout in PM&R. However, findings from previous studies using grounded theory to explore burnout in health care providers share similarities with ours. For example, studies in primary care and stroke medicine also suggested systemic issues such as excessive workloads, shifting job roles, and misalignment between personal and institutional values, lead to burnout [31,32]. Likewise, professional dissonance in these contexts also led to feelings of demoralization and a sense of being undervalued [31]. Among nurses, distressing relational aspects of health care also featured in conceptualizations of burnout, particularly during personal interactions, shaping care provided [33], and social support played a mitigating role, affecting emotional responses and some components of attitudinal reactions [33]. In our study, Canadian physiatrists highlighted how they felt ill-prepared for the emotional demands of the work of PM&R, and professionally isolated amid a culture of expected forbearance in the face of emotional distress, where support for their own distress was limited.

Limited literature addresses the origins and sustaining factors of burnout in PM&R. A recent US qualitative study identified similar themes to ours, including the impact of perfectionism, stigma around disclosing distress, and burdensome administrative tasks. Participants advised younger physiatrists to maintain "agency," set boundaries, and practice self-compassion to prevent burnout. These comparisons are interesting but derive from the unusual context of the COVID-19 pandemic, and are not in relation to origins or factors sustaining burnout in Canadian physiatrists [34].

With regards to the overarching concept identified in our study, of hierarchy in medicine, this is a well-recognized phenomenon [35], but has not been described previously in relation to burnout in PM&R. Hierarchy in medicine has been described as a means of imbuing power to make and enforce decisions based on knowledge and role that allows for autonomy and self-governance in the profession. Hierarchies in medicine can be conceptualized in both negative, "dysfunctional," and positive, "functional," terms. In dysfunctional medical hierarchies, those of lower status in the hierarchy are subject to lower health [36], trainees lack psychological safety and agency, and can be subject to outright abuse. In this way, hierarchies can be harmful, not just to those in low-status positions, that is, medical students, residents, but also to patients, collaboration among multiprofessional teams, which is highly relevant in PM&R.

However, it is important to note that hierarchy may also be beneficial, at the macro level, with enhanced performance and evolvability of modular systems [37], but also for the meso (institution, team), and micro (individual) levels, with role and

XSL•FO
RenderX

expectation clarity, enabling a social order that serves a greater purpose, which can support a culture of efficiency, and patient safety where the most responsible clinician is the one making critical decisions regarding patient care, particularly when risk is high, and time is pressured. However, to do this in a way that supports social justice and compassionate care likely requires openness and clarity around normative values and conscious and accountable oversight [38]. If not, hierarchy can also be the basis upon which a hidden curriculum, that is, social norms, can develop, including those around punishment, reward, and progression upwards within the hierarchy. Indeed, medical students appear to actively seek to decipher which behaviors their supervisors value most and to preferentially cultivate these [35]. The implications for the development of compassion, or less desirable behaviors [39], at the individual and organizational levels thus seem clear.

Our theoretical framework, the looking-glass self, provided a helpful lens to explore the origins and factors sustaining burnout in PM&R. Physiatrists in our study relayed how the culture of academic elitism in medicine can lead to adverse social comparisons, imposter syndrome, self-criticism, and shame. In keeping with the looking-glass self precept, Gilbert defines shame as a form of undesired self and a form of social disconnection, but differentiates self-shame from social shame perceived to arise from the judgment (and rejection) of others [40]. Unfortunately, shame is a well-described phenomenon in medical education, which can relate to mistreatment, substandard academic performance, social comparisons, and can be reinforced by hidden curricula. Bynum [41] suggests that shame within medical education can be prevented by explicitly eliminating shame-based teaching strategies, creating a culture of inclusion, and facilitating a mindset of growth in learners. Among internal medicine trainees in the US, exposure to a hidden curriculum (unprofessional conduct, including humiliation by colleagues and patients, disrespectful behavior towards patients) has been found to correlate significantly with all 3 core domains of burnout [42]. Indeed, research into hidden curricula in medical education has previously identified that assessment criteria (which drive what medical students see as being prioritized) do not emphasize compassion as a competency being assessed [43]. We found that self-compassion is seen by Canadian physiatrists as antithetical to the medical training they received. This is unfortunate, as training in self-compassion can help health care providers better self-care, improve compassion for others, and may be protective against burnout with moderate to large effect sizes [44], whilst Compassionate Mind Training may be particularly beneficial in addressing shame-based disorders associated with burnout [45].

Strengths and Limitations

This is the first qualitative study examining the origins and sustaining factors of burnout in PM&R. Strengths include a large, diverse Canadian sample, rigorous methods, and varied team perspectives. Limitations include the majority of participants being full-time urban physiatrists from Ontario, a lack of regional analysis, and a cross-sectional design, which may miss formative early-career experiences and long-term patterns in burnout development. No PM&R residents being interviewed is a notable limitation because many of participants indicated origins of burnout date back to experiences in medical school and residency. We used Canadian PM&R leaders to identify participants, consistent with purposive sampling but potentially introducing selection bias. Self-reported interviews may also be affected by recall or social desirability biases. As a preliminary grounded theory, the framework is at an early stage, may not capture all nuances of burnout, and requires further empirical validation. Findings are specific to Canadian PM&R and may not generalize elsewhere.

Future Directions

Future research should explore the views and experiences of PM&R residents and specifically seek to delve deeper into how intersectional factors influence burnout among physiatrists from marginalized backgrounds within Canada and elsewhere. Research should consider how aspects such as immigration status, ethnic minority status, religious beliefs, and sexual orientation intersect with burnout experiences. Marginalized physiatrists might face unique stressors related to their identity that impact their professional experiences differently than those of their peers. For instance, immigrants might deal with additional challenges related to cultural integration and professional validation, while members of minority ethnic or religious groups might encounter systemic biases or discrimination. It would also be very useful to explore how regional variation in normative organizational culture, health care systems, and wellness practices influences the experiences of physiatrists with respect to hierarchy, hidden curricula, bureaucracy, and burnout.

Conclusions

The origins of burnout in Canadian physiatrists relate to a hierarchical performance paradox that appears to be seeded early in their careers. Budding Canadian physiatrists cite altruistic motivations for joining the specialty, but it is difficult to maintain compassion for oneself and others amidst a hidden curriculum that prioritizes academic elitism over compassionate care. Risk of burnout among Canadian physiatrists is described as peaking in residency but is also ever-present due to the emotionally demanding nature of the work for which they feel ill-prepared, fragmented health care systems, and societal bureaucracy around disability. Burnout can be a source of shame for Canadian physiatrists and is perceived as having the potential for severe professional repercussions. Canadian physiatrists value positive peer appraisal and emotional support (ie, a supportive community); this provides a potential inroad to addressing some of the factors that are responsible for burnout in the specialty.

Acknowledgments

Generative artificial intelligence was not used in any phase of this research study.

XSL•FO
RenderX

Data Availability

Study data will be made available on reasonable request to the corresponding author.

Funding

This work was funded by a research fellowship from Associated Medical Services (Canada; no grant number to report). The funder had no role in the study design; collection, analysis, and interpretation of data; writing of the paper; and decision to submit for publication.

Authors' Contributions

RS, AF, LR, MB, EC, CSG, SM, and KK conceptualized the study. EC conducted data collection. RS, EC, SP, and KK undertook all analyses. CSG, SM, and MB provided senior methodological guidance. RS and KK scripted the manuscript. All authors reviewed, edited, and approved the manuscript.

Conflicts of Interest

Coauthor SM is the current Editor in Chief of the JMIR Rehabilitation and Assistive Technology. Senior author, KK, is a current member of the editorial board of JMIR Rehabilitation and Assistive Technology. None of the authors has any personal financial interests related to the subject matter discussed in this manuscript.

Multimedia Appendix 1
Semistructured interview topic guide.
[DOCX File , 16 KB - rehab_v13i1e80499_app1.docx]

References

1. Freudenberger HJ. Staff burn-out. J Soc Issues 1974;30(1):159-165. [doi: 10.1111/j.1540-4560.1974.tb00706.x]
2. Maslach C. A multidimensional theory of burnout. In: Cooper CL, editor. Theories of Organizational Stress. England: Oxford University Press; 1998:68-85.
3. Rotenstein LS, Torre M, Ramos MA, Rosales RC, Guille C, Sen S, et al. Prevalence of burnout among physicians: a systematic review. JAMA 2018;320(11):1131-1150 [FREE Full text] [doi: 10.1001/jama.2018.12777] [Medline: 30326495]
4. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med 2018;283(6):516-529 [FREE Full text] [doi: 10.1111/joim.12752] [Medline: 29505159]
5. West CP, Dyrbye LN, Erwin PJ, Shanafelt TD. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. Lancet 2016;388(10057):2272-2281. [doi: 10.1016/S0140-6736(16)31279-X] [Medline: 27692469]
6. Makowski MS, Trockel M, Paganoni S, Weinstein S, Verduzco-Gutierrez M, Kinney C, et al. Occupational characteristics associated with professional fulfillment and burnout among US physiatrists. Am J Phys Med Rehabil 2023;102(5):379-388. [doi: 10.1097/PHM.0000000000002216] [Medline: 37076955]
7. Bateman EA, Viana R. Burnout among specialists and trainees in physical medicine and rehabilitation: a systematic review. J Rehabil Med 2019;51(11):869-874 [FREE Full text] [doi: 10.2340/16501977-2614] [Medline: 31608964]
8. Simpson R, Robinson L. Rehabilitation after critical illness in people with COVID-19 infection. Am J Phys Med Rehabil 2020;99(6):470-474 [FREE Full text] [doi: 10.1097/PHM.0000000000001443] [Medline: 32282359]
9. Negrini S, Selb M, Kiekens C, Todhunter-Brown A, Arienti C, Stucki G, 3rd Cochrane Rehabilitation Methodology Meeting participants. Rehabilitation definition for research purposes. a global stakeholders' initiative by cochrane rehabilitation. Neurorehabil Neural Repair 2022;36(7):405-414 [FREE Full text] [doi: 10.1177/15459683221093587] [Medline: 35574944]
10. International classification of functioning, disability and health: ICF. World Health Organization. 2001. URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2026-01-21]
11. O'Brien R, Wyke S, Guthrie B, Watt G, Mercer S. An 'endless struggle': a qualitative study of general practitioners' and practice nurses' experiences of managing multimorbidity in socio-economically deprived areas of Scotland. Chronic Illn 2011;7(1):45-59. [doi: 10.1177/1742395310382461] [Medline: 20974642]
12. Sliwa JA, Clark GS, Chiodo A, Kinney CL, Raddatz MM, Francisco GE, et al. Burnout in diplomates of the American board of physical medicine and rehabilitation-prevalence and potential drivers: a prospective cross-sectional survey. PM R 2019;11(1):83-89. [doi: 10.1016/j.pmrj.2018.07.013] [Medline: 30703291]
13. Teel J, Reynolds M, Bennett M, Roden-Foreman JW, McShan E, Hamilton R, et al. Secondary traumatic stress among physiatrists treating trauma patients. Bayl Univ Med Cent Proc 2019;32(2):209-214 [FREE Full text] [doi: 10.1080/08998280.2018.1559694] [Medline: 31191130]
14. Francisco GE, Chae JC, DeLisa JA. Physiatry as a primary care specialty. Am J Phys Med Rehabil 1995;74(3):186-192. [doi: 10.1097/00002060-199505000-00002] [Medline: 7779328]
15. Mann S. What life care planners need to know about the professional discipline of physician physiatrist. J Life Care Plann 2019;17(1). [doi: 10.70385/001c.151587]

XSL•FO
RenderX

16. Heath H, Cowley S. Developing a grounded theory approach: a comparison of Glaser and Strauss. Int J Nurs Stud 2004;41(2):141-150. [doi: 10.1016/s0020-7489(03)00113-5] [Medline: 14725778]
17. Charmaz K. Constructivist grounded theory. J Posit Psychol 2016;12(3):299-300. [doi: 10.1080/17439760.2016.1262612]
18. Charmaz K. Grounded Theory for Qualitative Research: A Practical Guide. Thousand Oaks, CA: Sage Publications, Inc; 2015:53-84.
19. Charmaz K. Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. Thousand Oaks, CA: Sage; 2006.
20. Chun Tie Y, Birks M, Francis K. Grounded theory research: a design framework for novice researchers. SAGE Open Med 2019;7:2050312118822927 [FREE Full text] [doi: 10.1177/2050312118822927] [Medline: 30637106]
21. Cooley CH. The looking-glass self. In: Human Nature and the Social Order. New York: Charles Scribner's Sons; 1902:126-128.
22. Corbin J. Grounded theory. J Posit Psychol 2017;12(3):301-302. [doi: 10.1080/17439760.2016.1262614]
23. Physical medicine and rehabilitation profile. Canadian Medical Association. 2019. URL: https://www.cma.ca/sites/default/files/2019-01/physical-med-rehab-e.pdf [accessed 2026-01-21]
24. Woike BA. Content coding of open-ended responses. In: Handbook of Research Methods in Personality Psychology. New York: The Guilford Press; 2007:292-307.
25. Belgrave LL, Seide K. Coding for grounded theory. In: The SAGE Handbook of Current Developments in Grounded Theory. Thousand Oaks, CA: Sage Publications, Inc; 2019:167-185.
26. Vollstedt M, Rezat S. An introduction to grounded theory with a special focus on axial coding and the coding paradigm. In: Compendium for Early Career Researchers in Mathematics Education. Cham, Switzerland: Springer; 2019:81-100.
27. Thornberg R, Charmaz K. Grounded theory and theoretical coding. In: The SAGE Handbook of Qualitative Data Analysis. Thousand Oaks, CA: SAGE Publications; 2014:153-169.
28. Aldiabat KM, Le Navenec CL. Data Saturation: the mysterious step in grounded theory method. TQR 2018;23(1):246-261. [doi: 10.46743/2160-3715/2018.2994]
29. Goundar PR. Researcher positionality: ways to include it in a qualitative research design. Int J Qual Methods 2025;24:16094069251321251. [doi: 10.1177/16094069251321251]
30. Hafferty FW, O'Donnell JF. The Hidden Curriculum in Health Professional Education. Hanover, NH: Dartmouth College Press; 2015.
31. Agarwal SD, Pabo E, Rozenblum R, Sherritt KM. Professional dissonance and burnout in primary care: a qualitative study. JAMA Intern Med 2020;180(3):395-401 [FREE Full text] [doi: 10.1001/jamainternmed.2019.6326] [Medline: 31904796]
32. Lee JV, Scott A, Osbun J, Zipfel G. Impact of stroke call on career satisfaction and burnout for academic neurointerventionalists: a grounded theory model. World Neurosurg 2021;151:e552-e564. [doi: 10.1016/j.wneu.2021.04.091] [Medline: 33933697]
33. Rafii F, Oskouie F, Nikravesh M. Factors involved in nurses' responses to burnout: a grounded theory study. BMC Nurs 2004;3(1):6 [FREE Full text] [doi: 10.1186/1472-6955-3-6] [Medline: 15541180]
34. Amano A, Makowski MS, Trockel MT, Menon NK, Wang H, Sliwa J, et al. A qualitative study of strategies to improve occupational well-being in physical medicine and rehabilitation physicians. Am J Phys Med Rehabil 2024;103(8):674-684. [doi: 10.1097/PHM.0000000000002555] [Medline: 38838100]
35. Vanstone M, Grierson L. Thinking about social power and hierarchy in medical education. Med Educ 2022;56(1):91-97. [doi: 10.1111/medu.14659] [Medline: 34491582]
36. Marmot MG. Status syndrome: a challenge to medicine. JAMA 2006;295(11):1304-1307. [doi: 10.1001/jama.295.11.1304] [Medline: 16537740]
37. Mengistu H, Huizinga J, Mouret J, Clune J. The evolutionary origins of hierarchy. PLoS Comput Biol 2016;12(6):e1004829 [FREE Full text] [doi: 10.1371/journal.pcbi.1004829] [Medline: 27280881]
38. de Zulueta P. Developing compassionate leadership in health care: an integrative review. J Healthc Leadersh 2016;8:1-10 [FREE Full text] [doi: 10.2147/JHL.S93724] [Medline: 29355200]
39. Ceñido JF, Obanor ON. Losing face: narcissism in medical training. J Psychiatry Psychiatr Disord 2020;04(04):188-190. [doi: 10.26502/jppd.2572-519x0105]
40. Gilbert P. Shame, humiliation, guilt, and social status: the distress and harms of social disconnection. In: Compassion Focused Therapy. UK: Routledge; 2022:122-163.
41. Bynum WE, Varpio L, Lagoo J, Teunissen PW. 'I'm unworthy of being in this space': the origins of shame in medical students. Med Educ 2021;55(2):185-197. [doi: 10.1111/medu.14354] [Medline: 32790934]
42. Billings ME, Lazarus ME, Wenrich M, Curtis JR, Engelberg RA. The effect of the hidden curriculum on resident burnout and cynicism. J Grad Med Educ 2011;3(4):503-510 [FREE Full text] [doi: 10.4300/JGME-D-11-00044.1] [Medline: 23205199]
43. Wright S, Boyd VA, Ginsburg S. The hidden curriculum of compassionate care: can assessment drive compassion? Acad Med 2019;94(8):1164-1169. [doi: 10.1097/ACM.0000000000002773] [Medline: 31033597]

XSL•FO
RenderX

44. Neff KD, Knox MC, Long P, Gregory K. Caring for others without losing yourself: an adaptation of the mindful self-compassion program for healthcare communities. J Clin Psychol 2020;76(9):1543-1562. [doi: 10.1002/jclp.23007] [Medline: 32627192]
45. Gilbert P. Compassionate mind training: key themes. In: Compassion Focused Therapy. Milton Park UK: Routledge; 2022:273-312.

Abbreviations

OR: odds ratio
PM&R: Physical Medicine and Rehabilitation

Edited by J Sarvestan; submitted 11.Jul.2025; peer-reviewed by P Puram, S Mitchell; comments to author 21.Nov.2025; accepted 15.Jan.2026; published 26.Feb.2026.

Please cite as:
Simpson R, Cohen E, Posa S, Wasilewski M, Feinstein A, Bayley M, Robinson L, Munce S, Steele Gray C, Kokorelias K
Understanding the Origins and Factors of Burnout in Physical Medicine and Rehabilitation: Grounded Theory Analysis
JMIR Rehabil Assist Technol 2026;13:e80499
URL: https://rehab.jmir.org/2026/1/e80499
doi:10.2196/80499
PMID:

©Robert Simpson, Eva Cohen, Stephanie Posa, Marina Wasilewski, Anthony Feinstein, Mark Bayley, Larry Robinson, Sarah Munce, Carolyn Steele Gray, Kristina Kokorelias. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Barriers and Facilitators in the Implementation of Digital, Blended, and Analog Evidence-Based Osteoarthritis Care: International Qualitative Expert Interview Study

Theresa Sophie Busse[1,2], Prof Dr; Julia Nitsche[3], Dr rer medic; Ina Carola Otte[2,4], Prof Dr; Nino Chikhradze[2,5], Dr rer medic; Barbara Jömann[2], Dr med; Horst Christian Vollmar[2], Prof Dr med; Jennifer Bosompem[2], MSc

[1]Digital Health, Department of Human Medicine, Faculty of Health, Witten/Herdecke University, Alfred-Herrhausen-Str. 50, Witten, Germany

[2]Institute of General Practice and Family Medicine (AM RUB), Medical Faculty, Ruhr University Bochum, Bochum, Germany

[3]Didactics and Educational Research in Health Science, Department of Human Medicine, Faculty of Health, Witten/Herdecke University, Witten, Germany

[4]Health Services Research, Institute for Diversity Medicine, Medical Faculty, Ruhr University Bochum, Bochum, Germany

[5]Institute for Diversity Medicine, Medical Faculty, Ruhr University Bochum, Bochum, Germany

Corresponding Author:
Theresa Sophie Busse, Prof Dr
Digital Health, Department of Human Medicine, Faculty of Health, Witten/Herdecke University, Alfred-Herrhausen-Str. 50, Witten, Germany

Abstract

Background: Evidence-based nonsurgical osteoarthritis (OA) care programs improve patient outcomes and may reduce the burden of disease. Despite their demonstrated effectiveness, their implementation into routine health care remains limited across health care systems. Understanding factors that facilitate or hinder implementation is essential to support broader adoption and long-term sustainability worldwide.

Objective: This study aimed to explore factors that facilitate or hinder the implementation of evidence-based nonsurgical OA care programs. Specifically, it examined implementation across different health care settings, implementation phases, and analog, digital, and blended care models to identify transferable insights for implementation in different national and regional contexts.

Methods: We conducted a qualitative expert interview study with 12 international experts involved in evidence-based OA care programs implemented across 6 countries. Semistructured interviews explored experiences with program development, implementation, and further development of analog, digital, and blended care models. Interview transcripts were analyzed using qualitative content analysis according to Kuckartz, using deductive and inductive category development.

Results: Key barriers included limited funding and reimbursement structures, insufficient political support, organizational complexity, and legal or regulatory requirements. Important facilitators comprised strong leadership, interdisciplinary collaboration, structured communication, adaptation to local health care contexts, and the use of existing international program structures. Participants further emphasized that implementation requirements varied across the preparation, initial implementation, and consolidation phases, highlighting the need for phase-specific implementation strategies.

Conclusions: Successful implementation of evidence-based nonsurgical OA care programs requires sustainable funding, stakeholder engagement, context-specific adaptation, and structured implementation planning across all implementation phases. Existing international program structures and collaboration may facilitate implementation across health care systems while reducing development efforts. Future implementation initiatives should consider phase-specific requirements and local contextual factors to improve long-term adoption and sustainability.

(JMIR Rehabil Assist Technol 2026;13:e85941) doi:10.2196/85941

KEYWORDS

physiotherapy; blended care; implementation science; osteoarthritis; osteoarthrosis; musculoskeletal diseases; conservative treatment

Introduction

Prevention and health promotion approaches are becoming increasingly important in health care. One area in which such approaches have shown promising results is the management of knee and hip osteoarthritis (OA). This musculoskeletal disease affects up to 40% of people older than 65 years [1], leads to high costs [2], and causes a burden on patients.

XSL•FO
RenderX

For mild to moderate OA, guideline-recommended treatment is nonsurgical and combines patient education, exercises, and weight reduction [3-5]. These components are increasingly delivered through structured OA treatment programs, spanning analog, digital, and blended care models implemented at regional, national, and international levels. Such programs may be developed specifically for OA or adapted from interventions targeting other musculoskeletal conditions. Evidence suggests that these approaches are safe, effective, cost-efficient, and suitable for routine care [3].

Despite growing evidence and international interest, the integration of OA programs into routine health care remains limited [6]. Consequently, patient access outside research settings is often restricted [6,7]. The limited uptake reflects the continuing challenge of translating research findings into clinical practice [4,7].

Despite growing evidence supporting OA care programs, important knowledge gaps remain regarding how these interventions can be successfully implemented across different health care systems, delivery models, and implementation stages. To address this gap, we conducted qualitative interviews with experts involved in established OA care programs. The study addresses the following research question: What factors influence the implementation of evidence-based OA treatment programs across different health care settings, implementation stages, delivery formats (digital, analog, and blended), and stakeholder groups?

We explored barriers and facilitators to implementation across health care settings, implementation stages, delivery formats, and stakeholder groups.

By synthesizing expert perspectives across multiple international programs, this study aims to provide practice-oriented insights into implementing evidence-based OA care in routine practice and contributes to the broader implementation science literature [8].

Methods

The reporting of the study results follows the COREQ (Consolidated Criteria for Reporting Qualitative Research) framework [5].

Study Design

The qualitative approach was chosen due to its suitability for exploring contextual and implementation-related factors that cannot be fully captured through quantitative methods. We conducted semistructured interviews with experts in the field of treatment programs for knee and hip OA as well as back pain.

To identify relevant programs, we first carried out a rapid review focusing on conservative treatment programs for knee and hip OA [7]. Based on this rapid review, we assembled a list of existing approaches and consulted international experts to gain insights into national characteristics and contextual influences.

In addition to OA programs, the review identified one back pain program that had been developed through adaptation of an established OA program. We included this program because it represented a unique example of how implementation strategies and program structures can be transferred across musculoskeletal conditions. Its inclusion enabled exploration of implementation processes beyond a single diagnosis while maintaining a focus on comparable nonsurgical, exercise-based care models.

The selected interventions additionally provided insights into how treatment concepts have been adapted across countries and how care models differ between analog, digital, and blended care [9].

Program Overview

The included programs differed in their form of delivery, implementation strategies, and national contexts. Better Management of Patients with OsteoArthritis served as the foundation for the subsequent development of GLA:D (Good Life with osteoArthritis in Denmark) [10,11], while Joint Academy emerged as a digital adaptation designed to improve accessibility and overcome barriers to participation in conventional OA care [12]. Studies accompanying the GLA:D program showed a significant reduction in the risk of being unable to work due to knee or hip OA, as well as improvements in pain intensity, physical resilience, and everyday activity [6,11].

Although differing in form of delivery and organizational structure, the programs shared a common evidence-based treatment philosophy.

Recruitment and Participants

JB and TSB designed the recruitment strategy to target participants with specific expertise in implementing OA treatment programs for knee and hip. Eligible participants were professionals who had been directly involved in the development, coordination, implementation, or scientific evaluation of one of the identified programs. The interviewers (JB and TSB) had no prior professional relationship with the participants before recruitment except for one participant, with whom they had had occasional contact during a joint project in the past. This did not affect the interview situation.

The participants were informed that both researchers are currently working on the facilitating and inhibiting factors in implementation research and about the insights they had gained from the rapid review mentioned above.

In the beginning, JB and TSB compiled a list of all OA treatment programs for knee and hip that had been identified worldwide in the previously mentioned review. JB and TSB then identified and contacted the program coordinators responsible for the development and implementation of these programs. In addition, they expanded recruitment through expert recommendations provided by interview participants (snowball sampling). Recruitment continued until additional interviews were considered unlikely to provide substantially new insights regarding the implementation processes under investigation.

The research team invited all identified candidates via email to participate in the study. Candidates were asked to complete a short questionnaire regarding sociodemographic information.

XSL•FO
RenderX

In total, the research team contacted 30 experts and one additional contact email address for a specific treatment program. Twelve experts agreed to participate in the study. Recruitment and interviews were conducted between August and October 2023. The final sample was considered appropriate because it captured experts from multiple countries, health care settings, and implementation roles, providing sufficient variation to address the study aim.

Data Collection

Interviews were conducted in English or German, depending on the interviewee's language preference. Accordingly, we developed the interview guides in both languages and ensured full content equivalence across both versions.

The interview guides were reviewed by an orthopedic surgeon for technical adequacy and by a team of research assistants from the Institute of General Practice and Family Medicine at Ruhr University Bochum during a qualitative research workshop. This review process focused on content relevance, comprehensibility, and alignment with the study objectives. Revisions were made iteratively based on the feedback received and discussed within the group of authors.

The semistructured interview guides (see Multimedia Appendix 1) focused on 4 key thematic areas:

1. Background information about the treatment program and the expert's role within it;
2. Description of the implementation process;
3. Experiences in program development since implementation; and
4. Considerations, expectations, and wishes for future development of the program.

Interviews were conducted by 2 female researchers (TSB: doctoral degree in health sciences; JB: MSc in health economics), both trained in qualitative research and employed at the Institute of General Practice and Family Medicine at Ruhr University Bochum at the time of the study. In most cases, both researchers attended the interviews—one leading the interview and the other monitoring the process and taking notes. TSB and JB took turns in this role to ensure consistent quality.

Neither interviewer was involved in the development or implementation of any of the included OA treatment programs. This helped to reduce potential conflicts of interest and enabled participants to discuss implementation experiences from an external perspective. Throughout data collection, the interviewers documented contextual observations and reflected on potential influences on the interview process in field notes.

All interviews were conducted via videoconferencing (using Zoom by Zoom Communications), audio-recorded, pseudonymized, and transcribed by a professional transcription service. German interviews were transcribed verbatim in German, just as English interviews were transcribed verbatim in English. The interviews were always analyzed in the original language. German quotations selected for publication were translated into English by the research team, and the translations were checked by 2 different authors.

In one case, the formal interview was preceded by a detailed preliminary conversation, which the participant had requested before agreeing to the interview. During this conversation, the participant described his role in the program. The participant consented to audio recording only after this preliminary exchange and had limited time available. To avoid redundancy and respect these constraints, the thematic first block of the interview guide was omitted during the recording. This interview was still included in the analysis, as 3 researchers (JB, JN, and TSB) independently agreed that the data provided sufficient contextual richness. The preliminary conversation was not included as primary interview data.

Data Analysis

We analyzed the interview transcripts using qualitative content analysis following the approach described by Kuckartz [13]. Interviews were transcribed shortly after completion of data collection and prepared for analysis. The formal coding and category development were conducted between September and November 2024 using MAXQDA software by VERBI GmbH (TSB: MAXQDA 2024; JN: MAXQDA 2022). The time lapse between data collection and analysis is attributable to the authors' new professional positions and the subsequent conclusion of the project, which prevented them from working on the project earlier, and is not due to changes in the dataset. Coding was performed in English. Two researchers (TSB and JN), both holding doctoral degrees in health care science and trained in qualitative research, independently coded all interviews. Discrepancies were discussed during regular analytical meetings. Where differences in interpretation occurred, both researchers revisited the original transcript segments and agreed on a final coding decision through discussion and consensus.

We applied a structured content-analytic approach combining deductive and inductive category development. Initial main categories were derived deductively from the interview guide, while additional categories were generated inductively from the data. Throughout the analysis, the coding framework was iteratively refined and continuously compared across interviews to ensure analytical consistency.

To enhance credibility, a third researcher (JB) with experience in qualitative content analysis reviewed the final category system and code tree for plausibility, coherence, and completeness. Feedback from this review informed the final interpretation and refinement of the category system, which was finalized in March 2025.

Methodological rigor was supported through independent coding, regular consensus meetings, and review of the final category system by a third researcher. Field notes were used to contextualize interview data and support interpretation.

Ethical Considerations

The study received a favorable ethics vote from the Ethics Committee of Ruhr University Bochum (register number 21-7438) in February 2022.

Interested individuals received detailed study information outlining the study's objectives, scope, and participation

XSL•FO
RenderX

requirements. To formally participate, candidates were required to sign a declaration of consent. Participants returned the signed consent forms via email.

Participation in the interviews was voluntary. Participants could at any time—even without giving a reason—decline to participate or withdraw their consent and leave the study without suffering any adverse consequences as a result.

Participants did not receive compensation for the interviews.

Results

Participant Overview

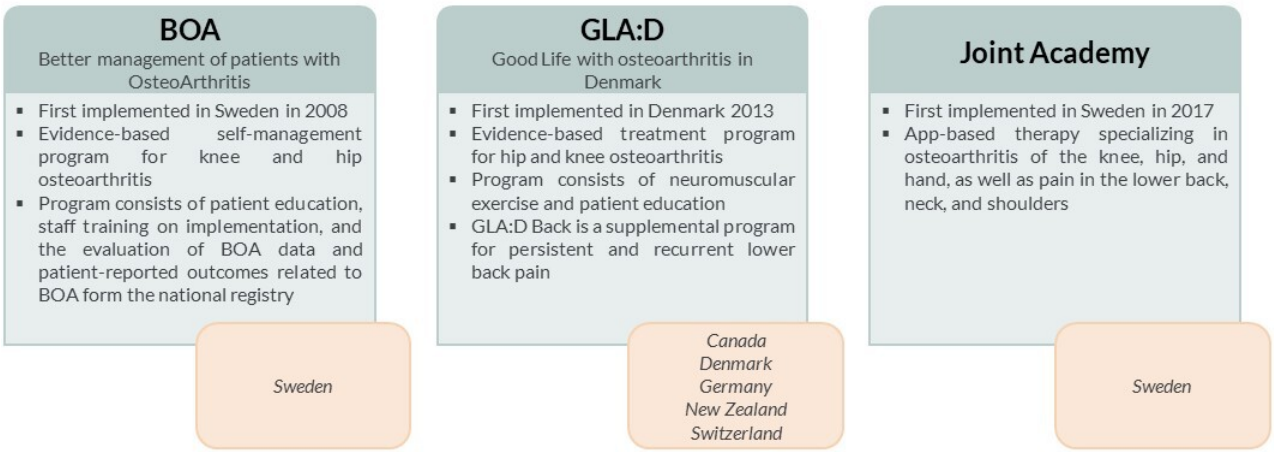
The recruitment process resulted in 12 completed interviews with experts involved in 8 different treatment programs. The interviews ranged from 33.42 to 65.02 minutes, with an average duration of 49.5 (SD 9.42) minutes.

The majority of experts (8 of 12) were affiliated with the GLA:D network, making it the most represented program in the sample. Specifically, experts participated from GLA:D country programs in Germany (n=2), New Zealand (n=2), Switzerland (n=2), Canada (n=1), and Denmark (n=1). Another expert from the Danish GLA:D Back program (n=1) took part in an interview. In addition, experts from the Joint Academy Program (n=2) and the Better Management of Patients with OsteoArthritis program in Sweden (n=1) were interviewed.

Figure 1 provides an overview of the programs in which the experts were involved, detailing their structural characteristics and the national context in which the interviewed experts had worked.

The age of participants as of 2024 ranged from 35 to 81 years, with an average age of 53.4 years (SD: 13.35). Three experts identified as female and 9 as male. To protect the anonymity of the participants, no further sociodemographic data are reported in this publication.

Figure 1. Overview of associated osteoarthritis treatment programs (own representation). BOA: Better Management of Patients with OsteoArthritis; GLA:D: Good Life with osteoArthritis in Denmark.



Results of the Data Extraction

Category System

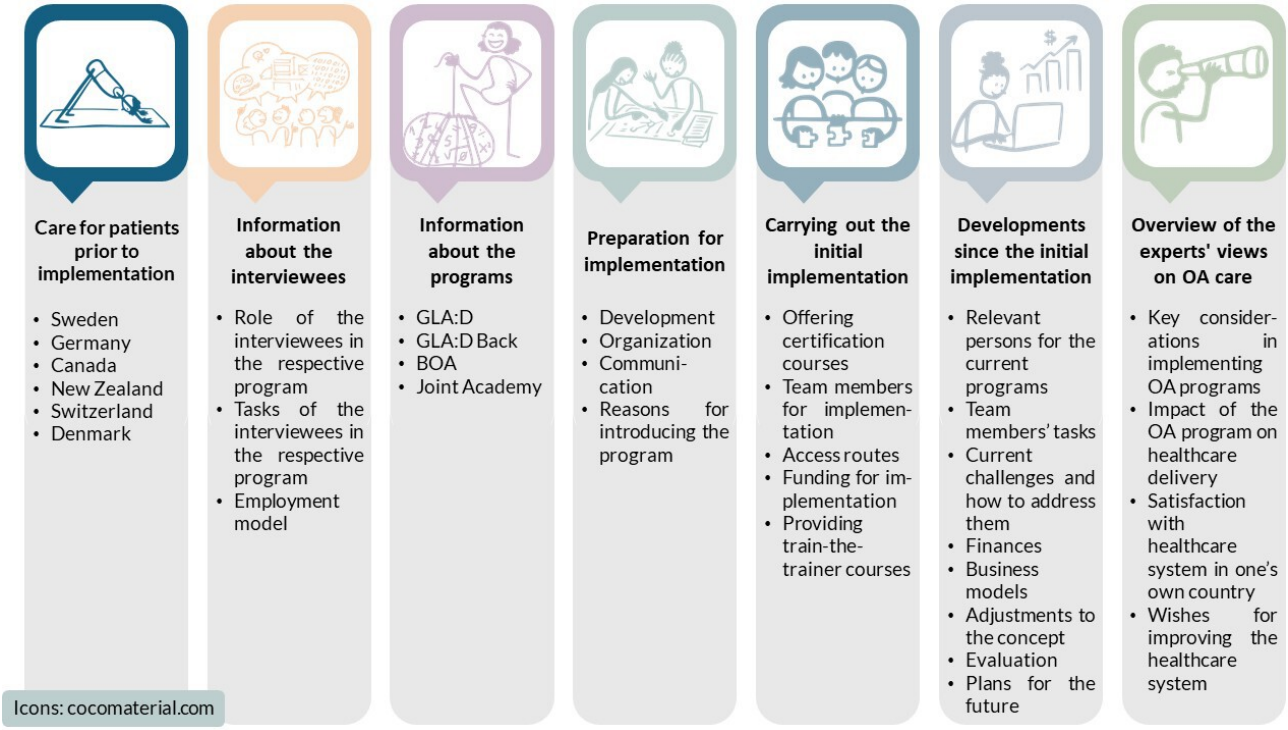
During the analysis, the research team assigned 1043 codes to 7 main categories covering contextual information on care provision before implementation, interviewee and program characteristics, implementation processes, program development, and overarching reflections on OA care.

A detailed overview of the category system (first level) is provided in Figure 2. Although the overall coding combined deductive and inductive category development, the 3 implementation phases emerged inductively from the interview data and were subsequently used to structure the presentation of the findings.

A detailed overview of the category system (first level) is provided in Figure 2. Although the overall coding combined deductive and inductive category development, the 3 implementation phases emerged inductively from the interview data and were subsequently used to structure the presentation of the findings.

XSL•FO
RenderX

Figure 2. Main categories and subcategories of the qualitative analysis (own representation). Icons sourced from Coco Material [14] published under CC0 1.0 Universal license [15]. BOA: Better Management of Patients with OsteoArthritis; GLA:D: Good Life with osteoArthritis in Denmark.



Across programs and countries, experts described implementation as a dynamic process consisting of 3 interrelated phases: preparation, early implementation, and consolidation. Although individual programs differed in their design and organizational context, similar implementation challenges and enabling factors emerged across settings. The following sections present the findings according to these implementation phases.

These implementation phases provide the framework for the presentation of the results in the following sections. Despite variations in program design (see Figure 1), a collective reporting approach addresses the research questions comprehensively.

Preparation and Development

Reasons for Introducing the Program

Across programs, implementation was primarily motivated by perceived deficiencies in existing OA care pathways. Experts described the introduction of OA programs as an opportunity to improve access to evidence-based treatment, reduce regional inequalities, and standardize care delivery. Established implementation models, particularly GLA:D, frequently served as inspiration and provided practical guidance for local adaptation.

> *The main goal is to give a more equal healthcare to patients with OA in [Country]. Because we knew back in the days it was a little bit different depending on where you're living if you're getting an education, a training or not. So, I think equality was the first goal.* [Interviewee 9]

Experts reported that, in several programs, the COVID-19 pandemic accelerated the adoption or expansion of digital and blended forms of delivery and that these models were subsequently valued for their scalability and potential to increase access to care.

Adaptations and Development

The development process ranged from entirely new programs to adaptations of evidence-based programs. One program, for example, began as a small university pilot that underwent multiple iterations before public rollout. Development tasks included designing analog materials (eg, screening tools, educational content, training modules, preparation materials) and digital components for blended care, as well as setting up data registries.

With the help of train-the-trainer models, therapists were able to deliver the program independently.

In the case of GLA:D adaptations, the Danish team was consulted for linguistic and cultural fit. One interviewee advised building on existing programs:

> *So don't reinvent the wheel yourselves. Even though you might not be happy with all parts of it, I would really encourage you to take on something that is already developed and structured.* [Interviewee 10]

Program adjustments considered country-specific conditions and multicultural environments. Cost-effective equipment, intensive simple training units, and well-developed educational concepts were also highlighted as helpful elements.

Initial Team and Actors Involved

During preparation, work was mainly driven by small core teams consisting of research managers, developers, or project leads. Legal advisors, patients, and therapists provided feedback. Experts emphasized that practitioners from the relevant field should be involved to ensure practical understanding:

XSL•FO
RenderX

> *But you do need to have someone driving it who has a really good understanding of the research and understanding of how that works, the business model, but also the whole training program. [...] And don't get them to do it for free.* [Interviewee 2]

Organizational Tasks

Organizational efforts involved contracts, bureaucracy, university resource allocation, and securing legal permissions, as well as GLA:D license agreements with Danish developers for data transfer.

> *Plus, in the [Country] healthcare context, in the vast majority of cases, universities do not have the resources to enter a program, no matter how good it is, into the healthcare sector.* [Interviewee 5]

Financing

Experts emphasized that securing sustainable financing for program development and initial establishment was a prerequisite for implementation. While funding models differed between countries and health care systems, experts consistently emphasized the importance of reimbursement mechanisms, institutional support, and early engagement with payers. Reliance on short-term project funding was widely viewed as insufficient for long-term implementation.

> *And because we have had good experiences in [Country] with the 18 units that we bill and charge to the health insurance companies, we have decided to adopt this structure identically, but this has meant that we first carry out four times two exercise units in the first four weeks, followed by four times one exercise unit, so that in the end we also have a total of 18 units, which we then bill to the health insurance companies and payers.* [Interviewee 11]

Communication

Communication activities focused on building legitimacy and awareness among key stakeholders. Experts highlighted the importance of engaging professional associations, health care providers, patient representatives, and policymakers early in the process. Public visibility through media appearances, conferences, and awareness campaigns was perceived as an important facilitator of implementation.

> *So I think, looking back (...) the kick-off was really great for us with this TV appearance, even if it took an incredible amount of time and energy to get everything off the ground.* [Interviewee 1]

Carrying Out the Initial Implementation

During rollout, responsibilities expanded to include trained therapists (chiropractors, physiotherapists, physicians, and nurses), scientists, software developers, and regional coordinators. Interviewees emphasized the importance of creating local ownership among clinicians, who frequently acted as ambassadors and facilitators within their professional networks:

> *And on the first, I would say, five, ten courses we were also very focused on ensuring that the new clinicians that had received the course were also ambassadors for the program in their local community.* [Interviewee 12]

Rollouts varied from regional pilots scaling nationally to single-insurer collaborations or multiregion launches.

Initial funding during the rollout phase came from associations, grants, or public sources, though high license fees and missing reimbursements hindered progress.

Consolidation and Ongoing Development

Members of the Program Team and Actors Involved

After implementation, operational responsibility increasingly shifted from the original developers to coordinators, database managers, analysts, and regional contacts. Initial developers stepped back as physiotherapists and coordinators took the operational lead.

> *But then gradually their role was less and less, [...] in the early days they were really important.* [Interviewee 12]

The recruitment of therapists emphasized openness without pressure.

> *So it's usually not so difficult to get a physiotherapist to try. But don't try to convince anyone [...]* [Interviewee 8]

Further Development of the Program

The teams initiated an exchange with physicians on points of criticism regarding the program and regular strategy meetings of the team. Annual reports and health services research supported this process. They facilitated active exchange with patients, providers, health policymakers, and payers.

Feedback From Patients and Providers

Following implementation, experts reported generally positive feedback from both patients and providers. Perceived benefits included improved access to evidence-based care, greater standardization of treatment practices, enhanced professional knowledge among participating clinicians, and low equipment requirements. However, providers also identified barriers related to workload, reimbursement structures, organizational demands (lack of experience with group training and duplication of documentation), and the practical challenges of delivering group-based interventions.

> *But we're getting more and more orthopedic surgeons who see the value of conservative management [...]* [Interviewee 2]

Adjustments Based on the Feedback and Research Findings

Following feedback and research findings, the teams introduced quality assurance measures such as recertification, knowledge checks, and additional support services (eg, support team, coaching, webinars, and information materials), refined program content, revised didactics, and adapted teaching materials.

Some interviewees nevertheless reported a lack of support during implementation.

XSL•FO
RenderX

Organizational Tasks

Team tasks included coordination, organizing, and ensuring data collection. Programs that established registries documented a wide range of clinical parameters at program entry and exit, enabling pre-post comparisons and cross-regional and cross-national analysis. The data ensure quality and help to demonstrate effectiveness to policymakers. In addition, programs with digital components supported adherence monitoring for both educational content and exercise sessions.

Financing

After implementation, financial issues primarily concerned long-term sustainability for registries, quality assurance, administration, and stable reimbursement. The interviewees reported that while therapists faced only minimal costs (eg, for room rental or digital access and the initial train-the-trainer course), initiators bore expenses for the registry, digital design, and administration. It was possible to cover costs through research projects, risk capital, private investments, or self-funding by patients. Payment models varied widely, including national coverage, regional schemes, private funding, coverage by health insurance, or mixed models. Interviewees emphasized the perceived legitimacy of the program:

> *Because if it's available publicly, people think it's got to work better. So, they will pay privately. But if it's not available publicly they go: 'Okay, obviously it's not a good thing.'* [Interviewee 3]

It was mentioned that no funding had been available for many years. To improve financing, interviewees recommended integrating the programs into the public health care system, negotiating with professional associations and health insurance companies, and conducting evaluations to demonstrate effectiveness. Positive developments were underway in some regions:

> *We were probably a year too early with trying to launch the program, because there's been significant funding change and now there is funding for osteoarthritis programs developed by physios from the state and the government.* [Interviewee 3]

Perceived Limitations

Interviewees pointed out that the current structure might meet the needs of highly active patients and that programs lacked long-term motivation methods since lifelong training should be supported. One interviewee reported that personal trainers were introduced to bridge this gap.

Perceived Effects on Care

Experts reported a variety of factors that influence the impact of the respective programs on the care of patients with OA. They highlighted the international exchange for continuous improvement, changes in perceptions of conservative and surgical care, as well as earlier, more consistent, or even initial treatment, and positive treatment outcomes.

> *And it's not such a big benefit for the few who already provide very, very good care. [...] But I think there are many who provide care that is far from the quality standard [...]* [Interviewee 6]

Experts perceived that the programs contributed to earlier access to conservative treatment, greater consistency in OA management, and improved patient outcomes. Several interviewees also reported that patients appeared better prepared when surgery eventually became necessary.

> *[...] it's very important that the patients are in the best possible shape [...] because that will make rehabilitation after surgery better and easier and they will reach a better result of the surgery as well.* [Interviewee 7]

Interviewees underlined the value of group treatment in addressing the shortage of skilled workers and praised the low costs and accessibility of digitally designed programs, especially for patients for whom analog services are not attractive. They additionally pointed out their relevance for highly educated patients, those with multiple comorbidities, those living in urban areas with limited access to programs, as well as those living in rural areas with the same problem. This also has advantages for therapists in view of increased workplace flexibility:

> *And there is a lot of young physiotherapists that like not to be in the office every day in the week [...]. Or you can make your own schedule and still get things going.* [Interviewee 8]

A Glimpse Into the Future

Despite perceived successes, experts identified several ongoing challenges, including insufficient reimbursement structures, continued reliance on surgical care pathways, limited implementation of evidence-based recommendations, and administrative burdens. To address these issues, interviewees emphasized the need for stronger political support, expanded digital services, improved professional education, and sustainable financing models.

> *We need to replace some of our traditional ways with digital ways which are smarter, which are faster, which are lower cost to the system and the patient. And we need to find smarter ways and other ways, many of them digital, to find the resources without getting more money.* [Interviewee 7]

Future priorities included increasing public visibility, strengthening quality assurance mechanisms, expanding programs to new settings and populations, and generating further evidence through registry and health services research.

Discussion

Principal Findings

This study provides in-depth insights into international implementation processes for evidence-based care programs for musculoskeletal disorders, with a primary focus on knee and hip OA and 1 related back pain program derived from an OA model. The findings were obtained from 12 expert interviews. Beyond documenting persistent challenges across preparation, initial implementation, and consolidation phases, our findings reveal a critical disconnect: while international

XSL•FO
RenderX

networks and evidence-based models like GLA:D are widely valued, systematic knowledge exchange and joint solutions remain rare. In some instances, international cooperation even required additional coordination and administrative effort. Nonetheless, the possibility of adapting and translating existing program materials, rather than developing entirely new ones, was seen as a major advantage. Westerlund et al [16] highlight that implementation science encompasses a growing range of frameworks, models, and theories aimed at bridging the gap between knowledge generation and its application in health care practice [9]. Evidence suggests differing perceptions between researchers (knowledge producers) and policymakers (knowledge users) [8]. Knowledge producers tended to underestimate their influence on policy while overemphasizing barriers to implementation [8]. While opinions differ on whether this gap can be bridged by theoretical approaches in implementation science, it can be assumed that the application of theory is explicit and verifiable but might lead to problems being overlooked that are not addressed in theory. Applying common sense for implementations is often based on implicit assumptions and beliefs [13]. Nilsen underlines that implementation needs to be connected to "the real world of implementation practice" [13].

Multiframework Interpretation

Connecting Knowledge, Barriers, and Implementation Practice

To situate the disconnect between available knowledge, perceived barriers, and actual implementation practice, we draw on 3 complementary implementation frameworks following Nilsen's typology of (1) process models guiding the implementation, (2) determinant frameworks, classic theories, and implementation theories explaining determinants of implementation outcomes, and (3) frameworks evaluating implementation [13]. These frameworks did not guide data collection or coding but served to contextualize the generated findings.

Process Perspective: Knowledge-to-Action

The Knowledge-to-Action (KTA) Framework [10] is frequently used in different health care settings for the design, delivery, and evaluation of implementation activities [11]. The action cycle includes central, iterative phases that can be interpreted as conceptually corresponding to the phases presented in the inductively derived categories of our study. Our phase (1) preparation matches the KTA steps 1 - 3 (problem identification, knowledge adaptation, barrier assessment), since experts described their motivations (equity, COVID-driven digitalization) and development tasks (cultural adaptation, train-the-trainer materials). Our phase (2) early implementation matches KTA steps 4 - 5 (tailoring, execution), since regional pilots expanded through local champions and coordination gaps persisted. Phase (3) consolidation matches KTA steps 6 - 7 (evaluation, sustainment), where registries enabled pre-post comparison, although short-term research funding threatened sustainability. The KTA cycle breaks down the phases into specific steps and shows in step 4 (tailoring implementation) that systematic cross-national knowledge exchange was rare, and this gap in the action cycle might prevent scale-up.

Determinant Perspective: Consolidated Framework for Implementation Research

Nilsen underlines that implementation needs to be connected to "the real world of implementation practice" [17]. One suitable framework for the determinant perspective is the Consolidated Framework for Implementation Research (CFIR), originally developed in 2009 [18] and revised in 2022 [19]. It identifies constructs influencing implementation success and effectiveness in 5 main areas (innovation, outer setting, inner setting, individuals, and implementation process), as well as 3 areas in the addendum (antecedent assessments, implementation outcomes, and innovation outcomes).

In the innovation domain, the evidence base of the programs needs to be mentioned (construct of innovation evidence base). The possibility of continuing to use existing evidence-based OA treatment programs and the possibility of adaptation to different (regional) contexts fit the construct of innovation adaptability. Ensuring the program's advantage over other types of treatment (eg, conservative vs surgical treatment; digital vs analog vs blended care) fits the construct of innovation relative advantage. Regarding the outer setting domain, financing emerged as a dominant barrier across phases (construct financing). This prominence of financing across all implementation phases confirms previous implementation research identifying sustainable funding as one of the strongest determinants of successful implementation. Project funding ended post-efficacy proof, while reimbursement negotiations and license costs hindered rollout. Research costs are identified as a frequently mentioned hurdle to implementation [20]. This reflects the broader challenge of insufficient funding for implementation research in health care, where many projects support development but fail to ensure integration into routine care, a problem widely documented across different areas of the health care system [18]. Political support was mentioned as essential but slow (construct policies and law). Political challenges discussed in the interviews relate primarily to national reimbursement policies and regulatory frameworks. As described in the review by Gesulga et al [21], political support constitutes a major challenge when it comes to implementation.

In the inner setting domain, interviewees noted that structured communication channels significantly enhanced program satisfaction and quality, especially in the consolidation phase (construct communications).

In the individuals domain, it became clear that an initially well-connected team with knowledge from the field is particularly important (construct implementation team members). Interviewees emphasized the importance of motivation (construct motivation), technical expertise (eg, in project management, research, and therapy) (construct capability), and existing networks for effective implementation (construct opportunity).

In the implementation process domain**,** the gradual rollout from the regional to national level through local pioneers reflects the CFIR constructs (construct planning). Engaging stakeholders such as national opinion leaders, formally appointed implementation leads, and local pioneers among therapists

XSL•FO
RenderX

accelerated awareness (construct engaging). Registry analyses, annual reports, and feedback loops informed adjustments to content, didactics, and quality assurance measures (construct reflecting and evaluating).

Our findings extend previous CFIR-based work by illustrating how the relevance of specific determinants changes across different implementation phases.

Evaluation Perspective: Reach, Effectiveness, Adoption, Implementation, Maintenance Framework

The RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) Framework conceptualizes the impact on public health across its name-giving key dimensions, highlighting the importance of evaluating outcomes at the individual and organizational levels [22]. Interviewees prioritized reach (eg, patient access, offering programs in various cost systems), while digital formats addressed accessibility for rural/urban patients, and group training mitigated workforce shortages. Participants perceived improvements in precomparisons and postcomparisons and better preparation for surgery after the programs. Registries enabled cross-national comparisons, supporting policy advocacy from the participants' point of view. Program adoption is hindered by limited therapist uptake, which is further challenged by heavy workloads, reimbursement gaps, and group logistics. Implementation was influenced by train-the-trainer fidelity. Regarding maintenance, recertification ensured consistency, although smaller practices struggled. Feedback loops (patient satisfaction, provider critiques) drove refinements such as recertification and didactics.

Comparison With Existing Guidelines

In 2019, the OA Research Society International published an international consensus guideline on priorities for the effective implementation of OA care, emphasizing, among other objectives, the need for standardized implementation approaches that align with international best practices [23]. Our findings suggest that several of these objectives remain only partially achieved, particularly regarding sustainable financing models (objective 3: operational funding for programs) and consistent reimbursement structures (objective 12: reimbursement of out-of-pocket costs for participants), as well as systematic implementation support and training for clinicians (objective 5: skills, confidence, and training of health care professionals). The experts perceived a persistence of implementation challenges since 2019 and attributed this to slow policy responses, unchanged or fragmented funding arrangements, and insufficient alignment between health system incentives and guideline recommendations.

Synthesis of the Interpretation

Taken together, the 3 frameworks shed light on complementary aspects of implementation. While the KTA framework helps us understand implementation as a dynamic process spanning various phases, CFIR explains the determinants that influence the success of implementation within these phases. RE-AIM complements this perspective by focusing on long-term adoption and maintenance. Our synthesis underscores Nilsen's approach [13], because the combined application of these frameworks and guidelines enables a more comprehensive understanding of the implementation of evidence-based OA care programs in different health care systems.

Limitations

This study represents an important exploratory step toward understanding the implementation of evidence-based OA treatment programs in different international contexts. However, some limitations must be considered.

By focusing on experts, we were able to extract subject-specific expertise regarding the practical implementation of OA programs. On the other hand, the exclusive reliance on expert interviews restricts the range of perspectives, particularly as patient views are therefore not included.

Regarding the selection of interviewees, the overrepresentation of experts affiliated with GLA:D-related programs may have influenced the thematic emphasis of the findings. In addition, the international sample was recruited through purposive and snowball sampling, which may have favored well-established professional networks and highly visible programs while excluding less prominent initiatives and perspectives.

Another limitation relates to the fact that most participants were directly involved in the development, implementation, or management of the programs under investigation. Their close engagement may have influenced the way implementation successes and challenges were described, potentially resulting in a more favorable assessment of the programs.

Although the interview guide covered the main aspects of implementation, it was not explicitly developed using a comprehensive implementation framework. Consequently, some contextual aspects, such as regional or federal policy influences, may not have been explored in equal depth.

By focusing exclusively on qualitative interviews, the study did not include methodological triangulation (eg, additional quantitative data). Consequently, the results could not be supported by multiple types of evidence.

Similarly, the findings were not externally validated through participants or additional stakeholder validation in the sense of member checking. While participant validation is not considered essential in all qualitative research traditions, its absence may have limited the opportunities to further enhance the confirmability of the findings.

Although purposive sampling allowed us to maximize variation in implementation experiences across countries and program types, transferability of the findings depends on similarities between health care systems, reimbursement structures, regulatory environments, and organizational contexts. The findings should therefore be interpreted in relation to the specific health care contexts in which the included programs were implemented.

Future research should address specific gaps identified in this study, including the cost-effectiveness of implementation strategies, the contextual adaptation of programs to diverse health system contexts, and implementation in resource-limited settings.

XSL•FO
RenderX

Conclusion

This qualitative study explored facilitators and barriers influencing the international implementation of evidence-based nonsurgical OA programs from the perspective of experts (n=12) involved in established international programs.

Participants emphasized the importance of international collaboration, established program structures, and access to implementation experience. They highlighted the importance of stakeholder engagement, effective communication, and interdisciplinary collaboration throughout the implementation process. Furthermore, participants emphasized that context-appropriate combinations of digital, analog, and blended care models may improve accessibility, scalability, and long-term sustainability of OA programs. Fragmented reimbursement structures and short-term funding were consistently described as key barriers to sustainable implementation.

These insights provide practical guidance for researchers, clinicians, and policymakers seeking to implement evidence-based OA care programs across diverse health care systems while acknowledging the importance of adapting implementation strategies to local contexts.

Acknowledgments

We would like to thank all experts for their time and valuable insights. Many thanks to Professor Kai Bernsmann, who came up with the idea for the original application and supported us in refining the guidelines with his orthopedic expertise. We would like to thank Georgius Agricola Stiftung Ruhr for funding this research. Special thanks to our research assistant for supporting the recruitment process.

ChatGPT (GPT 5.5, GPT 5.6; OpenAI) and DeepL Text translation (DeepL SE) were used to improve the wording.

Funding

This research was funded by the Georgius Agricola Stiftung Ruhr.

Authors' Contributions

JB and TSB developed the concept of the interview study. Both authors developed the guidelines for the interview and conducted the interviews. Both authors wrote the original manuscript of the publication, drafted the included figures and tables, and were responsible for editing the manuscript.
TSB and JN performed the content analysis of the interviews. JB reviewed the results of the content analysis.
NC, BJ, and HCV contributed to the publication through funding acquisition, support in reviewing the interview guidelines, as well as reviewing and editing the manuscript.
ICO contributed to the publication through funding acquisition, reviewing and editing the manuscript. ICO was further responsible for project administration together with TSB and provision of resources for JB and TSB.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide in English.
[PDF File, 223 KB - rehab_v13i1e85941_app1.pdf]

References

1. Dawson J, Linsell L, Zondervan K, et al. Epidemiology of hip and knee pain and its impact on overall health status in older adults. Rheumatology (Oxford) 2004 Apr;43(4):497-504. [doi: 10.1093/rheumatology/keh086] [Medline: 14762225]
2. Krankheitskapitel und Alter in Euro je Einwohner/-in der jeweiligen Altersgruppe [Article in German]. Statistisches Bundesamt (Destatis). URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankheitskosten/Tabellen/krankheitsklassen-alter.html [accessed 2025-03-15]
3. S3-Leitlinie Gonarthrose [Article in German]. AWMF-Leitlinienregister. URL: https://register.awmf.org/de/leitlinien/detail/187-050 [accessed 2026-04-06]
4. S2k-Leitlinie Koxarthrose [Report in German]. : Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC); 2019 URL: https://register.awmf.org/assets/guidelines/033-001l_S2k_Koxarthrose_2019-07_1-abgelaufen.pdf [accessed 2026-04-06]
5. Cunningham J, Doyle F, Ryan JM, et al. Primary care-based models of care for osteoarthritis; a scoping review. Semin Arthritis Rheum 2023 Aug;61:152221. [doi: 10.1016/j.semarthrit.2023.152221] [Medline: 37327762]
6. Basedow M, Esterman A. Assessing appropriateness of osteoarthritis care using quality indicators: a systematic review. J Eval Clin Pract 2015 Oct;21(5):782-789. [doi: 10.1111/jep.12402] [Medline: 26083547]

XSL•FO
RenderX

7. Hurley MV, Walsh NE. Effectiveness and clinical applicability of integrated rehabilitation programs for knee osteoarthritis. Curr Opin Rheumatol 2009 Mar;21(2):171-176. [doi: 10.1097/BOR.0b013e3283244422] [Medline: 19339929]
8. Bosompem J, Rasche P, Chikhradze N, et al. Promoting and inhibiting factors for the implementation of evidence-based non-invasive treatment programs for patients with knee- and/or hip-osteoarthritis: a rapid review. BMC Health Serv Res 2026 Mar 4;26(1):330. [doi: 10.1186/s12913-026-14219-5] [Medline: 41776556]
9. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
10. Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA:D™): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. BMC Musculoskelet Disord 2017 Feb 7;18(1):72. [doi: 10.1186/s12891-017-1439-y] [Medline: 28173795]
11. Roos EM, Barton CJ, Davis AM, et al. GLA:D to have a high-value option for patients with knee and hip arthritis across four continents: good Life with osteoArthritis from Denmark. Br J Sports Med 2018 Dec;52(24):1544-1545. [doi: 10.1136/bjsports-2017-098904] [Medline: 29514823]
12. Dahlberg LE, Grahn D, Dahlberg JE, Thorstensson CA. A web-based platform for patients with osteoarthritis of the hip and knee: a pilot study. JMIR Res Protoc 2016 Jun 3;5(2):e115. [doi: 10.2196/resprot.5665] [Medline: 27261271]
13. Kuckartz U. Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung [Book in German], 4th edition: Beltz Juventa; 2018.
14. License. Coco Material. URL: https://cocomaterial.com/license [accessed 2026-10-14]
15. CC0 1.0 universal deed. Creative Commons. URL: https://creativecommons.org/publicdomain/zero/1.0/ [accessed 2026-10-14]
16. Westerlund A, Sundberg L, Nilsen P. Implementation of implementation science knowledge: the research-practice gap paradox. Worldviews Evid Based Nurs 2019 Oct;16(5):332-334. [doi: 10.1111/wvn.12403] [Medline: 31603625]
17. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53. [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
18. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4:50. [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
19. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
20. Field B, Booth A, Ilott I, Gerrish K. Using the Knowledge to Action Framework in practice: a citation analysis and systematic review. Implement Sci 2014 Nov 23;9:172. [doi: 10.1186/s13012-014-0172-2] [Medline: 25417046]
21. Gesulga JM, Berjame A, Moquiala KS, Galido A. Barriers to electronic health record system implementation and information systems resources: a structured review. Procedia Comput Sci 2017;124:544-551. [doi: 10.1016/j.procs.2017.12.188]
22. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
23. Ellen ME, Lavis JN, Horowitz E, Berglas R. How is the use of research evidence in health policy perceived? A comparison between the reporting of researchers and policy-makers. Health Res Policy Syst 2018 Jul 20;16(1):64. [doi: 10.1186/s12961-018-0345-6] [Medline: 30029647]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
COREQ: Consolidated Criteria for Reporting Qualitative Research
GLA:D: Good Life with osteoArthritis in Denmark
KTA: Knowledge-to-Action
OA: osteoarthritis
RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance

Edited by S Munce; submitted 15.Oct.2025; peer-reviewed by A Ramani-Chander, GGD Alencar; revised version received 28.Aug.2026; accepted 15.Sep.2026; published 23.Oct.2026.

Please cite as:
Busse TS, Nitsche J, Otte IC, Chikhradze N, Jömann B, Vollmar HC, Bosompem J
Barriers and Facilitators in the Implementation of Digital, Blended, and Analog Evidence-Based Osteoarthritis Care: International Qualitative Expert Interview Study
JMIR Rehabil Assist Technol 2026;13:e85941
URL: https://rehab.jmir.org/2026/1/e85941
doi:10.2196/85941

XSL•FO
RenderX

© Theresa Sophie Busse, Julia Nitsche, Ina Carola Otte, Nino Chikhradze, Barbara Jömann, Horst Christian Vollmar, Jennifer Bosompem. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design Recommendations for Virtual Reality–Based Upper Limb Exercises From People With Tetraplegia and Spinal Cord Injury Rehabilitation Specialists: Focus Group Study

Andrew Goodsell[1], BSc, MSc; Mariel Purcell[2], BCh, BAO, MB; Matthieu Poyade[3], MSc, MPhil, PhD; Louise Cownie[2], BSc, MSc; Co-designers[2,4]; Lorna Paul[1], MPhil, PhD

[1]School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom

[2]Queen Elizabeth National Spinal Injuries Unit, NHS Greater Glasgow and Clyde, Glasgow, United Kingdom

[3]School of Innovation & Technology, Glasgow School of Art, Glasgow, United Kingdom

[4]Spinal Injuries Scotland, Glasgow, United Kingdom

Corresponding Author:
Lorna Paul, MPhil, PhD
School of Health and Life Sciences
Glasgow Caledonian University
Cowcaddens Road
Glasgow, G4 0BA
United Kingdom
Phone: 44 141 331 8108
Email: Lorna.Paul@gcu.ac.uk

Abstract

Background: The global incidence of spinal cord injury (SCI) is between 10 and 80 new cases per million people each year. This equates to between 250,000 and 500,000 injuries worldwide per year. In the United Kingdom, approximately 4400 people per year sustain an SCI. People with tetraplegia report upper limb function as their highest priority for improvement after SCI. Using immersive virtual reality (VR) headsets, physical rehabilitation exercises can be completed in engaging digital environments. Immersive VR therefore has the potential to increase the amount of therapy undertaken, leading to improvements in arm and hand function. There is little evidence supporting immersive VR as exercise in SCI, especially while patients with SCI are undergoing acute rehabilitation. In SCI research, co-design of new interventions is not a widely adopted approach, yet people with tetraplegia want to contribute with their expert knowledge on their experiences of SCI.

Objective: This study aims to explore the lived experiences of people with tetraplegia and specialist SCI therapists related to acute upper limb rehabilitation and identify design considerations for VR-based interventions targeting the upper limb.

Methods: We conducted 7 online focus groups using Microsoft Teams: 4 with people with tetraplegia (n=15; age range, 36-65 years) and 3 with occupational therapists and physiotherapists specializing in SCI rehabilitation (n=11). Participants were asked to discuss their experiences and expertise about acute SCI upper limb rehabilitation and their opinions and ideas on the use of VR for upper limb rehabilitation. The transcripts were analyzed using content analysis, enabling the proposition of design characteristics of a VR-based intervention for upper limb exercise.

Results: The study identified 5 major themes describing the clinical features, treatment, and recovery of people with SCI during the acute stage of SCI, their motivations for participating in therapy, and suggestions for the design of a VR intervention in treating the upper limbs following SCI.

Conclusions: The themes identified in this study allow the elicitation of software requirements for a bespoke immersive VR platform for upper limb rehabilitation following SCI. They can also contribute to a better understanding of the advantages of using VR as an adjunct to upper limb rehabilitation. Additionally, participants used their expertise to suggest factors that would enable the development of a usable and effective intervention, as well as identifying potential pitfalls and software features to avoid during intervention development. These findings can be used to design accessible VR applications for use by people with tetraplegia and their therapists.

(JMIR Rehabil Assist Technol 2026;13:e66832) doi:10.2196/66832

XSL•FO
RenderX

KEYWORDS

acute/sub-acute tetraplegia; co-design; focus groups; secondary care; spinal cord injury; upper limb rehabilitation; virtual reality

Introduction

A spinal cord injury (SCI) results from an insult to the spinal cord and impacts nearly every aspect of a person's life. People with SCI usually have a degree of permanent neurological disability, including muscular paralysis and impaired sensory and autonomic function below the level of their injury [1]. The global incidence of SCI is between 10 and 80 new cases per million people each year [2], with 4400 new cases per year in the United Kingdom [3]. The lifetime costs of SCI are estimated to total £1.43 billion (US $1.97) for 1270 cases per year in the United Kingdom [4].

The level of SCI is a major factor determining the degree of independence a person will achieve following injury [5]. Typically, the higher the level of injury, the lower the level of independence [6]. Damage to the cervical spinal cord results in tetraplegia, where function of all 4 limbs is impaired [7]. As well as physical and sensory issues, people with SCI are at risk of pressure ulcers, urinary tract infections, spasticity, autonomic dysreflexia, depression, neuropathic pain, difficulty breathing, and circulatory problems [8]. This multitude of impairments following SCI is associated with lower quality of life [9].

Immediately following injury, patients with SCI require acute medical management. This is followed by a period of rehabilitation as an inpatient. The aim of rehabilitation post SCI is to optimize recovery, restore and maximize function by leveraging activity-dependent plasticity for neurological recovery, learn compensatory techniques for lost function, and prevent secondary complications [10]. Central nervous system plasticity, where neural pathways can be altered in response to an injury and promote recovery, can occur spontaneously and/or as a result of rehabilitation [11]. In the last 30 years, the primary aim of physical rehabilitation of SCIs has moved away from compensatory strategies to neurological recovery, particularly in people with incomplete SCI [12]. Incomplete SCI refers to the preservation of sensory and/or motor function below the level of injury, as opposed to complete injuries, which do not have preservation of sensation or movement [7]. Rehabilitation is lifelong, but it is recommended that rehabilitation starts early after injury to maximize functional status and clinical outcomes [10].

Reducing reliance on care and achieving higher levels of independence with respect to activities of daily living is a major goal for patients and therapists [13]. Improving motor function, especially of the upper limbs, through rehabilitation [14] enables patients to carry out tasks, such as dressing, bladder and bowel care, transferring in and out of a wheelchair, and feeding which may otherwise require a carer [15-17]. Even small improvements in upper limb function can have large effects on a patient's independence [18,19]. It is understandable, therefore, that improving upper limb function is a high priority for people with tetraplegia [13,20-22].

Improvements in upper limb function can be achieved through Activity-Based Therapy (ABT) [23]. ABT, a rehabilitation approach, refers to any intervention that involves high-intensity, repetitive exercises that target activity-dependent plasticity in spinal circuits [24]. The improvements in upper limb function from ABT have greater effects on quality of life when compared to traditional physical interventions targeted above the level of injury [19].

Virtual reality (VR) technology has the potential to deliver upper limb therapy by providing engaging interventions for patients as part of their rehabilitation [25,26]. Immersive VR delivers to the user a real-time digital environment that incorporates multimodal sensory paradigms, including proprioception, kinesthesia, stereopsis (the perception of depth via a stereoscopic display), and spatial audio, along with the ability to interact within the digital environment via head and hand tracking [27]. In these environments repetitive upper limb movements, like ABT, can be performed. VR-based therapy can facilitate greater adherence to therapy [28] and increase access to the most effective rehabilitation strategies for people with neurological disorders [27]. However, currently there are few studies that have investigated the use of VR in rehabilitation of the upper limbs, with most focusing on the chronic stage after injury.

VR-based upper limb therapy has the potential to provide engaging interactive environments within which patients can undertake the same type of movement repetitions as they would in conventional therapy. Due to its immersive qualities, patients may be more motivated to engage in therapy and may be able to complete more repetitions than they would compared to conventional treatment. A total of 6 reviews published between 2019 and 2024 have evaluated the evidence for VR for upper limb rehabilitation following SCI. These reviews included between 3 and 7 studies involving the upper limb. The overall findings from these reviews suggested that the evidence base is limited and that immersive VR appears to be more beneficial than non- or low-immersive VR [28-31]. However, these studies also suggested that exercises delivered in immersive VR could be used as an adjunct to conventional rehabilitation, which requires further investigation.

The VR devices, software, and methodology surrounding implementation used for upper limb therapy vary widely. VR-based interventions also vary in terms of their intended clinical use, fidelity, and cost, and the design and production of the software have rarely incorporated the lived experiences of SCI experts (people with SCI and health care professionals). The input of end users into the development of software, particularly in health care, is important in delivering successful products [32-34]. This process allows highly customized and bespoke projects to be developed, ensuring that the service or intervention is targeted toward the needs of patients [33].

This study forms the first part of a cocreation process for the design of a suite of VR-based exercises for upper limb rehabilitation in the acute SCI rehabilitation context. The aim of this study was to explore the lived experiences of people with tetraplegia and specialist SCI therapists related to acute upper

XSL•FO
RenderX

limb rehabilitation and to co-design immersive VR-based upper limb activities.

Methods

Overview

In this qualitative study, 7 focus groups were conducted online from August to October 2022 to gather opinions and insights from the lived experiences of people with tetraplegia and, separately, SCI specialist therapists with experience of upper limb rehabilitation.

Recruitment (Procedure)

A convenience sample of people with (1) tetraplegia and (2) therapists specializing in SCI rehabilitation were recruited. Although the VR intervention was to be developed for inpatient SCI rehabilitation, in order to fully explore lived experiences of participants with tetraplegia, the study recruited people who had been discharged from hospital having completed rehabilitation.

Participants with tetraplegia were recruited through the third sector organization Spinal Injuries Scotland. Recruitment material was also shared on social media channels (Twitter and LinkedIn). Inclusion criteria for participants with tetraplegia were a diagnosis of nonprogressive tetraplegia, aged ≥18 years, able to attend focus groups and contribute in spoken English, and having been discharged from hospital after receiving upper limb rehabilitation in a spinal injuries unit as part of their treatment. Exclusion criteria were any comorbidities that could preclude participants from attending and contributing to focus groups.

Therapists (occupational therapists and physiotherapists) were recruited through similar social media channels and a WhatsApp group of United Kingdom and Ireland–based SCI rehabilitation specialists. Inclusion criteria for therapists were being aged 18 years or older and able to attend and contribute to focus groups in spoken English.

Respondents to recruitment material emailed AG to register interest in participating in the study. Participant information sheets were provided to these interested parties by email. Participants provided informed consent prior to taking part in the study by either printing and physically signing a digital consent form or completing the form digitally. Participants without the physical ability to sign consent forms provided consent by designating a witness to sign on their behalf [35].

After providing informed consent, participants were asked to complete a demographic information form; for people with tetraplegia, this included their age, sex, Spinal Cord Independence Measure self-reported score [36], and for therapists, this included sex, occupation, and years of professional experience.

Focus Groups

Focus groups were chosen over individual interviews to promote discussion and ideation between focus group participants. Question schedules were created to enable semistructured discussion between participants (Multimedia Appendices 1 and 2). The questions were devised to enable future software design and software requirements specification. The difference between the question schedules and conduct of each focus group was minimal. Participants with tetraplegia focus groups aimed to explore the topic of VR through the lived experience of participants with tetraplegia, and therapist focus groups aimed to explore the topic through the lens of clinical practice. Each focus group had 2 facilitators, AG (male) and LP (female). A total of 3 focus groups of therapist participants and 4 focus groups of participants with tetraplegia were conducted. The duration of focus groups ranged between 45 and 75 minutes.

Participants with tetraplegia were asked to recall their experiences of the acute stage of inpatient SCI rehabilitation. They were then asked to discuss upper limb impairments caused by SCI and how these relate to prehension, strength, and range of motion. Therapists were instead asked about the current practice surrounding delivery of upper limb therapy. Both groups were introduced to the concept of VR as a commercial product for immersive entertainment. They were shown a video promoting a high-immersion head-mounted display being used by able-bodied people to play interactive games [37]. Both cohorts were then asked to discuss their opinions on VR as an adjunct to usual upper limb rehabilitation and the potential facilitators and barriers to its use in the acute rehabilitation setting. Finally, specifications for the design of a VR system to promote upper limb exercise were gathered from both groups.

Analysis of Transcripts

Focus groups were audio-recorded, and recordings were transcribed verbatim following pseudonymization of participants. Transcriptions were generated manually to enable familiarization with the data. Content analysis was used to count the responses and determine the meaning of participants' discussions via code generation [38]. Therapist and tetraplegia focus groups were analyzed separately. After codes were generated by AG, they were independently reviewed by LP, M Poyade, and M Purcell. AG, LP, M Poyade, and M Purcell discussed codes to come to a consensus on identification of codes and naming of themes in the theme hierarchy. Quotes are presented with participant identifiers and their focus group number. Data were managed using NVivo (version 20; Lumivero) [39].

Codes (units of meaning) were typically sentences, phrases, or longer portions of text. Shorter phrases and sometimes single words were coded if they related specifically to recommendations for the design of the VR games. Initially, codes were organized under categories deduced from the schedule of questions. Subthemes were abstractions of groups of similar codes and were created during coding. Major themes were then generated from groups of subthemes and replaced the organizational categories. Transcripts were not returned to participants for comments or corrections.

Reflexivity

Relationships with study participants were not established prior to study commencement. LP had previous experience facilitating focus groups and interviewing research participants, but AG did not. Neither facilitator worked in SCI rehabilitation and, as

XSL•FO
RenderX

such, reduced the likelihood of inhibiting frank discussion about patient care with participants [40].

Results

Participants (Demographics)

A total of 18 participants with tetraplegia and 14 therapists expressed interest in the study. Of these, 3 therapists and 3 participants with tetraplegia did not attend the focus groups; thus, 15 people with tetraplegia and 11 therapists took part (Table 1). We conducted 3 focus groups with therapists and 4 focus groups with participants with tetraplegia. A total of 7 participants with tetraplegia did not complete the demographic information forms. Reasons for not completing the information form included technical difficulty completing the form (reported only by participants with tetraplegia). A total of 3 therapists did not report their years of professional experience. Therapists reported not having enough time or opted out of form completion, with no reason given. Half of the participants with tetraplegia had experienced immersive VR at least once, but none had used VR extensively.

Table 1. Summary demographic information for participants with tetraplegia and therapist participants.

Characteristic	Value
Participants with tetraplegia (n=15)	
SCIM[a] score[b] (n=8), mean (SD)	17.9 (6.4)
Age (years; n=8), mean (SD)	54.5 (8.6)
Sex (n=13), n	
Male	10
Female	5
Experience with VR[c] (n=8), n	
None	4
Some	4
Therapist participants (n=11)	
Professional background (n=11), n	
Occupational therapists	7
Physiotherapists	4
Mean years of professional experience[d] (n=8)	14.8 (SD 8.6)
Sex (n=11), n	
Male	2
Female	9

[a]SCIM: Spinal Cord Independence Measure.

[b]Only 8 out of 15 participants with tetraplegia completed the SCIM and age information; 13 participants reported sex. Lower SCIM scores indicate lower independence (total score range, 0-75). Only 15 out of 19 subscales were included, comprising 6 self-care items and 9 mobility items.

[c]VR: virtual reality.

[d]A total of 3 therapists did not report their years of professional experience.

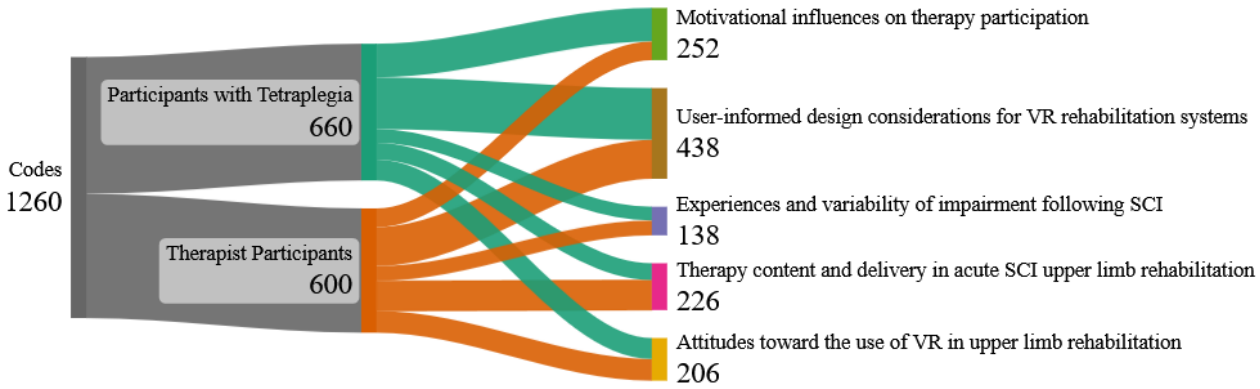
Major Themes and Subthemes

Overview

Analysis of the focus group transcripts identified 5 major themes shared by the therapists and participants with tetraplegia: descriptions of impairment, descriptions of acute rehabilitation, factors affecting participation in therapy, perspectives on VR, and suggestions for VR design. A total of 35 subthemes were generated from 660 and 600 individual codes from the tetraplegia and therapist groups transcripts, respectively (Figure 1).

XSL•FO
RenderX

Figure 1. Visual summary of the number of major themes (listed in the right column) identified from the 1260 codes (shown in the left column) generated by content analysis of focus groups with participants with tetraplegia and therapists. SCI: spinal cord injury; VR: virtual reality.



Major Theme 1: Experiences and Variability of Impairment Following Spinal Cord Injury

A total of 5 subthemes were identified that described impairments resulting from SCI immediately after injury and during the acute stage of rehabilitation (Table 2). Participants described profound loss of upper limb movement following injury, with variable patterns of recovery.

Table 2. Summary of subthemes for the major theme "experiences and variability of impairment following spinal cord injury."

Subtheme	Therapist codes, n	Participant codes, n	Total, n
Impairment of upper limb movement	17	28	45
Other impairments	27	8	35
Impairment of function, including activities of daily living	9	19	28
Heterogeneous presentations and demographics	16	2	18
Impairment of upper limb strength	3	9	12

> *My arms, I wasn't able to move them at all after I had my accident.* [PwT27, FG 3]

Impairments were described as highly variable across individuals and influenced by the level and severity of SCI.

> *I suppose [the degree of upper limb impairment] depends on the severity of your tetraplegic-ness, you know. Some people might, might be really bad and some people might still have a bit of movement.* [PwT27, FG 3]

Both groups, but especially therapists, discussed "other impairments." These included spasms, pain, and mental health issues, with therapists also raising age-related comorbidities and complications.

Therapists specifically raised the concept of "heterogeneous presentations and demographics" in acute SCI, with a growing prevalence of central cord syndrome in older adults, presenting as greater weakness in the upper limbs (than lower limbs) and greater involvement of distal muscles [7].

> *You could have somebody who's got, you know, full shoulder movement, full elbow movement, but they've got some weakness in the hands. Or the opposite, you know you could have someone who has got good hands but reduced movement at the shoulder.* [TP3, FG 1]

Major Theme 2: Therapy Content and Delivery in Acute Spinal Cord Injury Upper Limb Rehabilitation

A total of 7 subthemes described the nature of inpatient rehabilitation following SCI (Table 3). Upper limb therapy was described as a combination of functional and remedial activities, with a strong focus on maintaining range of movement and developing task-specific skills.

XSL•FO
RenderX

Table 3. Summary of subthemes for the major theme "therapy content and delivery in acute spinal cord injury upper limb rehabilitation."

Subtheme	Therapist codes, n	Participant codes, n	Total, n
Upper limb therapy	92	51	143
Patient-centered care	31	0	31
Assistive devices, orthotics, and neck stabilization	0	16	16
Barriers to effective therapy	12	0	12
Psychological adjustment and treatment	4	9	13
Bed rest and ventilation	0	5	5
Symptom management	6	0	6

> *[…] we do work on range of movement, you know shoulder, elbow, wrist, fingers.* [TP7, FG 2]
>
> *There were various things that we used—we had these games where we had little pegs that we had to move around the board, or also another one with a bucket full of the marbles.* [PwT16, FG 4]

Therapists also discussed upper limb functional improvement, particularly where patients have incomplete injuries:

> *If they're an incomplete injury we might be looking for more of a normal movement. […] Whereas if they're more of a complete, it's obviously looking at function and how we can use what they've got—what muscles they've got working to functionally achieve the task.* [TP9, FG 2]

Both groups highlighted the common use of "assistive devices" during rehabilitation, which included arm and hand stabilization using orthotics and neck stabilization, which hindered movement.

> *I had to use the Active Hand, which are the Velcro gloves so I could hold things.* [PwT16, FG 4]
>
> *I was 12 weeks in a neck brace, so I ended up with a very, very weak neck.* [PwT13, FG 3]

Under therapist-only subthemes, patient-centered care was achieved through collaborative goal setting. Barriers to therapy delivery included delayed referral and limited resources. Psychological adjustment was described by both groups, with participants with tetraplegia emphasizing loss of independence and therapists focusing on expectation management.

> *We would set those goals collaboratively with the patient based on a little bit of, kind of advice and guidance from us as to what we anticipate their outcomes might be.* [TP5, FG 1]
>
> *Unfortunately, the later we get someone sometimes the more issues we can have with contractures and upper limb pain and things, which then makes it harder for us then to, to rehab them.* [TP9, FG 2]
>
> *Trying to deal with everything and I found it really hard to be having to rely on other people for your personal care just day in, day out for everything.* [PwT21, FG4]
>
> *In their mind they want [a goal] to be achievable. So it, again comes down to communication and managing those expectations of patients and their families.* [TH9, FG2]

"Bed rest and ventilation" and "symptom management" were less commonly raised in both groups.

Major Theme 3: Motivational Influences on Therapy Participation

A total of 10 subthemes described factors influencing participation in therapy. Subthemes were then classified as "motivational" or "demotivational" (Table 4). Motivational influences included task enjoyment, goal setting, peer interaction, and visible improvement. Participants with tetraplegia described greater engagement with game-like and functional activities compared to repetitive remedial exercises.

XSL•FO
RenderX

Table 4. Summary of subthemes for the major theme "motivational influences on therapy participation." Subthemes are categorized as either promoting motivation or reducing motivation.

Subtheme	Motivational or demotivational	Therapist codes, n	Participant codes, n	Total, n
Useful or enjoyable rehabilitation	Motivational	1	39	40
Goal setting and therapist support	Motivational	6	31	37
Understanding tasks and prognosis	Motivational	30	0	30
Peer support and competition	Motivational	5	11	16
Seeing improvement	Motivational	0	10	10
Boring, repetitive, or difficult rehabilitation	Demotivational	5	41	46
Limiting symptoms	Demotivational	22	6	28
Low mood	Demotivational	11	6	17
Downtime	Demotivational	0	15	15
Limited resources and lack of supervision	Demotivational	10	3	13

> *We had some games to play, which is quite good fun. Like maybe like playing dominoes. Uh, so fiddling about with them things and also trying to build things with Lego was quite motivational and we had a tower of blocks that was probably my favourite thing.* [PwT26, FG 4]

In contrast, "boring, repetitive, or difficult rehabilitation," which included pointless, tedious, and humiliating activities, was demotivational.

> *You know what it's like sometimes you go and do a bit of exercise, you know you're pulling on a TheraBand and [imitates snoring] 'boring!'.* [PwT1, FG 1]

Sometimes strategies that were intended to make therapy more enjoyable, such as creative activities using putty or interactive cycling, were frustrating for patients.

> *[…] most days doing the exercises with the putty and so on, it felt like a really kind of childish, I don't know. Like, 'what's this doing?'* [PwT21, FG 4]

> *And [patients] found that really—you know they were cycling [FES cycling] away, cycling away along this road but they never went anywhere. And they got really frustrated by that.* [TP9, FG 2]

Some activities were too difficult or too easy for participants with tetraplegia, which resulted in them disengaging from therapy.

> *I also had a board with different sizes of bolts and nuts that I had to screw and I hated that because I thought, you know, 'I can't do this! Why are they making me do this when I cannot do it? This is so humiliating'.* [PwT16, FG 4]

Therapist discussions described patients were more motivated if they understood their injury, prognosis, and purpose of therapy ("understanding tasks and prognosis").

> *[Patients need] to have a good understanding of their injury and what the task is working towards. If they don't then often they're: "What's the point?" or "I don't really know why I'm just rolling this putty out". Umm "I'm not gonna get any better" or "How's this gonna strengthen?". So how much they understand about their injury is really important in their engagement.* [TP5, FG 1]

Both groups also raised that "peer support and competition" could be motivational.

> *Some of our patients…. a kind of element of challenge or competition almost with themselves in some of the more repetitive remedial tasks, um, can be really beneficial—like if they get 10 pegs one day - the next time they want them to get 15.* [TP4, FG 1]

> *I think being in the group as well gave you a bit more, you know, just a bit more reason to sort of get on with it and people would be chatting and you had that, you know, nice social aspect as well.* [PwT27, FG3]

Demotivational factors included repetitive or infantilizing tasks, symptom burden, low mood, and comparisons with faster-progressing peers.

> *Other people were whizzing through, and you just sat there doing like one at a time. And it was just demoralizing.* [PwT21 FG 4]

> *[…] a lot of the time I was just knackered, it turned out that I had an underlying infection due to the metalwork in the back of my neck.* [PwT1, FG1]

> *A lot of pain can be difficult for them to kind of adhere to upper limb therapy (or any therapy).* [TP4, FG 1]

> *I mean, some people in the unit—when I was there for 12 months didn't do anything. And they just lay in bed. They gave up.* [PwT4, FG 2]

Disengagement could occur if therapists could not always be present during therapy ("limited resources and lack of supervision").

> *[…] you know people who never do what's asked of them unless someone's standing over them making them do it.* [TP12, FG 3]

XSL•FO
RenderX

> *You know, they did their best to keep things fresh and, and try introducing, but obviously they had a limited amount of resources available to them, but equally limited staff time.* [PwT13, FG 3]

Finally, participants with tetraplegia identified "downtime," especially at the weekend, which meant there was nothing to do, and, as such, was demotivational.

> *The weekends at the spinal unit are, are very long and you just you know everything finishes on a Friday and you do nothing till Monday and you just feel I could I could be doing [something].* [PwT27, FG 3]

Major Theme 4: Attitudes Toward the Use of VR in Upper Limb Rehabilitation

A total of 8 subthemes captured attitudes toward VR (Table 5). Participants described VR as a potentially engaging and meaningful adjunct to therapy, particularly for individuals who struggle to engage with conventional rehabilitation.

Table 5. Summary of subthemes for the major theme "attitudes toward the use of virtual reality in upper limb rehabilitation." Subthemes are categorized as having a positive or negative sentiment.

Subtheme	Positive or negative sentiment	Therapist codes, n	Participant codes, n	Total, n
VR[a] could/should be engaging and meaningful	Positive	14	31	45
VR could/should be used for rehabilitation	Positive	9	18	27
Younger patients will like VR	Positive	9	0	9
VR could alleviate symptoms	Positive	5	4	9
Consumer equipment is not suitable	Negative	29	15	44
VR could cause adverse effects/be inappropriate	Negative	17	33	50
Resources are limited	Negative	15	0	15
Training is required	Negative	7	0	7

[a]VR: virtual reality.

> *So I think VR would have a very big use in a spinal injury centre there's a whole range of things it could cover. It's exciting it would be fun to use with people.* [TP8, FG 3]

> *I think [VR] is great because as I said earlier, there are other people [in hospital] who, you know, their mindset is totally negative. […] they need that stimulation, that encouragement, and VR might be the way.* [PwT4, FG 2]

> *[VR could] make any training or rehab more, more interesting, more exciting, more varied. Yes, I think… huge potential.* [PwT13, FG 3]

Both groups said that VR could encourage engagement in therapy by reducing symptoms such as pain and improve mood ("VR could alleviate symptoms").

> *And [VR is] good, actually help people to get involved in activities without even imagining the pain they are in.* [PwT19, FG 3]

> *I do quite like that element of escapism for them, I think. It could psychologically really help them as well.* [TP3 FG 3]

Conversely, concerns were raised from both groups regarding equipment suitability, adverse effects, staffing capacity, and training requirements. Head-mounted displays were described as potentially heavy and unsafe for individuals with high-level injuries.

> *If [devices] are too labour intensive or take a while to set up or need someone there then when we're short-staffed as we have unfortunately been for quite a long time, um, you tend to go for things that are quicker and easier to use if you've only got quite a short treatment session.* [TH9, FG2]

> *The other thing as well is our very high injury levels, [who are] struggling just holding their head up—or that kind of movement with the head [TP8 demonstrates neck rotation] a lot—it can tire them.* [TP8, FG 3]

Both groups warned that "VR could cause adverse effects" such as nausea, or that the specific activities in VR that could remind them of their SCI.

> *I think when you're already feeling self-conscious about yourself as an individual and how you move… VR might have a bit of a negative effect on people as well.* [TP5, FG 1]

> *When you take [the VR headset] off and you go 'Oh shit. It's true, this is where I am'.* [PwT21, FG 4]

Therapists speculated that their younger patients would be more suited to VR games ("younger patients will like VR").

> *So I think the younger people will probably get [VR], you know they'll understand a lot better how it works.* [TP4 FG 1]

Major Theme 5: User-Informed Design Considerations for VR Rehabilitation Systems

A total of 5 subthemes described participant recommendations for VR system design (Table 6). Suggested activities ranged from functional or vocational tasks, suggested mostly by therapists, to leisure, sporting, and creative environments, which were suggested predominantly by participants with tetraplegia.

XSL•FO
RenderX

Preferences varied widely, reflecting differences in interests and functional abilities.

Table 6. Summary of subthemes for the major theme "user-informed design considerations for virtual reality rehabilitation systems."

Subtheme	Therapist codes, n	Participant codes, n	Total, n
Activities and scenarios	86	150	236
Usability and training	32	33	65
Measuring performance and achievement	38	27	65
Movements	21	23	44
Immersion, presence, and engagement	11	17	28

> *[…] it's down to the individual what interests they have and you know what would benefit them the most and… I mean, metalwork is obviously not gonna be the same as scuba diving and, you know, skiing, or playing the piano. One's having loads of fun, one's more sort of creative and you're constructing something in VR. Yeah. So yeah, just depends.* [PwT4, FG 2]

Therapists also discussed providing a range of different activities for their patients. Both groups recommended incorporating concepts such as relaxation and escapism into the design ("immersion, presence, and engagement").

> *So we are really trying to provide a variety of different type of games that we offer but obviously when you've got a patient who's had a long (and all of our spinal injury patients have quite a long) admission with us, it's like the variety that's key.* [TP14 FG 3]

> *What about making really realistic sound, […] soaring noises, painting noises of the paint going on to the wall and those kind of things where it actually feeds back extra reality to the experience?* [PwT16, FG 4]

Additionally, participants with tetraplegia and therapists suggested outdoor spaces such as hills, glens, mountains, underwater spaces, outer space, and jungles would be good environments for VR activities.

> *[I have] a static bicycle at home. I get a bit bored because I'm always in the same place, while if you could have the feeling that you're actually outside and going places that might make it more interesting.* [PwT16, FG 4]

Both groups said that VR games should enable a range of movements, from fine to gross ("movements"), including reaching, grasping, and wrist flexion and extension.

> *[VR] could be something that could be really useful for larger gross movements but also for really drilling down to repetition of very fine motor movements.* [TP12, FG 3]

Therapists also recommended the quality of movements in VR was of high importance.

> *Sometimes it's not the number of movements, it's quality as well. So, you know, we'd be looking at the quality of someone's reach, not just that they are able to reach, but maybe not recruiting the right muscle groups to perform that task.* [TP9, FG 2]

Both groups gave recommendations for "measuring performance and achievement" of users. These included giving encouragement and providing feedback to increase engagement. Therapists discussed providing incremental levels of difficulty to accommodate progression and achievement.

> *[…] like you get a big tick or a big thumbs up or something. Not something annoying but something that shows you've reached the level of the activity that's actually doing something.* [PwT27, FG 4]

Therapists recommended incorporating easily comprehensible progress reports and feedback to users to help with motivation.

> *[…] and patients love seeing kind of that—we can put it—display it on a graph for them to kind of say like this is what you did like a month ago and this is what you're doing now.* [TP8, FG 3]

Therapists and participants with tetraplegia recommended that the movements should be configurable for users, and that training should be provided to patients and therapists ("usability and training").

> *So there's gonna have to be sort of that adjustability with the system to make it relevant to that patient's rehab and what muscles you want to work on and that, yeah.* [TP7, FG 2]

Furthermore, under this subtheme, it was highlighted that adding straps to the controllers might be required, and for those with limited hand function, participants with tetraplegia also requested interaction mechanisms using head-movement, gaze, hand tracking, and voice input.

> *I think you'd have to be really aware of even things like how people could grip any hand controls, you know, just the simple things like holding them, putting headsets on, how much assistance was needed to make it as practical as possible.* [PwT21, FG 4]

Ethical Considerations

Ethical approval for this study was granted by the Psychology, Social Work, and Allied Health Sciences (PSWAHS) Research Ethics Committee of Glasgow Caledonian University (reference: HLS/PSWAHS/21/247). Participants with tetraplegia were recruited via a gatekeeper (Spinal Injuries Scotland). Therapist participants were recruited via social media. Both groups were provided with participant information sheets that detailed the

XSL•FO
RenderX

aims of the study and, if they chose to take part, the requirements of participation and the information collected about them. This information would be stored on Glasgow Caledonian University's OneDrive for 5 years and shared only with the study team. All information would be anonymized after consent was given, and participants assigned an ID number to comply with the General Data Protection Regulation (GDPR). Prior to participation, all participants provided informed consent either by printing and physically signing a digital consent form or by completing the form digitally. Participants who were unable to physically sign the consent form provided consent through a designated witness who signed on their behalf [35]. The study abided by the guidelines set out in the Declaration of Helsinki. Participants were not compensated for their participation and were informed they were free to withdraw from the study at any time without giving a reason. Participants could have their information removed from the study but were informed that this would not be possible after anonymization.

Discussion

Principal Results

The aim of this qualitative study was to explore the lived experiences of people with tetraplegia and SCI specialist therapists in terms of upper limb rehabilitation to inform the design of a VR-based intervention. A total of 5 major themes and their associated subthemes were identified that enabled the specification of software requirements, which have subsequently been implemented in a prototype VR intervention. The focus groups indicated that VR has the potential to help improve function and build strength, as well as learning skills, by adapting to loss of function. This suggests that end users of potential VR-based interventions understand its possible benefits, as well as pitfalls. The results also show that end users see a place for co-designed interventions in the acute/subacute setting and showed interest in VR, which is critical for "buy-in" for the adoption of new technologies.

Major Themes 1 and 2

The findings indicate that VR-based interventions must accommodate substantial variability in upper limb impairment following SCI. To account for diverse user profiles and needs, software must adapt to the personal rehabilitation needs of the user. Although some therapists anticipated uptake among younger patients, and therefore specific targeting toward this demographic, restricting designs to this group would misalign with the clinical diversity of the population. Design should therefore prioritize broad accessibility and usability across impairment, age groups, and levels of technological expertise.

Major Theme 3

This study highlights how motivation and engagement are shaped by both individual and service-level factors during acute/subacute rehabilitation. The findings align with established barriers to engagement in therapy in the SCI literature, including comorbidities (eg, urinary tract infections, pressure sores, fatigue, pain, and low mood), alongside the resource limitations of health care providers, periods of "downtime," variation in individual rehabilitation needs, and patients' desire for more control over their therapy [41-44]. Patients with SCI were generally described by both groups as highly motivated, with engagement closely linked to perceived progress and the pursuit of greater independence. This finding may be influenced by sampling bias, as individuals willing to participate in research may be more motivated than the broader inpatient population. Motivational drivers suggest that interventions incorporating visible progress indicators and structured challenges may be particularly effective. Such features could also increase acceptability of novel interventions, including VR therapy, when used as an adjunct to conventional rehabilitation.

Major Theme 4

Participants with tetraplegia described VR as an assistive technology to enhance function and therefore improve independence, a factor shown to be of high importance to people with SCI [45]. In addition, VR was perceived as a tool to give participants with tetraplegia independence during their rehabilitation. Participants' emphasis on autonomy and meaningful activity suggests that VR may be most effective when it enables users to exercise choice and engage in personally relevant tasks. These factors could also enable therapists in administering or prescribing therapy by facilitating movements that are typically difficult to achieve in usual treatment or that are more appealing to people with tetraplegia.

The repetitive nature of upper limb rehabilitation presents an opportunity for VR to embed prescribed movements within more engaging tasks. VR could maintain motivation by using repetitive input to drive engaging game mechanics, such as firing a bow and arrow, hitting boxing targets, or completing puzzles. Progress-based feedback and structured challenges may further promote sustained participation.

Users could be engaged in gameplay and rewarded in immersive scenarios for repeating movements. This concept relies on the immersive capabilities of VR to provide positive distraction, which is a term inferred from participants' discussion regarding "escapism," from typical therapy while still enabling prescribed movement.

The acute context of rehabilitation is relevant, as patients must adapt to both sudden functional loss and prolonged hospitalization. Participants highlighted the importance of experiences that offer psychological relief from this environment via relaxation and escapism, which are achievable with immersive VR systems [46].

Resource limitations were identified as a barrier to therapy delivery by both groups. Participants implied that a VR-based system could mitigate these limitations by enabling self-guided, semi-independent practice. This could allow therapists to allocate time more efficiently while maintaining a supervisory role in therapy sessions.

The group-based nature of rehabilitation suggests that future VR systems could benefit from supporting multiuser networked environments to encourage peer-to-peer interactions.

Major Theme 5

Both therapists and participants with tetraplegia made many suggestions about what type of activities and upper limb

XSL•FO
RenderX

movements they would like to see implemented in VR. There was no consensus on particular types of VR scenarios with which patients with tetraplegia would engage, reflecting the heterogeneity of the tetraplegic population. This supports the development of flexible systems capable of offering varied activities and environments to their users.

VR is most likely to add value when repetitive therapeutic movements are embedded within adaptive, gamified tasks. These findings are consistent with existing research on rehabilitation during acute/subacute tetraplegia [45,47-50]. Participants' emphasis on independence, motivation, and shared decision-making reinforces the important of human-centered design principles. Unique to this population is the complexity of SCI, which remains a central factor for consideration during intervention development.

These findings add to the growing body of research concerning the recommendations for implementing VR for people with upper limb impairments and their rehabilitation [51-54]. VR interventions must be designed to take advantage of their attributes, which include real-time feedback and immersive environments. These designs must be informed by the specific populations who are intended to use these systems [55-58].

Recommendations for the Design of a VR Upper Limb Rehabilitation Intervention for People With Tetraplegia

The following recommendations outline key design considerations for VR-based upper limb rehabilitation interventions targeting people with tetraplegia. They are structured around practical software development domains to support systematic intervention design and informed by the themes identified in the focus group data.

Domain 1: Accessibility

VR applications should be usable by the widest possible range of people with tetraplegia and accommodate the wide variation in upper limb function of users for control of the system. They should be usable to the same extent for all users, irrespective of upper limb function, and support patient autonomy while maintaining the therapeutic role of clinicians. VR applications should be usable across age groups and comorbidities associated with tetraplegia, allow all prescribed movements to be specified and configured by the user, and must be comfortable to use. They should target as many movement domains as possible and support movement or activity that is difficult to achieve in standard upper limb treatment.

Domain 2: Implementation

VR applications should be feasible, fit into existing clinical workflows, and aim to alleviate resource limitations. They should support collaborative goal setting between patients and therapists and minimize physical and psychological adverse effects associated with, and attributable to, VR. Applications should consider the appropriate intensity and duration of an intervention and prevent overexertion during use. They should also be able to monitor usage and performance and be supported by appropriate training for end users to ensure safe implementation and effective integration into clinical services.

Domain 3: Immersion and Engagement

VR applications should increase and maintain engagement beyond what is achieved in usual treatment by encouraging repetitions of prescribed movements through meaningful virtual activities and environments. They should avoid factors that cause disengagement from therapy by providing an optimal degree of challenge for each user, ensuring that the experience is neither frustrating nor too difficult or too easy. Applications should provide motivational feedback on performance, such as scores and progress, and make use of the immersive and enjoyable qualities of VR technology to ensure users feel present within virtual environments and benefit from escapism or positive distraction.

Limitations

Participants with tetraplegia sometimes had difficulty recalling their time as inpatients, typically when years or decades had passed following discharge. However, it was important to prioritize the perspective of individuals with a full retrospective view of inpatient care.

Separate focus groups were conducted for therapists and participants with tetraplegia to remove the effect of the relationships or perceived power differences between patients and clinicians. For example, when discussing opinions about particular types of upper limb therapy, participants with tetraplegia may have been reluctant to express negative opinions about care if therapists were present. Interaction between participants with tetraplegia and therapists may, however, have been beneficial in terms of ideating suggestions for the design of VR scenarios.

Data saturation was not achieved as the final focus groups in both groups yielded additional insights, though at a lower frequency compared to earlier groups. However, the information power of the focus groups was improved by revising the question schedules after each focus group. In particular, follow-up questions and prompts were added to explore new areas of discussion that had not been covered by previous groups. The inclusion of different rehabilitation disciplines with experience of working in different specialist centers across the United Kingdom ensured diverse groups were assembled, which was a strength of this study and allowed for more varied discussion.

Conclusions

This study reported the perspectives and suggestions of people with tetraplegia and SCI therapists about the use of VR for upper limb rehabilitation in the acute setting. As a result, it presents valuable first-hand information about the design of VR technology for SCI rehabilitation. A total of 5 themes were generated that designers and developers of immersive applications for rehabilitation can use to create their applications. People with tetraplegia have a large variation in upper limb impairments, for which many different movements are prescribed by therapists. This variation must be accommodated by applications in a usable and accessible way. Applications must also implement features that motivate patients to attend and adhere to therapy, such as providing useful feedback and appealing to the broad range of patients' interests, and also

XSL•FO
RenderX

avoid factors that can cause disengagement from therapy. The findings of this study facilitated the cocreation of a novel immersive VR intervention for upper limb rehabilitation for people with tetraplegia while they are inpatients.

Future Work

The next part of this project is to use the results from this study to produce software requirements and a software architecture for a suite of VR games for upper limb therapy. This study enabled the development and deployment of the proposed software and its continued improvement.

Acknowledgments

The authors would like to thank all the participants (co-designers) who took part in this study, Spinal Injuries Scotland for helping recruit participants, and Stoke Mandeville Spinal Research for funding this PhD through their Assistive Technology for Upper Limb Rehabilitation PhD Studentship.

Authors' Contributions

The “Co-designers” members are Ray Brown, Lauren Cope, Graham Cullen, Sheena Egan, Amanda Howat, Donald Hutton, John Lynch, Samantha Maguire, and Sylvia Merino. AG, LP, M Poyade, and M Purcell designed and conceptualized this study. AG and LP facilitated the focus groups. AG and LC were responsible for project administration, including recruitment of focus group participants. AG, LP, M Poyade, and M Purcell performed the analysis. AG and LP wrote and prepared the original draft of the manuscript. AG, LP, M Purcell, M Poyade, and LC reviewed and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Schedule of Questions (PwT).
[DOCX File , 23 KB - rehab_v13i1e66832_app1.docx]

Multimedia Appendix 2
Schedule of Questions (Therapists).
[DOCX File , 25 KB - rehab_v13i1e66832_app2.docx]

References

1. Hutson TH, Di Giovanni S. The translational landscape in spinal cord injury: focus on neuroplasticity and regeneration. Nat Rev Neurol 2019;15(12):732-745. [doi: 10.1038/s41582-019-0280-3] [Medline: 31728042]
2. Thompson C, Mutch J, Parent S, Mac-Thiong J. The changing demographics of traumatic spinal cord injury: an 11-year study of 831 patients. J Spinal Cord Med 2015;38(2):214-223 [FREE Full text] [doi: 10.1179/2045772314Y.0000000233] [Medline: 25096709]
3. New spinal cord injury data revealed. Spinal Injuries Association. URL: https://www.spinal.co.uk/news/new-spinal-cord-injury-data-revealed/ [accessed 2026-01-27]
4. McDaid D, Park A, Gall A, Purcell M, Bacon M. Understanding and modelling the economic impact of spinal cord injuries in the United Kingdom. Spinal Cord 2019;57(9):778-788 [FREE Full text] [doi: 10.1038/s41393-019-0285-1] [Medline: 31086273]
5. Ditunno JF, Cohen ME, Hauck WW, Jackson AB, Sipski ML. Recovery of upper-extremity strength in complete and incomplete tetraplegia: a multicenter study. Arch Phys Med Rehabil 2000;81(4):389-393. [doi: 10.1053/mr.2000.3779] [Medline: 10768525]
6. Mateo S, Di Marco J, Cucherat M, Gueyffier F, Rode G. Inconclusive efficacy of intervention on upper-limb function after tetraplegia: a systematic review and meta-analysis. Ann Phys Rehabil Med 2020;63(3):230-240 [FREE Full text] [doi: 10.1016/j.rehab.2019.05.008] [Medline: 31233828]
7. Rupp R, Biering-Sørensen F, Burns S, Graves D, Guest J, Jones L. International standards for neurological classification of spinal cord injury: revised 2019. Top Spinal Cord Inj Rehabil 2021;27(2):1-22 [FREE Full text] [doi: 10.46292/sci2702-1] [Medline: 34108832]
8. Bickenbach J, Officer A, Shakespeare T, von Groote P, International Spinal Cord Society, World Health Organization. International Perspectives on Spinal Cord Injury. Geneva: World Health Organization; 2013.
9. Adriaansen J, Ruijs L, van Koppenhagen CF, van Asbeck FWA, Snoek G, van Kuppevelt D. Secondary health conditions and quality of life in persons living with spinal cord injury for at least ten years. J Rehabil Med 2016;48(10):853-860 [FREE Full text] [doi: 10.2340/16501977-2166] [Medline: 27834436]

XSL•FO
RenderX

10. Burns AS, Marino RJ, Kalsi-Ryan S, Middleton JW, Tetreault LA, Dettori JR. Type and timing of rehabilitation following acute and subacute spinal cord injury: a systematic review. Global Spine J 2017;7(3 Suppl):175S-194S [FREE Full text] [doi: 10.1177/2192568217703084] [Medline: 29164023]
11. Lynskey JV, Belanger A, Jung R. Activity-dependent plasticity in spinal cord injury. J Rehabil Res Dev 2008;45(2):229-240 [FREE Full text] [doi: 10.1682/jrrd.2007.03.0047] [Medline: 18566941]
12. Musselman KE, Shah M, Zariffa J. Rehabilitation technologies and interventions for individuals with spinal cord injury: translational potential of current trends. J Neuroeng Rehabil 2018;15(1):40 [FREE Full text] [doi: 10.1186/s12984-018-0386-7] [Medline: 29769082]
13. Anderson KD. Targeting recovery: priorities of the spinal cord-injured population. J Neurotrauma 2004;21(10):1371-1383. [doi: 10.1089/neu.2004.21.1371] [Medline: 15672628]
14. Roy RR, Harkema SJ, Edgerton VR. Basic concepts of activity-based interventions for improved recovery of motor function after spinal cord injury. Arch Phys Med Rehabil 2012;93(9):1487-1497. [doi: 10.1016/j.apmr.2012.04.034] [Medline: 22920448]
15. Savic G, Frankel HL, Jamous MA, Soni BM, Charlifue S. Participation restriction and assistance needs in people with spinal cord injuries of more than 40 year duration. Spinal Cord Ser Cases 2018;4:28 [FREE Full text] [doi: 10.1038/s41394-018-0056-9] [Medline: 29619249]
16. Lu X, Battistuzzo CR, Zoghi M, Galea MP. Effects of training on upper limb function after cervical spinal cord injury: a systematic review. Clin Rehabil 2015;29(1):3-13. [doi: 10.1177/0269215514536411] [Medline: 25575932]
17. Kramer JLK, Lammertse DP, Schubert M, Curt A, Steeves JD. Relationship between motor recovery and independence after sensorimotor-complete cervical spinal cord injury. Neurorehabil Neural Repair 2012;26(9):1064-1071 [FREE Full text] [doi: 10.1177/1545968312447306] [Medline: 22647878]
18. Kalsi-Ryan S, Beaton D, Curt A, Popovic MR, Verrier MC, Fehlings MG. Outcome of the upper limb in cervical spinal cord injury: profiles of recovery and insights for clinical studies. J Spinal Cord Med 2014;37(5):503-510. [doi: 10.1179/2045772314Y.0000000252] [Medline: 25229734]
19. Quel de Oliveira C, Refshauge K, Middleton J, de Jong L, Davis GM. Effects of activity-based therapy interventions on mobility, independence, and quality of life for people with spinal cord injuries: a systematic review and meta-analysis. J Neurotrauma 2017;34(9):1726-1743. [doi: 10.1089/neu.2016.4558] [Medline: 27809702]
20. Huh S, Ko H. Recovery target priorities of people with spinal cord injuries in Korea compared with other countries: a survey. Spinal Cord 2020;58(9):998-1003. [doi: 10.1038/s41393-020-0457-z] [Medline: 32246088]
21. Simpson LA, Eng JJ, Hsieh JT, Wolfe DL, Spinal Cord Injury Rehabilitation Evidence Scire Research Team. The health and life priorities of individuals with spinal cord injury: a systematic review. J Neurotrauma 2012;29(8):1548-1555 [FREE Full text] [doi: 10.1089/neu.2011.2226] [Medline: 22320160]
22. Namrata Y. Priorities of spinal cord injured population – a survey. Am J Appl Psychol 2017;6(6):183. [doi: 10.11648/j.ajap.20170606.17]
23. Behrman A, Ardolino E, Harkema S. Activity-based therapy: from basic science to clinical application for recovery after spinal cord injury. J Neurol Phys Ther 2017;41 Suppl 3(Suppl 3 IV STEP Spec Iss):S39-S45. [doi: 10.1097/NPT.0000000000000184] [Medline: 28628595]
24. Kazim SF, Bowers CA, Cole CD, Varela S, Karimov Z, Martinez E. Corticospinal motor circuit plasticity after spinal cord injury: harnessing neuroplasticity to improve functional outcomes. Mol Neurobiol 2021;58(11):5494-5516. [doi: 10.1007/s12035-021-02484-w] [Medline: 34341881]
25. Scalise M, Bora TS, Zancanella C, Safa A, Stefini R, Cannizzaro D. Virtual reality as a therapeutic tool in spinal cord injury rehabilitation: a comprehensive evaluation and systematic review. J Clin Med 2024;13(18):5429 [FREE Full text] [doi: 10.3390/jcm13185429] [Medline: 39336916]
26. Vinciguerra C, Federico A. Virtual rehabilitation in spinal cord injury patients: realities and future perspectives: a systematic review. Curr Phys Med Rehabil Rep 2024;12(4):425-434. [doi: 10.1007/s40141-024-00470-6]
27. Schiza E, Matsangidou M, Neokleous K, Pattichis CS. Virtual reality applications for neurological disease: a review. Front Robot AI 2019;6:100 [FREE Full text] [doi: 10.3389/frobt.2019.00100] [Medline: 33501115]
28. Leemhuis E, Esposito RM, De Gennaro L, Pazzaglia M. Go virtual to get real: virtual reality as a resource for spinal cord treatment. Int J Environ Res Public Health 2021;18(4):1819 [FREE Full text] [doi: 10.3390/ijerph18041819] [Medline: 33668438]
29. De Miguel-Rubio A, Rubio MD, Alba-Rueda A, Salazar A, Moral-Munoz JA, Lucena-Anton D. Virtual reality systems for upper limb motor function recovery in patients with spinal cord injury: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020;8(12):e22537. [doi: 10.2196/22537] [Medline: 33270040]
30. Orsatti Sánchez BA, Diaz Hernandez O. Efficacy of virtual reality in neurorehabilitation of spinal cord injury patients: a systematic review. Rev Mex Ing Bioméd 2021;42(2):90-103 [FREE Full text] [doi: 10.17488/RMIB.42.2.8]
31. de Araújo AVL, Neiva JFDO, Monteiro CBDM, Magalhães FH. Efficacy of virtual reality rehabilitation after spinal cord injury: a systematic review. Biomed Res Int 2019;2019:7106951 [FREE Full text] [doi: 10.1155/2019/7106951] [Medline: 31828120]

XSL•FO
RenderX

32. Bird M, McGillion M, Chambers EM, Dix J, Fajardo CJ, Gilmour M. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021;7(1):12 [FREE Full text] [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]
33. Bourke JA, Bragge P, River J, Sinnott Jerram KA, Arora M, Middleton JW. Shining a light on the road towards conducting principle-based co-production research in rehabilitation. Front Rehabil Sci 2024;5:1386746 [FREE Full text] [doi: 10.3389/fresc.2024.1386746] [Medline: 38660394]
34. Bate P, Robert G. Experience-based design: from redesigning the system around the patient to co-designing services with the patient. Qual Saf Health Care 2006;15(5):307-310 [FREE Full text] [doi: 10.1136/qshc.2005.016527] [Medline: 17074863]
35. The Medicines and Healthcare products Regulatory Agency, Health Research Authority. Joint statement on seeking consent by electronic methods. UK Government & HRA/MHRA. 2018. URL: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/hra-mhra-econsent-statement-sept-18.pdf [accessed 2026-01-27]
36. Fekete C, Eriks-Hoogland I, Baumberger M, Catz A, Itzkovich M, Lüthi H. Development and validation of a self-report version of the spinal cord independence measure (SCIM III). Spinal Cord 2013;51(1):40-47. [doi: 10.1038/sc.2012.87] [Medline: 22890418]
37. SteamVR featuring the HTC Vive. Valve Corporation. 2016. URL: https://youtu.be/qYfNzhLXYGc [accessed 2026-01-27]
38. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
39. NVivo qualitative data analysis software. QSR International Pty Ltd. 2018. URL: https://lumivero.com/products/nvivo/ [accessed 2026-01-27]
40. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
41. Vissers M, van den Berg-Emons R, Sluis T, Bergen M, Stam H, Bussmann H. Barriers to and facilitators of everyday physical activity in persons with a spinal cord injury after discharge from the rehabilitation centre. J Rehabil Med 2008;40(6):461-467 [FREE Full text] [doi: 10.2340/16501977-0191] [Medline: 18509562]
42. Tallqvist S, Eskola K, Täckman A, Kauppila A, Koskinen E, Anttila H. Facilitators and barriers in the rehabilitation process described by persons with spinal cord injury: a deductive-inductive analysis from the Finnish spinal cord injury study. Ann Med 2023;55(2):2303398 [FREE Full text] [doi: 10.1080/07853890.2024.2303398] [Medline: 38232950]
43. Whalley Hammell K. Experience of rehabilitation following spinal cord injury: a meta-synthesis of qualitative findings. Spinal Cord 2007;45(4):260-274. [doi: 10.1038/sj.sc.3102034] [Medline: 17310257]
44. Lindberg J, Kreuter M, Taft C, Person L. Patient participation in care and rehabilitation from the perspective of patients with spinal cord injury. Spinal Cord 2013;51(11):834-837. [doi: 10.1038/sc.2013.97] [Medline: 23999110]
45. Sousa SS, Andrade MJ, Fernandes CS, Barbeiro SR, Teixeira VT, Pereira RS. Healthcare experience of people with acute spinal cord injury: a phenomenological study. Nurs Rep 2023;13(4):1671-1683 [FREE Full text] [doi: 10.3390/nursrep13040138] [Medline: 38133114]
46. Williamson SD, Aaby AO, Ravn SL. Psychological outcomes of extended reality interventions in spinal cord injury rehabilitation: a systematic scoping review. Spinal Cord 2025;63(2):58-65. [doi: 10.1038/s41393-024-01057-7] [Medline: 39789357]
47. Kirchberger I, Sinnott A, Charlifue S, Kovindha A, Lüthi H, Campbell R. Functioning and disability in spinal cord injury from the consumer perspective: an international qualitative study using focus groups and the ICF. Spinal Cord 2010;48(8):603-613. [doi: 10.1038/sc.2009.184] [Medline: 20065983]
48. Krysa JA, Gregorio MP, Pohar Manhas K, MacIsaac R, Papathanassoglou E, Ho CH. Empowerment, communication, and navigating care: the experience of persons with spinal cord injury from acute hospitalization to inpatient rehabilitation. Front Rehabil Sci 2022;3:904716 [FREE Full text] [doi: 10.3389/fresc.2022.904716] [Medline: 36188987]
49. Hearn JH, Finlay KA, Fine PA, Cotter I. Neuropathic pain in a rehabilitation setting after spinal cord injury: an interpretative phenomenological analysis of inpatients' experiences. Spinal Cord Ser Cases 2017;3:17083 [FREE Full text] [doi: 10.1038/s41394-017-0032-9] [Medline: 29423289]
50. Kerstin W, Gabriele B, Richard L. What promotes physical activity after spinal cord injury? An interview study from a patient perspective. Disabil Rehabil 2006;28(8):481-488. [doi: 10.1080/09638280500211932] [Medline: 16513581]
51. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018;10(11):1237-1251.e1 [FREE Full text] [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
52. Herrera V, Albusac J, Schez-Sobrino S, Reyes-Guzmán A, Glez-Morcillo C. Dynamic adjustment of interactive objects in virtual environments for upper limb rehabilitation: a patient-centred solution. : The Eurographics Association; 2024 Presented at: Proceedings of the Spanish Computer Graphics Conference (CEIG); June 17-19, 2024; A Coruña. [doi: 10.1007/s10055-024-01078-w]
53. Herrera V, Albusac J, Castro-Schez JJ, González-Morcillo C, Monekosso DN, Pacheco S. Creating adapted environments: enhancing accessibility in virtual reality for upper limb rehabilitation through automated element adjustment. Virtual Reality 2025;29(1). [doi: 10.1007/s10055-024-01078-w]

XSL•FO
RenderX

54. Lewis GN, Rosie JA. Virtual reality games for movement rehabilitation in neurological conditions: how do we meet the needs and expectations of the users? Disabil Rehabil 2012;34(22):1880-1886. [doi: 10.3109/09638288.2012.670036] [Medline: 22480353]
55. Creed C, Al-Kalbani M, Theil A, Sarcar S, Williams I. Inclusive AR/VR: accessibility barriers for immersive technologies. Univ Access Inf Soc 2023;23(1):59-73. [doi: 10.1007/s10209-023-00969-0]
56. Mott M, Cutrell E, Gonzalez Franco M, Holz C, Ofek E, Stoakley R. Accessible by design: an opportunity for virtual reality. Proc IEEE Int Symp Mix Augment Real Adjunct 2019:451-454. [doi: 10.1109/ismar-adjunct.2019.00122]
57. Dudley J, Yin L, Garaj V, Kristensson PO. Inclusive immersion: a review of efforts to improve accessibility in virtual reality, augmented reality and the metaverse. Virtual Reality 2023;27(4):2989-3020. [doi: 10.1007/s10055-023-00850-8]
58. Vlahovic S, Suznjevic M, Skorin-Kapov L. A framework for the classification and evaluation of game mechanics for virtual reality games. Electronics 2022;11(18):2946. [doi: 10.3390/electronics11182946]

Abbreviations

ABT: Activity-Based Therapy
GDPR: General Data Protection Regulation
PSWAHS: Psychology, Social Work, and Allied Health Sciences
SCI: spinal cord injury
VR: virtual reality

Edited by A Scano; submitted 08.Apr.2025; peer-reviewed by MM Ahmed Al Nattah, J Thrift, H Heppner; comments to author 01.Oct.2025; revised version received 17.Dec.2025; accepted 17.Jan.2026; published 11.Feb.2026.

Please cite as:
Goodsell A, Purcell M, Poyade M, Cownie L, Co-designers, Paul L
Design Recommendations for Virtual Reality–Based Upper Limb Exercises From People With Tetraplegia and Spinal Cord Injury Rehabilitation Specialists: Focus Group Study
JMIR Rehabil Assist Technol 2026;13:e66832
URL: https://rehab.jmir.org/2026/1/e66832
doi:10.2196/66832
PMID:

©Andrew Goodsell, Mariel Purcell, Matthieu Poyade, Louise Cownie, Co-designers, Lorna Paul. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and Satisfaction of a Multidomain Information and Communication Technology–Based Rehabilitation Management Program in Patients With Postacute Stroke: Multicenter Formative Evaluation

Won Kee Chang[1], MD, MS; Yun-Sun Jung[1], PhD; Ji-Soo Choi[1], MD, MS; Won-Seok Kim[1], MD, PhD; Min Kyun Sohn[2], MD, PhD; Sungju Jee[2], MD, PhD; Yong-Il Shin[3], MD, PhD; Sung-Hwa Ko[3], MD, PhD; Nam-Jong Paik[1], MD, PhD

[1]Department of Rehabilitation Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 82 Gumi-ro 173 Beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do, Republic of Korea

[2]Department of Rehabilitation Medicine, Chungnam National University College of Medicine, Chungnam National University Hospital, Daejeon, Republic of Korea

[3]Department of Rehabilitation Medicine, Pusan National University School of Medicine, Pusan National University Yangsan Hospital, Yangsan, Gyeongsangnam-do, Republic of Korea

Corresponding Author:
Nam-Jong Paik, MD, PhD
Department of Rehabilitation Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 82 Gumi-ro 173 Beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do, Republic of Korea

Abstract

Background: Information and communication technology (ICT) can help bridge the postacute stroke care gap by supporting home-based rehabilitation, but evidence on the usability and adoption of such programs in Asian health care settings is limited.

Objective: This study aimed to describe the usability and satisfaction of a tablet-based, multidomain ICT-based rehabilitation management program in patients with postacute stroke, and to explore descriptively how participants with and without acceptable usability differed.

Methods: This multicenter prospective cohort study enrolled patients with postacute stroke who were discharged home within 30 days of onset from 3 university hospitals. Participants used a tablet-based program providing medication guidance, rehabilitation content in 4 therapy domains, clinic scheduling, and social welfare information. Usability was defined according to ISO 9241-11:2018 and assessed with the System Usability Scale (SUS), with a prespecified threshold of 68. Satisfaction (6 items, 5-point Likert scale), self-reported usage, and open-ended responses were collected at 1 month after discharge and 3 months after onset. Mann-Whitney *U* tests, Cohen *d* effect sizes, and Firth penalized logistic regression compared participants with and without acceptable usability.

Results: Of 24 enrolled patients (mean age 62.5, SD 9.1 years; 17 male), 22 completed the 1-month, and 8 the 3-month evaluation. The mean 1-month SUS score was 60.34 (SD 17.48), below the threshold of 68, and 6 of 22 participants (27%) scored ≥68. Patients younger than 65 years scored higher than older patients (68.3 vs 50.8; *P*=.02). Participants scoring ≥68 were younger (56.2 vs 64.2 years), had higher cognitive scores (27.5 vs 24.2) and fewer depressive symptoms (1.3 vs 6), and more often used the program for more than 20 minutes per session (6/6, 100% vs 7/16, 44%; *P*=.03). Satisfaction ranged from 3.68 to 4.23 out of 5 and was higher in all 6 domains among participants with lower ambulatory function, although no comparison reached significance (averaged across items, 4.29 vs 3.70; *P*=.09). Completers and noncompleters of the 3-month evaluation did not differ in baseline characteristics or 1-month SUS score (*P*=.97).

Conclusions: In this formative evaluation, a multidomain ICT-based rehabilitation management program showed suboptimal usability alongside favorable satisfaction. With only 6 participants reaching the usability threshold, the observed differences in age, cognitive function, and depressive symptoms are hypotheses for adequately powered studies rather than determinants. End users helped define the program content through focus group interviews, but the design itself was never tested with users before deployment, which plausibly explains the dissociation between favorable content satisfaction and poor interface usability. User involvement in requirements definition is necessary but not sufficient.

Trial Registration: ClinicalTrials.gov NCT04720820; https://clinicaltrials.gov/study/NCT04720820

(JMIR Rehabil Assist Technol 2026;13:e98098) doi:10.2196/98098

XSL•FO
RenderX

KEYWORDS

stroke rehabilitation; telerehabilitation; transitional care; usability; patient satisfaction; digital health; information and communication technology; formative evaluation

Introduction

Stroke remains the second leading cause of disability-adjusted life years globally, imposing a substantial burden on both individuals and health care systems [1,2]. Following acute treatment and initial hospitalization, patients with subacute stroke who are discharged home frequently face a critical gap in continuity of care. Systematic reviews have demonstrated that community-dwelling survivors with stroke experience substantial unmet needs across multiple domains, including rehabilitation, self-care, cognition, mood, and health education [3,4]. Nearly half of survivors with stroke in the community report at least one unmet rehabilitative need, with the highest proportions observed in the domains of anxiety and depression, self-care, and pain management [4]. Critically, the number of unmet needs is an independent negative predictor of health-related quality of life, and higher levels of unmet needs have been associated with greater functional dependency and poorer community reintegration [3-5]. These findings underscore the urgent need for systematic approaches that extend rehabilitative management beyond the hospital setting to support survivors with stroke during the vulnerable postdischarge period.

Information and communication technology (ICT) has emerged as a promising approach for bridging this postacute care gap by extending rehabilitation services into the home environment [6,7]. Telerehabilitation and mobile health (mHealth) platforms can facilitate home-based exercise programs, enable remote monitoring of functional recovery, and deliver patient education [8]. A Cochrane review of telerehabilitation services for stroke found that remote-delivered interventions achieved comparable outcomes to in-person rehabilitation for motor function, activities of daily living, and quality of life [6]. More recently, app-based and tablet-based programs have been developed to deliver multidomain rehabilitation content, including physical, cognitive, and language therapy, with several pilot studies reporting high acceptability and feasibility among survivors with stroke [6,9]. The Comprehensive Post-Acute Stroke Services (COMPASS) trial in the United States demonstrated that integrating digital health tools into a transitional care pathway improved care coordination and patient-reported outcomes [8]. An AI-driven cognitive telerehabilitation platform was shown to be noninferior to therapist-supervised rehabilitation in patients with subacute stroke in a recent randomized controlled trial [10]. Despite this expanding evidence base, most ICT-based rehabilitation programs have focused on a single domain (eg, upper limb motor function or cognition), and comprehensive multidomain programs that address the diverse postdischarge needs of survivors with stroke remain scarce.

A fundamental prerequisite for the successful implementation of any ICT-based rehabilitation program is adequate usability, as poor usability directly undermines treatment adherence, sustained engagement, and ultimately clinical effectiveness [11,12]. In stroke populations, usability presents a particular challenge because patients exhibit highly variable characteristics in age, cognitive function, motor impairment, and prior technology experience, all of which may influence the ability to adopt and effectively use digital health tools [9]. A systematic review of barriers to mHealth access among survivors with stroke identified older age, cognitive impairment, low digital literacy, and limited caregiver support as key impediments to technology adoption [11]. Broderick et al [9] further demonstrated that technology usability and acceptability were significant determinants of adoption in a cohort of survivors with stroke using a self-directed upper limb rehabilitation system, with stroke severity mediating the relationship. Similarly, a blended-care telerehabilitation platform developed for survivors with stroke was rated highly usable by inpatients and therapists, yet adherence to the prescribed training varied widely between participants, indicating that usability and sustained engagement are not equivalent [13]. However, few studies have systematically examined how specific patient characteristics relate to usability outcomes in the context of comprehensive, multidomain ICT-based rehabilitation programs for postacute stroke care.

In this formative evaluation, we aimed to describe the usability and satisfaction of a tablet-based, multidomain rehabilitation management program developed for patients with postacute stroke, encompassing rehabilitation content across physical, swallowing, cognitive, and language domains, along with medication guidance, clinic scheduling, and social welfare information. The program was provided to patients discharged home after acute stroke, and usability was assessed at 1 month after discharge and 3 months after stroke onset. As a secondary, exploratory objective, we described how participants who did and did not reach the conventional threshold for acceptable usability differed in demographic and clinical characteristics. Given the small size of this cohort, this second objective was framed as hypothesis-generating: our aim was to identify candidate characteristics for adequately powered future studies rather than to establish determinants of usability.

Methods

Study Design

This was a multicenter prospective cohort study conducted as a substudy of the Korean Model for Post-Acute Comprehensive Rehabilitation (KOMPACT) project, a pragmatic randomized controlled trial investigating the efficacy of early supported discharge (ESD) for patients with postacute stroke in Korea [4]. The study was conducted at 3 university hospitals: Seoul National University Bundang Hospital (SNUBH; Seongnam), Chungnam National University Hospital (CNUH; Daejeon), and Pusan National University Yangsan Hospital (PNUYH; Yangsan). This substudy was a single-arm prospective cohort evaluation and was not itself randomized; the parent randomized trial (ClinicalTrials.gov NCT04720820) is reported separately [4].

XSL•FO
RenderX

Conceptual Framework and Usability Assessment

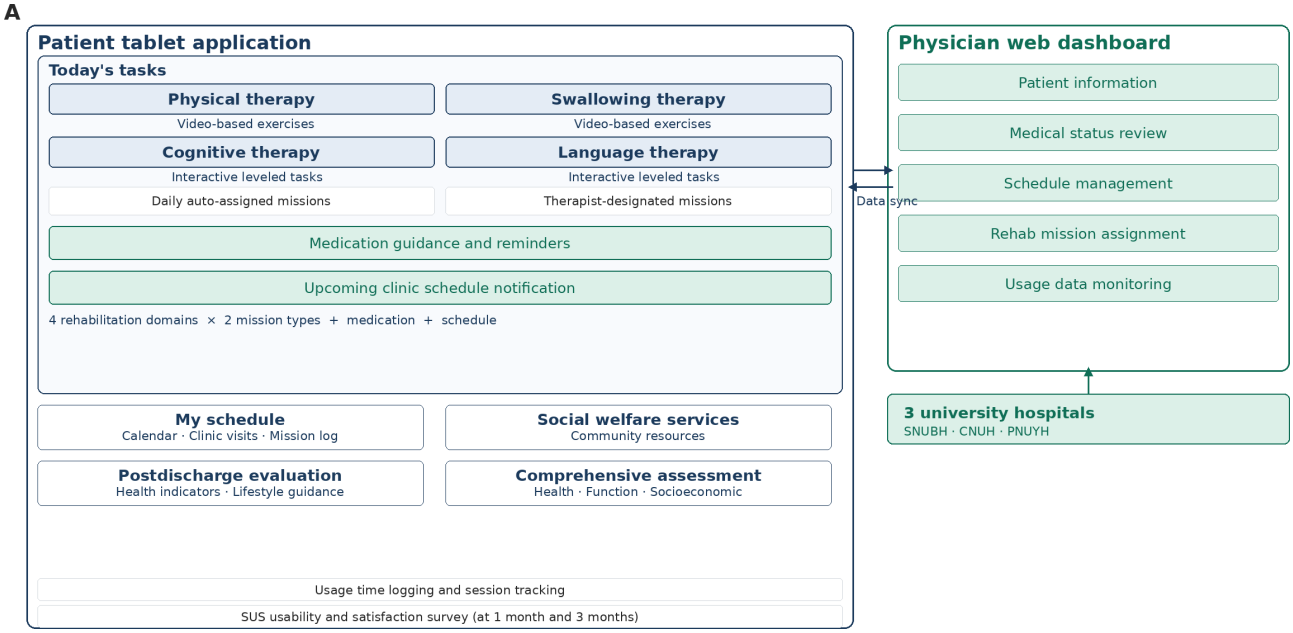
We adopted the ISO 9241-11:2018 definition of usability: the extent to which a system can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [14]. The users were adults recovering from mild-to-moderate stroke (modified Rankin Scale [mRS] 1‐3; functional ambulatory category [FAC] score of ≥3) living at home with a caregiver, with heterogeneous cognitive function and technology experience, together with their caregivers. The goals were to perform prescribed home exercises across physical, occupational, swallowing, and language domains, to manage medications and appointments, and to access rehabilitation and welfare information. The environment was the participant's home during the first 3 months after stroke, using a loaned tablet without on-site technical support.

Usability was measured with the System Usability Scale (SUS), a validated 10-item questionnaire yielding a score from 0 to 100 [15,16]. A score of ≥68 was prespecified as the threshold for acceptable usability; this value approximates the mean of the reference distribution from general software evaluations and the "OK"-to-"good" boundary of the adjective rating scale [16,17], and is the benchmark most often applied to stroke rehabilitation technologies, permitting comparison with that literature. Because it was not derived from digital health applications, and a meta-analysis of digital health apps found a differing distribution [18], results are also reported as the full SUS distribution. The SUS captures the satisfaction component of ISO 9241‐11 principally and does not directly measure effectiveness or efficiency.

Development of the ICT-Based Program

The tablet-based program was developed by a multidisciplinary team of rehabilitation physicians, nurses, physical and occupational therapists, speech-language pathologists, and medical social workers (Figure 1). It was installed on a 10.4-inch Android tablet (Galaxy Tab S6 Lite; Samsung Electronics) provided to participants at discharge and comprised 5 domains: (1) medication guidance with reminders; (2) outpatient appointment notifications; (3) rehabilitation modules for physical, occupational, dysphagia, and language therapy with instructional videos; (4) social welfare and community resource information; and (5) a comprehensive evaluation module tracking functional status. A web-based physician dashboard allowed clinicians to review usage data and functional status at outpatient visits. Screenshots are shown in Multimedia Appendix 1.

Figure 1. Schematic overview of the information and communication technology–based rehabilitation management program. The system comprises a patient application and a physician web dashboard synchronized through a central database across the 3 hospitals. The patient app delivers 4 rehabilitation domains (physical, cognitive, swallowing, and language) through auto-assigned and therapist-designated missions, alongside scheduling, educational, and assessment modules; the dashboard supports remote monitoring and personalized mission allocation. Usability and satisfaction were assessed at 1 month after discharge and 3 months after onset. CNUH: Chungnam National University Hospital; PNUYH: Pusan National University Yangsan Hospital; SNUBH: Seoul National University Bundang Hospital; SUS: System Usability Scale.



The program was designed and built at SNUBH; the 2 other centers contributed participants but were not involved in development. Before specification, focus group interviews (FGIs) were held separately with 4 stakeholder groups: 5 physical and occupational therapists; 5 rehabilitation physicians; 5 survivors with stroke, and 5 accompanying family caregivers; and 2 medical social workers. Each group met once for 60 to 90 minutes in December 2020, before parent-trial enrollment began in March 2021, and was moderated by the first author. Participants described difficulties after discharge, unmet rehabilitation and information needs, and the functions a home-based tablet program should provide. The content domains subsequently implemented were derived from these needs and reconciled with the clinical requirements of the KOMPACT transitional care model. Session notes were consolidated by team consensus into a list of required functions; verbatim transcripts were not retained, and no formal qualitative coding was applied. The FGIs were a development activity rather than

XSL•FO
RenderX

a research study and are reported to document the provenance of the program's content, not as qualitative findings.

Interface and interaction design (screen architecture, navigation depth, text and icon sizing, and the mapping of content onto screens) was then determined by the clinical investigators with HEALTHCONNECT Co, Ltd (Seoul, Republic of Korea), an external health information technology company contracted for the project, and implemented in a single specification-and-build cycle; the same company hosted and operated the platform throughout the study. The design was not reevaluated with survivors with stroke or caregivers before deployment, and no prototype usability testing was conducted. In terms of the human-centered design cycle of ISO 9241-210:2019 [19], user involvement was concentrated in understanding the context of use and specifying requirements, whereas evaluation of the design solution with end users was deferred until after deployment; this study constitutes that first evaluation cycle.

Participants

Participants were recruited from the 3 university hospitals between March 2021 and June 2023. The inclusion criteria were as follows: (1) clinical diagnosis of acute stroke (ischemic or hemorrhagic, excluding transient ischemic attack); (2) expected discharge to home within 30 days of stroke onset; (3) mRS score of 1 to 3; (4) FAC score of 3 or greater; and (5) no deficit of consciousness (National Institutes of Health Stroke Scale [NIHSS] items 1a, 1b, and 1c all scored 0). Exclusion criteria included medical instability, indwelling urinary catheter, prohibited oral intake, uncontrolled pain, psychobehavioral problems, severe cognitive deficit, and inability to use the ICT-based program. The study was planned to enroll 30 participants, 10 at each of the 3 centers. Enrollment was constrained by the COVID-19 pandemic, during which admissions for acute stroke rehabilitation and the opportunity to approach eligible patients before discharge were both reduced, and recruitment closed at 24 participants when enrollment for the parent trial ended in June 2023. Because participants were identified within the workflow of the ongoing parent trial, screening logs recording the numbers screened, found eligible, approached, and declining participation were not maintained systematically across the 3 sites and could not be reconstructed retrospectively.

Procedures

Prior to discharge, participants were familiarized with the ICT-based program through a structured orientation session conducted by the research team. Participants were encouraged to use the program for at least 30 minutes per day until 3 months after stroke onset. Outcome assessments were performed at 2 time points: 1 month after discharge and 3 months after stroke onset. Participants had the option to discontinue usage at 1 month after discharge if they did not wish to continue, at which time the tablet was returned and a final evaluation was conducted.

These assessment time points were inherited from the parent KOMPACT trial, in which 3 months after onset was selected to correspond to the end of the period of most rapid functional recovery and to the point at which responsibility for transitional care passes to community services. We acknowledge that this time point lies within the subacute phase of stroke recovery, so that usability was measured during a period of ongoing functional change. Because participants were discharged home at varying intervals after onset (all within 30 days by protocol), the interval between the 2 assessments necessarily varied between participants and could be short for participants discharged late; the observed intervals are reported in the Results.

Outcome Measures

The primary outcome was the SUS. Secondary outcomes included a satisfaction survey consisting of 6 items assessed on a 5-point Likert scale (program quality, fulfillment of postdischarge health care needs, overall satisfaction, willingness to recommend, problem-solving assistance, and intention for future use), usage patterns, content module evaluation (frequency of use and perceived helpfulness of each rehabilitation module), and structured open-ended responses.

The satisfaction instrument was developed by the research team for the parent KOMPACT trial and adapted for this study. It has not undergone formal psychometric validation, and no reliability or validity data are available for it; scores are therefore reported item by item rather than as a composite.

All usage metrics, including the number of sessions per day, the number of sessions per week, the session-duration category, and the use of each content module were self-reported by participants or their caregivers on a structured questionnaire administered at each assessment visit. Although the physician dashboard displayed usage information during outpatient visits, server-side usage logs were not exported or retained for research purposes; consequently, no objectively recorded usage data were available for analysis. Session duration was recorded on a 4-level ordinal scale (1,≤20 minutes; 2, 21-40 minutes; 3, 41 - 60 minutes; and 4,>60 minutes) and is reported as a frequency distribution. No continuous measure of session duration in minutes or hours was collected.

Analysis of Open-Ended Responses

Narrative data comprised brief free-text answers to 3 structured open-ended prompts administered at each assessment visit (helpful aspects of the program, unhelpful aspects, and suggestions for improvement), recorded verbatim in Korean by the assessor at the outpatient visit. Responses were analyzed by inductive content categorization rather than by a formal qualitative methodology. All responses were read and categorized by a single investigator (WKC), who grouped statements into inductively derived categories; no second coder was involved, and interrater reliability was therefore not assessed. Category frequencies are reported as the number of participants contributing at least 1 statement to each category. Responses were categorized in the original Korean, and illustrative statements were translated into English by a bilingual member of the research team. Because responses were brief, elicited by fixed prompts, and categorized by 1 coder, no claim is made to thematic saturation or to analytic triangulation, and the results are reported descriptively as structured participant feedback.

XSL•FO
RenderX

Statistical Analysis

Descriptive statistics were used to summarize demographic and clinical characteristics, SUS scores, satisfaction scores, and usage patterns. Mann-Whitney *U* tests were used to compare SUS and satisfaction scores between subgroups defined by age (<65 vs≥65 years), sex, ambulatory function (FAC 3 vs FAC 4 - 5), cognitive status (Korean Mini-Mental State Examination [K-MMSE] score≥24 vs<24), and program usage duration (1-mo vs 3-mo users). The Fisher exact test was used for categorical comparisons.

Participants were dichotomized into SUS score≥68 and SUS score<68 groups. Cohen *d* effect sizes were calculated for continuous variables (0.2 small, 0.5 medium, and 0.8 large). Because only 6 participants had SUS scores of ≥68, the events-per-variable (EPV) ratio (outcome events per candidate predictor) was 0.86, far below the conventional minimum of 10, so multivariable modeling was not supportable. Firth penalized likelihood logistic regression [20,21] was therefore used to estimate odds ratios for each predictor individually (age, sex, K-MMSE, Patient Health Questionnaire-9 [PHQ-9], mRS, FAC, and Korean version of the modified Barthel index [K-MBI]); these estimates are reported for completeness, not as tests of prediction. Paired 1-month and 3-month scores were compared with the Wilcoxon signed-rank test. Completers and noncompleters of the 3-month evaluation were compared on baseline, usage, usability, and satisfaction variables. In an exploratory analysis of response style, each SUS item was converted to its 0-to-4 contribution score, mean contributions were computed separately for the 5 positively and 5 negatively worded items, and their difference was compared across cognitive and age strata. Analyses used Python 3.10 (SciPy, statsmodels) and the logistf package in R 4.3.2. Significance was set at $P<.05$.

Ethical Considerations

The study protocol was reviewed and approved by the institutional review board of each participating institution: SNUBH (B-2012/654 - 308), CNUH (2020-12-094), and PNUYH (04-2021-004). Written informed consent was obtained from all participants or their legal guardians prior to enrollment. The study was conducted in accordance with the Declaration of Helsinki.

Participants and caregivers were informed during consent about what data the program would record, how they would be stored and used, who could access them, and for how long they would be retained; the consent form was approved by the institutional review boards. Because the platform was built and hosted by HEALTHCONNECT Co, Ltd, we distinguish safeguards the research team could verify from those within the contractor's responsibility. Verified from the system and its documentation: data were transmitted over an encrypted (HTTPS) connection; access to the clinician web system required individually issued credentials with complex passwords; accounts were provisioned by institution and professional role, so that each clinician could retrieve only the patient list of their own institution; and account creation, modification, and password reset were restricted to a single administrator account at the coordinating center. Server configuration, encryption at rest, access logging, and any certification of the hosting environment were the contractor's responsibility and were not documented to the research team; we make no claim about them. For the research dataset, direct identifiers were held in a separate file under the principal investigator and replaced by study codes. Data were stored in password-protected files on encrypted external storage in a locked research office, released to approved third parties only in nonidentifiable form, and are retained for 3 years after study completion under Article 15 of the Enforcement Rule of the Bioethics and Safety Act, after which documents containing personal information are destroyed under Article 16 of the Enforcement Decree of the Personal Information Protection Act. Participants were reimbursed for travel. The approvals cited above pertain to the cohort study. The FGIs were conducted at SNUBH in December 2020 as a program-development activity and did not collect or retain individually identifiable research data.

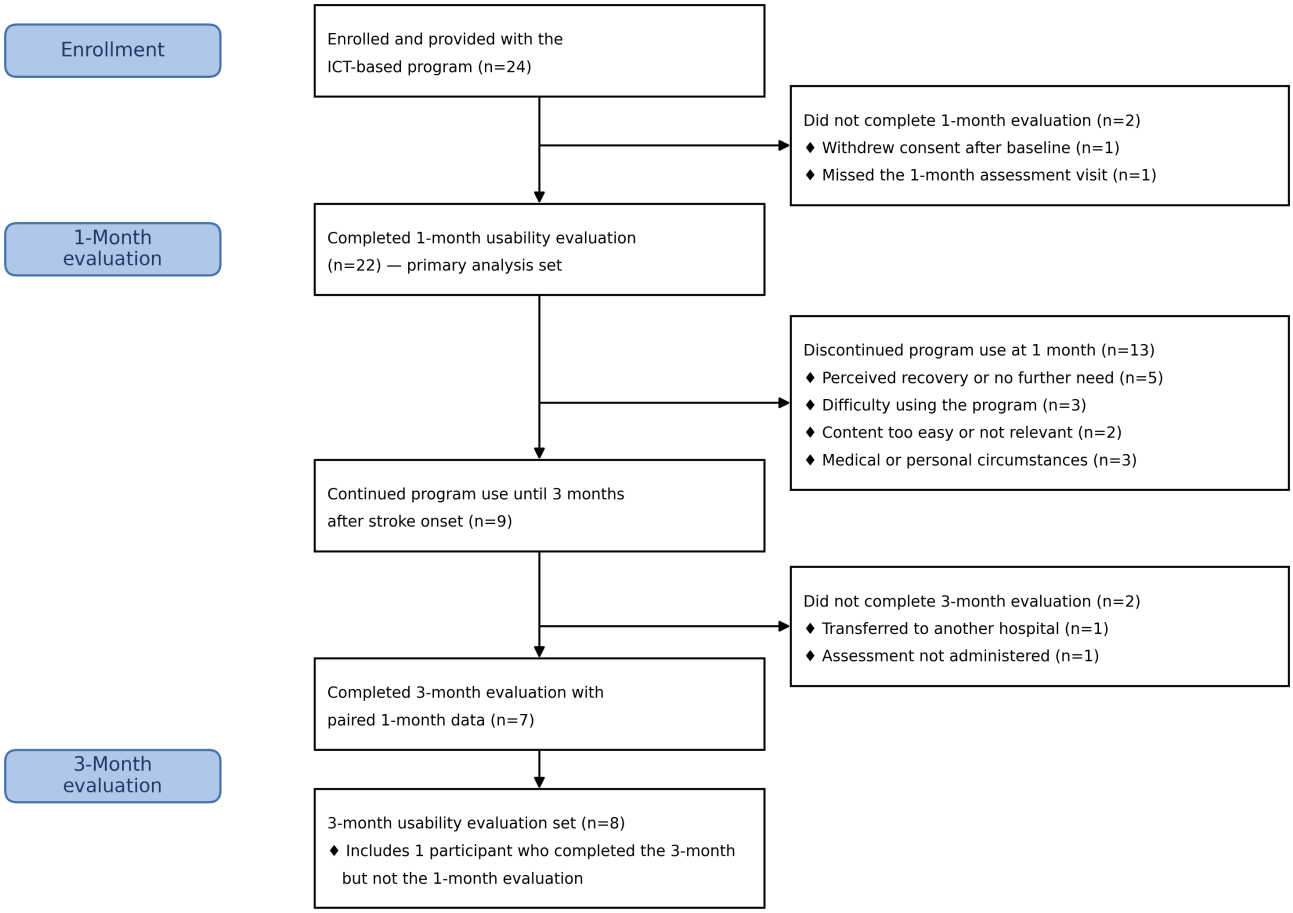
Results

Participant Flow

In total, 24 participants were enrolled and provided with the program (SNUBH, n=4; CNUH, n=10; and PNUYH, n=10). Two did not complete the 1-month evaluation (1 withdrew consent after baseline; 1 missed the visit), leaving 22 in the primary analysis set. At 1 month, 13 of 22 (59%) discontinued program use, citing perceived recovery or no further need (n=5), difficulty using the program (n=3), content that was too easy or not relevant (n=2), and competing medical or personal circumstances (n=3). Of the 9 who continued, 7 completed the 3-month evaluation (1 was transferred to another hospital; 1 was not assessed at the scheduled visit). Including the participant who missed the 1-month but attended the 3-month assessment, the 3-month set comprised 8 participants, 7 of whom had paired data (Figure 2). Numbers screened, eligible, approached, and declining were not recorded systematically.

XSL•FO
RenderX

Figure 2. Participant flow through the study. Reasons for discontinuation were obtained from structured open-ended responses recorded at the 1-month and 3-month assessment visits. ICT: information and communication technology.



Screening logs (numbers screened, eligible, approached, and declining) were not maintained systematically across the 3 study sites. ICT: information and communication technology.

Participant Characteristics

A total of 24 patients were enrolled across the 3 participating hospitals (SNUBH, n=4; CNUH, n=10; PNUYH, n=10). The mean age was 62.5 (SD 9.1) years, and 17 of 24 (71%) patients were male. The majority had ischemic stroke. Regarding functional status, 9 (38%) patients had mRS scores of 1, 11 (46%) had mRS scores of 2, and 4 (17%) had mRS scores of 3. The FAC distribution was 13 (54%) patients with FAC scores of 3 and 11 (46%) with FAC scores of 4 to 5. A total of 16 (67%) patients had a K-MMSE score within the normal range (≥24), whereas 8 (33%) had scores suggestive of cognitive impairment.

Content Module Usage Patterns

Among the rehabilitation content modules, the physical therapy module was the most frequently used, with 15 of 22 (68%) participants reporting its use. The daily exercise missions and therapist-designated missions were rated as “very useful” or “useful” by 15 of 22 (68%) participants at 1 month. The cognitive and language therapy modules were less frequently used, with 9 to 11 of 22 participants (41% to 50%) reporting their use. Participants expressed concerns regarding the appropriateness of the difficulty level of rehabilitation content, with several noting that exercises were too easy for their functional level. At the initial screen assessment at 1 month, 14 of 22 (64%) participants reported being able to review daily tasks, upcoming schedules, and medications at a glance, while 3 of 22 (14%) found the interface difficult to navigate.

System Usability

The mean SUS score at 1 month postdischarge was 60.34 (SD 17.48), which fell below the acceptable usability threshold of 68. Six of 22 (27%) participants had SUS scores of 68 or above. Among the 10 individual SUS items, item 5 (“I found the various functions in this program are well integrated”) received the highest mean score (3.95), while item 6 (“I thought there was too much inconsistency in this program”) received the lowest (reverse-scored: 2.14), indicating perceived inconsistency as a relative weakness.

Subgroup analysis revealed significant differences in SUS scores by age group. Patients younger than 65 years (n=12) scored significantly higher than those aged 65 years and older (n=10): 68.3 (SD 13.2) versus 50.8 (SD 17.7*; P*=.02). Male patients scored numerically higher than female patients (61.2 vs 57.9), and patients with normal cognitive function (K-MMSE≥24; n=14) scored higher than those with cognitive impairment (K-MMSE<24; n=8): 63.9 versus 54.1 (*P*=.18), though these differences did not reach statistical significance (Table 1 and Figure 3).

XSL•FO
RenderX

Table . Patient characteristics and outcomes of participants with acceptable (SUS[a] score≥68) and unacceptable (SUS score<68) usability at 1 month after discharge (N=22).

Variables	Overall (N=22)	SUS score≥68 (n=6)	SUS score<68 (n=16)	*P* value
Demographics				
Age (years), mean (SD)	62 (9.1)	56.2 (6.3)	64.2 (9.1)	.06[b]
Age <65 years, n (%)	12 (55)	5 (83)	7 (44)	.16[c]
Male, n (%)	16 (73)	5 (83)	11 (69)	.63[c]
Study site				
SNUBH[d], n (%)	3 (14)	2 (33)	1 (6)	.13[e]
CNUH[f], n (%)	10 (45)	1 (17)	9 (56)	—
PNUYH[g], n (%)	9 (41)	3 (50)	6 (38)	—
Clinical characteristics				
K-MMSE,[h] mean (SD)	25.1 (4.5)	27.5 (3.8)	24.2 (4.6)	.13[b]
K-MBI,[i] mean (SD)	82.7 (12)	81.2 (16.5)	83.3 (10.5)	.91[b]
FAC,[j] mean (SD)	3.6 (0.7)	3.8 (0.8)	3.5 (0.6)	.33[b]
mRS,[k] mean (SD)	1.9 (0.7)	1.8 (0.8)	1.9 (0.7)	.94[b]
PHQ-9,[l] mean (SD)	5.1 (5.7)	1.3 (2.3)	6 (6)	.17[b]
Program usage (self-reported)				
Sessions per day, mean (SD)	2.4 (0.9)	2.3 (0.8)	2.4 (1)	.97[b]
Sessions per week, mean (SD)	3.9 (2)	4.5 (2.3)	3.7 (1.9)	.53[b]
Duration per session[m]				.03[b]
≤20 min, n (%)	9 (41)	0 (0)	9 (56)	—
21 - 40 min, n (%)	5 (23)	2 (33)	3 (19)	—
41 - 60 min, n (%)	4 (18)	2 (33)	2 (13)	—
>60 min, n (%)	4 (18)	2 (33)	2 (13)	—
Sessions >20 min, n (%)	13 (59)	6 (100)	7 (44)	.046[c]
Satisfaction (5-point Likert)				
Program quality, mean (SD)	4.23 (0.81)	4.50 (0.55)	4.12 (0.89)	.43[b]
Needs fulfillment, mean (SD)	3.82 (1.05)	4.33 (1.03)	3.62 (1.02)	.17[b]
Overall satisfaction, mean (SD)	4.14 (0.77)	4.67 (0.52)	3.94 (0.77)	.06[b]
Willingness to recommend, mean (SD)	4.05 (0.95)	4.17 (1.17)	4 (0.89)	.56[b]
Problem-solving help, mean (SD)	4.05 (0.90)	4.33 (1.03)	3.94 (0.85)	.36[b]
Future use intention, mean (SD)	3.68 (1.43)	3.83 (1.47)	3.62 (1.45)	.73[b]

[a]SUS: System Usability Scale.

[b]Mann-Whitney *U* test.

[c]Fisher exact test.

[d]SNUBH: Seoul National University Bundang Hospital.

XSL•FO
RenderX

[e]Permutation-based exact test for the 3×2 table.

[f]CNUH: Chungnam National University Hospital.

[g]PNUYH: Pusan National University Yangsan Hospital.

[h]K-MMSE: Korean Mini-Mental State Examination.

[i]K-MBI: Korean version of the modified Barthel index.

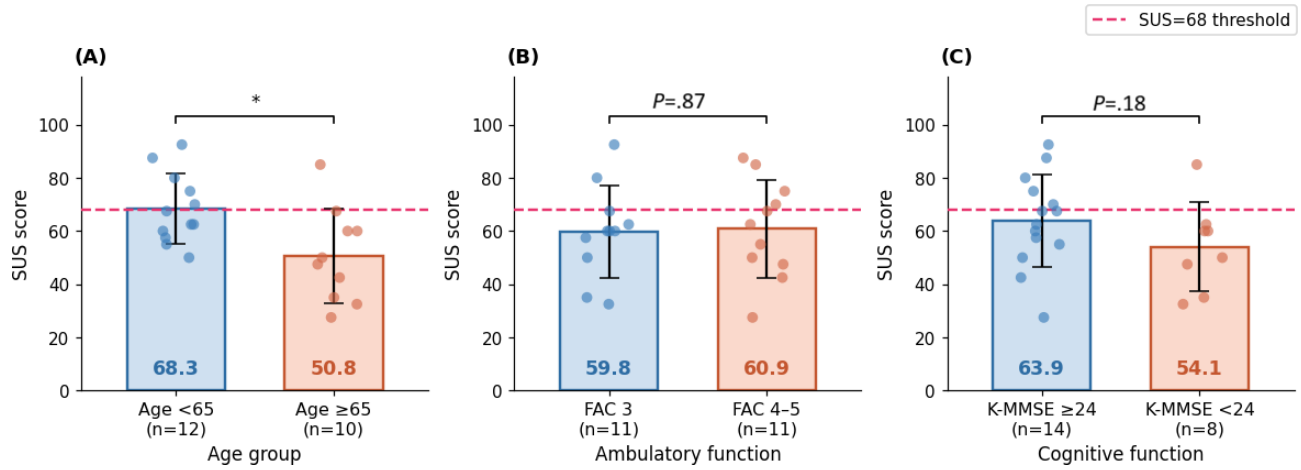
[j]FAC: functional ambulatory category.

[k]mRS: modified Rankin Scale.

[l]PHQ-9: Patient Health Questionnaire-9.

[m]Duration per session was recorded as a 4-level ordinal category and compared with the Mann-Whitney *U* test on the ordinal scale; the dichotomized comparison (>20 min vs ≤20 min) is shown separately.

Figure 3. System Usability Scale scores by subgroup at 1 month after discharge (A) Age group: patients aged younger than 65 years scored significantly higher than those aged 65 years or older (*P*=.02). (B) Ambulatory function: no significant difference between functional ambulatory category 3 and functional ambulatory category 4 to 5 groups. (C) Cognitive function: patients with Korean Mini-Mental State Examination scores of 24 or higher scored numerically higher than those with Korean Mini-Mental State Examination scores below 24 (*P*=.18). The red dashed line indicates the System Usability Scale=68 acceptable usability threshold. Error bars represent SDs. FAC: functional ambulatory category; K-MMSE: Korean Mini-Mental State Examination; SUS: System Usability Scale.



Characteristics of Participants With and Without Acceptable Usability

Participants were dichotomized into a SUS score≥68 group (n=6) and a SUS score<68 group (n=16) to describe how they differed. Compared with the SUS score<68 group, the SUS score≥68 group was younger (56.2, SD 6.3 vs 64.2, SD 9.1 years; *P*=.06), had higher K-MMSE scores (27.5, SD 3.8 vs 24.2, SD 4.6; *P*=.13), had lower PHQ-9 scores (1.3, SD 2.3 vs 6, SD 6; *P*=.17), and reported longer sessions: all 6 (100%) participants in the SUS score≥68 group reported sessions longer than 20 minutes, compared with 7 of 16 (44%) in the SUS score<68 group (*P*=.03 by Mann-Whitney *U* test on the ordinal duration category; *P*=.046 by Fisher exact test on the dichotomized category; Table 1). Effect size analysis further demonstrated that age (Cohen *d*=−0.91), depressive symptoms (Cohen *d*=−0.79), cognitive function (Cohen *d*=+0.71), and ambulatory function (FAC; Cohen *d*=+0.48) had the largest magnitudes of difference between groups (Multimedia Appendix 2). Firth penalized likelihood logistic regression identified no statistically significant associations; all 95% profile likelihood CIs crossed 1.0 (Multimedia Appendix 3). Given that only 6 participants had acceptable usability, these analyses cannot support inference about determinants and are reported descriptively.

Item-Level SUS Responses

In an exploratory analysis of item-level responses, the difference between mean contributions from positively worded and negatively worded SUS items did not differ appreciably between participants with K-MMSE scores <24 and≥24 (0.58, SD 0.77 vs 0.94, SD 1.13; *P*=.76), and K-MMSE score was not correlated with this polarity difference (Spearman ρ=0.06; *P*=.80). A parallel analysis by age group was likewise unrevealing (≥65 years, 0.62 vs <65 years, 0.97; *P*=.67; Multimedia Appendix 4).

Satisfaction Survey

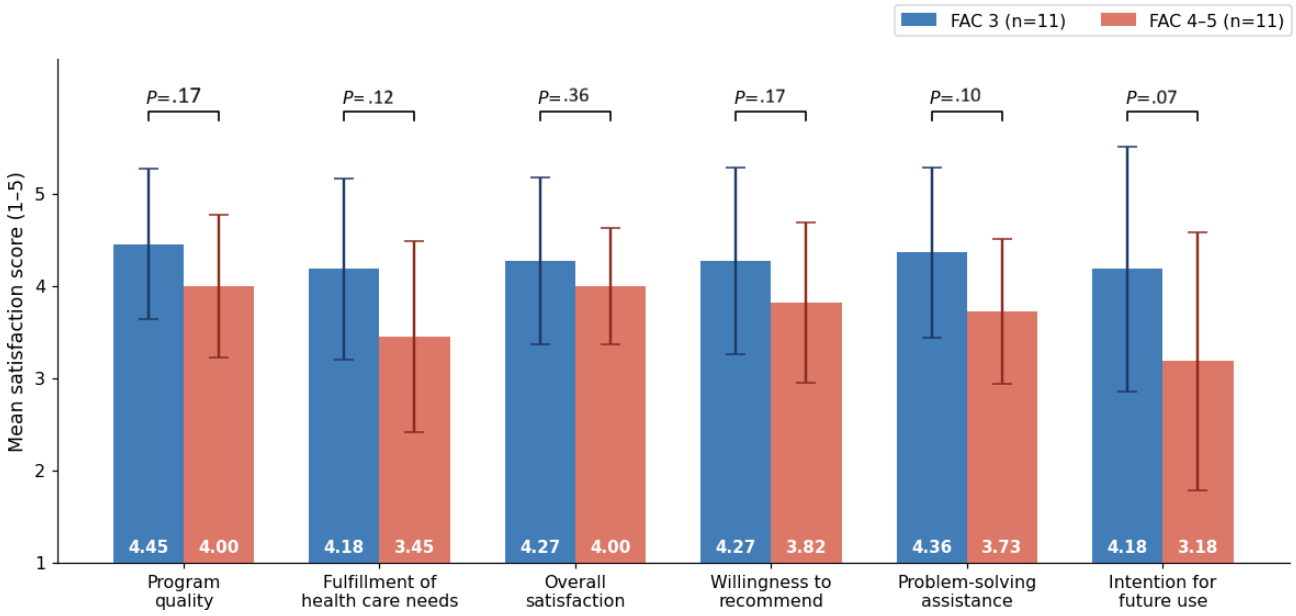
Satisfaction scores at 1-month postdischarge ranged from 3.68 to 4.23 out of 5 across the 6 survey items. The highest-rated item was program quality (mean 4.23, SD 0.81), followed by overall satisfaction (mean 4.14, SD 0.77), willingness to recommend (mean 4.05, SD 0.95), and problem-solving assistance (mean 4.05, SD 0.90). The lowest-rated item was intention for future use (mean 3.68, SD 1.43), reflecting uncertainty about long-term engagement.

Participants with FAC 3 (n=11) reported higher satisfaction than those with FAC 4 or 5 (n=11) in every domain (Figure 4): program quality 4.45 (SD 0.82) versus 4 (SD 0.77), fulfillment of health care needs 4.18 (SD 0.98) versus 3.45 (SD 1.04), overall satisfaction 4.27 (SD 0.90) versus 4 (SD 0.63), willingness to recommend 4.27 (SD 1.01) versus 3.82 (SD 0.87),

XSL•FO
RenderX

problem-solving assistance 4.36 (SD 0.92) versus 3.73 (SD 0.79), and intention for future use 4.18 (SD 1.33) versus 3.18 (SD 1.40; *P*=.07-.36). Averaged across items, satisfaction was 4.29 (SD 0.74) versus 3.70 (SD 0.75; *P*=.09). No individual domain reached significance; the consistent direction across all 6 domains is the observation of interest and suggests that patients with greater unmet rehabilitation needs perceived greater benefit. Satisfaction did not differ by age group or sex.

Figure 4. Satisfaction survey scores at 1 month after discharge by ambulatory function (functional ambulatory category 3 vs functional ambulatory category 4 - 5). Patients with lower ambulatory function (functional ambulatory category 3) showed consistently higher satisfaction across all 6 domains. Error bars represent SDs. FAC: functional ambulatory category.



Comparison Between 1-Month and 3-Month Evaluations

The median interval between the 1-month and the 3-month assessment was 38 (IQR 15 - 47; range 0 - 56) days. In 3 of 8 participants, the interval was 16 days or less, and in 1 participant, both assessments were administered on the same day. Among the 7 participants with paired data, the mean SUS score was 61.1 (SD 13.5) at 1 month and 56.1 (SD 16) at 3 months (Wilcoxon signed-rank *P*=.50); across all 8 participants evaluated at 3 months, the mean SUS score was 54.4 (SD 18.2). Because of the short and variable interval, these values should not be interpreted as evidence of a change in usability over time. Satisfaction scores at 3 months were generally consistent with the 1-month results, ranging from 3.38 to 4.12 out of 5. Participants who maintained usage until 3 months reported longer daily usage times than the overall cohort at 1 month, suggesting a self-selection effect in which more engaged users continued the program.

Attrition Analysis

Within the 22-participant primary analysis set, the 7 participants who completed the 3-month evaluation were compared with the 15 who did not (Multimedia Appendix 5). No appreciable differences were observed in age (63.1, SD 13 vs 61.5, SD 7.1 years; *P*=.86), K-MMSE score (25.4, SD 4.8 vs 25, SD 4.6; *P*=.75), K-MBI score (82.1, SD 15.4 vs 83, SD 10.7; *P*=.92), FAC, mRS, self-reported usage, or 1-month satisfaction. One-month SUS scores were essentially identical between groups (61.1, SD 13.5 vs 60, SD 19.5; *P*=.97), and the proportion with an SUS score of ≥68 was 2 of 7 (29%) versus 4 of 15 (27%). Depressive symptoms showed the largest numerical difference, being higher among completers (PHQ-9 8.2, SD 7.1 vs 3.5, SD 4.5; *P*=.14), although PHQ-9 data were available for only 15 of 22 participants. Given the very small subgroup sizes, these comparisons are underpowered, and the absence of statistically significant differences should not be interpreted as evidence that completers and noncompleters were equivalent.

Open-Ended Responses

Inductive content categorization of the narrative feedback identified 3 primary categories. Positive feedback highlighted the usefulness of the rehabilitation exercise videos, educational content for caregivers on safe transfer techniques and activities of daily living assistance, and the convenience of having a centralized schedule and medication reminder. Negative feedback centered on the cumbersome and nonintuitive user interface, small text and icon sizes that were difficult to read (particularly for older participants), rehabilitation content difficulty levels perceived as too low for patients with mild functional impairment, and difficulty using the program independently. Suggestions for improvement included tailoring rehabilitation content according to each patient's functional level, increasing text and icon sizes, implementing a more intuitive navigation structure, providing detailed step-by-step exercise guidance, and incorporating real-time feedback from therapists through a 1:1 consultation service (preferred by 10 of 22 participants, 45%, at 1 month).

XSL•FO
RenderX

Discussion

Principal Findings

This study evaluated the usability and satisfaction of an ICT-based rehabilitation management program for postacute stroke rehabilitation within the KOMPACT project across 3 university hospitals in Korea. The principal findings were as follows: (1) the overall SUS score of 60.3 fell below the accepted benchmark of 68, indicating suboptimal usability, and only 6 of 22 (27%) participants had scores at or above that benchmark; (2) satisfaction scores were nevertheless generally favorable, ranging from 3.7 to 4.2 out of 5; (3) participants with and without acceptable usability differed descriptively in age, cognitive function, and depressive symptoms, although with 6 events these contrasts cannot be interpreted as determinants; (4) patients with lower ambulatory function reported greater satisfaction, suggesting that the program may be most valuable for those with greater unmet rehabilitation needs; and (5) engagement was low throughout, with 13 of 22 (59%) participants discontinuing at the earliest permitted opportunity, most often because of perceived recovery, difficulty of use, or content that did not match their functional level.

Where User Involvement Stopped

Satisfaction with the program's content was favorable while interface usability fell below the acceptable threshold, and the development history offers a plausible account of this dissociation. End users were involved substantively in defining what the program should contain: FGIs with survivors with stroke, caregivers, therapists, physicians, and social workers established the content domains, and the open-ended responses in this study suggest that those domains were well chosen. Participants repeatedly identified the exercise videos, caregiver education, and consolidated schedule and medication information as useful, endorsing precisely the content specified during the interviews.

End users were not involved in how that content would be operated. Screen architecture, navigation, text and icon sizing, and module arrangement were settled by the clinical and development teams and deployed without a cycle of usability testing with representative users. Participants' complaints corresponded to this division: what they praised was content; what they criticized (small text and icons, cumbersome navigation, inability to use the program unaided, and exercise difficulty not graded according to functional level) was almost entirely interface and interaction design.

Requirements-stage focus groups cannot substitute for design-stage testing, because focus group participants reason about a system they have not yet operated: they can state what they need but not whether a given screen makes that need easy to satisfy. Every interface problem reported here belongs to the class that a single round of think-aloud testing with a few representative users would surface quickly and cheaply. The pattern we followed (thorough requirements elicitation, a single build, and evaluation deferred until after deployment) is common in publicly funded health technology development, and its cost is visible in our data.

We therefore frame this study not as a summative evaluation of a finished product but as the first, belated iteration of a front-loaded human-centered design cycle. User involvement in requirements definition, however thorough, does not discharge the obligation to test the design solution with users before multicenter deployment; development plans for multidomain rehabilitation platforms should budget for at least 1 round of prototype testing round between specification and rollout.

Comparison With Prior Work

Benchmarking a multidomain platform against SUS scores from single-purpose applications is not a like-for-like comparison. Our mean SUS score of 60.3 falls in the "marginal" to "OK" range [15-17]. Considerably higher scores have been reported under far more supported conditions: the START program, a blended-care platform with daily therapist chat support evaluated in 8 inpatients at a Swiss rehabilitation clinic, achieved a mean SUS score of 86.6 (SD 10.9) [13]. That evaluation differs from ours in almost every element of the context of use (supervised inpatient setting, continuous therapist availability, exercise content confined to a single domain, and a 1-week test period), illustrating how far a SUS score depends on the circumstances of use rather than the system alone. The present program spans 4 rehabilitation domains, scheduling, medication management, and welfare information in one platform, and the navigational and cognitive demands of such architectures exceed those of single-purpose applications, so lower usability scores can be expected even without interface deficiencies. Published usability scores for stroke rehabilitation technologies span a wide range accordingly: an immersive virtual reality mirrored-hand system reported a mean SUS score of 56.67 (SD 13.88) in 15 survivors with stroke [22], 30 points below the START figure, and a scoping review of 111 usability evaluations identified 28 distinct instruments, with most studies combining 2 to 6 [23]. However, this reasoning must not explain away genuine design deficiencies: our qualitative data point to concrete, remediable interface problems, and both the inherent complexity of a multidomain design and the absence of predeployment usability testing contributed to the observed scores. Finally, the benchmark of 68 derives from general software evaluations; a meta-analysis of digital health apps reported a benchmark mean of 68 with an SD of 12.5 [18], so the threshold itself is a source of interpretive uncertainty.

Discordance Between Usability and Satisfaction

The discordance between the below-threshold SUS score (60.3) and favorable satisfaction (3.7‐4.2 out of 5) merits interpretation. We had considered whether the alternating positive and negative wording of SUS items might have depressed scores among participants with reduced cognition, but an exploratory item-level analysis did not support this: differential responding by item polarity was similar across cognitive strata and uncorrelated with K-MMSE (Multimedia Appendix 4). The discordance is better explained by the constructs themselves. Interface usability and content satisfaction are distinct: low SUS scores reflect difficulty operating the tablet interface, whereas high satisfaction reflects the perceived value of the clinical content. This pattern is consistent with, but was not tested against, the technology

XSL•FO
RenderX

acceptance model (TAM), in which perceived usefulness and perceived ease of use are distinct antecedents of adoption intention [24]; neither construct was measured here, and we offer TAM only as a post hoc interpretation. A Korean nationwide survey of people with disabilities found usage intention driven primarily by perceived usefulness (β=.662; *P*<.001) rather than ease of use (β=.203; *P*<.001) [24]. Future evaluations should assess usability, perceived usefulness, and content value as distinct constructs.

Engagement and Adoption Barriers

Engagement was low even at its peak: the physical therapy module, the most-used component, was accessed by only 15 of 22 (68%) participants, the cognitive and language modules by 41% to 50%, and 13 of 22 (59%) stopped at the earliest permitted opportunity. Read with the reasons given for discontinuing, this points to 3 barriers. First, operability: 3 participants stopped because they could not use the program without help, 1 attributing this to a caregiver's unavailability. Second, relevance: participants who had recovered ambulatory function found the exercise content too easy and perceived no continuing need, echoed by the higher satisfaction among those with lower ambulatory function. Third, sustained motivation: the program had no reengagement mechanism, therapist contact between visits, or progress feedback. Each implies a different remedy: caregiver-aware interaction design; adaptive content graded to functional level; and structured re-engagement through scheduled therapist contact, feedback, and prompting. High usability does not guarantee engagement: the START program, with a SUS of 86.6, reported adherence of 17% to 80%, with 3 of 8 patients completing fewer than a quarter of planned sessions, partly because content was too easy [13]. That a highly usable program met the same relevance and motivation barriers suggests they are not simply consequences of poor interface design. Formative evaluations of multidomain platforms should therefore report module-level engagement and reasons for discontinuation alongside summary usability scores.

Patient Characteristics and Usability

The large effect size for age (Cohen *d*=−0.91) is consistent with systematic reviews identifying older age as a primary barrier to mHealth adoption after stroke [11,12]. Patients younger than 65 years reached a mean SUS score of 68.3, meeting the threshold, whereas those 65 and older scored 50.8; the association did not reach significance in Firth regression, most likely because of the critically limited event count rather than a true absence of association. This divide reflects reduced familiarity with touchscreens, sensory deficits, and age-related cognitive decline [25], and argues for larger text and icons (as participants requested), simpler navigation, and caregiver co-use models.

Depressive symptoms (PHQ-9; Cohen *d*=−0.79) and cognitive function (K-MMSE; Cohen *d*=+0.71) also showed meaningful effect sizes favoring the acceptable-usability group, consistent with the literature on patient-level barriers [11,26]. Poststroke depression, affecting roughly 30% of survivors, reduces engagement in rehabilitation [26]; the KOMPACT trial found that ESD improved depressive symptoms [4], suggesting possible synergy between ICT platforms and ESD services. The K-MMSE difference (27.5 vs 24.2) supports cognitive screening before ICT prescription and adaptive interface modes for mild cognitive impairment.

The absence of significant associations in Firth regression (Multimedia Appendix 3) is most parsimoniously attributed to the low event count (n=6; EPV=0.86), under which even large effects are undetectable; every penalized odds ratio nevertheless pointed in the same direction as the corresponding effect size. We therefore make no claim that these variables are determinants of usability. The convergent effect sizes for age (Cohen *d*=−0.91), depressive symptoms (Cohen *d*=−0.79), cognitive function (Cohen *d*=+0.71), and ambulatory function (Cohen *d*=+0.48) are consistent with prior work [9,11] and provide hypotheses for future testing, whereas functional independence (K-MBI; Cohen *d*=−0.17) and stroke severity (mRS; Cohen *d*=−0.06) showed negligible differences. Adequately powered confirmation would require 50 to 100 participants in the acceptable-usability group (EPV≥10).

Satisfaction and Unmet Rehabilitation Needs

The higher satisfaction reported by patients with lower ambulatory function (FAC 3), consistent in direction across all 6 domains although not statistically significant in any of them, likely reflects greater unmet rehabilitation needs, for whom the multidomain program provided meaningful support across rehabilitation exercises, scheduling, and educational content during a vulnerable recovery period. Conversely, patients with better ambulatory function (FAC 4 - 5) not only perceived the exercise content as insufficiently challenging but may also have found the overall program content less clinically relevant to their remaining rehabilitation needs. Patients who have already regained near-normal ambulatory capacity may have fewer unmet needs that the current program's content domains adequately address, potentially limiting the program's perceived utility in this subgroup beyond convenience features such as scheduling reminders and medication guidance. These findings highlight that the clinical value of a multidomain ICT platform is likely to vary substantially across the functional spectrum and reinforce the importance of developing adaptive, personalized content that is dynamically matched to each patient's functional level and evolving rehabilitation goals [10]. Notably, qualitative data indicated that caregivers actively engaged with educational content, consistent with systematic evidence that dyad-focused telehealth interventions improve outcomes and caregiver preparedness [27,28]. In the Korean cultural context, where family-centered care is deeply embedded, explicitly incorporating caregiver-specific modules and guided co-use protocols into future platform iterations is particularly warranted.

Future Directions

Future evaluations of this platform should incorporate a usability instrument developed specifically for health information technology, such as the Health Information Technology Usability Evaluation Scale (Health-ITUES), which assesses perceived usefulness, ease of use, user control, and impact as separate dimensions and for which a validated Korean version is available [29,30]; a usability cut-point for this instrument has also been proposed [31]. Applying such an instrument alongside

XSL•FO
RenderX

the SUS would allow the interface and content dimensions that appear to diverge in our data to be measured separately rather than inferred. Prospective application of a TAM-based instrument would likewise permit direct testing of the interpretation offered above. Beyond measurement, the next iteration of the platform should be co-designed with survivors with stroke and caregivers through iterative usability testing, should deliver content adaptively according to functional level, and should incorporate structured reengagement mechanisms; server-side usage logging should be prespecified and retained so that objective engagement metrics can be analyzed alongside self-report.

Limitations

This study has several limitations. First, the sample was small (n=24 enrolled, n=22 analyzed) and did not reach the planned 30 participants, principally because recruitment coincided with the COVID-19 pandemic. With only 6 participants achieving acceptable usability, the EPV ratio precluded standard logistic regression; Firth penalized estimates and the reported effect sizes should be interpreted cautiously given wide CIs and extreme group imbalance, and the findings are hypothesis-generating. Second, the single-arm design and short follow-up (maximum of 3 months after onset) did not permit assessment of clinical effectiveness; although telerehabilitation can achieve outcomes comparable to in-person rehabilitation [6], whether this program produces clinical benefit remains unverified.

Third, attrition was substantial: 8 of 22 participants were evaluated at 3 months and 7 had paired data. Completers and noncompleters did not differ appreciably in baseline characteristics or 1-month SUS score (Multimedia Appendix 5), arguing against selective retention of those who found the program usable, but the comparison is underpowered. Fourth, the interval between assessments was short and variable (median 38, range 0‐56 days), coinciding in 1 participant, so the 3-month data are not an independent longitudinal observation. Fifth, usage data were self-reported and subject to recall and social-desirability bias; server-side logs were not retained. Sixth, the satisfaction instrument was developed by the research team and not psychometrically validated. Seventh, open-ended responses were brief, prompt-directed, and categorized by a single investigator who was also involved in program development, without a second coder, reliability assessment, or member checking; they should be read as structured feedback rather than qualitative findings. Eighth, screening data were not recorded systematically, so the representativeness of the cohort cannot be assessed. Ninth, the platform was built and hosted by an external contractor (HEALTHCONNECT Co, Ltd), and infrastructure-level security (server configuration, encryption at rest, access logging, and any certification) was not documented to the research team; investigators commissioning comparable platforms should specify these arrangements contractually at the outset. Tenth, authentication on the patient tablet required only hospital registration number and name, with an auto-login option; this deliberately lowered the access burden for older users but affords weaker protection, and future deployments should adopt stronger patient authentication and session timeouts. Eleventh, the FGIs were a development activity: transcripts were not retained, no formal coding was applied, and all sessions were moderated by a single investigator on the design team, so the requirements phase cannot be independently appraised, and the correspondence we describe between specified and endorsed content is a plausible account rather than a tested proposition. Finally, the SUS captures principally the satisfaction component of ISO 9241‐11; a fuller evaluation would require task-based performance testing, error analysis, and objectively logged interaction data.

Conclusions

This formative evaluation showed that a multidomain ICT-based rehabilitation management program for postacute stroke achieved suboptimal usability alongside generally favorable satisfaction among community-dwelling survivors with stroke, and that engagement declined rapidly after the first month. Participants with acceptable usability were descriptively younger, with better cognitive function and fewer depressive symptoms, but the number of such participants was far too small to identify determinants, and these contrasts should be treated as hypotheses for adequately powered future studies. Patients with greater unmet rehabilitation needs derived higher satisfaction, and the discordance between usability and satisfaction underscores the importance of evaluating interface usability and content value as separate constructs in digital health platforms. The clinical effectiveness of this program remains to be established in future controlled trials. The principal implementation lesson concerns where user involvement was placed: survivors with stroke and caregivers helped define the program content through FGIs, and content satisfaction was correspondingly favorable, but the design solution itself was never tested with users before deployment, and it was interface and interaction design that participants found difficult. User involvement in requirements definition is necessary but not sufficient; multidomain rehabilitation platforms intended for older survivors with stroke should undergo at least 1 cycle of prototype usability testing with end users before multicenter deployment, alongside adaptive content delivery, caregiver couse protocols, and structured reengagement mechanisms.

Generative AI Use

During the preparation of this manuscript, the authors used a large language model (Claude, Anthropic) as an AI-assisted technology. Its use was limited to editing English text and formatting to journal style under the direction and review of the authors. The study design, data analysis, interpretation of results, and final wording were determined solely by the authors. No AI tool was used to generate, fabricate, or alter data or results. The authors reviewed all AI-assisted output, take full responsibility for the scientific integrity of the content, and confirm that AI-assisted technologies are not listed as authors.

XSL•FO
RenderX

Funding

This research was supported by a grant from the Korea National Institute of Health, the Korea Disease Control and Prevention Agency (2020ER630602). The funder had no role in the design or conduct of the study, the collection, analysis, or interpretation of the data, the preparation of the manuscript, or the decision to submit it for publication.

Data Availability

Pseudonymized data supporting this study's findings are available from the corresponding author upon reasonable request, subject to approval from the local institutional review board and completion of a data-sharing agreement.

Authors' Contributions

Conceptualization: WKC, W-SK, MKS, Y-IS, NJP
Data curation: WKC, YSJ, SJ, SHK
Formal analysis: WKC, JSC
Investigation: WKC, YSJ, WSK, MKS, SJ, Y-IS, SHK, N-JP
Methodology: WKC
Visualization: WKC, JSC
Writing – original draft: WKC
Writing – review & editing: WKC, YSJ, JSC, WSK, MKS, SJ, YIS, SHK, NJP
Funding acquisition: NJP
Project administration: NJP
Supervision: NJP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the information and communication technology–based rehabilitation management program interface.
[DOCX File, 711 KB - rehab_v13i1e98098_app1.docx]

Multimedia Appendix 2
Forest plot of Cohen *d* effect sizes comparing patient characteristics between participants with acceptable (System Usability Scale ≥68) and unacceptable (System Usability Scale <68) usability.
[DOCX File, 163 KB - rehab_v13i1e98098_app2.docx]

Multimedia Appendix 3
Forest plot of Firth penalized likelihood logistic regression estimates for individual variables in relation to acceptable usability (System Usability Scale ≥68).
[DOCX File, 236 KB - rehab_v13i1e98098_app3.docx]

Multimedia Appendix 4
Exploratory analysis of item-level System Usability Scale responses by item polarity and cognitive status at 1 month (n=22).
[DOCX File, 16 KB - rehab_v13i1e98098_app4.docx]

Multimedia Appendix 5
Comparison of participants who did and did not complete the 3-month evaluation.
[DOCX File, 20 KB - rehab_v13i1e98098_app5.docx]

References

1. Murray CJL, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the global burden of disease study 2010. Lancet 2012 Dec 15;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61689-4] [Medline: 23245608]
2. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. The Lancet 2020 Oct;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9]

XSL•FO
RenderX

3. Langhorne P, Baylan S, Early Supported Discharge Trialists. Early supported discharge services for people with acute stroke. Cochrane Database Syst Rev 2017 Jul 13;7(7):CD000443. [doi: 10.1002/14651858.CD000443.pub4] [Medline: 28703869]
4. Chang WK, Jung YS, Choi JS, et al. Pragmatic multicenter randomized controlled study on early supported discharge after stroke in Korea: the KOMPACT study. Ann Phys Rehabil Med 2025 Nov;68(8):102025. [doi: 10.1016/j.rehab.2025.102025] [Medline: 40972310]
5. Leigh JH, Kim WS, Sohn DG, Chang WK, Paik NJ. Transitional and long-term rehabilitation care system after stroke in Korea. Front Neurol 2022;13:786648. [doi: 10.3389/fneur.2022.786648] [Medline: 35432175]
6. Laver KE, Adey-Wakeling Z, Crotty M, Lannin NA, George S, Sherrington C. Telerehabilitation services for stroke. Cochrane Database Syst Rev 2020 Jan 31;1(1):CD010255. [doi: 10.1002/14651858.CD010255.pub3] [Medline: 32002991]
7. Wang Y, Yang F, Shi H, Yang C, Hu H. What type of transitional care effectively reduced mortality and improved ADL of stroke patients? A meta-analysis. IJERPH 2017 May 10;14(5):510. [doi: 10.3390/ijerph14050510]
8. Duncan PW, Abbott RM, Rushing S, et al. COMPASS-CP: an electronic application to capture patient-reported outcomes to develop actionable stroke and transient ischemic attack care plans. Circ Cardiovasc Qual Outcomes 2018 Aug;11(8):e004444. [doi: 10.1161/CIRCOUTCOMES.117.004444] [Medline: 30354371]
9. Broderick M, O'Shea R, Burridge J, Demain S, Johnson L, Bentley P. Examining usability, acceptability, and adoption of a self-directed, technology-based intervention for upper limb rehabilitation after stroke: cohort study. JMIR Rehabil Assist Technol 2023 Aug 21;10:e45993. [doi: 10.2196/45993] [Medline: 37603405]
10. Kim S, Park SW, Jeong T, Kang MS, Kim DY. AI-driven cognitive telerehabilitation for stroke: a randomized controlled trial. Front Neurol 2025;16:1636017. [doi: 10.3389/fneur.2025.1636017] [Medline: 40895105]
11. Niyomyart A, Ruksakulpiwat S, Benjasirisan C, et al. Current status of barriers to mhealth access among patients with stroke and steps toward the digital health era: systematic review. JMIR Mhealth Uhealth 2024 Aug 22;12(1):e54511. [doi: 10.2196/54511] [Medline: 39173152]
12. Veras M, Auger LP, Sigouin J, et al. Ethics and equity challenges in telerehabilitation for older adults: rapid review. JMIR Aging 2025 Aug 13;8(1):e69660. [doi: 10.2196/69660] [Medline: 40802966]
13. Ziller C, Gäumann S, Lüscher S, et al. Telerehabilitation following stroke: development of training content and evaluation of an app-based training program. JMIR Rehabil Assist Technol 2026 Mar 31;13:e77090. [doi: 10.2196/77090] [Medline: 41915681]
14. ISO 9241-11:2018. ergonomics of human-system interaction—part 11: usability: definitions and concepts. International Organization for Standardization. 2018. URL: https://www.iso.org/standard/63500.html [accessed 2026-03-01]
15. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry: Taylor & Francis; 1996.
16. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123. [doi: 10.5555/2835587.2835589]
17. Lewis JR, Sauro J. Item benchmarks for the system usability scale. J Usability Stud 2018;13(3):158-167. [doi: 10.5555/3294033.3294037]
18. Hyzy M, Bond R, Mulvenna M, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
19. ISO 9241-210:2019 - ergonomics of human-system interaction—part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2026-10-17]
20. Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993;80(1):27-38. [doi: 10.1093/biomet/80.1.27]
21. Heinze G, Schemper M. A solution to the problem of separation in logistic regression. Stat Med 2002 Aug 30;21(16):2409-2419. [doi: 10.1002/sim.1047] [Medline: 12210625]
22. Hsieh YW, Howe TH, Lee MT, Tai RY, Chen CC. Design and usability evaluation of an immersive virtual reality mirrored hand system for upper limb stroke rehabilitation. Sci Rep 2025 Feb 17;15(1):5781. [doi: 10.1038/s41598-025-90698-6] [Medline: 39962160]
23. Ouédraogo F, Demers M, Soares Menezes KVR, et al. Tools and methods for assessing the usability and related aspects of usability of extended reality and telerehabilitation technologies in stroke rehabilitation: a scoping review. J NeuroEngineering Rehabil 2026;23(1):44. [doi: 10.1186/s12984-025-01861-z]
24. Kim JH, Kim J, Youn BY. Using a technology acceptance model to explore the intention to use digital health technologies among people with disabilities: cross-sectional survey study. J Med Internet Res 2025 Nov 20;27:e79595. [doi: 10.2196/79595] [Medline: 41264861]
25. Verma A, Towfighi A, Brown A, Abhat A, Casillas A. Moving towards equity with digital health innovations for stroke care. Stroke 2022 Mar;53(3):689-697. [doi: 10.1161/STROKEAHA.121.035307] [Medline: 35124973]
26. Towfighi A, Ovbiagele B, El Husseini N, et al. Poststroke depression: a scientific statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2017 Feb;48(2):e30-e43. [doi: 10.1161/STR.0000000000000113] [Medline: 27932603]
27. Bartoli D, Petrizzo A, Vellone E, Alvaro R, Pucciarelli G. Impact of telehealth on stroke survivor-caregiver dyad in at-home rehabilitation: a systematic review. J Adv Nurs 2024 Oct;80(10):4003-4033. [doi: 10.1111/jan.16177] [Medline: 38563582]

XSL•FO
RenderX

28. Bakas T, McCarthy MJ, Miller EL. Systematic review of the evidence for stroke family caregiver and dyad interventions. Stroke 2022 Jun;53(6):2093-2102. [doi: 10.1161/STROKEAHA.121.034090] [Medline: 35264010]
29. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: validation study. JMIR Mhealth Uhealth 2018 Jan 5;6(1):e4. [doi: 10.2196/mhealth.8851] [Medline: 29305343]
30. Lee J, Schnall R. Validity and reliability of the Korean version of the health information technology usability evaluation scale: psychometric evaluation. JMIR Med Inform 2022 Jan 24;10(1):e28621. [doi: 10.2196/28621] [Medline: 35072630]
31. Loh KP, Liu J, Ganzhorn S, Sanabria G, Schnall R. Establishing a usability cut-point for the health information technology usability evaluation scale (Health-ITUES). Int J Med Inform 2022 Apr;160:104713. [doi: 10.1016/j.ijmedinf.2022.104713] [Medline: 35144102]

Abbreviations

CNUH: Chungnam National University Hospital
COMPASS: Comprehensive Post-Acute Stroke Services
EPV: events-per-variable
ESD: early supported discharge
FAC: functional ambulatory category
FGI: focus group interview
Health-ITUES: Health Information Technology Usability Evaluation Scale
ICT: information and communication technology
K-MBI: Korean version of the modified Barthel index
K-MMSE: Korean Mini-Mental State Examination
KOMPACT: Korean Model for Post-Acute Comprehensive Rehabilitation
mHealth: mobile health
mRS: modified Rankin Scale
NIHSS: National Institutes of Health Stroke Scale
PHQ-9: Patient Health Questionnaire-9
PNUYH: Pusan National University Yangsan Hospital
SNUBH: Seoul National University Bundang Hospital
SUS: System Usability Scale
TAM: technology acceptance model

Edited by A Mastropietro; submitted 13.Apr.2026; peer-reviewed by S Hoermann, V Mishra; comments to author 28.Jul.2026; revised version received 10.Sep.2026; accepted 01.Oct.2026; published 23.Nov.2026.

Please cite as:
Chang WK, Jung YS, Choi JS, Kim WS, Sohn MK, Jee S, Shin YI, Ko SH, Paik NJ
Usability and Satisfaction of a Multidomain Information and Communication Technology–Based Rehabilitation Management Program in Patients With Postacute Stroke: Multicenter Formative Evaluation
JMIR Rehabil Assist Technol 2026;13:e98098
URL: https://rehab.jmir.org/2026/1/e98098
doi:10.2196/98098

© Won Kee Chang, Yun-Sun Jung, Ji-Soo Choi, Won-Seok Kim, Min Kyun Sohn, Sungju Jee, Yong-Il Shin, Sung-Hwa Ko, Nam-Jong Paik. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Validation of Single-Camera MediaPipe BlazePose for Knee Joint Angle Measurement Across Multiple Functional Tasks: Concurrent Validity, Interrater Reliability, and Exploratory Test–Retest Reliability Study

Xiaotian Yue[1], DPT; YiXuan Gu[1]; Bo Yu[1,2], MD, PhD

[1]Department of Rehabilitation, School of International Medical Technology, Shanghai Sanda University, 2727 Jinhai Road, Pudong New District, Shanghai, China

[2]Orthopaedic Rehabilitation Unit, Department of Rehabilitation Medicine, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, No. 100, Haining Road, Shanghai, China

Corresponding Author:
Bo Yu, MD, PhD
Department of Rehabilitation, School of International Medical Technology, Shanghai Sanda University, 2727 Jinhai Road, Pudong New District, Shanghai, China

Abstract

Background: Single-camera markerless pose estimation may support accessible knee movement assessment, but performance can depend on task, joint angle, and detection quality, and evidence across complementary measurement properties remains limited.

Objective: This study aimed to evaluate agreement of MediaPipe BlazePose Heavy knee-flexion estimates with goniometer-set static target positions and with Kinovea 2D video annotation for dynamic tasks, to assess interrater agreement of the Kinovea annotations, and to explore repeated-session performance in a small retest subsample.

Methods: Squat, sit-to-stand, and bilateral single-leg stance tasks were performed by 15 healthy adults during standardized lateral-view smartphone recording. Static knee flexion was assessed at 4 goniometer-set targets (30°, 60°, 90°, and 120°). Native hip-knee-ankle included angles were converted to clinical knee flexion (0° at full extension). Agreement was assessed using the intraclass correlation coefficient (ICC), specifically $ICC_{2,1}$, signed bias and 95% limits of agreement, mean absolute error (MAE), and root mean square error (RMSE). Primary intervals used participant-cluster bootstrapping, and limits of agreement accounted for repeated measurements. Dynamic primary estimates were conditional on BlazePose target-frame confidence ≥0.3; low-confidence observations were also reported as detection failures and included in an all-output stress test. Four participants returned after 21 days to 29 days (median 25 d) for exploratory repeated-session analyses.

Results: Dynamic measurement was available in 177 of 210 observations (84.3%); 33 of 210 observations (15.7%) failed the confidence criterion. Conditional task-specific ICCs were 0.773 (95% CI 0.564 to 0.875; n=71, 15 participants) for squat, 0.632 (95% CI 0.455 to 0.726; n=58, 12 participants) for sit-to-stand, and 0.136 (95% CI −0.078 to 0.482; n=48, 15 participants) for combined single-leg stance. Task-specific availability was present for 71 of 75 observations (95%), 58 of 75 observations (77%), and 48 of 60 observations (80%), respectively. Without the quality gate, squat ICC was 0.428, and RMSE was 25.43°. Static agreement with set targets resulted in an ICC of 0.874 (95% CI 0.770 to 0.953), with RMSE of 15.49° and angle-dependent signed error. Kinovea interrater ICC was 0.999 (95% CI 0.998 to 0.999) overall, which did not establish 3D accuracy. Repeated-session estimates with the 4 participants remained exploratory.

Conclusions: With dynamic measurement available in 177 of 210 observations (84.3%), agreement among retained outputs depended on task, flexion angle, and detection quality. The findings support further feasibility and screening-level evaluation under controlled acquisition with per-recording quality checks; they do not support replacement of clinical goniometry, precise longitudinal decision-making, absolute 3D kinematic accuracy, or generalization to clinical populations or other smartphones.

(JMIR Rehabil Assist Technol 2026;13:e102399) doi:10.2196/102399

KEYWORDS

markerless motion capture; pose estimation; MediaPipe BlazePose; knee kinematics; concurrent validity; interrater reliability; test–retest reliability; telerehabilitation; Kinovea; goniometry

XSL•FO
RenderX

Introduction

Background

Accurate measurement of knee joint kinematics is fundamental to the assessment and management of musculoskeletal conditions, including anterior cruciate ligament reconstruction recovery and knee osteoarthritis rehabilitation. Clinicians routinely monitor knee flexion and extension range of motion (ROM) to guide treatment progression, set discharge criteria, and evaluate surgical outcomes. In conventional clinical practice, the universal goniometer is used for static ROM measurement and offers acceptable intrarater reliability when used by trained clinicians [1]. However, the goniometer requires in-person contact, is limited to static or quasistatic postures, and cannot capture the dynamic joint behavior that characterizes functional tasks such as squatting, sit-to-stand transfers, and single-leg balance—tasks that more closely reflect the demands of daily living and return-to-activity benchmarks.

Laboratory-based, marker-based, motion-capture systems (eg, Vicon and Qualisys) are commonly used reference methods for 3D joint-kinematic analysis. However, they require dedicated space, specialized hardware, marker placement by trained personnel, and substantial postprocessing expertise [2,3]. These requirements limit their availability for routine or remote assessment, where neither an in-person goniometer nor laboratory motion capture is readily deployable.

The rapid development of markerless pose estimation technologies has created new possibilities for accessible, low-cost movement analysis. Deep learning–based frameworks can estimate joint positions from standard video without physical markers. Stenum et al [4] demonstrated the feasibility of automated gait analysis from single-camera 2D pose estimation. Among single-camera 2D approaches, Gupta et al [5] compared MediaPipe Pose with Kinovea 2D video annotation for knee flexion during gait, reporting a Pearson correlation of 0.941 and a mean absolute error (MAE) of 5.88°. Langley et al [6] compared both smartphone-based Deep Vision and Kinovea with 3D motion capture in patients with knee osteoarthritis. For Deep Vision, limits of agreement for peak flexion and ROM were within or equal to the study's clinical thresholds. Koul et al [7] reported high intrarater reliability for knee flexion measurements using DETROM (intraclass correlation coefficient [ICC]=0.98) in healthy adults. Agreement with a universal goniometer was evaluated separately using Bland-Altman analysis. Akturk et al [8] applied MediaPipe to joint-angle and joint-moment estimation during sit-to-stand transfers, extending single-camera pose estimation to a functional task.

Multicamera and 3D markerless systems have also been evaluated during functional tasks. Lima et al [9] evaluated OpenCap, a multicamera system for 3D kinematic reconstruction, across functional tasks. For single-leg squatting, root mean square error (RMSE) ranged from 2.9° to 13.6° across joint angles. Peak-angle intersession ICCs across joints were 0.60 to 0.92 for single-leg squatting and 0.62 to 0.93 for the jump phase of a countermovement jump. Yoma et al [10] reported mean between-day root mean square differences <5° and mean ICCs of 0.77 to 0.83 across kinematic variables for Theia3D during single-leg squat and landing tasks, although hip flexion errors exceeded 5°. These findings provide context for single-camera studies, but differences in tasks, populations, and analysis methods limit direct performance comparisons.

Single-camera markerless measurements can be affected by self-occlusion and out-of-plane motion [3]. This study therefore focused on task-dependent agreement, the consistency of the 2D reference annotations, and repeated-session uncertainty. These measurement properties require evaluation before longitudinal or remote uses can be considered.

To our knowledge, few studies have evaluated an open-source single-camera pose-estimation system across multiple knee-related functional tasks in one cohort while examining complementary measurement properties, including agreement with a clinical instrument and a 2D video reference, interrater agreement of that 2D reference, and exploratory repeated-session performance. The selection and reporting of measurement properties were informed by COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) terminology [11].

Objectives

The purpose of this study was to evaluate the MediaPipe BlazePose Heavy model for knee-flexion measurement in healthy young adults. We examined (1) interrater agreement of Kinovea 2D video annotation; (2) static agreement of BlazePose with goniometer-set target flexion across 4 angles; (3) dynamic agreement of BlazePose with Kinovea during squat, sit-to-stand, and bilateral single-leg stance; and (4) exploratory repeated-session agreement of BlazePose, Kinovea, and their between-method differences. We expected performance to vary by task, joint angle, and detection quality. This study did not evaluate absolute 3D kinematic accuracy, clinical populations, or cross-device equivalence.

Methods

Study Design

This cross-sectional validation study evaluated agreement of smartphone-based MediaPipe BlazePose knee-flexion estimates with goniometer-set targets for static positions and with Kinovea 2D video annotations by 2 trained raters for dynamic functional tasks. Secondary analyses assessed interrater agreement of the Kinovea annotations and exploratory repeated-session performance in a 4-participant subsample. Kinovea was treated as a 2D reference, not as an absolute biomechanical standard.

Participants

We recruited 15 healthy adults (8 men, 7 women; mean age 20.8, SD 0.6, range 20 - 22 years; mean height 171.1, SD 9.4 cm; mean weight 70.7, SD 15.7 kg; mean BMI 23.9, SD 3.4, range 18.8 - 31.3 kg/m²) via convenience sampling from a university community. Inclusion criteria were age 18 years to 45 years, no current lower-extremity pain, and ability to perform a full-depth squat independently. Exclusion criteria included a history of lower-extremity surgery, current lower-extremity injury, or any musculoskeletal condition limiting knee range of motion. The study was approved by the Academic Ethics and

XSL•FO
RenderX

Integrity Review Committee of Shanghai Sanda University (approval number 2026006). All participants provided informed consent prior to data collection. The sample size of 15 was determined by feasibility considerations (participant availability and resource constraints within a single-semester data collection window) rather than an a priori power calculation; this is acknowledged as a limitation, and the achieved precision of ICC estimates is reported via 95% CIs throughout.

A subset of 4 participants (S01–S04) returned for an exploratory second session 21 days to 29 days after their initial session (median 25 d). The second session used the same static and dynamic task protocol. The dataset comprised 15 initial-session datasets plus 4 repeated-session datasets, yielding 342 videos processed through the BlazePose pipeline (228 static and 114 dynamic). The primary cross-sectional static analysis used the 180 Session 1 trials from S01-S15; the 48 Session 2 static trials were reserved for the exploratory repeated-session analysis.

Equipment and Setup

All videos were recorded with the same iPhone 17 Pro rear camera mounted on a tripod in a lateral view at knee height and 2.0 m from the participant. The device was selected because it was available and readily accessible to the research team; using one device throughout kept acquisition hardware constant. Recordings used 1920×1080 H.264 video. Participants S03-S15 and all repeated-session recordings were captured at 60 frames per second (approximately 23 Mbps), whereas S01-S02 were captured at 120 frames per second (approximately 43 Mbps) during the pilot phase. Codec profile, pixel format (yuv420p), and color space (BT.709) were verified with FFprobe. Floor markings standardized participant and camera position. The study did not compare smartphones, lenses, device-specific image processing, or frame rates.

We placed 3 small reflective markers on each participant's test limb to support manual Kinovea annotation: the greater trochanter (hip), the lateral femoral epicondyle (knee), and the lateral malleolus (ankle). The BlazePose pipeline used MediaPipe's native keypoint detection and did not rely on these external markers. An anterior superior iliac spine marker was placed during some pilot trials but was not used in the final analysis pipeline. Participants wore dark-colored tight-fitting trousers or compression pants and a short-cropped top and were barefoot. The background was a plain pale wall with even lighting and no harsh shadows.

Data Collection Protocol

Static Validation

For static validation, participants were seated at 4 target knee-flexion angles (30°, 60°, 90°, and 120°). Target positions were set with a digital limb goniometer (GemRed model 82311‐00; transparent polycarbonate arms; manufacturer-stated resolution 0.1°, accuracy ±1°, and range 0°-180°). The fulcrum was placed at the lateral femoral epicondyle, the stationary arm was aligned with the greater trochanter, and the moving arm was aligned with the lateral malleolus. For each of the 3 trials at each target, the goniometer was reapplied, and the angle was reread. The limb was adjusted so that the displayed value, rounded to the nearest degree, equaled the target; floor tape served only as a positional guide. Thus, all recorded integer values equaled their targets after rounding, corresponding to displayed readings within approximately ±0.5° of the target at positioning, not exact independent measurements of the recorded posture. Positioning was performed by a physiotherapist (DPT) or a trained student assessor, and a research assistant (Rater B) verified the setup. This check was not an independent criterion measurement. The goniometer was removed before recording, and participants held the position for approximately 5 seconds. The angle was not remeasured after recording. There were 180 Session 1 trials (15 participants × 4 targets × 3 trials) and 48 additional trials from 4 retest participants. The comparison therefore assessed agreement with goniometer-set target positions, not simultaneous criterion validity.

Dynamic Validation

We selected 3 functional tasks to represent a range of clinically relevant knee movements and to align with prior markerless motion capture validation literature [4,12]: squat, sit-to-stand, and single-leg stance.

Squat

The participants performed 5 consecutive repetitions at a self-selected pace, with feet shoulder-width apart and arms extended forward for balance. This task produced the greatest clinical knee flexion and a clearly identifiable target frame.

Sit-to-Stand

The participants performed 5 consecutive sit-to-stand transitions from a standard-height chair (approximately 45-cm seat height) at a self-selected pace, with arms extended forward for balance.

Single-Leg Stance

The participants stood on 1 leg for 15 seconds per trial; 2 trials were performed per leg (left and right), yielding 4 single-leg stance trials per participant. The supporting (standing) leg was the annotated limb for both BlazePose and Kinovea measurements. Beginning with S03, participants were instructed to position the nonstance leg so that it did not occlude the supporting leg in the camera view.

Assessment

Each participant completed 6 dynamic videos: 1 squat video containing 5 repetitions, 1 sit-to-stand video containing 5 repetitions, and 4 single-leg stance videos. For squat and sit-to-stand, raters identified the frame of maximum clinical knee flexion, equivalent to the minimum native hip-knee-ankle included angle. For single-leg stance, raters selected a steady midstance frame and measured the supporting leg, which remained near full extension (approximately 0°-20° of clinical flexion). This produced 14 target-frame observations per participant (5 squat, 5 sit-to-stand, 2 left single-leg stance, and 2 right single-leg stance) and 210 observations across 15 participants. The right knee was the default measurement side for bilateral squat and sit-to-stand; the supporting limb was measured for each single-leg stance trial.

XSL•FO
RenderX

Automated Pose Estimation: MediaPipe BlazePose

Overview

All videos were processed using MediaPipe BlazePose [13] via the PoseLandmarker Task API (mediapipe 0.10.33, Python 3.9.6) with the Heavy model (pose_landmarker_heavy.task). The model outputs 33 body landmarks per frame. Processing was configured in VIDEO running mode (frame-by-frame with temporal tracking) with a single-pose constraint (num_poses=1) and a default minimum pose-detection confidence of 0.5. All processing was performed on a local workstation (Apple M2 Max).

For each detected frame, the pipeline first computed the native 2D hip-knee-ankle included angle at the knee from pixel-coordinate vectors. Native-included angles approach 180° at full extension and decrease with flexion. For reporting and all analyses, this value was converted to clinical knee flexion as 180° − included angle, so that 0° represents full extension and larger values represent greater flexion. Right-knee landmarks 24, 26, and 28 and left-knee landmarks 23, 25, and 27 were used. Per-frame confidence was defined as the minimum visibility score of the hip, knee, and ankle landmarks. Unless otherwise specified, positive signed error denotes greater clinical flexion from BlazePose than from the comparator.

Detection Threshold Protocol Deviation

The default minimum pose detection confidence was 0.5. We reprocessed 5 static videos with insufficient detections at 0.3: S09_static_120deg_2, S10_static_120deg_2, S10_static_30deg_3, S04_retest_static_120deg_2, and S04_retest_static_120deg_3. This threshold reduction was a protocol deviation and was not applied to the remaining videos. The first 3 videos contributed to the Session 1 static analysis, and the latter 2 contributed to paired Session-2 static analyses.

The primary analyses retained the threshold-rescued videos. Sensitivity analyses repeated the Session 1 static analysis after excluding the 3 affected Session 1 videos and repeated the exploratory static repeated-session analysis after excluding the 2 affected pairs.

Automated Video-Minimum Extraction (Exploratory Use Only)

The pipeline also generated one automated video-minimum native-included angle per dynamic recording after excluding values below the implemented 10° anatomical-plausibility bound. This corresponds to the maximum reported clinical flexion after conversion. Raw-peak confidence was the minimum hip, knee, and ankle landmark visibility score at that automated video-minimum frame. These video-level automated minima were not used for the primary time-matched concurrent agreement analysis. They were used only in the exploratory n=24 all-output repeated-session stress test described in the following sections, without a raw-peak confidence gate.

Confidence-Based Frame Filtering

For static videos, frames with confidence ≥0.8 contributed to the filtered trial mean. For the primary dynamic BlazePose-versus-Kinovea comparison, the BlazePose clinical flexion value at the Kinovea-selected target time was classified as a detection failure when target frame confidence was <0.3. This criterion yielded 33 failures among 210 available observations; the conditional primary agreement analysis used the remaining 177. An all-output sensitivity stress test retained all 210 numeric outputs to characterize performance in the absence of this quality gate; it was not treated as a replacement for the conditional analysis. The target frame cutoff was established during analysis, not prespecified before data analysis. The April 16 interim analysis of S01-S06 included all 84 observations; the April 27 full-sample script introduced the 0.3 cutoff as an operational flag for low-confidence pose estimation failure. No calibration nor documented rationale for the numerical cutoff was available. It was distinct from the 0.3 pose detection parameter used to reprocess 5 static videos. A stricter target frame cutoff of ≥0.5 was also examined as a sensitivity analysis.

Kinovea 2D Video Reference

Two trained raters independently annotated task-specific target frame knee angles using Kinovea 2025.2. Each rater selected maximum clinical knee flexion for squat and sit-to-stand and a steady midstance frame for single-leg stance then digitized the greater trochanter, lateral femoral epicondyle, and lateral malleolus using the reflective markers as guides. The native 2D-included angle was converted to clinical knee flexion as 180° − included angle. Kinovea served as a 2D video-analysis reference; it was not assumed to provide absolute 3D joint kinematics.

Raters were blinded to each other's annotations and to BlazePose output. Their mean clinical flexion angle was used as the Kinovea 2D reference for dynamic agreement, and their mean target frame time stamp was used for BlazePose time matching. High interrater agreement reduces uncertainty from rater inconsistency but does not remove systematic 2D reference error.

BlazePose–Kinovea Time Matching

For the primary dynamic comparison, BlazePose was evaluated at the Kinovea-selected target time rather than at BlazePose's automated video minimum. For each of the 210 observations, the mean of the 2 rater time stamps was calculated, and the closest time stamp in the BlazePose per-frame output was selected. The native-included angle at that frame was converted to clinical flexion before analysis. This separates target-time selection by the 2D reference from BlazePose angle estimation at that time.

Descriptive Data Quality Grading

Each of the 342 videos received a descriptive A-D quality assurance grade based on the proportion of frames with confidence ≥0.8: Grade A (≥80%), Grade B (50% - 79%), Grade C (20% - 49%), or Grade D (<20%). These trial-level labels did not determine inclusion, weighting, or statistical inference and are not equivalent to target frame detection success. For a post hoc descriptive acquisition sequence summary, participants were grouped as S01-S06, S07-S10, and S11-S15; these groups were used only to describe target frame confidence and failure rates, without inferential testing. Detailed grade and detection

XSL•FO
RenderX

quality distributions are reported in Tables S1-S5 in Multimedia Appendix 1 .

Statistical Analysis

The original analyses used Python 3.9.6, pingouin for ICCs, and scipy.stats. The revision-stage participant cluster analyses used Python 3.12.14, NumPy 2.3.5, pandas 2.2.3, and SciPy 1.16.2. $ICC_{2,1}$ was recomputed from the same ANOVA formula and verified against the frozen results. Figures were generated using matplotlib.

Concurrent Validity

Agreement was summarized using the 2-way random effects, single-measure, absolute agreement ICC, $ICC_{2,1}$. Signed bias and limits of agreement (LoA) used the Bland-Altman framework [14]; MAE and RMSE quantified error magnitude. Static bias was BlazePose minus goniometer target clinical flexion, dynamic bias was BlazePose minus Kinovea clinical flexion, and interrater bias was Rater A minus Rater B. Pearson *r* described association not absolute agreement. ICC interpretation followed Koo and Li [15].

Primary cross-sectional inference retained the observation-level ICC and error estimands. We calculated percentile 95% CIs by resampling eligible participants with replacement, retaining all paired observations from each sampled participant (10,000 resamples; master seed 20260923; stable analysis-specific seeds). No trials were sampled independently. Dynamic agreement remained conditional on target frame confidence ≥0.3: Squat and combined single-leg stance each had 15 contributing participants, whereas sit-to-stand had 12. Participants with no retained observations remained in availability denominators. Repeated-measures LoA were calculated as bias ±1.96 times the square root of the sum of within- and between-participant variance components, allowing unequal repeats [16]. Negative estimated between-participant variance components were set to zero. Bootstrap CIs were also computed separately for bias and each LoA endpoint (Tables S16-S19 in Multimedia Appendix 1). Task-specific dynamic and angle-specific static results guided interpretation; pooled results are descriptive mixtures. Participant-mean analyses are supplementary and estimate agreement after averaging, not single-observation agreement. Participant-by-task or participant-by-angle cells from the same individual were not treated as independent for primary inference. Four-participant repeated-session results retained unadjusted descriptive intervals and were not used for primary inference.

For static analysis, the filtered-mean native BlazePose angle was converted to clinical flexion and compared with the goniometer target. Angle-specific results were reported to characterize error across 30°, 60°, 90°, and 120° of target flexion. For dynamic analysis, time-matched BlazePose clinical flexion was compared with the mean Kinovea clinical flexion value. Squat, sit-to-stand, and combined single-leg stance estimates were the main task-level results; side-specific single-leg stance and pooled cross-task estimates were secondary descriptive results.

Interrater Reliability

Interrater agreement of the Kinovea 2D annotations was assessed across 210 paired clinical flexion measurements using $ICC_{2,1}$, bias and 95% LoA, MAE, RMSE, and Pearson *r*. Estimates were reported overall and by task. The a priori descriptive benchmark for acceptable interrater ICC was 0.90. Between-task significance tests were not performed because trial-level implementations would not account for repeated observations within participants.

Exploratory Repeated-Session Analyses

Four participants (S01-S04) completed a second session after observed intervals of 21 days to 29 days (median 25 d). All repeated-session analyses were exploratory and addressed distinct measurement questions rather than constituting a single reliability claim.

First, BlazePose repeated-session agreement was evaluated for 48 paired static trial means and, as an all-output stress test, for 24 paired automated video-minimum dynamic values (4 participants × 6 videos). The dynamic estimand used 1 automated video minimum per recording rather than a Kinovea-selected time-matched frame and applied no raw-peak confidence gate; 10 of the 24 pairs had at least one raw-peak confidence value <0.3.

Second, repeated-session agreement of the Kinovea 2D reference was explored using 56 paired task/trial clinical flexion observations (4 participants × 14 observations).

Third, conditional BlazePose-versus-Kinovea agreement was described for Session 2. The same target-frame confidence criterion used in Session 1 retained 35 of 56 available observations. This 4-participant analysis was not treated as a replication study.

Fourth, the between-method clinical flexion difference was calculated in each session for 34 observations valid at both sessions. The ICC of these paired differences was used as an exploratory indication of whether the observed between-method offset might be consistent across sessions. Given the 4-participant sample and repeated observations, the estimate was not interpreted as proof of either stability or instability.

Repeated-session analyses used $ICC_{2,1}$, signed bias and 95% LoA, MAE, and RMSE. The analysis unit for these descriptive calculations was the repeated task/trial observation nested within 4 participants; the F-distribution ICC intervals did not adjust for participant-level clustering. Estimates were therefore presented with their denominators and uncertainty and were not used as headline clinical evidence.

Interpretation Thresholds

ICC values were described using the categories from Koo and Li [15]: <0.50, poor; 0.50 - 0.75, moderate; 0.75 - 0.90, good; and >0.90, excellent. Following Langley et al [6], ±5° for signed mean difference and ±10° for LoA were used only as descriptive contextual anchors for those same metrics, not as universal clinical acceptability thresholds. MAE and RMSE were reported as error-magnitude measures without assigning them a universal cutoff. Statistical association and agreement were interpreted

XSL•FO
RenderX

separately, and no multiplicity-adjusted confirmatory inference was claimed for exploratory strata.

Ethics Approval

This study was approved by the Academic Ethics and Integrity Review Committee of Shanghai Sanda University (approval number 2026006, April 14, 2026). The study was conducted in accordance with the Declaration of Helsinki. All participants provided informed consent prior to enrollment.

Results

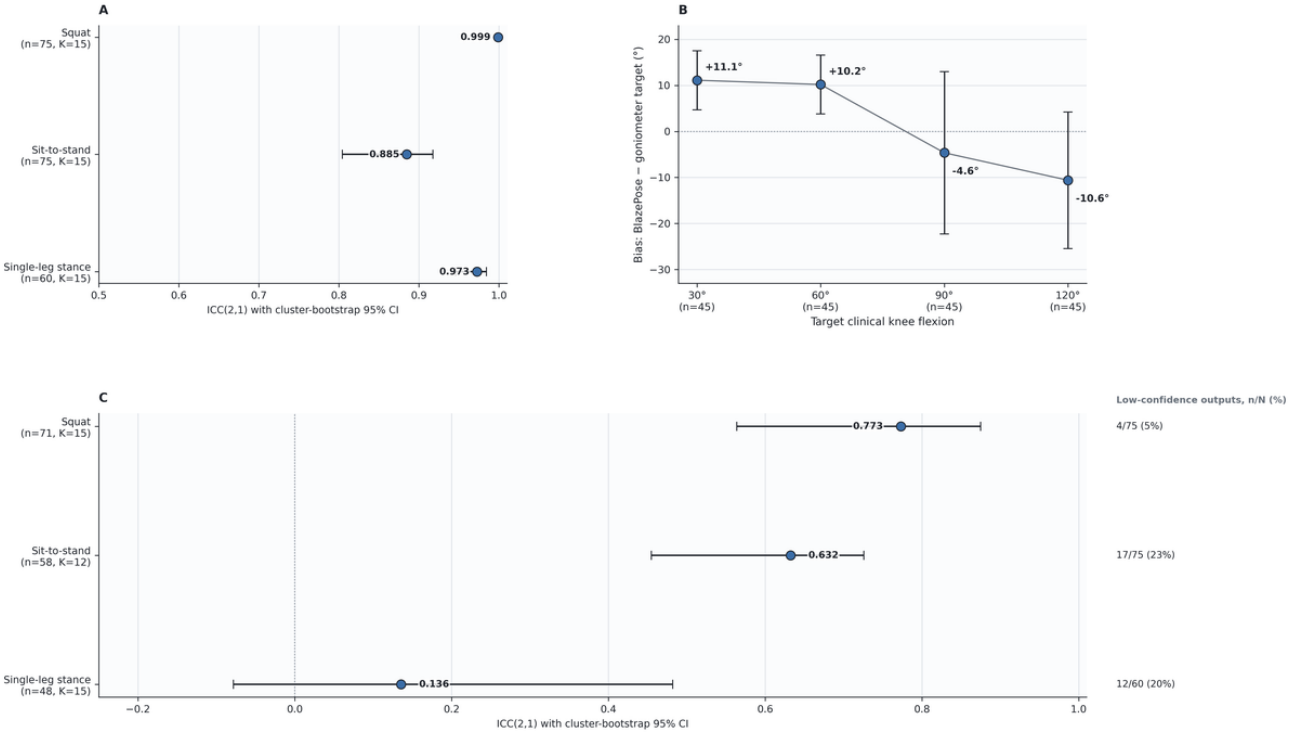
Participants

The cross-sectional analysis included 15 healthy adults (8 men, 7 women; mean age 20.8, SD 0.6, range 20 - 22 years; mean height 171.1, SD 9.4 cm; mean weight 70.7, SD 15.7 kg; mean BMI 23.9, SD 3.4, range 18.8 - 31.3 kg/m²). All 15 participants completed the full data collection protocol without reported adverse events. Four participants (S01-S04; 2 men, 2 women) returned for the exploratory repeated-session assessment after 21 days to 29 days (median 25 d).

For the primary cross-sectional analyses, 180 static trials (15 participants × 12 conditions) and 210 dynamic Kinovea-annotated target-frame observations were available.

Figure 1 summarizes the task-specific interrater agreement, angle-specific static signed error, and task-specific conditional dynamic agreement with detection failure rates. Detailed results are presented in Tables 1–5 and Multimedia Appendix 1.

Figure 1. Task- and angle-specific measurement results: (A) task-level interrater reliability between Kinovea Raters A and B, with points showing the intraclass correlation coefficient (ICC)$_{2,1}$, horizontal bars showing participant-cluster bootstrap 95% CI, K indicating the number of eligible participants, and n indicating the number of observations; (B) static BlazePose bias relative to goniometer target flexion, with points showing mean bias and error bars showing ±1 SD of paired differences not CIs and positive values indicating greater flexion estimated using BlazePose; (C) task-level dynamic agreement with the Kinovea 2D reference, conditional on target-frame BlazePose confidence ≥0.3, with the adjacent column reporting low-confidence outputs using all attempts as denominators. Single-leg stance combines left and right sides. All angles are clinical knee flexion, with 0° at full extension.



XSL•FO
RenderX

Table . Analysis domains, available and analyzed observations, and independent participants.

Analysis[a]	Comparison/estimand	Available observations, n	Analyzed observations, n	Independent participants, n	Role
Static agreement	BlazePose versus goniometer target	180	180	15	Main cross-sectional
Dynamic conditional agreement[b]	BlazePose versus Kinovea 2D reference; target-frame confidence ≥0.3	210	177	15 overall; 12 - 15 by task[c]	Main cross-sectional; task-specific results lead
Interrater agreement	Kinovea Rater A versus Rater B	210	210	15	Main cross-sectional
BlazePose static repeated-session agreement	Session 1 versus Session 2	48	48	4	Exploratory
BlazePose dynamic video-minimum stress test[d]	Session 1 versus Session 2; automated video minimum, no confidence gate	24	24	4	Exploratory stress test
Kinovea repeated-session agreement	Session 1 versus Session 2	56	56	4	Exploratory
Session 2 conditional agreement[e]	BlazePose versus Kinovea 2D reference	56	35	4	Exploratory
Between-method difference across sessions	Session 1 versus Session 2 between-method differences	56	34 valid in both sessions	4	Exploratory

[a]Confidence interval and repeated measures LoA methods for cross-sectional analyses are described in the Statistical Analysis section and Tables S16-S19 in Multimedia Appendix 1.

[b]Excluded 33 of the 210 target-frame observations with BlazePose confidence <0.3 and reported them separately as detection failures.

[c]Primary dynamic cluster counts are 15 for squat, 12 for sit-to-stand, and 15 for combined single-leg stance.

[d]All-output automated video-minimum stress test; 10 of 24 pairs had at least one raw-peak confidence value <0.3.

[e]Retained 35 of the 56 observations.

Table . Interrater agreement of Kinovea 2D clinical knee-flexion annotations (n=210), with all angles representing clinical knee flexion (0° at full extension).

Task	Observations, n	Bias[a]	95% LoA[b,c]	MAE[d]	RMSE[e]	Pearson *r*[f]	ICC[g,h]$_{2,1}$ (95% CI[i])
Overall[j]	210	+1.15°	−3.51° to +5.81°	1.56°	2.62°	0.999	0.999 (0.998 to 0.999)
Squat	75	+0.09°	−1.25° to +1.43°	0.39°	0.68°	0.999	0.999 (0.997 to 0.999)
Sit-to-stand	75	+2.65°	−3.78° to +9.08°	3.13°	4.16°	0.936	0.885 (0.804 to 0.918)
Single-leg stance	60	+0.58°	−1.88° to +3.05°	1.05°	1.37°	0.977	0.973 (0.946 to 0.984)

[a]Bias: Kinovea Rater A clinical flexion−Rater B clinical flexion.

[b]LoA: limits of agreement.

[c]LoA use repeated-measures variance components; endpoint CIs are in Table S17 in Multimedia Appendix 1.

[d]MAE: mean absolute error.

[e]RMSE: root mean square error.

[f]Describes the association, not absolute agreement.

[g]ICC: intraclass correlation coefficient.

[h]2-way random-effects, single-measure, absolute-agreement ICC.

[i]ICC CIs use participant-cluster bootstrapping (15 participants per row).

[j]Secondary descriptive cross-task summary.

XSL•FO
RenderX

Table . Static agreement of BlazePose clinical knee flexion with goniometer target flexion (n=180; S01-S15).

Target flexion	Observations, n	Bias[a]	95% LoA[b,c,d]	MAE[e]	RMSE[f]	SD(diff)[g]	SEM[h,i,g]
Overall[j,k]	180	+1.51°[l]	−28.85° to +31.87°	10.01°	15.49°	15.46°	10.93°
30°	45	+11.11°	−1.73° to +23.94°	11.15°	12.79°	6.41°	4.53°
60°	45	+10.18°	−2.66° to +23.03°	10.82°	11.99°	6.40°	4.53°
90°	45	−4.64°	−39.36° to +30.07°	7.45°	18.02°	17.60°	12.45°
120°	45	−10.60°	−39.71° to +18.50°	10.60°	18.10°	14.84°	10.49°

[a]Bias: BlazePose clinical flexion−goniometer target flexion.

[b]LoA: limits of agreement.

[c]LoA endpoint CIs are in Table S17 in Multimedia Appendix 1.

[d]LoA use repeated-measures variance components (15 participants per row).

[e]MAE: mean absolute error.

[f]RMSE: root mean square error.

[g]SD(diff) and SEM remain descriptive trial-level summaries.

[h]SEM: standard error of measurement.

[i]SEM =SD(diff)/√2.

[j]Overall $ICC_{2,1}$=0.874 (95% CI 0.770 to 0.953).

[k]The ICC CI uses participant-cluster bootstrapping.

[l]Overall pooled bias is descriptive because angle-specific errors reverse direction.

Table . Conditional agreement of BlazePose with the Kinovea 2D clinical flexion reference at target frames with BlazePose confidence ≥0.3, with all values reflecting clinical knee flexion.

Task	Retained observations, n	Available observations, n	Failures from the available observations, n (%)	Bias[a] (95% LoA[b,c])[d]	MAE[e]	RMSE[f]	Pearson *r*[g]	$ICC^{h}_{2,1}$ (95% CI[i])
Squat (n=15 participants)	71	75	4 (5)	+6.17° (−6.45° to +18.78°)	7.10°	8.78°	0.885	0.773 (0.564 to 0.875)
Sit-to-stand (n=12 participants)	58	75	17 (23)	+6.43° (−1.82° to +14.67°)	6.68°	7.59°	0.896	0.632 (0.455 to 0.726)
Single-leg stance, both legs combined (n=15 participants)	48	60	12 (20)	+2.01° (−23.42° to +27.44°)	7.51°	12.88°	0.173	0.136 (−0.078 to 0.482)
Overall, secondary descriptive (n=15 participants)	177	210	33 (15.7)	+5.12°[j] (−11.30° to +21.55°)	7.07°	9.73°	0.987	0.982 (0.967 to 0.992)

[a]Bias: BlazePose−Kinovea 2D reference; positive values indicate greater flexion from BlazePose.

[b]LoA: limits of agreement.

[c]LoA use repeated-measures variance components.

[d]Bias and LoA endpoint CIs are in Tables S16-S19 in Multimedia Appendix 1.

[e]MAE: mean absolute error.

[f]RMSE: root mean square error.

[g]Describes an association, not absolute agreement.

[h]ICC: intraclass correlation coefficient.

[i]ICC CIs use participant-cluster bootstrapping.

[j]The overall pooled ICC is range-sensitive and is not a headline result.

XSL•FO
RenderX

Table . Exploratory repeated-session analyses with 4 participants (S01-S04; observed interval 21 - 29 d, median 25 d); pooled estimates may be influenced by between-task or between-angle range and are not headline clinical evidence.

Exploratory analysis	Observations, n	Bias definition	Bias (95% LoA[a])	MAE[b]	RMSE[c]	ICC[d]$_{2,1}$ (95% CI)[e]
BlazePose static, Session 1 versus Session 2	48	Session 2−Session 1	−7.30° (−42.65° to +28.04°)	10.32°	19.28°	0.782[f] (0.617 to 0.877)
BlazePose dynamic automated video-minimum stress test[g]	24	Session 2−Session 1	+18.92° (−82.04° to +119.89°)	34.71°	53.86°	0.522 (0.174 to 0.757)
Kinovea 2D reference, Session 1 versus Session 2	56	Session 2−Session 1	+1.19° (−12.05° to +14.42°)	5.63°	6.80°	0.991 (0.985 to 0.995)
BlazePose versus Kinovea, Session 2 conditional	35	BlazePose−Kinovea	+8.10° (−14.69° to +30.89°)	8.95°	14.03°	0.969 (0.889 to 0.988)
Between-method difference, Session 1 versus Session 2	34	Session 2 difference−Session 1 difference	+2.45° (−14.01° to +18.91°)	5.19°	8.63°	0.094 (−0.225 to 0.405)

[a]LoA: limits of agreement.

[b]MAE: mean absolute error.

[c]RMSE: root mean square error.

[d]ICC: intraclass correlation coefficient.

[e]Do not adjust for participant-level clustering.

[f]Excluding the 2 retest threshold-rescued static pairs changed the static repeated-session ICC from 0.782 to 0.766.

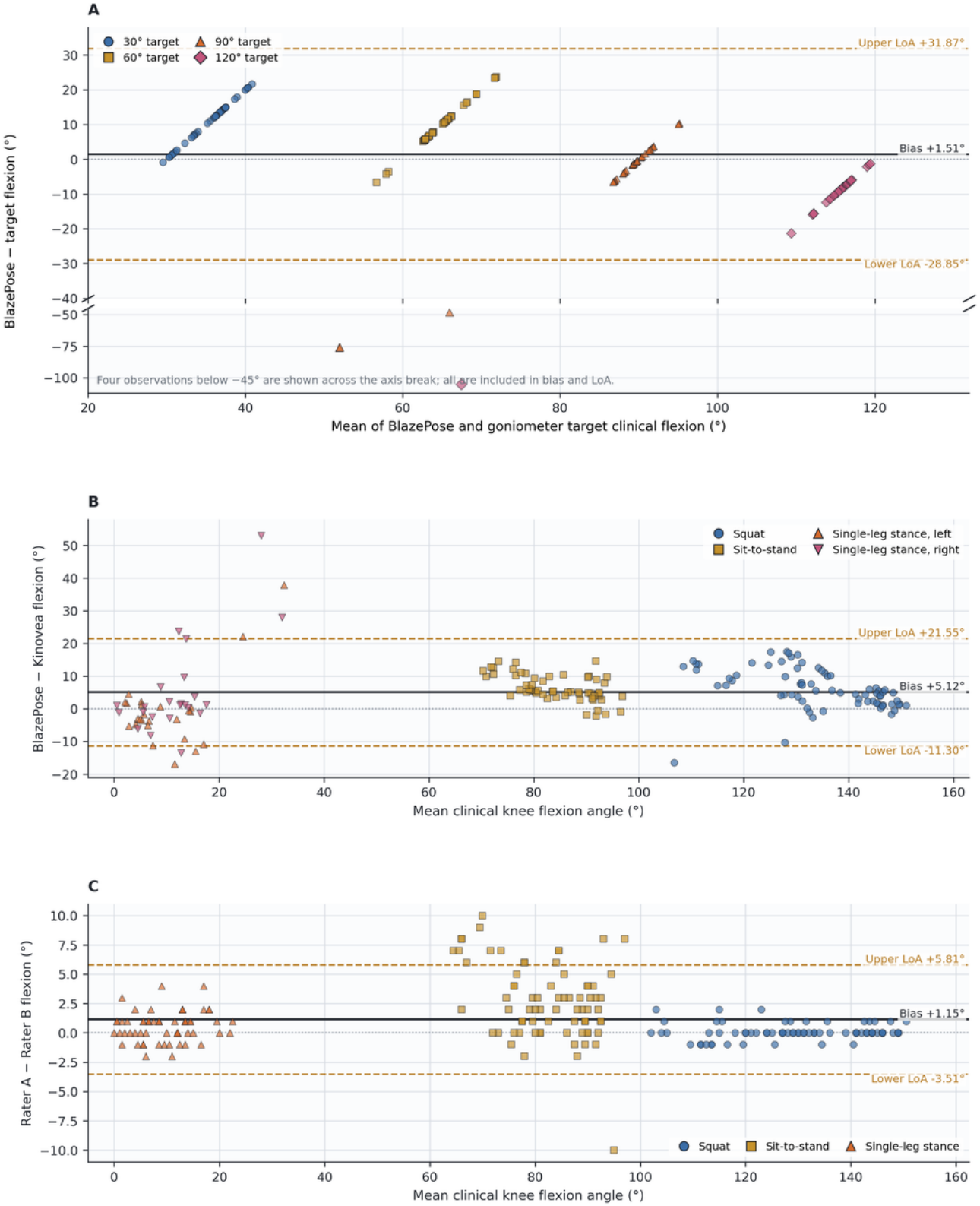
[g]Uses 1 automated video minimum per recording, applies no raw-peak confidence gate, and includes 10 of 24 pairs with at least one raw-peak confidence value <0.3.

Interrater Agreement of Kinovea 2D Annotations

Overall interrater agreement for the Kinovea 2D clinical flexion annotations was high ($ICC_{2,1}$=0.999, cluster-bootstrap 95% CI 0.998 to 0.999; MAE=1.56°; n=210 observations from 15 participants; Table 2; Figure 2C). This indicates consistent annotation between these 2 raters but does not establish absolute accuracy against 3D kinematics.

XSL•FO
RenderX

Figure 2. Bland-Altman plots using clinical knee flexion angles and showing the mean of paired angles on the x axis and the first-named measurement minus the second on the y axis, with positive differences indicating greater flexion from the first measurement: (A) BlazePose versus goniometer-set targets (n=180, 15 participants; bias +1.51°; 95% limits of agreement [LoA] –28.85° to +31.87°), with 4 observations <−45° shown across a broken y axis and included in all calculations; (B) BlazePose versus Kinovea 2D reference among outputs with target-frame confidence ≥0.3 (n=177, 15 participants overall; bias +5.12°; 95% LoA –11.30° to +21.55°); (C) Kinovea Rater A versus Rater B (n=210, 15 participants; bias +1.15°; 95% LoA –3.51° to +5.81°). The solid dark line is the mean bias, the dashed lines are repeated-measures 95% LoA from within- and between-participant variance components, and the dotted line is zero. These limits are not CIs; CIs for their endpoints are in Table S17 in Multimedia Appendix 1. Points are repeated observations. Pooled lines summarize mixtures of angles or tasks; they do not imply constant bias or uniform coverage across those conditions. Angle- and task-specific limits guide interpretation (Tables 2–4).



XSL•FO
RenderX

Task-specific ICCs were 0.999 for squat, 0.885 for sit-to-stand, and 0.973 for single-leg stance. The sit-to-stand point estimate of 0.885 was less than the a priori descriptive benchmark of 0.90, and its cluster-bootstrap 95% CI was 0.804 to 0.918. Sit-to-stand had the largest interrater MAE (3.13°) and a clinical-flexion bias of +2.65° for Rater A minus Rater B. The lower sit-to-stand ICC may reflect the more gradual transition around peak flexion, which makes the selection of a single target frame less distinct than in squat; this is a plausible explanation rather than a demonstrated mechanism.

Static Agreement With Goniometer-Set Targets

Across 180 static observations from 15 participants, $ICC_{2,1}$ was 0.874 (cluster-bootstrap 95% CI 0.770 to 0.953), MAE was 10.01°, and RMSE was 15.49° (Table 3; Figure 2A). These differences include uncertainty in target positioning and holding as well as BlazePose estimation and are reported without a universal clinical acceptability cutoff.

Signed error varied by target angle. BlazePose overestimated clinical flexion at 30° (+11.11°) and 60° (+10.18°) and underestimated it at 90° (−4.64°) and 120° (−10.60°). Dispersion was substantially wider at 90° and 120° than at 30° and 60°. This range compression did not support a single fixed correction across the tested angles and did not identify one causal mechanism.

Excluding the 3 static Session 1 videos reprocessed at the lower pose detection threshold changed static ICC from 0.874 to 0.870 and RMSE from 15.49° to 15.58°. This sensitivity analysis quantifies the effect of the protocol deviation; full results are reported in Multimedia Appendix 1 (Table S14).

Dynamic Conditional Agreement With the Kinovea 2D Reference

Dynamic measurement was available for 177 of 210 (84.3%) observations. The remaining 33 of 210 (15.7%) observations had target-frame confidence <0.3. Availability reached 71 of 75 (95%) observations for squat, 58 of 75 (77%) observations for sit-to-stand, and 48 of 60 (80%) observations for combined single-leg stance. Failures occurred for 5 of 30 (17%) observations on the left and 7 of 30 (23%) observations on the right. Participant counts and distributions are reported in Tables S1 and S5 in Multimedia Appendix 1.

Failures were concentrated in S10 (11/14, 79%), S13 (6/14, 43%), and S12 (5/14, 36%). Kinovea flexion medians for failed versus retained observations were 113.2° versus 131.0° for squat, 89.0° versus 81.0° for sit-to-stand, and 10.2° versus 8.8° for combined single-leg stance. These are descriptive contrasts not evidence of why detection failed; only 4 squat observations failed. Quartiles, ranges, and participant counts are in Table S5 in Multimedia Appendix 1.

Detection performance also varied by acquisition group. Target-frame median confidence and failures were 0.939 and 2 of 84 (2%) observations, respectively, for S01-S06; 0.643 and 14 of 56 (25%) observations, respectively, for S07-S10; and 0.694 and 17 of 70 (24%) observations, respectively, for S11-S15. No recorded device nor encoding variable changed uniquely at S07 or aligned with the persistent later-group decrease; the known 120-to-60 frames-per-second change occurred earlier, between S02 and S03. Because lighting, exact camera geometry, exposure and focus, attire, and background were incompletely recorded, the cause remains unresolved.

Task-specific ICCs were 0.773 (95% CI 0.564 to 0.875; n=71, 15 participants) for squat, 0.632 (95% CI 0.455 to 0.726; n=58, 12 participants) for sit-to-stand, and 0.136 (95% CI −0.078 to 0.482; n=48, 15 participants) for combined single-leg stance (Table 4). Squat and sit-to-stand had high Pearson correlations (r=0.885 and r=0.896), but their absolute-agreement ICC point estimates were good and moderate, respectively. BlazePose overestimated clinical flexion by +6.17° and +6.43°, respectively. The restricted near-extension range in single-leg stance reduced between-participant variance and made the ICC sensitive to modest error; occlusion, out-of-plane rotation, perspective shortening, side selection, and landmark localization are additional plausible contributors.

Across 177 retained observations, the pooled cross-task ICC was 0.982 (cluster-bootstrap 95% CI 0.967 to 0.992), with bias +5.12° and repeated-measures LoA −11.30° to +21.55° (Figure 2B). This is a secondary descriptive mixture because between-task angle range inflates between-observation variance; it does not imply uniform agreement across tasks.

Participant mean results are provided in Table S18 in Multimedia Appendix 1. These compare the average retained measurements per person and task, not individual observations. The numbers averaged varied because of detection failures. Earlier exploratory participant-by-task summaries in Tables S10-S11 in Multimedia Appendix 1 are retained with their distinct units and unadjusted intervals explicitly identified.

In the all-output stress test, retaining all 210 numeric outputs produced the largest deterioration for squat, for which the ICC decreased from 0.773 to 0.428 and RMSE increased from 8.78° to 25.43°. The sit-to-stand ICC decreased from 0.632 to 0.596, and the combined single-leg stance ICC decreased from 0.136 to 0.119. The pooled ICC decreased from 0.982 to 0.942. These estimates characterize performance without the quality gate and do not replace the conditional primary analysis.

At the stricter target-frame threshold of ≥0.5, availability was 158 of 210 (75.2%) observations; task-specific ICCs were 0.770 (n=70) for squat, 0.670 (n=50) for sit-to-stand, and 0.204 (n=38) for combined single-leg stance (Table S15 in Multimedia Appendix 1). This sensitivity analysis describes the availability-agreement tradeoff and does not establish an optimal threshold.

Exploratory Repeated-Session Findings

Kinovea 2D Reference Across Sessions

Across the 56 repeated task/trial observations collected 21 days to 29 days apart, the pooled Kinovea repeated-session ICC was 0.991 (95% CI 0.985 to 0.995), with an MAE of 5.63° (Table 5). This pooled value reflects the wide range across tasks and repeated observations of only 4 participants. These 4-participant repeated-session data were insufficient to determine whether Kinovea agreement differed by task.

XSL•FO
RenderX

BlazePose Across Sessions

Static repeated-session agreement was an ICC of 0.782 (95% CI 0.617 to 0.877; n=48), with Session 2-minus-Session 1 clinical-flexion bias of −7.30° and a wide LoA (−42.65° to +28.04°). The dynamic all-output automated video-minimum stress test yielded an ICC of 0.522 (95% CI 0.174 to 0.757; n=24), an MAE of 34.71°, and an LoA of −82.04° to +119.89°. At least one raw-peak confidence value <0.3 was present for 10 of the 24 dynamic pairs. These values describe an ungated stress test of repeated automated video minima for 4 participants, not time-matched repeated-session validity.

Session 2 Conditional Agreement

In Session 2, 21 of the 56 target-frame observations failed the confidence criterion, leaving 35 for conditional agreement analysis. The pooled cross-task ICC among retained observations was 0.969 (95% CI 0.889 to 0.988), with a bias of +8.10° and an MAE of 8.95°. Because this estimate pools tasks and derives from 4 participants after substantial selection, it is presented as an exploratory descriptive result rather than evidence that Session 1 performance was replicated.

Between-Method Difference Across Sessions

Among 34 observations valid in both sessions, the ICC for the paired BlazePose-minus-Kinovea clinical-flexion differences was 0.094 (95% CI −0.225 to 0.405). This wide interval from 4 participants is insufficient to support a stable fixed-offset assumption; it is also insufficient to establish that bias instability is a general property of the method. Larger repeated-session samples are required before calibration or longitudinal change interpretations can be evaluated.

Data Quality and Detection Reproducibility

Across all 342 videos, the descriptive quality-grade distribution was Grade A, 164/342 (48%); Grade B, 21/342 (6.1%); Grade C, 31/342 (9.1%); and Grade D, 126/342 (36.8%; Table S4 in Multimedia Appendix 1). These video-level quality assurance labels did not determine target-frame inclusion, weighting, or statistical inference and are not measures of target-frame detection success.

Detailed detection quality, retest, subject mean, and sensitivity results are provided in Multimedia Appendix 1.

Discussion

Principal Findings

Agreement of single-camera BlazePose knee-flexion estimates with the study references depended on task, flexion angle, and detection quality. With measurement availability of 177 of 210 observations (84.3%), task-specific conditional ICC point estimates ranged from good for squat to poor for single-leg stance. The primary cross-sectional CIs now account for repeated observations through participant-cluster resampling, and LoA incorporated within- and between-participant variance. These changes did not remove the uncertainty from only 12 to 15 contributing participants per dynamic task. Static differences changed direction across target angles and included target-setting and holding uncertainty. Kinovea interrater agreement did not establish 3D accuracy, and 4-participant repeated-session findings remained exploratory.

Kinovea as a 2D Reference

High Kinovea interrater agreement indicates that 2 raters produced similar 2D annotations and reduced 1 source of uncertainty in the BlazePose comparison. It did not establish Kinovea's accuracy against 3D kinematics or show that every between-method difference originates from BlazePose. Both methods used a single sagittal view and may have shared projection, landmark, and out-of-plane error.

Task- and Quality-Dependent Dynamic Agreement

Task-specific estimates are more informative than the pooled cross-task ICC because pooling squat, sit-to-stand, and near-extension single-leg stance greatly expanded the angle range. Squat and sit-to-stand showed strong relative association but only good-to-moderate absolute agreement, illustrating why correlation cannot substitute for agreement metrics. In the clinical-flexion space, BlazePose overestimated flexion by approximately 6° for these tasks, and every task-specific LoA interval extended beyond at least one of the descriptive ±10° anchors. If reproduced clinically, this error direction could make recovery appear greater than it is; the healthy sample cannot quantify that clinical consequence.

The conditional analysis also described performance only when target-frame confidence was adequate. Retaining all numeric outputs produced marked deterioration for squat, and failure rates varied by task, participant, and acquisition group. The unresolved confidence decrease after S01-S06 shows that a constant camera model and file format did not guarantee consistent detection. Any future deployment would therefore require explicit measurement availability reporting and a per-recording quality gate rather than relying only on accuracy among successful detections.

Angle-Dependent Static Error

Static signed error changed from flexion overestimation at 30° and 60° to underestimation at 90° and 120°. The angle-specific signed bias exceeded the descriptive ±5° anchor at 30°, 60°, and 120°, and every angle-specific LoA interval extended beyond at least one of the descriptive ±10° anchors. Together with the overall RMSE, this direction reversal and dispersion did not support a single calibration offset across the tested range. The result was agreement with goniometer-set targets in healthy participants not evidence that BlazePose can replace clinical goniometry. The goniometer was reapplied and reread before every trial but removed for recording and not reapplied afterward. Consequently, target-setting error, rounding, drift during the hold, and goniometer error cannot be separated from BlazePose error; the manufacturer-stated accuracy is not an in-study calibration estimate.

Plausible contributors included mismatch between image-based and palpated anatomical landmarks, perspective shortening, out-of-plane limb rotation, occlusion or crossing limbs, side-selection errors, and amplification of small landmark displacements when segments approach collinearity. These mechanisms are consistent with the observed angle and task

XSL•FO
RenderX

dependence but were not isolated experimentally and should not be stated as proven causes.

Single-Leg Stance: Restricted Range and Plausible Failure Mechanisms

Single-leg stance had poor concurrent agreement. Because the supporting knee occupied a narrow range near full extension, between-participant variance was small, and ICC was especially sensitive to modest measurement error. Occlusion by the nonstance limb, out-of-plane rotation, perspective, and landmark localization may add error in a single lateral view. Exploratory repeated-session findings from 4 participants did not resolve whether this uncertainty was task-related, method-related, or sample-related and did not establish that instability is intrinsic to the task.

Exploratory Repeated-Session Findings

The repeated-session analyses were based on 4 participants assessed 21 days to 29 days apart and cannot support definitive reliability or bias stability claims. The n=24 dynamic analysis was an ungated, automated, video-minimum stress test not a time-matched measure, and 10 of the 24 pairs included at least one low-confidence raw peak. The ICC for the between-method difference was imprecise and is best interpreted as insufficient evidence for a stable fixed offset not proof of universal bias drift. The 21 - to-29-day interval also permits real within-person change to contribute to disagreement, so measurement error and biological variation cannot be separated here.

Limitations

The cross-sectional sample was small (n=15) and recruited by convenience from healthy young adults at one university, and no a priori power calculation was performed. The repeated-session subsample contained 4 participants, producing imprecise estimates from repeated observations nested within individuals. Findings cannot be generalized to older adults, knee osteoarthritis, anterior cruciate ligament reconstruction, total knee arthroplasty, or other clinical populations. One readily available iPhone 17 Pro was used throughout to hold hardware constant; equivalence across phones, lenses, frame rates, and device-specific image processing was not tested.

The static comparator was a rounded goniometer-set target rather than an independent measurement of the recorded posture, and the dynamic comparator was Kinovea 2D annotation rather than marker-based 3D kinematics, so absolute biomechanical accuracy remains unknown. We reprocessed 5 static videos at a lower pose-detection threshold, which was a protocol deviation; sensitivity analyses excluding them changed the estimates only slightly. The 33 failures from the 210 dynamic detection observations were unevenly distributed, making the retained-observation agreement estimates conditional on adequate detection. Target-frame confidence decreased after S01-S06 and did not recover in S11-S15; because acquisition conditions were incompletely recorded, the cause could not be identified. Participant-cluster CIs and repeated-measures LoA addressed clustering in the primary cross-sectional analyses, but with only 12 to 15 contributing participants per dynamic task, all interval procedures remained approximate; bootstrap intervals may have had inadequate coverage. The LoA variance model assumes approximately normal differences, common within-participant variance within a stratum, and conditionally independent residual differences. Consecutive repetitions may violate the last assumption. Resampling whole participants preserves the observed dependence but does not validate those distributional assumptions or guarantee future coverage. Across-angle and cross-task pooled LoA are descriptive only. The target-frame cutoff was established during analysis without a documented calibration basis, and sensitivity analyses did not make it a validated deployment threshold. Four-participant repeated-session intervals remained unadjusted and descriptive. Detailed angle-, side-, participant-, and sensitivity-level estimates are exploratory and were not adjusted as multiple confirmatory tests. The study assessed 1 BlazePose model, sagittal-plane angles, and selected task frames only; frontal/transverse motion and temporal variables were not evaluated.

Clinical Implications and Future Directions

This evidence supports only feasibility or screening-level research use under controlled acquisition with explicit quality checks. It does not support replacing a clinical goniometer, making precise longitudinal treatment decisions, assuming that a fixed calibration is stable, or using the method as an unconditional remote clinical measurement. Because BlazePose overestimated clinical flexion in squat and sit-to-stand, an unverified reading could overstate apparent recovery or mask a residual deficit; measurements that would affect care should be confirmed with an established clinical assessment.

Priorities for future work are larger, prospectively sized, and multisite cohorts; validation against 3D motion capture; testing in relevant clinical populations; direct cross-device evaluation; and controlled experiments that vary camera geometry, lighting, attire, occlusion, and out-of-plane motion. Repeated-session studies should prespecify the estimand, retain participant-level sampling in the analysis, and report both measurement availability and agreement conditional on successful detection.

Conclusions

Single-camera MediaPipe BlazePose had dynamic measurement availability of 177 of 210 (84.3%) observations: This included 71 of 75 (95%) observations for squat, 58 of 75 (77%) observations for sit-to-stand, and 48 of 60 (80%) observations for single-leg stance. Among the retained observations, the corresponding ICCs for squat, sit-to-stand, and single-leg stance were 0.773, 0.632, and 0.136, respectively with participant-cluster intervals showing substantial uncertainty. Static agreement with goniometer-set targets varied by angle and included uncertainty in the reference positioning. Four-participant repeated-session findings did not establish stable calibration or reliable longitudinal change measurement. These data support further controlled feasibility and screening-level evaluation with per-recording quality checks not replacement of clinical measurement or routine remote clinical decision-making.

XSL•FO
RenderX

Acknowledgments

The authors thank all 15 participants, including the 4 participants who returned for the second session, and the research assistants who contributed to Kinovea video annotation and data collection logistics.

OpenAI Codex and Anthropic Claude Code were used to assist with coding, statistical analyses, preparation of tables and figures, manuscript drafting and revision, and responses to reviewers. XY and BY reviewed the methods and code, checked the tabulated results against the study data, and reviewed the interpretations. The authors take responsibility for the accuracy and integrity of the final manuscript.

Funding

This research received no external funding.

Data Availability

The anonymized summary datasets generated and analyzed during this study are available from the corresponding author upon reasonable request. The BlazePose analysis pipeline used the open-source MediaPipe PoseLandmarker (Heavy model, mediapipe v0.10.33), available from Google's MediaPipe documentation. Kinovea is freely available from the official Kinovea project website. Raw video data cannot be shared due to participant identifiability (faces visible in lateral-view recordings).

Authors' Contributions

Conceptualization: XY
Data curation: XY
Formal analysis: XY
Investigation: XY, YG
Methodology: XY
Project administration: XY
Resources: BY
Software: XY
Supervision: BY
Validation: XY
Visualization: XY
Writing – original draft: XY
Writing – review & editing: XY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary detection quality, exploratory repeated-session, participant-mean, sensitivity, and participant-cluster inference analyses.
[DOCX File, 45 KB - rehab_v13i1e102399_app1.docx]

References

1. Brosseau L, Balmer S, Tousignant M, et al. Intra- and intertester reliability and criterion validity of the parallelogram and universal goniometers for measuring maximum active knee flexion and extension of patients with knee restrictions. Arch Phys Med Rehabil 2001 Mar;82(3):396-402. [doi: 10.1053/apmr.2001.19250] [Medline: 11245764]
2. Colyer SL, Evans M, Cosker DP, Salo AIT. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. Sports Med Open 2018 Jun 5;4(1):24. [doi: 10.1186/s40798-018-0139-y] [Medline: 29869300]
3. Wade L, Needham L, McGuigan P, Bilzon J. Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. PeerJ 2022;10:e12995. [doi: 10.7717/peerj.12995] [Medline: 35237469]
4. Stenum J, Rossi C, Roemmich RT. Two-dimensional video-based analysis of human gait using pose estimation. PLoS Comput Biol 2021 Apr;17(4):e1008935. [doi: 10.1371/journal.pcbi.1008935] [Medline: 33891585]
5. Gupta A, Shrestha PL, Thapa B, Silwal R, Shrestha R. Knee flexion/extension angle measurement for gait analysis using machine learning solution "MediaPipe Pose" and its comparison with Kinovea ®. IOP Conf Ser: Mater Sci Eng 2023 Mar 1;1279(1):012004. [doi: 10.1088/1757-899X/1279/1/012004]

XSL•FO
RenderX

6. Langley B, Greig M, Board TN. Exploring the validity of smartphone based single camera markerless motion capture technology to quantify knee range of motion in patients with knee osteoarthritis. Physiotherapy 2026 Mar;130:101850. [doi: 10.1016/j.physio.2025.101850] [Medline: 41275724]
7. Koul M, Thanki R, Koul R, Fujarski S, Radhakrishnan S, Jodbhavi S. An AI-driven video based goniometer for knee joint range of motion (ROM) assessment: reliability and validity compared to traditional goniometry. Comput Biol Med 2025 Sep;196(Pt B):110848. [doi: 10.1016/j.compbiomed.2025.110848] [Medline: 40768930]
8. Akturk S, Derdiyok FB, Serbest K. Markerless joint angle estimation using MediaPipe with a rapid setup for joint moment calculation. Multimed Tools Appl 2026;85(1):38. [doi: 10.1007/s11042-026-21256-z]
9. Lima YL, Collings T, Hall M, Bourne MN, Diamond LE. Validity and reliability of trunk and lower-limb kinematics during squatting, hopping, jumping and side-stepping using OpenCap markerless motion capture application. J Sports Sci 2024 Oct;42(19):1847-1858. [doi: 10.1080/02640414.2024.2415233] [Medline: 39444219]
10. Yoma M, Herrington L, Starbuck C, Llurda L, Jones R. Between-day reliability of kinematic variables using markerless motion capture for single-leg squat and single-leg landing tasks. Int J Sports Phys Ther 2025;20(8):1160-1175. [doi: 10.26603/001c.141870] [Medline: 40756794]
11. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010 May;19(4):539-549. [doi: 10.1007/s11136-010-9606-8] [Medline: 20169472]
12. Bae K, Lee S, Bak SY, Kim HS, Ha Y, You JH. Concurrent validity and test reliability of the deep learning markerless motion capture system during the overhead squat. Sci Rep 2024 Nov 27;14(1):29462. [doi: 10.1038/s41598-024-79707-2] [Medline: 39604407]
13. Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, Grundmann M. BlazePose: on-device real-time body pose tracking. arXiv preprint arXiv. Preprint posted online on 2020. [doi: 10.48550/arXiv.2006.10204]
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
16. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17(4):571-582. [doi: 10.1080/10543400701329422] [Medline: 17613642]

Abbreviations

COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
ICC: intraclass correlation coefficient
LoA: limits of agreement
MAE: mean absolute error
RMSE: root mean square error
ROM: range of motion
SEM: standard error of measurement

Edited by A Mastropietro; submitted 26.May.2026; peer-reviewed by J McMordie, M Chakit; revised version received 24.Sep.2026; accepted 01.Oct.2026; published 03.Nov.2026.

Please cite as:
Yue X, Gu Y, Yu B
Validation of Single-Camera MediaPipe BlazePose for Knee Joint Angle Measurement Across Multiple Functional Tasks: Concurrent Validity, Interrater Reliability, and Exploratory Test–Retest Reliability Study
JMIR Rehabil Assist Technol 2026;13:e102399
URL: https://rehab.jmir.org/2026/1/e102399
doi:10.2196/102399

© Xiaotian Yue, YiXuan Gu, Bo Yu. Originally published in JMIR Rehabilitation and Assistive Technologies (https://rehab.jmir.org), 3.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technologies, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integrating Functional Cognitive Rehabilitation With Activity-Based Stroke Telerehabilitation: Intervention Design and Development Study

Patricia Finetto[1], MSOT, OTR/L; Michelle L Woodbury[1], PhD, OTR/L; Julianne Laura[1], MSOT, OTR/L; Kelly Rishe[1], MSOT, OTR/L; Stephanie Garner[1], OTD, OTR/L; Lisa M McTeague[2,3], PhD; Deena Schwen Blackett[4], PhD, CCC-SLP; Elizabeth W Twamley[5,6], PhD; Stephanie Aghamoosa[1], PhD

[1]Department of Health Sciences & Research, College of Health Professions, Medical University of South Carolina, 77 President Street, Charleston, SC, United States

[2]Department of Psychiatry and Behavioral Sciences, College of Medicine, Medical University of South Carolina, Charleston, SC, United States

[3]Ralph H. Johnson Veterans Affairs Medical Center, Charleston, SC, United States

[4]School of Communication Sciences and Disorders, College of Health Professions, University of Central Florida, Orlando, FL, United States

[5]VA San Diego Healthcare System, San Diego, CA, United States

[6]Department of Psychiatry, University of California San Diego, San Diego, CA, United States

Corresponding Author:
Stephanie Aghamoosa, PhD
Department of Health Sciences & Research, College of Health Professions, Medical University of South Carolina, 77 President Street, Charleston, SC, United States

Abstract

Background: Many stroke survivors experience chronic cognitive deficits that, together with motor impairments, negatively impact well-being and daily function. Critically, poststroke cognitive deficits can interfere with successful engagement in rehabilitation programs aimed at improving participation in home and community activities. To address this issue, we developed a new telehealth intervention that integrates brief, functional cognitive rehabilitation (COG) into an existing upper extremity-focused, activity-based occupational therapy (OT) telerehabilitation program.

Objective: To describe the development and specification of a telehealth intervention (COG+OT) that integrates functional cognitive rehabilitation into activity-based stroke telerehabilitation.

Methods: COG+OT was developed using a multi-disciplinary, theory-informed process guided by principles of functional cognitive rehabilitation. It is reported according to the Template for Intervention Description and Replication checklist for rehabilitation (TIDieR-Rehab). The intervention was developed by a team with expertise in stroke rehabilitation from the fields of OT, neuropsychology, clinical psychology, and speech-language pathology. Cognitive rehabilitation content was adapted from an evidence-based compensatory strategy program and condensed into a brief "primer" format. The COG component was iteratively integrated into an existing upper-extremity, activity-based OT telerehabilitation program to support engagement and transfer of strategies to real-world functional tasks. In this paper, we describe the intervention structure, content, delivery, and tailoring procedures.

Results: COG+OT is an 8-week, 13-session telehealth intervention delivered in three phases: (1) cognitive strategy training ("primer"; sessions 1-4), (2) review and guided application of strategies within OT activities (sessions 5 - 10), and (3) integration of strategies to support generalization to everyday contexts (sessions 11-13). The intervention combines individualized, strategy-based cognitive rehabilitation with goal-oriented OT using a titrated approach. Patient-facing materials, structured session workflows, and therapist-guided metacognitive strategies support tailoring and real-world application.

Conclusions: COG+OT is a theory-informed, user-centered intervention that provides a structured approach for integrating functional cognitive rehabilitation into stroke telerehabilitation. This work establishes a well-specified, replicable intervention model and supports future feasibility, effectiveness, and implementation studies.

Trial Registration: ClinicalTrials.gov NCT06555302; https://clinicaltrials.gov/study/NCT06555302

(JMIR Rehabil Assist Technol 2026;13:e84627) doi:10.2196/84627

KEYWORDS

telerehabilitation; cognition; cognitive rehabilitation; occupational therapy; neuropsychology; stroke

XSL•FO
RenderX

Introduction

Poststroke cognitive deficits occur in up to 70% of stroke survivors, most commonly in the domains of memory, attention, and executive function [1]. These multifaceted cognitive deficits are consequential for stroke recovery, as they are associated with lower mood and quality of life, reduced participation in daily activities, and more protracted disability [2,3]. Cognitive rehabilitation interventions are effective at remediating cognitive dysfunction across many clinical conditions [4], and growing evidence indicates that functional cognitive rehabilitation, defined as strategy-based training aimed at enhancing real-world cognitive functioning, is particularly useful for improving everyday functioning poststroke [5]. However, cognitive rehabilitation is underutilized in stroke care due to prioritization of physical over psychosocial and cognitive recovery [6], limited access to cognitive rehabilitation services, lack of a well-defined replicable intervention model, and the burden of conventional multiweek treatment schedules. Furthermore, it is typically delivered as a stand-alone intervention, separate from other therapies in which cognitive strategies could be practiced and applied, creating a missed opportunity for integrated rehabilitation that enhances generalization of skills into daily life routines. With the increasing recognition of cognitive health as critical for functioning and quality of life, poststroke therapies combining physical and cognitive rehabilitation are an urgent need.

In prior work, we developed a telerehabilitation occupational therapy (OT) program that integrates upper-extremity task practice [7] with metacognitive strategy training (ie, teaching individuals to plan, monitor, and evaluate their own performance) based on the Cognitive Orientation to daily Occupational Performance approach [8]. In a recent study in adult stroke survivors, we showed that this activity-based telerehabilitation OT program improved paretic upper extremity motor skills, activity performance, self-efficacy, and skill generalization [9]. However, because this program is cognitively demanding, with metacognition components (ie, the process of thinking about one's own thinking and adjusting accordingly [10]) that require collaborative goal setting, self-reflection, error analysis, and problem solving, stroke survivors with cognitive deficits may struggle to fully engage and benefit. This challenge served as the motivation for integrating functional cognitive rehabilitation into this activity-based OT program.

This new intervention, called COG+OT, pairs a brief, functional cognitive rehabilitation "primer" (COG; this abbreviation specifically refers to the brief cognitive rehabilitation component of this intervention) with the previously developed activity-based OT program. The rationale for combining these approaches is based on evidence that compensatory cognitive strategies enhance the attention, metacognition, and goal management skills needed for therapy engagement [5], and cognitive strategies generalize most effectively when practiced within meaningful everyday activities [4]. The cognitive rehabiliation intervention we adapted was based on compensatory cognitive training approaches that have demonstrated benefits for cognition, functional capacity, and quality of life in randomized controlled trials including individuals with acquired brain injury and other neuropsychiatric conditions [11]. Our central hypothesis is that delivering COG before and alongside OT may enhance cognitive readiness, thereby supporting engagement in activity-based OT and yielding greater gains in real-world activity performance and participation.

We recently completed a single-arm pilot feasibility trial of 8-week COG+OT with 20 adults with chronic stroke and upper extremity paresis. Initial pilot testing supported the feasibility and acceptability of the intervention, as reported previously [12]. Given these findings, the purpose of the current paper is to describe the COG+OT intervention to support transparency, replication, and future intervention evaluation. The primary objective of this manuscript is to detail the conceptualization, design, and specification of the COG+OT intervention using systematic reporting via the Template for Intervention Description and Replication checklist for rehabilitation trials (TIDieR-Rehab) [13].

Methods

Ethical Considerations

The intervention was developed and tested in a pilot feasibility trial that was approved by the Medical University of South Carolina Institutional Review Board (Pro00136303) and registered with ClinicalTrials.gov (NCT06555302). All participants provided informed consent.

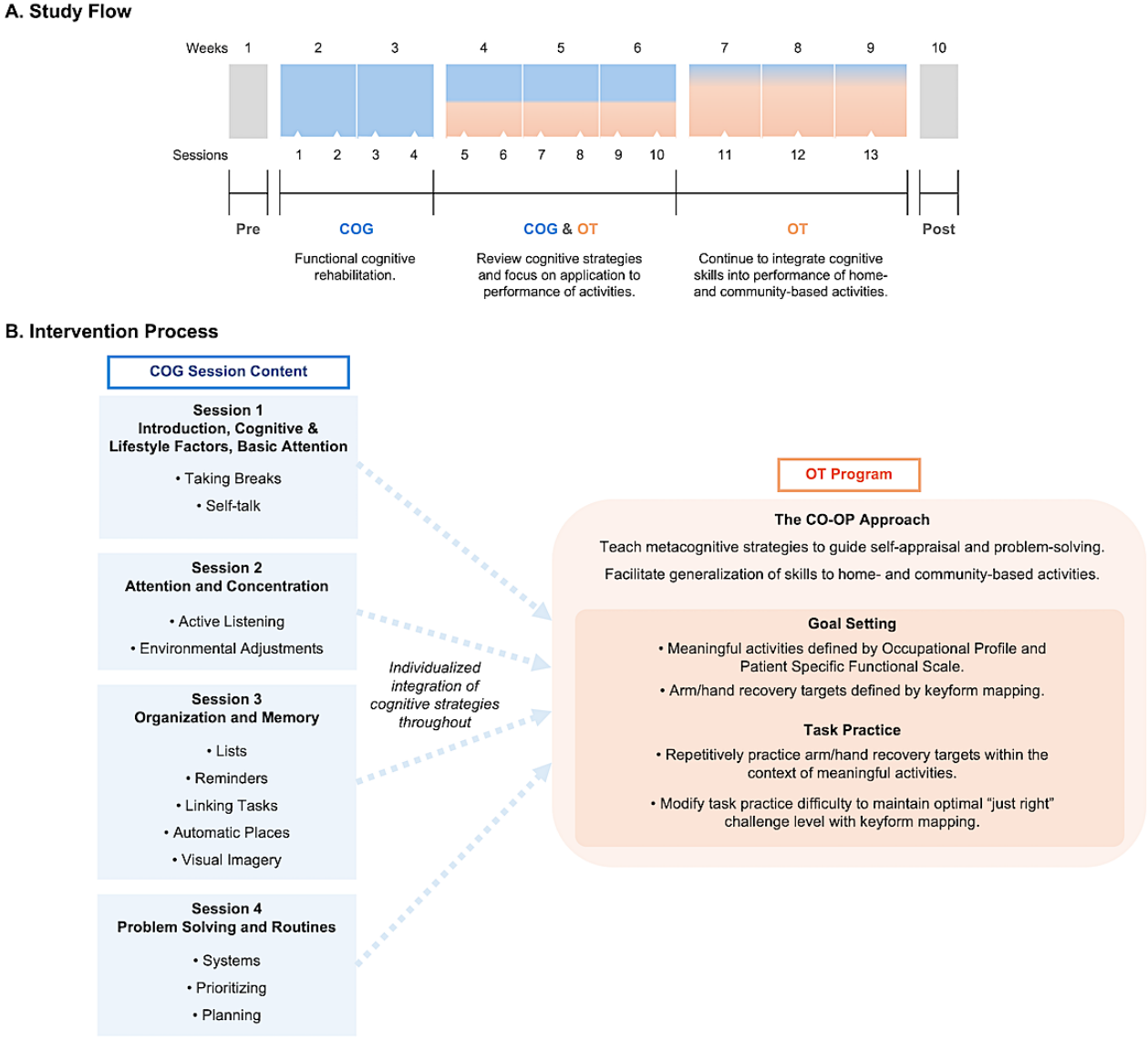
Intervention Development Process

Theoretical Rationale

The intervention embeds functional cognitive rehabiliation into an existing 9-session stroke telerehabilitation program [9] (OT) by using brief COG sessions as a "primer" to boost engagement in and benefit from OT (Figure 1A). Compensatory cognitive training was selected based on theoretical models emphasizing that functional cognitive improvement is best achieved through metacognitive strategy training, habit learning, environmental scaffolding (ie, external supports that facilitate performance), and contextualized practice rather than isolated restorative exercises [4,14]. These models highlight mechanisms such as teaching compensatory strategies (ie, working around deficits and capitalizing on intact abilities), enhancing self-regulatory and metacognitive processes, and applying strategies in meaningful real-world contexts to promote transfer. These principles align closely with a person-centered, goal-oriented, activity-based OT program.

XSL•FO
RenderX

Figure 1. Study flow and intervention overview. Panel A shows the procedures conducted during each week of the study, including outcomes assessments ("Pre" and "Post"), and the 3-phase intervention delivered across 13 sessions. Panel B lists the core topics covered in each of the 4 COG sessions (with additional optional topics such as calendar use available, when relevant), depicts the integration of these cognitive strategies into the pre-existing OT program, and provides an overview of the key components of the OT program. COG: cognitive rehabilitation; CO-OP: Cognitive Orientation to Daily Occupational Performance; OT: occupational therapy.



Relevance to the Existing OT Program

The existing OT program [9] is cognitively demanding, combining (1) upper-extremity task practice within the context of individualized activity goals identified using the American Occupational Therapy Association (AOTA) Occupational Profile [15] and the Patient-Specific Functional Scale [16] and (2) metacognitive strategy training based on the Cognitive Orientation to Daily Occupational Performance approach [8] (Figure 1B). Therapists encourage participants to use the "Goal-Plan-Do-Check" global cognitive strategy as a framework to plan, perform, and evaluate task performance. These elements require cognitive flexibility, self-monitoring, problem-solving, and sustained attention. Thus, the COG "primer" was intentionally designed to bolster functional cognition so participants could more effectively engage in the OT program. Specifically, COG session content focuses on learning and applying attention, organization, memory, and problem-solving strategies to facilitate performance of real-world tasks. Examples include planning a modified approach to challenging daily tasks, eliminating distractions to improve attention during functional activities, using reminders to implement learned strategies while completing activities, and engaging in a structured reflection process to self-analyze performance during the OT sessions.

Conceptual Intervention Design

Unlike conventional cognitive rehabilitation programs delivered over multiple weeks, we condensed the content into a brief format that was delivered using a titrated model to reduce participant burden and facilitate integration of cognitive strategies directly into OT. This design aimed to optimize generalization by first introducing strategies, then applying them during functional activities. Titration occurred across three phases (Figure 1A): (1) individualized cognitive strategy instruction (sessions 1-4), (2) brief review and application of strategies within activity-based OT (sessions 5-10), and (3)

XSL•FO
RenderX

continued integration of strategies into OT to promote generalization and habit formation (sessions 11-13). Personalized homework assigned at each session supported skill generalization to home and community environments.

Adaptation of COG Materials

COG materials were adapted from the 8-week Motivationally Enhanced Compensatory Cognitive Training for Mild Cognitive Impairmentprogram (M Huckans, et al; Motivationally Enhanced Compensatory Cognitive Training for Mild Cognitive Impairment [ME-CCT for MCI]; unpublished treatment manual; 2013), an iteration of evidence-based Compensatory Cognitive Training and Cognitive Symptom Management and Rehabilitation Therapy (EW Twamley, et al; Cognitive Symptom Management and Rehabilitation Therapy [CogSMART] for Traumatic Brain Injury; unpublished treatment manual; 2010). Materials for ME-CCT/CogSMART are available on the CogSMART website [17] and were adapted with permission for research use and acknowledgment of the original sources. These interventions emphasize compensatory cognitive strategies and habit learning and have been delivered via telehealth in prior research studies and clinical care settings [18-20]. Study team members (occupational therapists [PF, SG, JL, KR, MLW] and a neuropsychologist [SA]) reviewed the ME-CCT manual, selected stroke-relevant content, and condensed it into four 1-hour sessions (Figure 1B). Intervention delivery was designed to be individualized in terms of time spent on each cognitive strategy (based on relevance), focus on improving vs optimizing skills (based on impairment level), and application to activities/settings (based on context and goals).

The intervention was developed by a multidisciplinary rehabilitation team with expertise in cognitive rehabilitation and stroke telerehabilitation, including a clinical neuropsychologist (SA), occupational therapists (PF, SG, JL, KR, MLW), a clinical psychologist (LMM), and a speech-language pathologist (DSB). The neuropsychologist led development of the COG component, including strategy selection and adaptation (SA in consultation with EWT), while the OT team led integration of cognitive strategies into the existing activity-based telerehabilitation program. The speech-language pathologist guided adaptation of patient-facing materials to accommodate individuals with mild-to-moderate aphasia.

Results

The intervention development process resulted in COG+OT, an 8-week, 13-session telerehabilitation program delivered across three phases: (1) cognitive strategy training (sessions 1-4), (2) guided application of cognitive strategies within activity-based OT (sessions 5-10), and (3) continued integration and generalization of strategies during OT sessions (sessions 11-13; Figure 1). The completed TIDieR-Rehab checklist [13] for the COG+OT intervention is reported in Checklist 1, and details are provided in the following sections.

Personnel

COG+OT is designed to be delivered by licensed occupational therapists with training in stroke rehabilitation and telehealth. In this implementation, therapists were experienced in both domains and received structured training in the COG component and its integration with the OT program as well as ongoing supervision via weekly group case consultation with lead clinicians (SA, MLW).

Setting

COG+OT is designed for delivery via synchronous telehealth using a Health Insurance Portability and Accountability Act (HIPAA)–compliant videoconferencing platform (in this study, doxy.me), enabling participation from home and community environments. This intervention uses participants' real-world contexts by incorporating personally meaningful activities and readily available household spaces and materials into therapy sessions. Therapists ensure camera positioning to allow clear visibility of movements and task performance. Assistance and training with the videoconferencing technology is provided as needed. When appropriate, caregivers may assist with technology setup or safety monitoring. This home-based delivery model is intended to enhance ecological validity, accessibility, and generalization of cognitive and motor strategies to everyday activities.

Target Population

The COG+OT intervention is intended for adults with subacute and chronic stroke experiencing upper-extremity motor impairment and co-occurring cognitive challenges who are able to participate in telehealth-based rehabilitation. The COG+OT model is intended to be broadly applicable to stroke survivors who would benefit from integrated cognitive and activity-based rehabilitation and is therefore applicable to individuals with mild-to-moderate cognitive deficits in various domains (eg, attention, memory, executive function). The intervention is also designed to accommodate individuals with mild-to-moderate aphasia through the use of adapted, "aphasia-friendly" materials [21,22] and multimodal communication supports (developed in consultation with the speech-language pathologist [DSB]). Consideration for safety (eg, fall risk) and the potential need for caregiver assistance are incorporated into intervention delivery.

Intervention Materials

Measures Used to Guide Individualized Delivery of the Intervention

Occupational Profile

At the pretreatment assessment visit, therapists complete an AOTA Occupational Profile [15] via semistructured interview to identify participants' occupational history, patterns of daily living, values, and participation priorities. Insights obtained through the occupational profile are used to inform individualized goal setting and intervention planning. As participants complete the 4 initial COG sessions, the occupational profile is used to guide collaborative conversations to help the therapist gain a deeper understanding of participants'

XSL•FO
RenderX

routines, habits, and roles in relation to their cognitive functioning.

Patient-Specific Functional Scale

At the pretreatment assessment visit, therapists administer the Patient Specific Functional Scale [16] to identify each participant's activity goals. Participants are asked to generate up to 5 meaningful activities that have been made difficult due to upper extremity motor impairment since having a stroke. Participants then rate their ability to perform each activity on a 0‐10 scale, with higher numbers indicating better self-perceived performance. Throughout the OT program, sessions focus on optimizing participant performance in the context of these goals. Therapists regularly reference and revisit the goals with the participant when identifying target activities for the treatment sessions. This scale is readministered at the posttreatment assessment to assess change in self-perceived performance of individualized activity goals.

Upper Extremity Motor Assessment

Prior to beginning the phase-2 COG+OT sessions (at the end of session 4), the participant's hemiparetic arm/hand impairment is measured using the Fugl-Meyer Assessment of the Upper Extremity [23], which is the gold-standard assessment of poststroke arm/hand motor impairment. In prior work, we modified this measure for use in telerehabilitation [9], reducing it to 22 items that can be assessed virtually. Administration involves the therapist observing the participant's performance of 22 movements (ie, moving the hemiparetic arm/hand through a series of postures), rating each on a 3-point scale (0=unable to perform, 1=partially performed, 2=near normal performance). Items are summed to a total score between 0-44 points, with lower scores indicating more severe motor impairment. Using a procedure described in detail previously [7], the participant's Fugl-Meyer item performance ratings are mapped onto an item hierarchy (ranging from easy to difficult) to indicate person-specific arm/hand movements that are optimally challenging targets for task practice therapy. This process, termed "keyform recovery mapping" (or "recovery ruler") [24], is an emerging evidence-based method that operationalizes how item difficulty and performance are used to select individualized treatment targets. To monitor progress and refine individualized movement targets, the Fugl-Meyer Assessment is readministered and mapped to the keyform at the intervention midpoint (end of session 7) and at the end of phase-2 COG+OT sessions (end of session 10). To evaluate change in motor function, it is administered again at the posttreatment assessment.

Materials Used to Implement the Intervention

Session Tracking Notes

Therapists document session activities and participant progress using study-developed Session Tracking Notes in REDCap. The notes differ slightly in structure across the 3 intervention phases, but each session's note contains the following content: (1) Subjective (participant report and direct quotes), (2) Objective Session Content (checklist of completed content, documentation of functional task practice activities, and notes on activity grading per therapist report), (3) Assessment (checklists capturing therapists' ratings of participant response to the intervention and barriers to progress), and (4) Notes/Follow-Up Plan, including assigned home activity program (therapist narrative). The narrative portions of the Subjective, Assessment, and Plan sections reflect best-practice rehabilitation documentation procedures.

COG Session Manual

A manual was developed to guide therapists during COG sessions (sessions 1-4). The manual outlines session procedures and content to be covered, provides examples that could be shared with participants, and includes guidance for adjusting the complexity of information to support participant comprehension in the presence of cognitive and/or language impairment.

Patient-Facing COG Worksheets

For each COG session, we created 1‐2-page patient-facing worksheets, written at an eighth-grade reading level, that can be screen-shared, mailed, and/or emailed to the participant to guide discussion. Each session's worksheet contains (1) an introduction section that included descriptions of core concepts, definitions of key terms, and educational content followed by (2) a "Let's Practice" section containing real-life examples, space for self-reflection, and discussion prompts focused on functional application of cognitive strategies. Participants (and/or caregivers when present) are encouraged to take notes and regularly reference the worksheets outside of sessions.

Intervention Process

Overview

The intervention consists of 13 one-on-one telehealth sessions (45‐60 minutes each) over 8 weeks (Figure 1A). The schedule is 2 sessions per week for the first 5 weeks, followed by 1 session per week for the final 3 weeks. The reduction in session frequency during the latter phase is intended to promote greater independence and encourage skill generalization and transfer to real-world activities. To accommodate participant scheduling challenges, we allow make-up sessions to be added to any intervention week and for sessions to extend into one additional week (9 weeks total), if needed. The approximate proportion of each session allocated to COG vs OT activities varies over the 3 phases of the intervention (described in the following sections).

Sessions 1-4 (COG-Only)

Each COG session focuses on specific topics, listed in Figure 1B. Although the COG content follows a structured sequence, it is intentionally flexible to allow therapists to individualize the emphasis based on participant needs (eg, relevance, interest). Similarly, the use of cognitive strategies to compensate for deficits vs optimize intact abilities is personalized based on cognitive impairment level and contextual demands. Using the patient-facing worksheets via screensharing, therapists first introduce the session topic and review the associated compensatory strategies. The therapist and participant then collaboratively identify which strategies are most relevant to the participant's goals and cognitive challenges. Remaining session time focuses on practicing these strategies and discussing how they might be applied in daily activities. Personalized

XSL•FO
RenderX

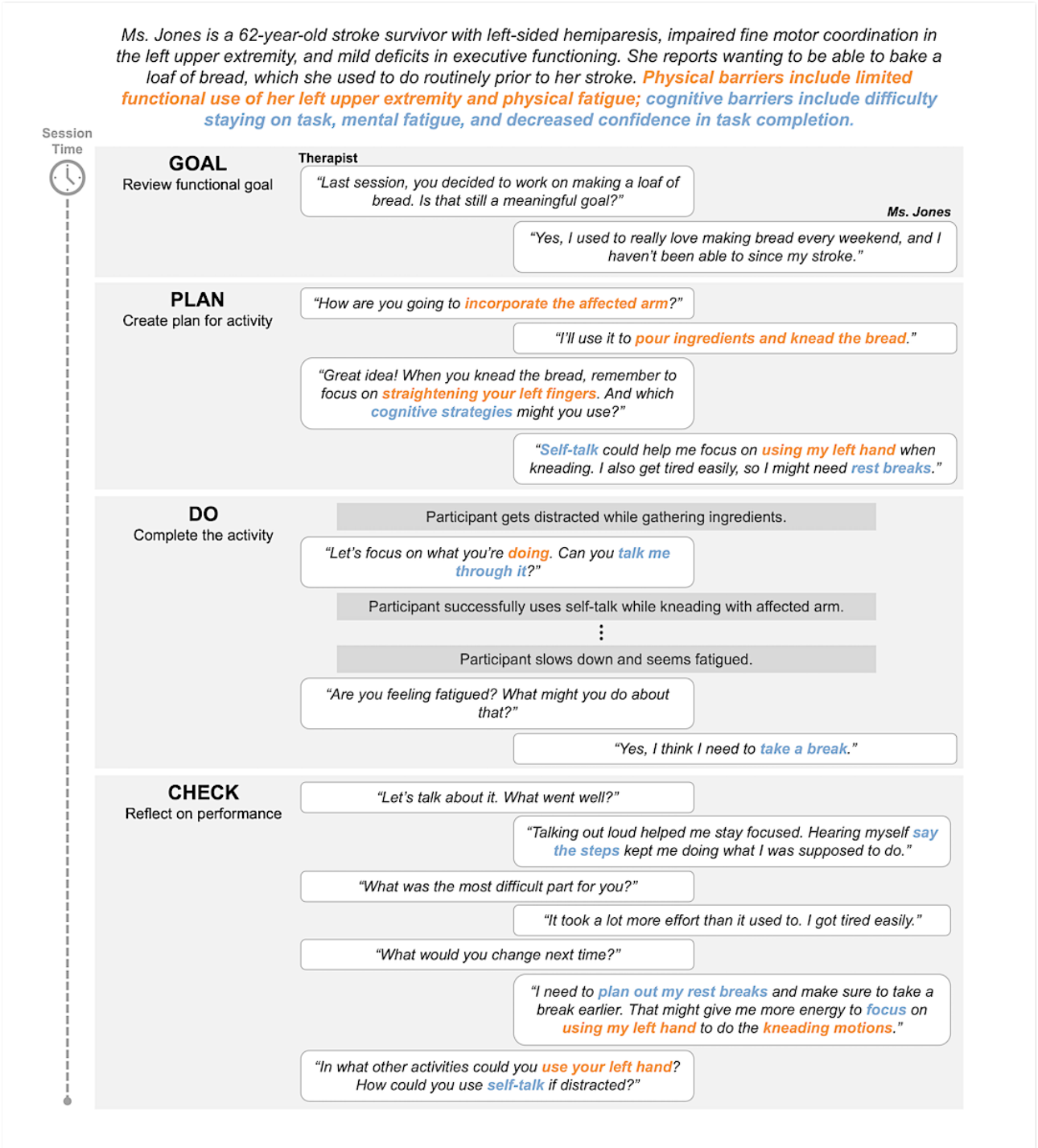
homework is assigned to reinforce real-world use of cognitive strategies and is reviewed with the therapist at the beginning of each following session. Throughout these sessions, therapists make note of strategies that emerge as most relevant or useful for each participant. Additional worksheets covering topics beyond the core sessions (eg, calendar use, additional memory strategies) are available and may be incorporated when relevant and if time permits. No new strategies are introduced after session 4.

Sessions 5–10 (COG+OT)

Each session begins with a brief review of the participant's use of learned cognitive strategies in their daily lives (10-15 minutes). The remainder of the session focuses on direct application of the strategies during OT activities. During the first COG+OT session (session 5), the telerehabilitation version of the Fugl-Meyer Assessment of the Upper Extremity keyform (described in *Intervention Materials*) is completed, and the therapist and participant review item performance scores to identify optimally challenging hemiparetic upper extremity movements to target in therapy. Together with the participant's goals from the Patient Specific Functional Scale and Occupational Profile, this information guides selection of meaningful in-home activities. Participants then practice these activities while being coached by the therapist using metacognitive strategy training and guided discovery based on the Cognitive Orientation to Daily Occupational Performance approach [8] (Figure 1B). Therapists use the "Goal-Plan-Do-Check" global cognitive strategy as a framework (ie, participants identify the task "goal," verbalize the "plan" to achieve the goal, attempt to "do" the task, and then self-reflect to "check" their own overall performance) while encouraging the participant to apply relevant cognitive strategies to support task completion. See Figure 2 for an example session.

XSL•FO
RenderX

Figure 2. Example COG+OT session (sessions 5-13). A representative scenario illustrates how individualized goals, motor targets, and cognitive strategies are integrated during task practice, with therapist-guided coaching supporting performance and reflection.



Sessions 11 - 13 (OT With Integrated COG)

Cognitive strategies are no longer explicitly reviewed during these sessions, but participants are encouraged to continue applying the strategies during OT functional task practice (using the same procedures as sessions 5 - 10). For example, a participant who had identified self-talk as a helpful strategy in earlier sessions would be coached to verbalize task steps during OT functional task practice to enhance attention and performance (see Figure 2). Therapists reinforce effective cognitive strategy use during target activities and facilitate generalization of new cognitive and motor skills to home and community contexts.

Intervention Fidelity

Multiple procedures are implemented to support consistent delivery of the intervention. All therapists receive training in both the COG component from the study neuropsychologist (SA) and in the OT program from the lead occupational therapist (MLW). A treatment manual was developed to standardize procedures, and standardized session workflows are used to guide intervention delivery and promote consistency across therapists. Intervention session activities and participant progress

XSL•FO
RenderX

are documented using Session Tracking Notes (described in *Intervention Materials*) within a secure data management platform (ie, REDCap). These data are used to track progress and guide subsequent sessions. Protocol deviations, barriers to progress, and adverse events are systematically recorded in REDCap and reviewed. If the planned content is not fully covered in the allotted time or when additional practice is needed, this is noted in REDCap and addressed in subsequent sessions to ensure each participant is exposed to all the core material.

Discussion

We developed COG+OT, a theory-informed intervention that integrates brief, functional cognitive rehabilitation into activity-based stroke telerehabilitation. This work addresses the well-documented challenge in poststroke care of siloed services for cognitive vs motor rehabilitation and lack of a well-defined multidisciplinary intervention protocol, which limits opportunities to apply cognitive strategies to meaningful activities and contexts. By embedding cognitive strategy training within meaningful, goal-directed activities, COG+OT is intended to support both engagement in rehabilitation and transfer of skills to everyday life.

An innovative design feature is the titrated integration of cognitive and motor components across 3 phases. The initial "primer" phase introduces compensatory cognitive strategies in a structured but individualized format, allowing participants to develop awareness and understanding of their cognitive challenges and strategies for addressing them. This is followed by guided application of these strategies within activity-based OT sessions and then continued integration to promote skill generalization into daily activities to form routines and habits. This progression aligns with contemporary models of functional cognitive rehabilitation, which emphasize metacognitive strategy training, contextualized practice, and repeated application across settings to facilitate generalization to real-world activities [4,14].

The use of a brief cognitive "primer" is intended to reduce burden while retaining core elements of evidence-based compensatory approaches, including personalization and strategy use in everyday contexts. Crucially, the COG sessions make cognitive strategies explicit for participants, complementing the implicit use of these strategies within the therapists' metacognition-based coaching. For example, therapists may introduce visual imagery during COG sessions as a strategy for planning and organization, then prompt participants to visualize each step of a task during OT sessions to support more accurate and efficient motor performance. Although the COG component draws from evidence-based materials (ie, ME-CCT/CogSMART), there are several other established cognitive rehabilitation programs that could have been considered. Approaches such as the Multicontext Approach, Goal Management Training, and other functional cognitive interventions [5,25,26] also include theoretically grounded elements (eg, metacognitive strategy training, guided practice, contextualized transfer) that may further strengthen components of COG+OT. Future iterations of the intervention may therefore benefit from selectively integrating features from these programs, particularly those that support generalization. In addition, explicit incorporation of motivational interviewing techniques may further enhance engagement, goal alignment, and sustained strategy use in daily life.

COG+OT was designed for real-world implementation through its telehealth delivery and use of participants' home environments to enhance ecological validity. Structured yet flexible materials support intervention fidelity while allowing for individualized tailoring. Additional features, including patient-facing worksheets, aphasia-friendly materials, and optional caregiver involvement increase inclusivity and make COG+OT amenable to implementation in a variety of settings (eg, outpatient clinics, community-based rehabilitation, hybrid telehealth models). Moreover, because compensatory cognitive strategies and metacognitive training are relevant across many neurologic and medical conditions, there is strong potential to adapt COG+OT for use in other populations (eg, individuals with traumatic brain injury, mild cognitive impairment, or other chronic conditions in which functional cognition supports daily participation) or clinical services (eg, speech therapy, physical therapy). Therefore, a key next step is to evaluate the feasibility, effectiveness, and implementation of COG+OT across diverse populations and settings. To support future dissemination, additional work is needed to develop standardized training and fidelity-monitoring procedures alongside evaluation of scalable delivery models and factors influencing real-world adoption.

Initial pilot testing supported the feasibility and acceptability of COG+OT telerehabilitation for stroke survivors, as reported previously [12]. The current manuscript extends that work by providing a detailed description of the intervention's theoretical basis, structure, materials, and delivery procedures to support replication and ongoing research. Analyses are underway to examine preliminary effects on cognitive, psychosocial, and participation-related outcomes, while qualitative feedback from participants and therapists will inform continued refinement of the intervention. Although aphasia-friendly adaptations were incorporated into COG+OT, additional research including greater representation of people with aphasia will be needed to evaluate their appropriateness. Future studies will also examine the added value of integrating functional cognitive rehabilitation into activity-based telerehabilitation, informing the design of future randomized controlled efficacy trials.

This paper focused on the systematic specification of the COG+OT intervention, including its theoretical basis, structure, and delivery procedures. Overall, COG+OT provides a structured and adaptable model for integrating functional cognitive rehabilitation into stroke telerehabilitation. This development work lays the groundwork for future refinement, evaluation, and broader implementation of COG+OT as an evidence-based, manualized intervention.

XSL•FO
RenderX

Acknowledgments

We sincerely thank our particicpants for dedicating their time and effort to this project. EWT gratefully acknowledges the support of a VA Rehabilitation Research, Development and Translation Research Career Scientist Award.

Generative AI Use

The generative AI tool Microsoft Copilot was used to assist with spelling, grammar, formatting, and language editing of the manuscript. All content was reviewed and approved by the authors, who take full responsibility for the final manuscript.

Funding

This project was supported, in part, by the National Institute of General Medical Sciences of the National Institutes of Health under grant number P30GM154630 and the South Carolina Clinical & Translation Research (SCTR) Institute, with an academic home at the Medical University of South Carolina NIH-NCATS grant number UL1 TR001450. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

No datasets were generated or analyzed as part of the current manuscript. This article describes the development and specification of the COG+OT intervention and does not report participant-level study results.

Authors' Contributions

Conceptualization: SA, MLW, LMM
Methodology: SA, MLW, LMM, EWT, DSB
Investigation: PF, JL, KR, SG, SA, MLW
Resources: MLW, SA, EWT, DSB
Writing – original draft: PF, SA, MLW
Writing – review & editing: KR, SG, DSB, LMM, EWT
Visualization: SA, PF, JL, MLW
Supervision: SA, MLW
Project administration: SA, MLW
Funding acquisition: SA, MLW, LMM.

Conflicts of Interest

None declared.

Checklist 1
TIDieR-Rehab checklist.
[DOCX File, 66 KB - rehab_v13i1e84627_app1.docx]

References

1. Rost NS, Brodtmann A, Pase MP, et al. Post-stroke cognitive impairment and dementia. Circ Res 2022 Apr 15;130(8):1252-1271. [doi: 10.1161/CIRCRESAHA.122.319951] [Medline: 35420911]
2. Stolwyk RJ, Mihaljcic T, Wong DK, Chapman JE, Rogers JM. Poststroke cognitive impairment negatively impacts activity and participation outcomes: a systematic review and meta-analysis. Stroke 2021 Jan;52(2):748-760. [doi: 10.1161/STROKEAHA.120.032215] [Medline: 33493048]
3. Stolwyk RJ, Mihaljcic T, Wong DK, Hernandez DR, Wolff B, Rogers JM. Post-stroke cognition is associated with stroke survivor quality of life and caregiver outcomes: a systematic review and meta-analysis. Neuropsychol Rev 2024 Dec;34(4):1235-1264. [doi: 10.1007/s11065-024-09635-5] [Medline: 38466357]
4. Cicerone KD, Goldin Y, Ganci K, et al. Evidence-based cognitive rehabilitation: systematic review of the literature from 2009 through 2014. Arch Phys Med Rehabil 2019 Aug;100(8):1515-1533. [doi: 10.1016/j.apmr.2019.02.011] [Medline: 30926291]
5. Zemsky C, Gillen G. Compensatory strategies to improve functional cognition post-stroke: a review for caregiver empowerment and burden reduction. Brain Sci 2025 Nov 30;15(12):41440092. [doi: 10.3390/brainsci15121297] [Medline: 41440092]
6. Bright FAS, Ibell-Roberts C, Featherstone K, et al. “Physical well-being is our top priority”: healthcare professionals’ challenges in supporting psychosocial well-being in stroke services. Health Expect 2024 Apr;27(2):e14016. [doi: 10.1111/hex.14016] [Medline: 38469645]

XSL•FO
RenderX

7. Woodbury ML, Anderson K, Finetto C, et al. Matching task difficulty to patient ability during task practice improves upper extremity motor skill after stroke: a proof-of-concept study. Arch Phys Med Rehabil 2016 Nov;97(11):1863-1871. [doi: 10.1016/j.apmr.2016.03.022] [Medline: 27117385]
8. Houldin A, McEwen SE, Howell MW, Polatajko HJ. The cognitive orientation to daily occupational performance approach and transfer: a scoping review. OTJR (Thorofare N J) 2018 Jul;38(3):157-172. [doi: 10.1177/1539449217736059] [Medline: 29083265]
9. Woodbury ML, Garner S, Hart E, et al. Preliminary effect of person-centered stroke telerehabilitation that combines Rasch analysis keyform mapping with cognitive strategy training. Arch Phys Med Rehabil 2025 Nov 26;S0003-9993(25):41314523. [doi: 10.1016/j.apmr.2025.11.009] [Medline: 41314523]
10. Sohlberg MM, Hamilton J, Turkstra LS. Transforming Cognitive Rehabilitation: Effective Instructional Methods: Guilford Publications; 2022.
11. Storzbach D, Twamley EW, Roost MS, et al. Compensatory cognitive training for Operation Enduring Freedom/Operation Iraqi Freedom/Operation New Dawn veterans with mild traumatic brain injury. J Head Trauma Rehabil 2017;32(1):16-24. [doi: 10.1097/HTR.0000000000000228] [Medline: 27022961]
12. Aghamoosa S, Rishe K, Laura J, et al. Functional cognitive rehabilitation as a primer to activity-based stroke telerehabilitation: feasibility, acceptability, and engagement. Brain Sci 2025 Nov 30;15(12):41440095. [doi: 10.3390/brainsci15121298] [Medline: 41440095]
13. Signal N, Gomes E, Olsen S, Alder G. Enhancing the reporting quality of rehabilitation interventions through an extension of the Template for Intervention Description and Replication (TIDieR): the TIDieR-Rehab checklist and supplementary manual. BMJ Open 2024 Nov 27;14(11):e084320. [doi: 10.1136/bmjopen-2024-084320] [Medline: 39608992]
14. Huckans M, Hutson L, Twamley E, Jak A, Kaye J, Storzbach D. Efficacy of cognitive rehabilitation therapies for mild cognitive impairment (MCI) in older adults: working toward a theoretical model and evidence-based interventions. Neuropsychol Rev 2013 Mar;23(1):63-80. [doi: 10.1007/s11065-013-9230-9] [Medline: 23471631]
15. AOTA Occupational Profile Template. Am J Occup Ther 2017;71(Supplement_2):7112420030p1. [doi: 10.5014/ajot.2017.716S12] [Medline: 29309016]
16. Sterling M, Brentnall D. Patient Specific Functional Scale. Aust J Physiother 2007;53(1):65. [doi: 10.1016/s0004-9514(07)70066-1] [Medline: 17326743]
17. CogSMART. URL: https://www.cogsmart.com/ [accessed 2026-10-07]
18. Lindamer L, Almklov E, Pittman JOE, et al. Multi-method study of the implementation of Cognitive Symptom Management and Rehabilitation Training (CogSMART) in real-world settings. BMC Health Serv Res 2022 Dec 17;22(1):1542. [doi: 10.1186/s12913-022-08941-z] [Medline: 36528588]
19. Stypulkowski K, Anquillare E, Twamley EW, Thayer RE. Feasibility of a telehealth compensatory cognitive training program for older adults with mild cognitive impairment. Clin Gerontol 2024;47(1):17-25. [doi: 10.1080/07317115.2023.2213694] [Medline: 37195804]
20. Shirley K, Campbell L, Maye J, et al. Feasibility of Motivationally Enhanced Compensatory Cognitive Training for older veterans with MCI. Innov Aging 2025 Dec 1;9(Supplement_2):3934. [doi: 10.1093/geroni/igaf122.3934]
21. King JM, Simmons-Mackie N. Communication supports and best practices: ensuring people with aphasia have an effective means of expressing needs and wishes. Top Lang Disord 2017 Dec;37(4):348. [doi: 10.1097/TLD.0000000000000135]
22. Shiggins C, Ryan B, O'Halloran R, et al. Towards the consistent inclusion of people with aphasia in stroke research irrespective of discipline. Arch Phys Med Rehabil 2022 Nov;103(11):2256-2263. [doi: 10.1016/j.apmr.2022.07.004] [Medline: 35905771]
23. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. Scand J Rehabil Med 1975;7(1):13-31. [Medline: 1135616]
24. Weaver JA, Papadimitriou C, McGuire A, et al. Co-creating and evaluating preimplementation outcomes of the recovery ruler for the coma recovery scale-revised: a concurrent mixed methods study. Arch Phys Med Rehabil 2025 Nov 10;S0003-9993(25):41224117. [doi: 10.1016/j.apmr.2025.10.019] [Medline: 41224117]
25. Rooha A, Sinchana S, Jacob NL, et al. Metacognitive strategies to optimise cognitive and metacognitive abilities among individuals with cognitive communication disorders and neurotypical adults: a scoping review. BMJ Open 2025 Nov 5;15(11):e099901. [doi: 10.1136/bmjopen-2025-099901] [Medline: 41198199]
26. Stamenova V, Levine B. Effectiveness of goal management training® in improving executive functions: a meta-analysis. Neuropsychol Rehabil 2019 Dec;29(10):1569-1599. [doi: 10.1080/09602011.2018.1438294] [Medline: 29540124]

Abbreviations

AOTA: American Occupational Therapy Association
COG: cognitive rehabilitation
CogSMART: Cognitive Symptom Management and Rehabilitation Therapy
HIPAA: Health Insurance Portability and Accountability Act
ME-CCT: Motivationally Enhanced Compensatory Cognitive Training

XSL•FO
RenderX

OT: occupational therapy
TIDieR-Rehab: Template for Intervention Description and Replication checklist for rehabilitation trials

Edited by J Sarvestan, S Munce; submitted 10.Oct.2025; peer-reviewed by L Mendes, L Sheehy, S Triscari; revised version received 29.Jul.2026; accepted 08.Sep.2026; published 27.Oct.2026.

Please cite as:
Finetto P, Woodbury ML, Laura J, Rishe K, Garner S, McTeague LM, Schwen Blackett D, Twamley EW, Aghamoosa S
Integrating Functional Cognitive Rehabilitation With Activity-Based Stroke Telerehabilitation: Intervention Design and Development Study
JMIR Rehabil Assist Technol 2026;13:e84627
URL: https://rehab.jmir.org/2026/1/e84627
doi:10.2196/84627

© Patricia Finetto, Michelle L Woodbury, Julianne Laura, Kelly Rishe, Stephanie Garner, Lisa M McTeague, Deena Schwen Blackett, Elizabeth W Twamley, Stephanie Aghamoosa. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telerehabilitation Following Stroke: Development of Training Content and Evaluation of an App-Based Training Program

Carina Ziller[1,2], MSc; Szabina Gäumann[1], MSc; Silya Lüscher[1], MSc; Nele Paulissen[1,2], MSc; Frank Behrendt[1,3], PhD; Zorica Suica[1,2], MSc; Björn Crüts[4], MSc; Luana Gamerschlag[4], BA; Katrin Parmar[1,5], PhD; Hans Ulrich Gerth[1,6], PhD; Leo H Bonati[1,7,8], PhD; Corina Schuster-Amft[1,3,9], PhD

[1]Research Department, Reha Rheinfelden, Salinenstrasse 98, Rheinfelden, Switzerland

[2]Physiotherapy Department, Reha Rheinfelden, Rheinfelden, Switzerland

[3]School of Engineering and Computer Science, Bern University of Applied Sciences, Biel, Switzerland

[4]Blended Clinic AI GmbH, Nuremberg, Germany

[5]Translational Imaging in Neurology (ThINk) Basel, Departments of Head, Spine and Neuromedicine and Biomedical Engineering, University Hospital, University of Basel, Basel, Switzerland

[6]Department of Medicine, University Hospital Münster, Münster, Germany

[7]Stroke Center and Department of Neurology, University Hospital Basel, Basel, Switzerland

[8]Department of Clinical Research, University of Basel, Basel, Switzerland

[9]Department for Sport, Exercise and Health, University of Basel, Basel, Switzerland

Corresponding Author:
Frank Behrendt, PhD
Research Department, Reha Rheinfelden, Salinenstrasse 98, Rheinfelden, Switzerland

Abstract

Background: Enhancing rehabilitation methods for patients with stroke is essential, particularly during the transition from inpatient to outpatient care. Digital applications are being developed to provide telerehabilitation programs. The existing virtual blended care platform Blended Clinic (Blended Clinic AI GmbH) offers app-based training for patients after a stroke and comprises 3 main components, including training, coaching, and monitoring.

Objective: This study assesses the usability and user experience of the novel Swiss Tele-Assisted Rehabilitation and Training (START) program within the Blended Clinic platform in patients after stroke and therapists.

Methods: The START program was developed within 3 workshops and an online survey. It contains 10 6-week exercise programs tailored to the levels of the modified Rankin Scale (mRS), 5 infographics, and 5 podcasts. Eight patients after stroke and 10 therapists took part in a single-center usability study. All participants were introduced to the Blended Clinic app and subsequently used it independently. The Blended Clinic platform, including the START program, was evaluated by both user groups based on the System Usability Scale (SUS) and the Mobile App Rating Scale (MARS-G). Additionally, feedback was collected, observations were documented, and program adherence metrics were calculated.

Results: The mean SUS scores were 87.2 (SD 10.8) for 8 patients and 83.3 (SD 11.3) for 10 therapists. The MARS-G scores were 3.9 (SD 0.5) for patients and 4.1 (SD 0.4) for therapists for categories A-D. User experience was rated 4.1 (SD 0.5) for patients, while device usability was 3.8 (SD 0.8) for patients and 4.2 (SD 0.5) for therapists. Adherence to the training schedule varied among patients (16.7%-80% of the planned sessions) and was rather low for many of the patients.

Conclusions: The START program, delivered via the Blended Clinic platform, was considered user-friendly and received good usability ratings from patients with stroke and therapists. Recommendations to enhance compliance are provided.

(JMIR Rehabil Assist Technol 2026;13:e77090) doi:10.2196/77090

KEYWORDS

telerehabilitation; stroke; app-based training; usability; blended care

Introduction

Stroke is a leading cause of acquired disability in adults worldwide, significantly impacting individuals' quality of life due to persistent motor and cognitive impairments [1]. Importantly, these impairments in patients with stroke are often accompanied by other issues like fatigue and depressive symptoms, significantly impacting quality of life and independence [2]. Stroke is a significant contributor to cognitive decline and dementia, being the second most common cause after Alzheimer disease [2,3].

XSL•FO
RenderX

Rehabilitation strategies for individuals after a stroke primarily focus on restoring impaired motor control to enhance the independence of those affected [4]. An interdisciplinary approach that considers individual needs is essential for the success of rehabilitation and the overall quality of life for patients [5]. Thus, the advancement of options for improved, sustainable, and effective follow-up care in stroke rehabilitation is urgently needed [6]. Technological innovations can play an essential role in efforts to maintain the motor and cognitive abilities of affected individuals at the highest possible level. By integrating new technologies into rehabilitation programs, recovery outcomes can be enhanced, while early intervention and rehabilitation are crucial for improving outcomes and enhancing recovery after a stroke [7].

In addition to face-to-face rehabilitation, telerehabilitation has emerged due to advancing digitalization in the health care sector. It has been found that telerehabilitation in physiotherapy is comparable to in-person rehabilitation for conditions such as stroke [8,9]. In line with that, studies have shown that telerehabilitation can effectively deliver therapeutic interventions, maintain patient engagement, and achieve similar outcomes to traditional therapy, making it a valuable option for patients who may have difficulty attending in-person sessions [10].

Various mobile health apps have been developed for patients after a stroke, mostly focusing on upper extremity function, secondary prevention, or language skills [11]. Additionally, according to a review by Rintala et al [12], most studies investigating these apps focused on patients with chronic stroke, who are mildly impaired [13]. Although existing mobile health apps have shown positive effects on various stroke-related impairments, a substantial gap remains in the development of apps specifically targeting goal-oriented exercises and also integrating educational components [14].

The newly developed Blended Clinic platform (Blended Clinic AI GmbH) consists of a mobile app for patients (Blended Clinic, Version 1.0.1 [iOS]; 1.3 [Android]), a web-based portal for therapists, and a database. The system is designed to provide home-based exercise programs for different patient groups by integrating telerehabilitation with a blended care approach. Blended care itself refers to a health care approach that combines traditional in-person services with digital or remote care options. It aims to provide a more flexible and comprehensive approach to health care delivery, enabling diverse patient populations, including individuals after stroke in the subacute and chronic phases, to receive care through multiple modalities [15]. To date, an earlier iteration of a training program for patients after a stroke, in conjunction with a previous version of the Blended Clinic platform, has been evaluated for user interaction with the app and for acceptability among patients with chronic stroke [16]. It was found to be highly acceptable, with strong user engagement, good task adherence, and small clinical improvements in quality of life [16].

Based on these positive findings, the new Swiss Tele-Assisted Rehabilitation and Training (START) program was developed. It builds on the same core technology but extends the intervention and addresses the research gap in 2 ways: first, by further developing the content with a stronger focus on the needs of patients with moderate-to-severe impairments; and second, by evaluating its use in a new clinical context, the subacute phase of stroke rehabilitation, particularly the transition phase from inpatient to outpatient rehabilitation.

The aim of the present study was to describe the development process of the new START program and to assess the perceptions of both therapists and patients after a stroke regarding the usability and functionality of certain components of the START program, provided via the updated Blended Clinic platform, during test sessions.

Methods

Hardware and Software Components of the Blended Clinic Platform

The Blended Clinic platform consists of a web-based portal for therapists and a mobile app for patients. It is intended to be installed on the mobile devices of patients and comprises components for training, coaching, and monitoring. The content for these components is provided by the respective program modules, such as START, which was developed and tested within this project.

1. Training: in conjunction with a targeted training program, exercises can be performed daily, individually, and to the necessary and feasible extent. Through the app, patients are provided with a personal training plan tailored to their individual impairments and rehabilitation goals, including exercise videos. The exercise content can be adjusted to the achieved functional level.
2. Coaching: the app ensures that individual training is supported by a personal therapist to maintain high motivation levels. Patients' questions about training and general topics related to stroke can be promptly addressed, both on-site and primarily through a digital chat function, with therapists available daily during working hours on weekdays.
3. Monitoring: the app also facilitates the collection of health-related, measurable parameters to support the ongoing individual optimization of the training plan. In this study, patient-reported outcome measures (PROMs), self-conducted blood pressure measurements, and activity tracking with a wrist-worn device (GENEActiv, Activinsights) were applied.

The START Program

The main objective of START was to create a training program to be used with the Blended Clinic app specifically designed for patients with moderate to severe motor impairments during the early subacute stages of rehabilitation. We developed the START content in line with recent exercise recommendations [17-20] and stroke guidelines [21,22]. While these guidelines offer valuable recommendations, they do not specify explicit exercise instructions or dosages.

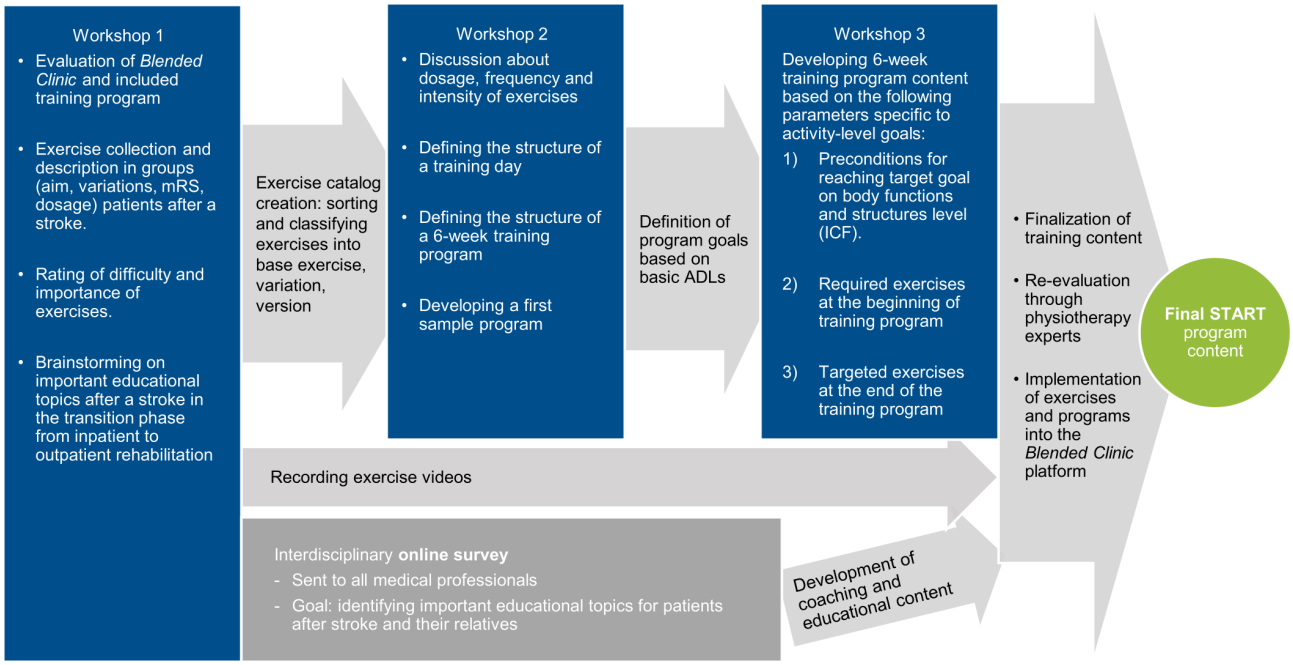
Three workshops were performed with a group of 7 clinicians, including 2 clinical physiotherapy neurological experts with over 25 years of professional experience and one research

XSL•FO
RenderX

physiotherapist with a doctoral degree, 3 clinical physiotherapists with at least 3 years of neurological experience, and 1 sports scientist. The participating therapists served as clinical instructors in the clinic and, in part, as lecturers and researchers.

In these workshops, important exercises and exercise variations for patients after stroke were identified and classified into different disability levels based on the modified Rankin Scale (mRS) [23,24] scores 1-5 (Figure 1). Exercises were further categorized into lower extremity, trunk exercises, balance and coordination exercises, and breathing exercises. In total, this resulted in the creation of 224 new exercises, resulting in a pool of around 300 exercises. In addition, for the presentation within the Blended Clinic platform, an introductory text outlining the aim of each exercise and a video instruction demonstrating how to perform the exercise were developed. Please refer to the supporting information for further details regarding the exercises (Multimedia Appendix 1).

Figure 1. Development process of the Swiss Tele-Assisted Rehabilitation and Training (START) program content and the main focus of each workshop. ADL: activities of daily living; ICF: International Classification of Functioning, Disability and Health; mRS: modified Rankin Scale; START: Swiss Tele-Assisted Rehabilitation and Training



Patient exercises were compiled to create 10 6-week programs, each with an overarching goal. The activity goals of such a program were based on abilities that are necessary to perform activities of daily living (ADL). The experts identified key components necessary for establishing individualized program goals, such as the ability to transfer to the toilet. For instance, a target exercise could involve standing for at least 60 seconds while engaging in arm activities. For this, a preliminary exercise might consist of "standing independently for 30 seconds without any arm activity."

Within each 6-week program, exercises progressed in terms of the number of sets, number of repetitions, and difficulty level. The exercise sessions, each lasting around 15 minutes and conducted daily, consisted of 3 phases, including a warm-up that typically involved a breathing or sensory stimulation exercise; a main component focusing on balance, strength, or endurance; and a cool-down that incorporated flexibility exercises. Additionally, care was taken to alternate strength, endurance, and balance training each day. Patients received information regarding training safety, content, meaning, and purpose of each exercise through the app.

In addition to the training programs, informational and educational materials for patients were developed. Relevant topics were identified by conducting an in-house interdisciplinary survey using the Findmind tool (Findmind). Based on the results from 90 health care professionals, several topics were identified, and content was developed. Podcasts were recorded covering the topics of fatigue, sleep, training, nutrition, and psychological well-being, along with the creation of 5 infographics. In addition to the regular motivational and informational messages sent via chat, this content constitutes an essential component of patient coaching and education and the blended care approach.

The START program was designed to be introduced in an inpatient rehabilitation setting and continue during the transition phase from inpatient to outpatient rehabilitation and then continue in the community. In the context of this usability study, not all the described content was used and investigated; rather, our focus was limited to an inpatient setting and specific components of START provided via the Blended Clinic platform.

Evaluation of Usability and Functionality of START

Participants

The evaluation of the usability and functionality of the app-based training was conducted as a single-center study primarily with patients after a stroke in their subacute stage and secondarily with therapists. The study took place in a neurological

XSL•FO
RenderX

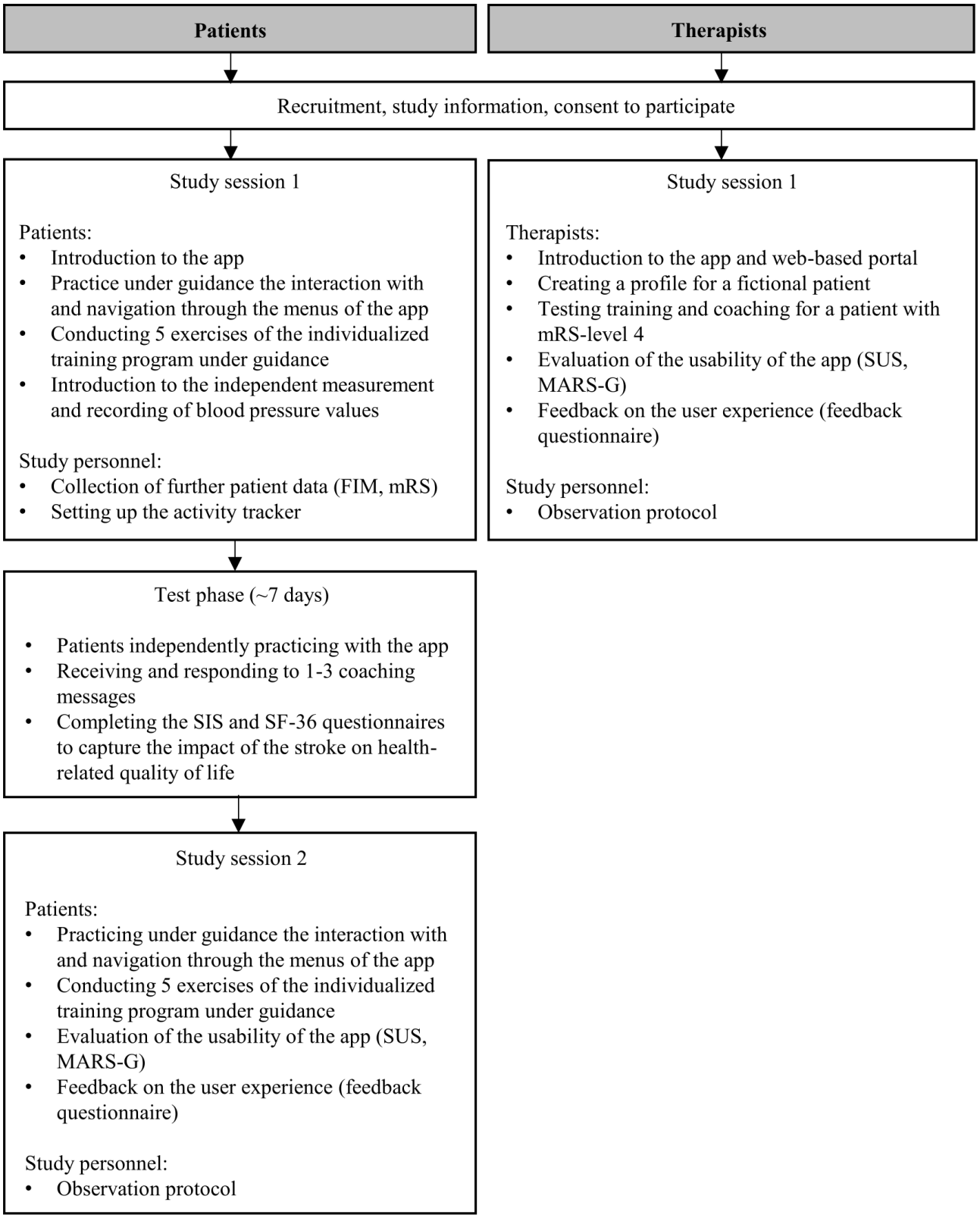
rehabilitation clinic in northwestern Switzerland. We planned to include 10 patients and 10 therapists, starting in December 2023. Inpatients and outpatients attending the clinic's day center were recruited by screening internal patient lists against the predefined eligibility criteria. The criteria comprised age ≥18 years, mRS 3-5, no severe aphasia, severe visual, neurological, cardiopulmonary, psychiatric, or orthopedic impairments that would prevent a patient from following the instructions, and no recent surgery or fractures. Therapists were included in the study if they held at least a bachelor's degree in physiotherapy, occupational therapy, speech and language therapy, or sports and exercise science, and if they had at least one year of experience in stroke rehabilitation. Therapists were invited via email to participate in the study.

Procedures

For the course of the study, please refer to the study flow chart (Figure 2). As demonstrated here, the therapists participated in a single session, whereas patients underwent 2 sessions, with a one-week testing phase in between.

XSL•FO
RenderX

Figure 2. Study flowchart. FIM: Functional Independence Measure, MARS-G: Mobile App Rating Scale; mRS: modified Rankin Scale; SF-36: 36-Item Short Form Survey Instrument; SIS: Stroke Impact Scale; SUS: System Usability Scale.



Patients had the opportunity to familiarize themselves with the app under supervision. During the introductory session with the study personnel, patients were instructed in the exercises assigned according to their mRS score and were trained in safe exercise execution. Besides, patients were also instructed on blood pressure measurements and on how to complete 2 questionnaires through the app (Stroke Impact Scale [SIS; 25,26], 36-Item Short Form Survey Instrument [SF-36; 27,28]). Patients subsequently tested the functionality of the app independently for a duration of one week. Before starting a daily exercise session, they received an automatic message reminding them to pay attention to their own safety during training. Furthermore, patients were provided with 2 information sheets, one outlining the criteria for discontinuing the exercises (eg,

XSL•FO
RenderX

feeling unwell) and the other containing general information on the training environment (eg, ensuring that a table and chair were available for support). Patients wore an activity tracker (GENEActiv, Activinsights) for the week. During the second session following the testing phase, patients were asked to perform predefined tasks within the app under the guidance of the study personnel, for instance, to follow exercise instructions and to write a message to the coach. Patients were observed to document this process, to evaluate their performance, and thus the comprehensibility of the necessary interactive steps within the app.

Therapists were also introduced to the use of the app and had the opportunity to test it at the beginning of the session. They subsequently were instructed to perform specific tasks on the web-based dashboard platform, such as selecting and sending a questionnaire. Their performance was also observed and documented. Finally, the therapists themselves evaluated the usability of the system.

Outcome Measures

For the subsequent assessment of usability, the System Usability Scale (SUS) [29], the German version of Mobile App Rating Scale (MARS-G) [30-32] and a customized feedback questionnaire (Multimedia Appendix 2) were used.

The widely used SUS is a questionnaire that assesses the usability of a system through a 10-item survey. Each item is rated on a 5-point Likert scale, allowing users to express their level of agreement or disagreement with various statements about the system.

MARS-G assesses the multidimensional quality of mobile health apps and comprises 4 components, including engagement (Section A), functionality (Section B), esthetics (Section C), and information quality (Section D). Each domain is rated on a 5-point Likert scale, where 1 indicates “insufficient,” 2 denotes “poor,” 3 represents “acceptable,” 4 signifies “good,” and 5 reflects “excellent.” Additionally, the MARS-G includes 2 sections for the subjective evaluation of characteristics within the domains of subjective app quality (Section E) and perceived impact of the app (Section F). The German version of MARS-G that we used has been supplemented with a further section (Section T) that focuses on the therapeutic benefits associated with the app [32]. For all sections, the maximum achievable score is 5 points [32]. The therapists were asked to complete the entire MARS-G, while the patients were only provided with Sections A, B, C, D, E, and F for evaluation.

The training experiences with and the usability of the app were assessed using a self-constructed, standardized feedback questionnaire (Multimedia Appendix 2). This questionnaire used a 5-point scale and included specific questions regarding the app and was pilot tested with one patient. The feedback questionnaire aimed to provide deeper insights into the participants’ opinions about the training platform. For patients, the questionnaire contained 8 questions related to user experience, along with a section comprising 3 questions regarding the use of the blood pressure monitor, activity tracker, and smartphones or tablets. In contrast, therapists completed a questionnaire consisting of 9 questions about user experience.

During the study sessions for patients and therapists, the study personnel also completed an observation protocol (see Results for task details). This protocol facilitated the evaluation of the usability and user experience for both patients and therapists by an external observer, assessing the extent to which the app is useful and applicable, as well as identifying any existing issues. The study personnel also evaluated the emotional reactions by observation of the patients while they were navigating through the app and while performing the exercises, using the following categories (present and not present): attention, joy, curiosity, astonishment, boredom, nervousness, confusion, anger, and rejection.

Data Analysis

The adherence of the participating patients to the START program protocol was derived from the execution of the planned training sessions, the number of blood pressure measurements, the completion of the questionnaires, and the duration of activity tracker usage. For all parameters, adherence was assessed in relation (ratio) to the intended use, with the following classifications defined: adherence of ≥75% of the planned activities was categorized as high adherence, 50% - 74% as moderate adherence, 25% - 49% as low adherence, and <25% as no adherence [33]. In this study, the tracker was used solely to assess whether the patients would tolerate wearing it, with a view to planning further studies and monitoring activity levels over a longer period.

- Training: number of completed training sessions versus planned training sessions. Patients received an exercise program consisting of 5 exercises each day. An adherent training day was defined as having performed at least 3 out of 5 exercises, quantified by spending a minimum of 20 seconds and a maximum of 5 minutes for each exercise on the app screen.
- Monitoring:
  - Blood pressure measurements: patients were instructed to conduct 3 blood pressure measurements each morning. Days on which the respective patient entered data themselves (eg, blood pressure) were classified as adherent.
  - Activity tracker: an adherent day was defined as a day on which the activity tracker was worn for a minimum of 5 hours on the first and last day or for at least 10 hours on each intervention day, consistent with established practices [34]. In this study, the focus was solely on understanding whether the trackers were accepted and worn, in preparation for a larger study.

All data were analyzed descriptively using JASP (version 0.18.0, The Jasp Team), R Studio (version 2023.06.0; Posit PBC) and Microsoft Excel (Excel 2019; Microsoft Corporation).

Ethical Considerations

The study received approval from the independent ethics committee of Northwestern and Central Switzerland (EKNZ, reference number: 2023 - 01783). All participants provided written informed consent and all participant data were de-identified. The participating therapists received compensation for participation of approximately $190.

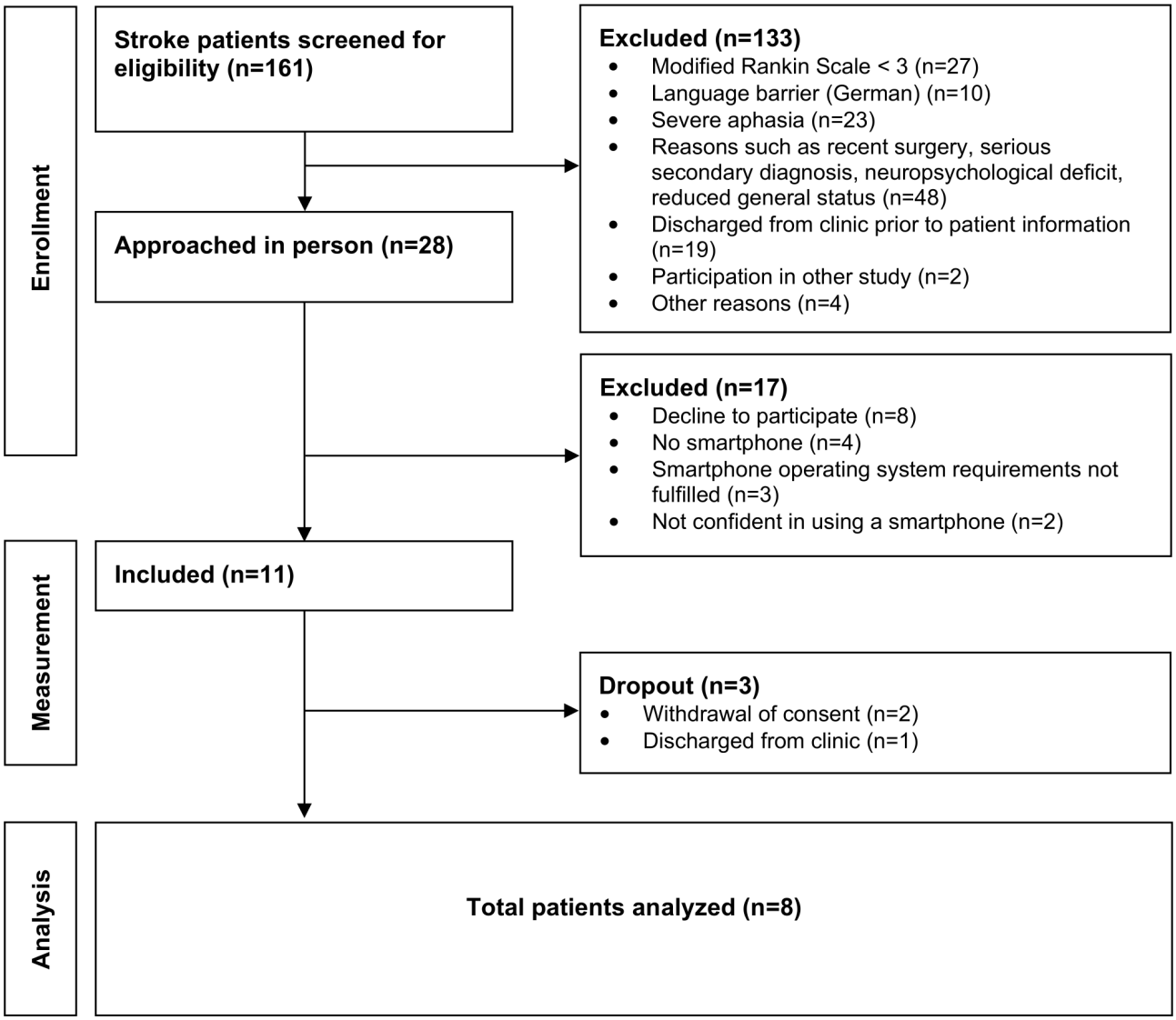
XSL•FO
RenderX

Results

Overview

From December 2023 to March 2024, a total of 161 patients with stroke were screened for eligibility (Figure 3). Eight patients (6 females) completed the study. The patient group had a mean age of 61.5 (SD 12.7) years with an average duration since the stroke of 180 days (265 days, range: 15-618 days) and a Functional Independence Measure (FIM) [24,35] score of 93.8 (SD 12.8; see Table 1).

Figure 3. Study patient flowchart.



XSL•FO
RenderX

Table . Patient characteristics.

Patient demographics and clinical characteristics	Patient number							
	1	2	3	4	5	6	7	8
Sex	Female	Male	Female	Female	Female	Female	Male	Female
Age (years)	77	57	72	74	56	59	38	59
Stroke type	Ischemic	Ischemic	Ischemic	Ischemic	Ischemic	Hemorrhagic	Hemorrhagic	Ischemic
Time since stroke (days)	26	58	28	15	48	44	618	600
FIM[a] total score (18-126)	96	103	119	90	84	90	91	77
FIM-motor (13-91)	65	75	86	64	59	65	68	—[b]
FIM-cognitive (5-35)	31	28	33	26	25	25	23	—
Modified Rankin Scale (0-5)	4	4	3	3	4	3	3	4
Smartphone operating system	iOS	iOS	Android	Android	iOS	Android	iOS	iOS

[a]FIM: Functional Independence Measure.

[b]Not available.

Ten therapists (8 females) with a mean age of 35.3 (SD 11.2) were enrolled. They were mostly physiotherapists (n=7) and occupational therapists (n=3).

Outcome Measures

The patients rated the usability of the app, including training content and monitoring, with an average SUS score of 86.6 out of 100, which is slightly higher than the score of the therapists, who reported a score of 83.3 out of 100 (Tables 2 and 3).

Table . Results from the System Usability Scale for all patients.

SUS score	Patient number								Mean (SD)	Median (IQR)
	1	2	3	4	5	6	7	8		
SUS[a] (total score)	97.5	87.5	80	100	95	70	87.5	75	86.6 (10.9)	87.5 (20.6)
SUS[a] (per item)										
Item 1	5	5	4	5	5	3	4	5	4.5 (0.8)	5 (1)
Item 2	2	2	2	1	1	3	1	5	2.1 (1.4)	2 (1.8)
Item 3	5	4	5	5	5	3	5	5	4.6 (0.7)	5 (0.8)
Item 4	1	1	3	1	1	1	1	1	1.3 (0.7)	1 (0)
Item 5	5	4	5	5	5	4	3	3	4.3 (0.9)	4.5 (1.8)
Item 6	1	1	2	1	1	3	1	5	1.9 (1.5)	1 (1.8)
Item 7	5	5	5	5	3	5	3	5	4.5 (0.9)	5 (1.5)
Item 8	1	2	1	1	1	1	1	1	1.1 (0.4)	1 (0)
Item 9	5	5	4	5	5	3	5	5	4.6 (0.7)	5 (0.8)
Item 10	1	2	3	1	1	2	1	1	1.5 (0.8)	1 (1)

[a]SUS: System Usability Scale.

XSL•FO
RenderX

Table . Results from the System Usability Scale (SUS) for all therapists.

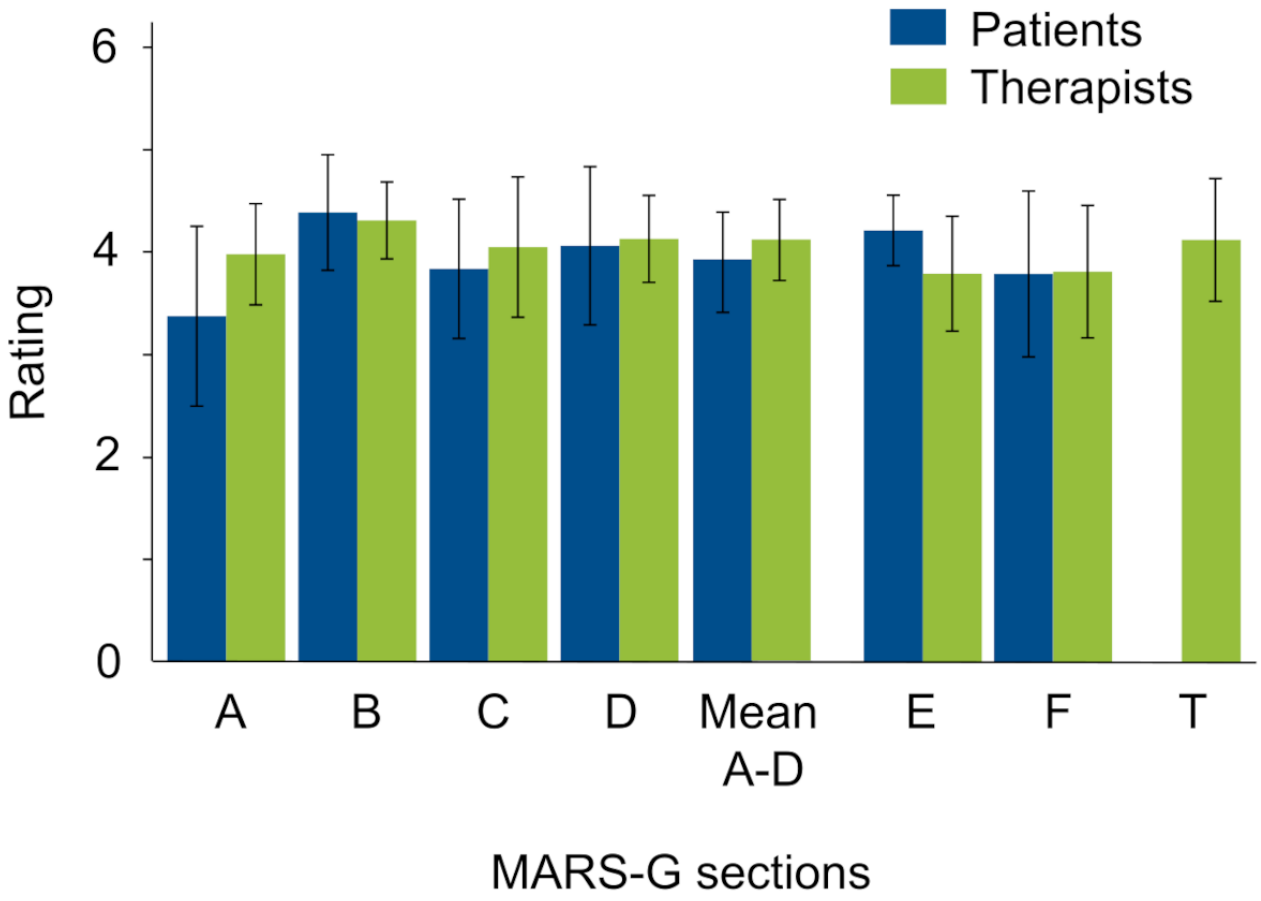
SUS[a] score	Therapist number										Mean (SD)	Median (IQR)
	1	2	3	4	5	6	7	8	9	10		
SUS[a] (total score)	95	97.5	62.5	90	90	80	77.5	85	65	90	83.3 (12.0)	87.5 (16.9)
SUS[a] (per item)												
Item 1	5	5	4	5	4	3	4	3	3	4	4.0 (0.8)	4 (2)
Item 2	1	1	2	1	1	1	2	1	1	1	1.2 (0.4)	1 (0.3)
Item 3	5	5	3	4	5	4	4	5	2	4	4.1 (0.9)	4 (1.3)
Item 4	1	1	1	1	1	1	1	1	3	1	1.2 (0.6)	1 (0)
Item 5	4	5	5	4	4	4	4	3	5	5	4.3 (0.7)	4 (1)
Item 6	1	1	4	1	2	2	1	1	1	2	1.6 (1.0)	1 (1)
Item 7	5	5	5	4	5	5	4	4	3	5	4.5 (0.7)	5 (1)
Item 8	1	1	3	1	1	2	2	1	1	1	1.4 (0.7)	1 (1)
Item 9	4	4	2	4	4	4	3	4	2	4	3.5 (0.8)	4 (4)
Item 10	1	1	4	1	1	2	2	1	3	1	1.7 (1.1)	1 (1)

[a]SUS: System Usability Scale.

For the MARS-G, both patients and therapists generally rated the user experience and usability of the app as acceptable to good, with the mean scores of Sections A-D being 3.9 for patients and 4.1 for therapists, as well as from the mean scores of the individual sections (Figure 4). The therapists rated no section below 4.0 points, thus categorizing all sections as good. Notably, they tended to rate the app slightly higher than the patients, particularly in the sections “Engagement” and “Aesthetics.” The section on “Functionality” received the highest ratings from patients and therapists. A table with the scores from patients and therapists is provided in the supporting information (Multimedia Appendix 3).

XSL•FO
RenderX

Figure 4. The German version of Mobile App Rating Scale (MARS-G) ratings of patients and therapists. A: engagement; B: functionality; C: esthetics; D information quality; E subjective quality; F: perceived impact; MARS-G: German version of Mobile App Rating Scale; T: therapeutic benefits.



The results of the feedback questionnaire indicated a generally positive perception of the app and the devices. The patients had mixed opinions regarding the handling of the devices in terms of self-monitoring blood pressure and wearing the activity tracker. The latter was partly perceived as somewhat bothersome. The operation of the app on the tablet or smartphone was deemed to be relatively straightforward.

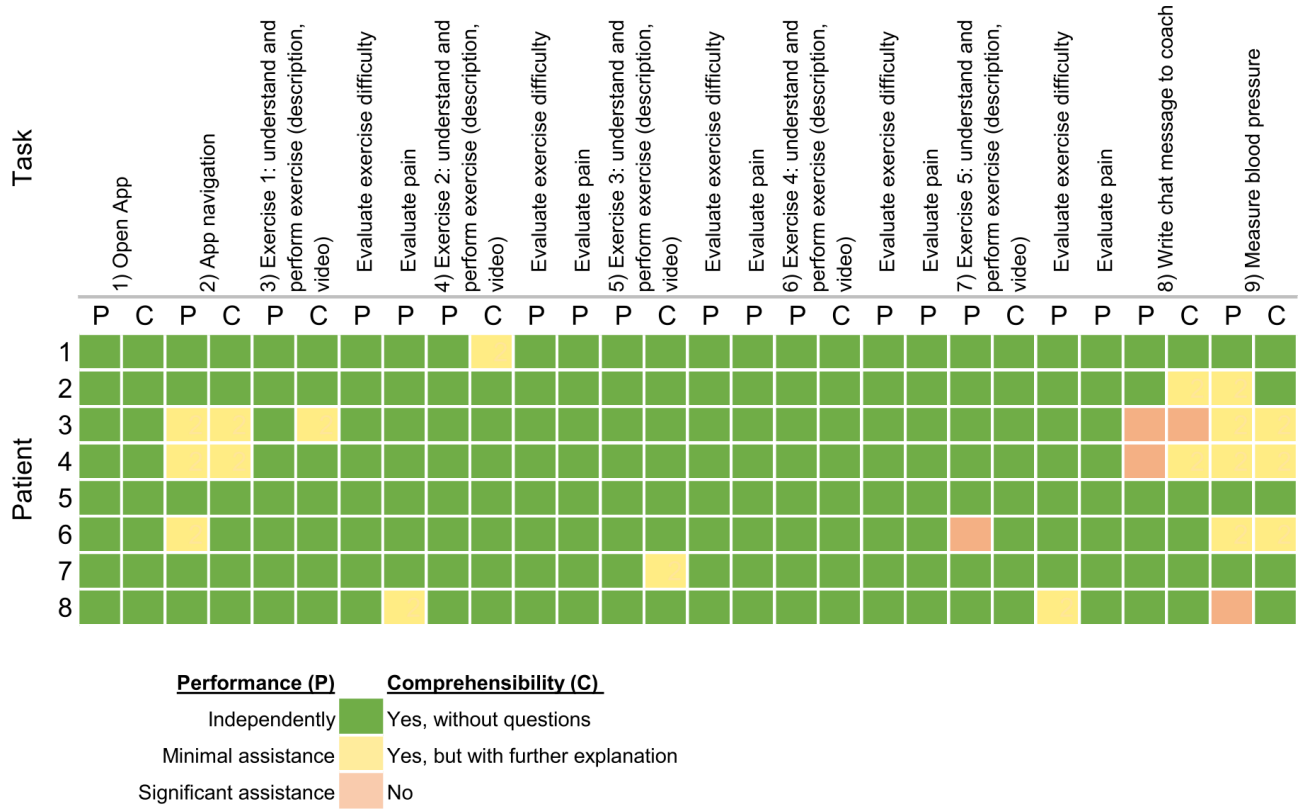
Among the therapists, there was a generally positive evaluation of the system. The question regarding the self-explanatory nature of the web-based dashboard for therapists received the lowest rating, while the question regarding technical functionality received the highest rating (Multimedia Appendix 4 for patients and therapists). Overall, all involved therapists expressed a desire to incorporate the app into their therapy planning.

Observation Protocols

Some patients encountered minor challenges in navigating the app's menus. A few needed substantial assistance in using the chat function to communicate with the coach or had follow-up questions. Patients required the most assistance with blood pressure measurement, with the majority needing at least some help during the measurement process (Figure 5). The emotions of attention, joy, curiosity, and astonishment were repeatedly mentioned. Except for one individual, all patients were engaged and attentive. Two-thirds of the patients exhibited joy throughout the observation. Curiosity and amazement were also noted in approximately half of the patients. In contrast, emotions such as rejection, anger, confusion, and nervousness were rarely observed, with 2 patients exhibiting these feelings.

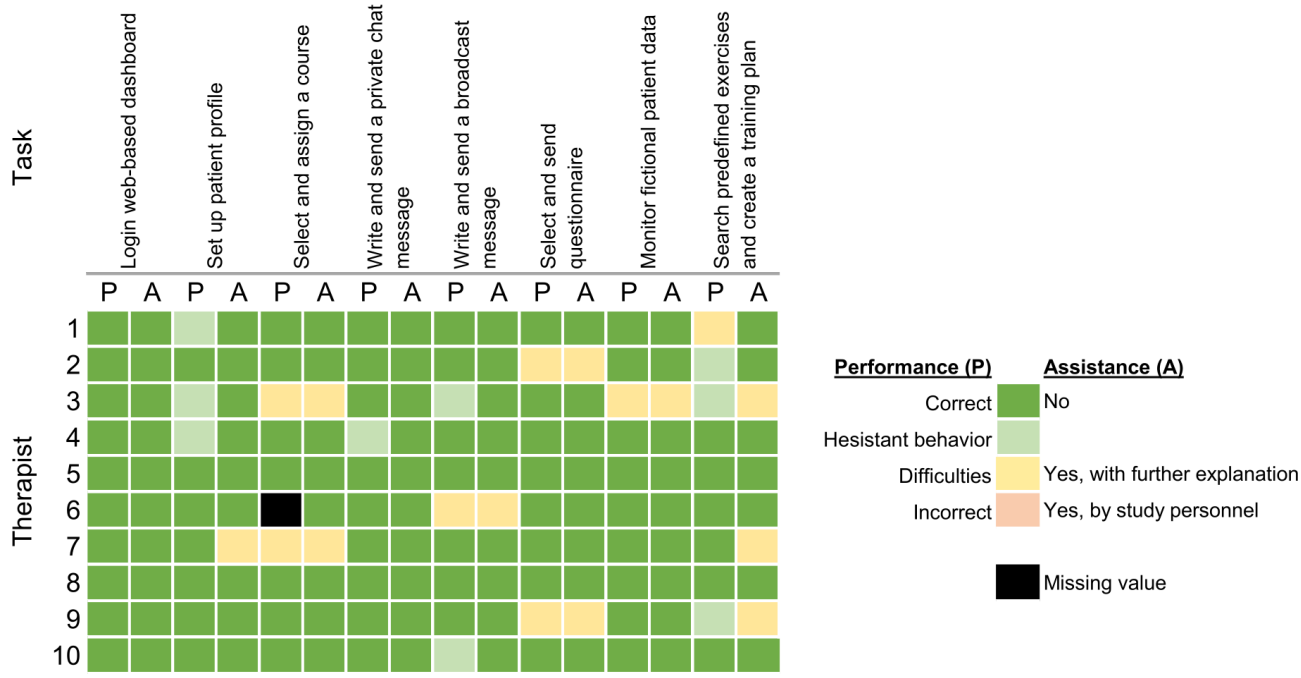
XSL•FO
RenderX

Figure 5. Evaluation of patients' task performance.



All therapists were able to perform the assigned tasks. Only a few needed some further explanations (Figure 6).

Figure 6. Evaluation of therapists' task performance.



Adherence to the START Program Protocol

Overall, patients' adherence to the protocol varied and was moderate for the test phase between study sessions. Adherence regarding training and monitoring differed distinctively among individual patients (Multimedia Appendix 5). Some patients consistently wore the activity tracker but did not adhere to the planned exercises or blood pressure measurements. There were no adverse events reported during the one-week trial phase.

- Training: adherence to training was low on average. Only one patient trained on most days (≥75% of planned training sessions), while 3 patients showed moderate adherence (between 50% and 74% of the planned sessions). One patient revealed low adherence to the training plan (between

XSL•FO
RenderX

25% and 49%), and 3 had no adherence (<25%). Two patients could not perform exercises on one day or 2 days due to technical problems. Furthermore, some patients skipped part of or all exercises in the app, as noticed by the short time spent on the exercise screens.

- Monitoring:
- Blood pressure measurements: only one patient performed the blood pressure measurement every day. Three patients measured blood pressure on more than half of the planned days.
- Activity tracker: most patients (n=6) showed high adherence (≥75% of planned days) wearing the activity tracker. The adherence of one patient was low (25% - 49%), and one patient did not wear the tracker at all. In the case of this latter patient, however, the reason was a skin allergic reaction. There was no reason for the patients to remove the tracker, whether for charging, sleeping, or bathing, which may have contributed to the generally high adherence regarding the wearing of the trackers.

Discussion

Principal Findings

The aim of this study was to introduce our newly developed START program and initially determine how patients after a stroke and therapists evaluate the usability and functionality of START within the Blended Clinic platform. The following key findings were identified:

- This study demonstrated that both patients and therapists consider the Blended Clinic platform, including the START program, to be user-friendly.
- Regarding the usability of the app for patients after a stroke, the study yielded generally positive results. However, there remains potential for improvement to enhance long-term compliance with the program by patients and to ensure ease of use for therapists in the future.

Usability Evaluation

Evaluating usability is crucial to engage both patients and therapists early in the development process and to incorporate their valuable feedback into ongoing product refinement [36]. The initial insights into the usability of the telerehabilitation platform Blended Clinic in combination with the START program demonstrated good user-friendliness. Most of the given tasks were completed by the patients without assistance, which reassures us that the intended independent use of the Blended Clinic app is achievable. This is supported by the high SUS scores (83.3 among therapists and 87.2 among the participating patients with stroke), indicating the Blended Clinic platform with the START program as user-friendly. In a meta-analysis of digital health apps, the benchmark was a mean score of 68 with an SD of 12.5 [37], which is considerably exceeded in this study.

The evaluation based on the MARS-G allows for the assessment of the overall quality of the app-based training program, as well as the identification of both strengths and weaknesses through the mean scores of the individual sections [30]. In general, both therapists and patients rated the user experience and usability of the app as acceptable to good. Notably, therapists tended to rate the app slightly higher than the patients, particularly in the sections on “Engagement” and “Aesthetics.” Patients appear to be somewhat more critical in their evaluation of the app. This discrepancy may be attributed to the differing requirements and expectations of the 2 groups regarding the app [38].

Although the generalizability and credibility of the results from the feedback questionnaire are limited, it provided valuable insights, for example, concerning the handling of the blood pressure monitor, which can be applied in future studies. The results of the feedback questionnaire also indicated a generally positive perception of the app and the devices by the patients, although several areas for potential improvement have been identified. The technical functionality was regarded as having development potential. However, technical problems are quite common in studies investigating mobile health apps, as described in a scoping review on problems and barriers related to the use of these apps [39].

The therapists’ feedback additionally provided valuable insights into the practical implementation of the app. The therapists expressed a consistently positive response to the question of whether they would like to use the system in their therapy. The ratings indicated a high level of acceptance, overall satisfaction, and the potential of the system as a telerehabilitation option during the rehabilitation process. Our results are consistent with previous research. Physiotherapists are generally open to digital health technologies but frequently emphasize barriers to their implementation. In the study by Martinsen et al [37], therapists expressed concerns particularly about technical aspects, reimbursement issues, and establishing and maintaining a therapeutic relationship with patients [40]. The Blended Clinic platform with the START program is intended to address these issues through a blended care approach, which combines digital elements and personal interaction. The therapists in our study highlighted other important considerations, such as integration into existing therapy workflows, the need for flexibility in scheduling, and potential barriers related to staffing or digital infrastructure.

Patient Adherence to Protocol

The results of the one-week test phase indicated low adherence to the START training exercises among patients, despite the exercises being designed to be manageable and requiring only 10 minutes of daily commitment. Several factors may have contributed to the low adherence rates.

Participation in the study was conducted alongside the patients’ regular therapy regimen, which may have impeded their motivation to fully adhere to the study protocol. In addition, the exercises were rather easy for some patients, which might have influenced their adherence. Exercise difficulty and progression should be evaluated in further investigations.

While the presence of instructional videos and access to a physiotherapist via chat provided valuable resources, patients may still have struggled with their training. It is important to consider that impaired cognitive abilities can hinder protocol adherence. A study involving patients with stroke showed that factors such as age and disarticulation significantly influence

XSL•FO
RenderX

rehabilitation exercise adherence [41]. From our perspective, it would be beneficial to assess and individualize the amount and method of external support and motivation required. Further strategies may have been needed to enhance patient engagement and adherence to the exercise program, for example, family involvement.

Regarding low adherence to blood pressure measurements, patients may have perceived little added value in measuring their blood pressure themselves since these were routinely taken by the care personnel during inpatient stay. Not all patients were equally affected after a stroke, and some may have had further difficulties using the device due to motor function limitations, uncontrolled movements, or even hemiplegia, which could explain the issues encountered during blood pressure measurement.

On the contrary, the adherence to wearing the activity tracker was quite high. Such trackers have high potential to deliver meaningful data for the health care professionals, for example, step counts or time spent on specific activity levels. This activity tracker data could also be included directly in the Blended Clinic platform to increase the motivation and physical activity of patients. Such an approach is described in the pilot study by Paul et al [39], where a gamified approach was used to enhance step count goals [42].

Grau-Pellicer et al [40] conducted a randomized controlled trial using a similar program in conjunction with a smartphone app [43]. The authors could show positive results, for example, an improved activity level in the intervention group, but there was a low adherence to app use. They stressed the importance of developing technologies adapted to patients with stroke that are easy to use in order to increase adherence [43]. In our study, we tested only part of the START program, a one-week period, with a selected patient group (mRS 3 - 5), as usability was of particular importance to us at this stage. In contrast, adherence rates are generally high (≥80%) in telerehabilitation studies for patients with stroke, although this aspect is often underreported [44].

However, a randomized controlled trial by Emmerson et al [42] found no superiority of a home-based exercise program delivered via electronic tablet compared to a paper-based version, underscoring the importance of remote therapeutic support [12]. Consistent with this, promising results have been reported when remote support was provided to enhance adherence to home exercises in individuals with musculoskeletal conditions [45].

Usability evaluations for a wide range of mobile health apps have been conducted across various target populations [46], including patients after a stroke [47], demonstrating beneficial effects for patients. In physical stroke rehabilitation, similar mobile apps were developed and evaluated, for example, ARMStrokes (Towson University) for upper extremity training [48], an app to increase physical activity [42], or a program involving caregivers [49]. These and other technologies show promising results, especially for improving upper extremity function [47].

However, despite the already available apps for patients after stroke, most apps have been developed outside Europe [11]. Furthermore, additional equipment, for example, a balance disc [50] or a virtual reality system [51], is often necessary. With START, we focused on individualized physiotherapeutic exercises that require no special equipment. Notably, our usability study deliberately included patients with moderate to severe disabilities—a group often excluded from similar research—thereby providing valuable insights into the applicability and accessibility of the intervention for a broader spectrum of stroke survivors.

Limitations

Our usability study has several limitations that should be considered when interpreting the results. The sample size was small; however, we could include the planned patients after a stroke and therapists to evaluate the START program.

Furthermore, the intervention period for patients was only one week, which may not have been sufficient to draw conclusions about long-term adherence for the holistic START program. Extended usage periods are warranted to gain a deeper understanding of the factors influencing adherence to digital home-based exercise programs.

We also lack more detailed information about the patients, such as cognitive function or digital literacy. In our study, patients and therapists were asked to engage with a technical system and evaluate its usability and user experience. For study participation, patients were required to own a smartphone or tablet. Therefore, we assume that patients possess at least basic technical knowledge and feel confident in their ability to operate the technology. On the other hand, these requirements proved to be a challenge, as we experienced in a prolonged recruitment phase. However, the proportion of older individuals who own a smartphone is expected to rise [52], ensuring that access to mobile apps will be available for the majority of patients after a stroke. It can therefore be assumed that in the future, a significant proportion of patients after stroke will possess the necessary prerequisites to use these digital opportunities.

The feedback collected from patients and therapists was quite positive, but we cannot exclude the possibility of socially desired responses when interpreting the subjective outcomes. Ried et al [51] propose 8 preventive measures to address this issue; one of them is indirect questioning. Future studies should consider such measures to minimize this bias.

These limitations underscore the need for further research. Based on the findings from our study, we have designed a feasibility study aimed at a more comprehensive evaluation of the entire START program with a larger sample size. This evaluation will encompass the transition between inpatient and outpatient care, as well as the holistic content of the program covering all mRS levels. Additionally, we have established other focal points concerning the outcome measures, such as performance-based outcome measures.

Conclusions

In summary, the Blended Clinic platform, including the START program, was initially considered user-friendly, with good

XSL•FO
RenderX

usability ratings from both patients with stroke and therapists. The available data provide an insight into patient adherence and the underlying challenges. Future efforts should focus on optimizing technical functionality, enhancing user motivation, and integrating training into the daily routines of patients to potentially improve adherence. This study's results and adjustments aimed at enhancing the motivational component and paving the way for a long-term evaluation within a multiweek clinical study, in which the transition between hospital stay and the use of the system in a home setting will be tested.

Acknowledgments

We sincerely thank all the patients and therapists who participated in this study.

Funding

This study is part of the "TeleRehaStroke" project funded by the Canton Aargau, Department of Health and Social Care, Division of Health. There is no reference number available.

Data Availability

Data collected during the usability study can be made available by the first author upon reasonable request.

Conflicts of Interest

BC is the lead developer and founder of the software application used in this study. LHB received a grant from the Canton of Aargau, Switzerland, for the conduct of the present study.

Multimedia Appendix 1
Exercise catalog.
[PDF File, 354 KB - rehab_v13i1e77090_app1.pdf]

Multimedia Appendix 2
Feedback questionnaires.
[PDF File, 188 KB - rehab_v13i1e77090_app2.pdf]

Multimedia Appendix 3
Mobile App Rating Scale (MARS-G) results.
[PDF File, 129 KB - rehab_v13i1e77090_app3.pdf]

Multimedia Appendix 4
Feedback questionnaire results of patients and therapists.
[PDF File, 130 KB - rehab_v13i1e77090_app4.pdf]

Multimedia Appendix 5
Adherence metrics for patients.
[PDF File, 93 KB - rehab_v13i1e77090_app5.pdf]

References

1. Mendis S. Stroke disability and rehabilitation of stroke: World Health Organization perspective. Int J Stroke 2013 Jan;8(1):3-4. [doi: 10.1111/j.1747-4949.2012.00969.x] [Medline: 23280261]
2. Pantoni L, Poggesi A, Inzitari D. Cognitive decline and dementia related to cerebrovascular diseases: some evidence and concepts. Cerebrovasc Dis 2009;27 Suppl 1(suppl 1):191-196. [doi: 10.1159/000200459] [Medline: 19342851]
3. Erkinjuntti T. Vascular cognitive deterioration and stroke. Cerebrovasc Dis 2007;24 Suppl 1(1):189-194. [doi: 10.1159/000107395] [Medline: 17971655]
4. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
5. Thelwell S. Rehabilitation for patients after stroke. Nurs Times 2013;109(35):20-22. [doi: 10.33667/2078-5631-2019-4-39(414)-35-39] [Medline: 24266239]
6. Duncan PW, Bushnell C, Sissine M, et al. Comprehensive stroke care and outcomes. Stroke 2021 Jan;52(1):385-393. [doi: 10.1161/STROKEAHA.120.029678] [Medline: 33349012]

XSL•FO
RenderX

7. Bernhardt J, Godecke E, Johnson L, Langhorne P. Early rehabilitation after stroke. Curr Opin Neurol 2017 Feb;30(1):48-54. [doi: 10.1097/WCO.0000000000000404] [Medline: 27845945]
8. Shem K, Irgens I, Alexander M. Getting started: mechanisms of telerehabilitation. In: Telerehabilitation: Principles and Practice: Elsevier; 2022. [doi: 10.1016/B978-0-323-82486-6.00002-2]
9. Laver KE, Adey-Wakeling Z, Crotty M, Lannin NA, George S, Sherrington C. Telerehabilitation services for stroke. Cochrane Database Syst Rev 2020 Jan 31;1(1):CD010255. [doi: 10.1002/14651858.CD010255.pub3] [Medline: 32002991]
10. Seron P, Oliveros MJ, Gutierrez-Arias R, et al. Effectiveness of telerehabilitation in physical therapy: a rapid overview. Phys Ther 2021 Jun 1;101(6):pzab053. [doi: 10.1093/ptj/pzab053] [Medline: 33561280]
11. Mendes Pereira C, Matos M, Carvalho D, et al. Building bridges between people with stroke, families, and health professionals: development of a blended care program for self-management. J Clin Med 2024 Jan 4;13(1):300. [doi: 10.3390/jcm13010300] [Medline: 38202307]
12. Rintala A, Päivärinne V, Hakala S, et al. Effectiveness of technology-based distance physical rehabilitation interventions for improving physical functioning in stroke: a systematic review and meta-analysis of randomized controlled trials. Arch Phys Med Rehabil 2019 Jul;100(7):1339-1358. [doi: 10.1016/j.apmr.2018.11.007] [Medline: 30529323]
13. Bindschedler A, Ziller C, Gerber EY, et al. Feasibility of an application-based outpatient rehabilitation program for stroke survivors: acceptability and preliminary results for patient-reported outcomes. Bioengineering (Basel) 2024 Jan 29;11(2):135. [doi: 10.3390/bioengineering11020135] [Medline: 38391621]
14. Saunders DH, Sanderson M, Hayes S, et al. Physical fitness training for stroke patients. Cochrane Database Syst Rev 2020 Mar 20;3(3):CD003316. [doi: 10.1002/14651858.CD003316.pub7] [Medline: 32196635]
15. Billinger SA, Arena R, Bernhardt J, et al. Physical activity and exercise recommendations for stroke survivors: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2014 Aug;45(8):2532-2553. [doi: 10.1161/STR.0000000000000022] [Medline: 24846875]
16. Liguori G, editor. ACSM's Guidelines for Exercise Testing and Prescription, 11th edition: Lippincott Williams & Wilkins; 2021.
17. Kim Y, Lai B, Mehta T, et al. Exercise training guidelines for multiple sclerosis, stroke, and Parkinson disease: rapid review and synthesis. Am J Phys Med Rehabil 2019 Jul;98(7):613-621. [doi: 10.1097/PHM.0000000000001174] [Medline: 30844920]
18. National Clinical Guideline for Stroke for the UK and Ireland. National Clinical Guideline for Stroke for the UK and Ireland. 2023. URL: https://www.strokeguideline.org [accessed 2025-10-09]
19. Living clinical guidelines for stroke management. Stroke Foundation. URL: https://informme.org.au/guidelines/living-clinical-guidelines-for-stroke-management [accessed 2025-10-09]
20. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988 May;19(5):604-607. [doi: 10.1161/01.str.19.5.604] [Medline: 3363593]
21. Bruno A, Akinwuntan AE, Lin C, et al. Simplified modified rankin scale questionnaire: reproducibility over the telephone and validation with quality of life. Stroke 2011 Aug;42(8):2276-2279. [doi: 10.1161/STROKEAHA.111.613273] [Medline: 21680905]
22. Mulder M, Nijland R. Stroke impact scale. J Physiother 2016 Apr;62(2):117. [doi: 10.1016/j.jphys.2016.02.002] [Medline: 26947003]
23. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. Stroke 1999 Oct;30(10):2131-2140. [doi: 10.1161/01.str.30.10.2131] [Medline: 10512918]
24. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [doi: 10.1097/00005650-199206000-00002] [Medline: 1593914]
25. Morfeld M, Kirchberger I, Bullinger M. SF-36 Fragebogen Zum Gesundheitszustand: Deutsche Version Des Short Form-36 Health Survey: Hogrefe Verlag; 2011. [doi: 10.1007/978-3-642-21700-5_14]
26. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry: Taylor & Francis; 1996:189-194. [doi: 10.1201/9781498710411]
27. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72. [doi: 10.2196/mhealth.5849] [Medline: 27287964]
28. Terhorst Y, Philippi P, Sander LB, et al. Validation of the mobile application rating scale (MARS). PLoS ONE 2020;15(11):e0241480. [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
29. Messner EM, Terhorst Y, Barke A, et al. The German version of the mobile app rating scale (MARS-G): development and validation study. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e14479. [doi: 10.2196/14479] [Medline: 32217504]
30. Hawley-Hague H, Horne M, Skelton DA, Todd C. Review of how we should define (and measure) adherence in studies examining older adults' participation in exercise classes. BMJ Open 2016 Jun 23;6(6):e011560. [doi: 10.1136/bmjopen-2016-011560] [Medline: 27338884]
31. Chan A, Chan D, Lee H, Ng CC, Yeo AHL. Reporting adherence, validity and physical activity measures of wearable activity trackers in medical research: a systematic review. Int J Med Inform 2022 Apr;160:104696. [doi: 10.1016/j.ijmedinf.2022.104696] [Medline: 35121356]

XSL•FO
RenderX

32. Granger CV, Hamilton BB, Keith RA, Zielezny M, Sherwin FS. Advances in functional assessment for medical rehabilitation. Top Geriatr Rehabil 1986 Apr;1(3):59-74. [doi: 10.1097/00013614-198604000-00007]
33. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
34. Hyzy M, Bond R, Mulvenna M, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
35. Saaei F, Klappa SG. Rethinking telerehabilitation: attitudes of physical therapists and patients. J Patient Exp 2021;8:23743735211034335. [doi: 10.1177/23743735211034335] [Medline: 34377773]
36. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
37. Martinsen L, Østerås N, Moseng T, Tveter AT. Usage, attitudes, facilitators, and barriers toward digital health technologies in musculoskeletal care: survey among primary care physiotherapists in Norway. JMIR Rehabil Assist Technol 2024 Sep 16;11:e54116. [doi: 10.2196/54116] [Medline: 39283661]
38. Xing F, Liu J, Mei C, et al. Adherence to rehabilitation exercise and influencing factors among people with acute stroke: a cross-sectional study. Front Neurol 2025;16:1554949. [doi: 10.3389/fneur.2025.1554949] [Medline: 40083453]
39. Paul L, Wyke S, Brewster S, et al. Increasing physical activity in stroke survivors using STARFISH, an interactive mobile phone application: a pilot study. Top Stroke Rehabil 2016 Jun;23(3):170-177. [doi: 10.1080/10749357.2015.1122266] [Medline: 27077973]
40. Grau-Pellicer M, Lalanza JF, Jovell-Fernández E, Capdevila L. Impact of mHealth technology on adherence to healthy PA after stroke: a randomized study. Top Stroke Rehabil 2020 Jul;27(5):354-368. [doi: 10.1080/10749357.2019.1691816] [Medline: 31790639]
41. Stephenson A, Howes S, Murphy PJ, et al. Factors influencing the delivery of telerehabilitation for stroke: a systematic review. PLoS ONE 2022;17(5):e0265828. [doi: 10.1371/journal.pone.0265828] [Medline: 35544471]
42. Emmerson KB, Harding KE, Taylor NF. Home exercise programmes supported by video and automated reminders compared with standard paper-based home exercise programmes in patients with stroke: a randomized controlled trial. Clin Rehabil 2017 Aug;31(8):1068-1077. [doi: 10.1177/0269215516680856] [Medline: 27920262]
43. Lambert TE, Harvey LA, Avdalis C, et al. An app with remote support achieves better adherence to home exercise programs than paper handouts in people with musculoskeletal conditions: a randomised trial. J Physiother 2017 Jul;63(3):161-167. [doi: 10.1016/j.jphys.2017.05.015] [Medline: 28662834]
44. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil 2019 Jan;100(1):115-127. [doi: 10.1016/j.apmr.2018.07.439] [Medline: 30171827]
45. Guo J, Smith T, Messing D, Tang Z, Lawson S, Feng JH. ARMStrokes: a mobile app for everyday stroke rehabilitation. Presented at: ASSETS '15: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility; Oct 26-28, 2015; Lisbon, Portugal p. 471-472. [doi: 10.1145/2700648.2811337]
46. van den Berg M, Crotty M Prof, Liu E, Killington M, Kwakkel G Prof, van Wegen E. Early supported discharge by caregiver-mediated exercises and e-Health support after stroke: a proof-of-concept trial. Stroke 2016 Jul;47(7):1885-1892. [doi: 10.1161/STROKEAHA.116.013431] [Medline: 27301941]
47. Burns SP, Terblanche M, Perea J, et al. mHealth intervention applications for adults living with the effects of stroke: a scoping review. Arch Rehabil Res Clin Transl 2021 Mar;3(1):100095. [doi: 10.1016/j.arrct.2020.100095] [Medline: 33778470]
48. Aphiphaksakul P, Siriphorn A. Home-based exercise using balance disc and smartphone inclinometer application improves balance and activity of daily living in individuals with stroke: a randomized controlled trial. PLoS ONE 2022;17(11):e0277870. [doi: 10.1371/journal.pone.0277870] [Medline: 36409753]
49. Sheehy L, Taillon-Hobson A, Sveistrup H, et al. Home-based nonimmersive virtual reality training after discharge from inpatient or outpatient stroke rehabilitation: parallel feasibility randomized controlled trial. JMIR Rehabil Assist Technol 2025 Mar 28;12:e64729. [doi: 10.2196/64729] [Medline: 40153779]
50. Seifert A. Digitale transformation in den haushalten älterer menschen. Z Gerontol Geriat 2022 Jul;55(4):305-311. [doi: 10.1007/s00391-021-01897-5] [Medline: 33909129]
51. Ried L, Eckerd S, Kaufmann L. Social desirability bias in PSM surveys and behavioral experiments: considerations for design development and data collection. Journal of Purchasing and Supply Management 2022 Jan;28(1):100743. [doi: 10.1016/j.pursup.2021.100743]
52. Gäumann S, Ziller C, Paulissen N, et al. START-the Swiss tele-assisted rehabilitation and training program to support transition from inpatient to outpatient care in the subacute phase after a stroke: feasibility, safety and performance evaluation. Front Digit Health 2024;6:1496170. [doi: 10.3389/fdgth.2024.1496170] [Medline: 39959919]

Abbreviations

ADL: activities of daily living
MARS-G: German version of Mobile App Rating Scale

XSL•FO
RenderX

mRS: modified Rankin Scale
SF36: 36-Item Short Form Survey Instrument
SIS: Stroke Impact Scale
START: Swiss Tele-Assisted Rehabilitation and Training program
SUS: System Usability Scale

Edited by S Munce; submitted 07.May.2025; peer-reviewed by L Sheehy, TA Wani, T Shimamoto; revised version received 19.Dec.2025; accepted 23.Dec.2025; published 31.Mar.2026.

Please cite as:
Ziller C, Gäumann S, Lüscher S, Paulissen N, Behrendt F, Suica Z, Crüts B, Gamerschlag L, Parmar K, Gerth HU, Bonati LH, Schuster-Amft C
Telerehabilitation Following Stroke: Development of Training Content and Evaluation of an App-Based Training Program
JMIR Rehabil Assist Technol 2026;13:e77090
URL: https://rehab.jmir.org/2026/1/e77090
doi:10.2196/77090

© Carina Ziller, Szabina Gäumann, Silya Lüscher, Nele Paulissen, Frank Behrendt, Zorica Suica, Björn Crüts, Luana Gamerschlag, Katrin Parmar, Hans Ulrich Gerth, Leo H Bonati, Corina Schuster-Amft. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Emotional Training via Telerehabilitation After Surgical Treatment for Facial Palsy: Prospective, Assessor-Blinded, 2-Arm Pilot Cohort Study

Matteo Guidetti[1], PT, PhD; Silvia Cupello[2,3], PT; Jacopo Reali[2], PT; Natale Vincenzo Maiorana[1], PsyD, PhD; Sara Marceglia[1], PhD; Rossella Pagani[3], MD; Federico Biglioli[4], MD; Alberto Priori[1,5], MD, PhD; Antonino Michele Previtera[2], MD

[1]Department of Health Sciences, Aldo Ravelli Center for Neurotechnology and Experimental Brain Therapeutics, University of Milan, Milan, Italy

[2]Department of Health Sciences, Rehabilitation Unit, ASST-Santi Paolo e Carlo University Hospital, University of Milan, Milan, Italy

[3]Department of Health Sciences, University of Milan, Via Antonio di Rudinì 8, Milan, Italy

[4]Department of Head and Neck, Division of Maxillofacial Surgery, San Paolo Hospital, University of Milan, Milan, Italy

[5]Department of Health Sciences, Clinical Neurology Unit, Azienda Socio-Sanitaria Territoriale Santi Paolo E Carlo, University of Milan, Milan, Italy

Corresponding Author:
Rossella Pagani, MD
Department of Health Sciences, University of Milan, Via Antonio di Rudinì 8, Milan, Italy

Abstract

Background: Peripheral facial nerve palsy is a debilitating condition that may necessitate surgical intervention. Although motor rehabilitation is considered essential, the most effective approach has not yet been determined.

Objective: This study aimed to evaluate the feasibility and effectiveness of emotional training, a novel telerehabilitation-based treatment, on motor, functional, and psychological outcomes in patients with unilateral facial palsy following triple innervation surgery.

Methods: A prospective, assessor-blinded, 2-arm pilot cohort study was conducted at the rehabilitation unit at University Hospital San Paolo, Milan, Italy, from January to October 2024. Participants (N=16) received 1 treatment session every 2 weeks over 20 weeks, each lasting 45 minutes, according to standard clinical procedures in place at the rehabilitation unit. Participants were nonrandomly assigned to either an in-person group (n=8) or an online group (ie, telerehabilitation; n=8) based on their ability to attend in-person sessions. The primary outcomes assessed at baseline (T0) and after treatment (T1) included facial symmetry (Sunnybrook Facial Grading System; SFGS), facial disability (Facial Disability Index; FDI), and anxiety levels (Beck Anxiety Inventory).

Results: Statistical analysis revealed significant improvements at T1 for both groups in the FDI social and well-being function subscale, Beck Anxiety Inventory, SFGS resting symmetry score, SFGS symmetry of voluntary movement score, SFGS composite score, SFGS with bilateral masseter contraction symmetry of voluntary movement score, and SFGS with bilateral masseter contraction composite score (P<.001 for all). Only the FDI physical function subscale showed a differential improvement at T1 for the in-person group treatment (ANOVA for time × treatment: F_1=14.356; P=.002; Holm-Bonferroni post hoc test: P<.001). Finally, a strong positive correlation was observed between the time elapsed from surgery to rehabilitation and SFGS composite score improvement at T1 (r=0.94; P=.005).

Conclusions: These results suggest that the online emotional training protocol is as feasible and effective as the in-person emotional training protocol in improving facial motor function, reducing anxiety, and enhancing facial expression spontaneity in patients who had undergone surgery for peripheral facial palsy. These findings support the validity of telerehabilitation approaches as a feasible, accessible, and sustainable alternative to conventional in-person therapy for facial nerve recovery.

(JMIR Rehabil Assist Technol 2026;13:e79520) doi:10.2196/79520

KEYWORDS

peripheral facial nerve palsy; physiotherapy; motor rehabilitation; telerehabilitation; Sunnybrook Facial Grading System

Introduction

Peripheral facial nerve palsy (FNP) is a neurological disorder that affects the muscles of one side of the face [1], leading to facial asymmetry; synkinesis; hemifacial spasm [2]; impaired eye and mouth closure; and functional deficits in daily activities, including vision, speech, and eating [3,4]. Peripheral FNP resulting from traumatic events (eg, cranial fractures) [5], oral

XSL•FO
RenderX

and maxillofacial surgery [6], or neoplastic conditions (eg, acoustic neuroma) [6] often presents with a highly variable prognosis [6,7] and may require invasive intervention, such as facial reanimation [8]. Following facial reanimation surgery, physiotherapy (PT) plays a critical role in motor and functional recovery [9]. PT interventions capitalize on principles of neuroplasticity and motor learning to restore facial symmetry [10,11]; preserve proprioception, muscle tone, and trophism [12]; reduce synkinesis; and enhance both spontaneous and voluntary facial movements [13].

Triple innervation is a novel surgical protocol [14,15] that uses 3 distinct neural inputs to restore movements and functions: the masseteric and hypoglossal nerves, which serve as quantitative input (ie, facilitating voluntary facial movements, such as smiling during teeth clenching and maintaining muscle tone), and the facial nerve contralateral to the damaged one, which functions as a qualitative input (ie, supporting spontaneous facial expressions such as smiling and blinking) [14,15]. Several studies suggest that triple innervation surgery is associated with complication rates comparable to those of other established techniques [14] while demonstrating improved outcomes in terms of facial symmetry, voluntary and spontaneous movements, and overall functional recovery both at short- and long-term follow-up [14,16,17]. Given the novelty of the surgical technique, no study has assessed the effect of PT following triple innervation yet. In addition, although several PT protocols, including massage [12,18], electrotherapy [19,20], exercise therapy [10,11], and biofeedback therapy [21,22], have demonstrated beneficial effects, no clear evidence supports the superiority of any of them [13,23]. Compared to other pathologies, facial motor rehabilitation is complicated by the intrinsic link between facial movements and emotional expression. The neural circuits governing facial expressions are closely connected to systems involved in emotion recognition and social communication [24-27], suggesting that interventions targeting both motor and emotional components may be particularly relevant in FNP. Moreover, aesthetic and functional impairments often lead to psychosocial distress, reduced self-esteem, and increased anxiety [28,29], reinforcing the need for rehabilitation approaches that also address emotional and cognitive aspects of facial movement. Within the context of PT interventions, emotional training has recently been proposed as a targeted approach to address these neurophysiological and clinical characteristics and optimize rehabilitation outcomes [30]. Emotional training conceptualizes movement control as a cognitive process involving selective attention to sensory cues and task demands [31,32] and integrates emotional, cognitive, and motor components to facilitate selective muscle recruitment, retraining of emotional expressiveness, and the execution of functional exercises oriented toward daily activities. This approach has been recently introduced as a rehabilitation approach for patients after triple innervation surgery in our center [30].

Recently, telerehabilitation has been proposed as a promising rehabilitation method for patients with neurological disorders, showing results comparable to those of conventional face-to-face rehabilitation [33]. While still promising, evidence remains limited for the motor rehabilitation of peripheral FNP [34,35]. Telerehabilitation represents a hospitalization-avoidant strategy that emphasizes the delivery of home-based rehabilitation, minimizing resource requirements and reducing direct supervision needs [36-38]. This approach may serve as a particularly viable treatment option for patients with FNP [36] as it minimizes the need for social exposure and interpersonal contact, which can be especially beneficial for patients facing psychosocial burden.

In this work, we primarily aimed to assess the feasibility and motor-functional and psychological effects of a telerehabilitation-based emotional training intervention [30] in patients with peripheral FNP who underwent triple innervation surgery and compare the results with those of an in-person emotional training approach.

Methods

Study Design and Setting

This prospective, assessor-blinded, 2-arm pilot cohort study was conducted at the rehabilitation unit at University Hospital San Paolo (Milan, Italy). On the basis of their ability to attend in-person sessions, each participant eligible for enrollment (N=16) engaged in either in-person PT treatment (ie, in-person emotional training; n=8) or in telerehabilitation treatment (ie, online emotional training; n=8). In both groups, treatments were led by trained physiotherapists. Motor outcomes (Sunnybrook Facial Grading System [SFGS] and SFGS with bilateral masseter contraction [SFGS with MassCon]), functional outcomes (Facial Disability Index; FDI), and psychological outcomes (Beck Anxiety Inventory; BAI) were assessed at baseline (T0) and at the completion of treatment (T1). Given the exploratory pilot nature of this study, which aimed primarily to assess feasibility and generate preliminary results to inform future confirmatory trials, and considering the rarity of the medical condition targeted by the inclusion criteria, no formal a priori sample size calculation was performed. However, the final sample (N=16) is consistent with those of previous pilot investigations conducted in comparable clinical populations [39,40]. A convenience sampling strategy was adopted, enrolling all consecutive eligible patients referred to the rehabilitation unit during the predefined recruitment period (November 2023 to November 2024). In addition, a prospective observational design was adopted due to several clinical (eg, emotional training is part of postoperative rehabilitation in our department) and logistical (eg, rarity of the target population and allocation to treatment according to patients' real-world logistical circumstances) reasons.

Ethical Considerations

The study protocol followed the Declaration of Helsinki. Considering that emotional training is common practice for patients undergoing triple innervation surgery at the rehabilitation unit of University Hospital San Paolo, and that patients with peripheral FNP would have been treated with the same protocol after triple innervation surgery despite not participating in the study, the research was considered as an observational study. Ethical approval was obtained from the Milan Area 1 Ethics Committee, protocol number 0048450/2022. All patients gave written informed consent

XSL•FO
RenderX

before participation. All data collected in this study were handled in accordance with strict privacy and confidentiality standards. Records were anonymized through the use of coded identifiers, and no personally identifiable information was retained in either the research database or the investigational application. Hard-copy materials (eg, consent forms) were securely stored in a locked cabinet in a restricted-access laboratory, while electronic data were maintained on protected institutional servers with controlled access. Access to the coded datasets was limited to authorized members of the research team. Participants were not provided with financial compensation for their involvement in the study.

Population

This study was conducted between January and October 2024. Participants were recruited from the hospital community based on the following inclusion criteria: (1) age between 18 and 75 years; (2) presence of unilateral facial palsy, confirmed through clinical history and anamnesis; (3) prior treatment with triple innervation surgery (as described in the work by Biglioli et al [14]); (4) a House-Brackmann scale grade of III or higher; (5) ability to understand and perform the tasks; and (6) capacity to provide informed consent. Exclusion criteria were (1) a diagnosis of major neurological, neuropsychological, or psychiatric disorders, confirmed through clinical history and anamnesis; (2) previous treatments for facial palsy other than triple innervation surgery [14]; and (3) changes in pharmacological therapy within 3 months prior to enrollment. Patients were instructed to maintain stable medication regimens throughout the study period and refrain from participating in additional rehabilitation programs.

Treatments

Overview

All patients received the standard 5-month emotional training rehabilitation program [30]. The treatment sessions lasted 50 minutes and were conducted once every 2 weeks over a 20-week period (total number of sessions=10; see Multimedia Appendix 1 for a structured session plan). The primary objective of both the in-person and online groups was to enhance the quality and symmetry of facial muscle movements, initially facilitated through teeth clenching, and progressively dissociating these movements from masseter activation [41-43]. This approach was chosen because one of the primary rehabilitative goals for patients who have undergone triple innervation surgery is to improve voluntary contraction of the mimic muscles by gradually decoupling them from masseter activation through repeated exercises and the progressive automation of movements [41-43]. According to the principles of triple innervation surgery [14], the masseteric nerve, which stimulates the masseter muscles, serves as the principal donor nerve for activating the facial nerve and, consequently, the facial mimic muscles. Therefore, postsurgical facial muscle contractions activate the masseter muscles. The progress of each patient was closely monitored by physiotherapists, who adjusted the complexity of exercises in response to individual progress. As facial movements improved, progressively more advanced exercises were introduced to stimulate emotional expression and functional facial movements in increasingly complex real-world contexts, providing continuous feedback to optimize movement patterns. Through consistent repetition and progressive automation, the secondary objective was to restore the functionality and spontaneity of facial expressions, enhancing facial perception and promoting cortical reorganization. As suggested by previous research [44], exercises were tailored to each patient's residual motor capabilities, functional limitations, and social challenges. The physiotherapists delivering the treatment underwent extensive training prior to its initiation and had prior experience in the rehabilitation of peripheral FNP.

In-Person Emotional Training Protocol

Before initiating the emotional training protocol, tissue flexibility in the impaired masseter, medial and lateral pterygoid, zygomaticus major, orbicularis oculi, and orbicularis oris muscles was improved using manual therapy techniques [45] such as internal and external stretching techniques and transverse massage [45,46]. This preliminary phase was necessary due to the presentation typically observed in soft tissues following facial palsy surgery [47].

As previously reported [30], emotional training comprises 3 phases. The first phase introduces exercises designed to increase sensory awareness, promote appropriate and selective muscle recruitment, and integrate both sensory perception and motor function [32]. Patients were asked to perform tactile and pressure perception exercises, identifying sensations at key points on the face (eg, the corners of the mouth and the eyebrows) [32]. When movement abilities allowed it, facial exercises were paired with the activation of the masseter muscle focusing on 2 movements of functional importance: eye closure and smiling (with either closed or open lips). Otherwise, patients were asked to perform motor imagery exercises, which have been shown to activate the same cortical and subcortical areas as during actual movement execution [48,49]. Patients mentally simulated facial movements focusing on the simulated sensations and were then asked to perform them and compare the simulated and actual sensations. As facial movements improved, exercises progressively intensified, with patients being asked to close their eyes or smile at different intensities of teeth clenching or to self-correct their movements according to their perceived facial symmetry.

The second phase introduces exercises to improve the spontaneity of emotional expression. Patients were asked to perform facial expressions corresponding to emotions defined by Osgood [50] for nonverbal communication (happiness, sadness, interest, disgust, fear, anger, and surprise) elicited through visual (eg, viewing triggering images or scenarios), taste (eg, tasting sour food), or memory (eg, recalling triggering experiences) stimuli. As facial movements improved, patients were asked to switch between emotional expressions (eg, from fear to anger).

The last phase of the protocol focuses on functional exercises aimed at restoring the ability to perform daily activities (eg, communicating effectively or drinking from a glass) using facial muscles. These exercises included, for example, activities to improve word articulation (eg, reading aloud a text or words with challenging sounds such as labial consonants [eg, "b" and

XSL•FO
RenderX

"p"]) or mastication (eg, chewing inside both cheeks or transferring food or liquid from one side to the other).

Online Emotional Training Protocol

Participants in the online treatment group received the same intervention protocol as those in the in-person treatment group in terms of the type, nature, frequency, and objectives of the exercises; however, the intervention was delivered remotely via video call, enabling the physiotherapist to directly monitor exercise correctness, safety, and adherence; provide immediate feedback; and adjust the intervention as needed. No session was recorded as supervision and feedback were provided live during each video call and adherence was tracked through supervised session attendance as the intervention did not include unsupervised exercises requiring additional logging or diaries. Due to the absence of in-person interaction with the physiotherapist, patients were initially trained during the assessment phase to perform manual relaxation techniques and facial exercises targeting eye closure and smiling with masseter muscle activation. These exercises were performed under the physiotherapist's virtual supervision, where the patient was guided to repeat the maneuvers, ensuring correct execution and understanding. Furthermore, throughout the treatment sessions, exercises were designed to encourage patients to focus attentively on the sensory feedback from their facial movements, with continuous feedback provided to assess progress. No further modifications were made to the in-person emotional training protocol compared to the online emotional training protocol.

Outcomes

SFGS, SFGS with MassCon, FDI, and BAI were administered both at baseline (T0) and the posttreatment time point (T1) by a single trained physiotherapist who was blinded to treatment group allocation but not to time points.

The SFGS is a widely used clinical scale designed to evaluate facial asymmetries and the presence of synkinesis comparing the paretic side of the face with the nonparetic side [51]. The scale consists of 3 components: the first (resting symmetry) involves assessing symmetry at rest, including the eye, cheek (nasolabial fold), and mouth; the second (symmetry of voluntary movement) rates symmetry during facial expressions (ie, during 5 standard facial expressions); and the third (synkinesis) grades the severity of synkinesis during these movements. Each component is assigned a score, and the weighted sum of these components generates a composite score ranging from 0 to 100, with a lower score indicating greater severity of facial palsy [51].

For this study, a modified version of the SFGS was used in which participants were instructed to perform the second component (symmetry of voluntary movement) while clenching their jaw (ie, activating the masseter muscles).

The FDI is a 10-item self-report questionnaire assessing 2 domains of facial disability: physical and psychosocial function [52]. It includes 2 subscales—one for physical function and the other for social and well-being function—each comprising 2 items. The scores for each subscale are converted to a 100-point scale, where a score of 100 denotes no impairment in function [52].

The BAI is a psychometrically validated, self-reported, 21-item scale commonly used to measure symptoms of anxiety [53]. The score for each item is summed, yielding a total score ranging from 0 to 63, categorized as minimal anxiety (0-7), mild anxiety (8-15), moderate anxiety (16-25), and severe anxiety (26-63) [53].

Statistical Analysis

Normal distributions of the dependent variables were assessed using the Shapiro-Wilk test for normality [54]. As all variables passed the test (P>.05), parametric statistical analysis was applied. To evaluate the effects of time, treatment, and the interaction between time and treatment, changes in the outcomes at each time point were examined using a 2-way repeated-measure ANOVA with treatment as the between-subject factor. Post hoc analysis was performed using the Holm-Bonferroni correction. Finally, the Pearson correlation coefficient was calculated to explore relationships among age, time from onset of facial palsy to surgery, time from surgery to rehabilitation, time from onset of facial palsy to rehabilitation, and the change in outcome scores (Δ; calculated as Δ score = T1 score − T0 score). A P value of less than .05 was considered statistically significant for all analyses. Data were analyzed using JASP (version 0.16.3) for Windows (JASP Team).

Results

Baseline characteristics of the sample are shown in Table 1 (see Figure 1 for the CONSORT [Consolidated Standards of Reporting Trials] flow diagram). At baseline, there were no significant differences in demographic variables or clinical assessments between the 2 treatment groups (P>.05 for all the analyses).

XSL•FO
RenderX

Table . Sample characteristics (N=16).

	Overall	in-person (n=8)	Online (n=8)	*P* value[a]
Sex (female), n (%)	10 (62.5)	4 (50.0)	6 (75.0)	.30
Age (y), mean (SD)	53.93 (16.63)	53.12 (18.30)	54.75 (16.02)	.88
Side, n (%)				.61
Right	7 (43.8)	3 (37.5)	4 (50.0)	
Left	9 (56.3)	5 (62.5)	4 (50.0)	
Cause, n (%)				.19
Neurinoma	12 (75.0)	5 (62.5)	7 (87.5)	
Meningioma	1 (6.3)	1 (12.5)	0 (0)	
Angioma	1 (6.3)	0 (0)	1 (12.5)	
Trauma	1 (6.3)	1 (12.5)	0 (0)	
Surgery	1 (6.3)	1 (12.5)	0 (0)	
Onset to surgery (mo), mean (SD)	21.04 (13.25)	25.70 (14.34)	11.74 (0.9)	.09
Surgery to rehabilitation (mo), mean (SD)	10.54 (4.5)	10.4 (5)	10.82 (5.16)	.52
Onset to rehabilitation (mo), mean (SD)	31.57 (13.49)	36.08 (14.7)	23 (4.27)	.13
SFGS[b] (resting symmetry score: 0-20; symmetry of voluntary movement score: 20-100; synkinesis score: 0-15; and composite score: 0-100), mean (SD)				
Resting symmetry score	12.19 (3.15)	11.25 (3.54)	13.13 (2.59)	.29
Symmetry of voluntary movement score	45.75 (12.39)	46 (14.81)	45.5 (10.46)	.99
Synkinesis score	1.38 (1.86)	1.38 (2.13)	1.38 (1.69)	.96
Composite score (%)	32.2 (13.2)	33.4 (15.1)	31 (11.8)	.79
SFGS with MassCon[c] (resting symmetry score: 0-20; symmetry of voluntary movement score: 20-100; synkinesis score: 0-15; composite score: 0-100), mean (SD)				
Symmetry of voluntary movement score	61.00 (10.78)	61.5 (12.64)	60.5 (9.42)	.83
Composite score (%)	47.3 (11.7)	48.5 (13.6)	46.1 (10.2)	.71
FDI[d], mean (SD)				
Physical function subscale (range 0-100)	53.75 (22.32)	45 (18.51)	62.5 (23.45)	.11
Social and well-being function subscale (range 0-100)	58.00 (25.42)	52 (26.01)	64 (25.02)	.46
BAI[e] total score (0-63), mean (SD)	16.81 (12.44)	20.13 (15.11)	13.5 (8.85)	.52

[a]Mann-Whitney *U* test (continuous variables) or chi-square test (categorical variables); in-person vs online group.

[b]SFGS: Sunnybrook Facial Grading System.

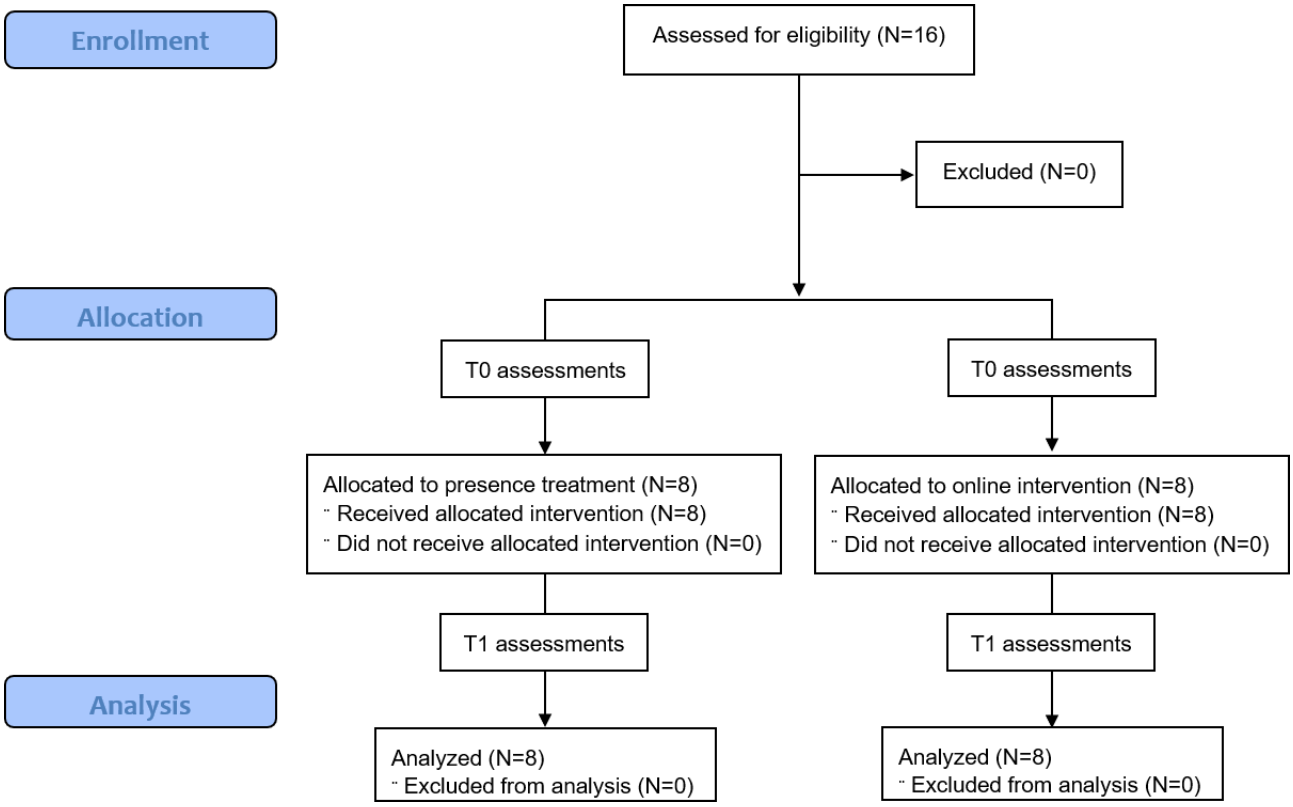
[c]MassCon: bilateral masseter contraction.

[d]FDI: Facial Disability Index.

[e]BAI: Beck Anxiety Inventory.

XSL•FO
RenderX

Figure 1. Flowchart of the study design. T0: baseline; T1: posttreatment time point.



Two-way ANOVA results are shown in Table 2. The analysis revealed no significant treatment effect for the outcomes considered (P>.05 for all the analyses; Table 2). A time × treatment effect was observed only for the FDI physical function subscale (F_1=14.356; P=.002), with post hoc analysis revealing a significant increase between T0 and T1 in the in-person treatment group (mean difference=–21.25, 95% CI −29.125 to −13.375; P<.001; Holm-Bonferroni post hoc test; Figure 2A). The effect of time was found for all the outcome variables except for the SFGS synkinesis score (P>.05). Specifically, a significant increase in score for the FDI social and well-being function subscale (mean difference=−11, 95% CI −15.163 to −6.837; F_1=32.114; P<.001) and decrease in the BAI total score (mean difference=5.375, 95% CI 2.357-8.393; F_1=14.592; P=.002; Figures 2B and 2C, respectively) were detected, a significant decrease in the SFGS resting symmetry score (mean difference=6.250, 95% CI 4.354-8.146; F_1=50.00; P<.001) and increase in the SFGS symmetry of voluntary movement score (mean difference=−17.250, 95% CI −21.111 to −13.389; F_1=91.810; P<.001) and SFGS composite score (mean difference=−0.239, 95% CI −0.288 to −0.191; F_1=111.430; P<.001) were observed (Figures 3A-3C, respectively), and a significant increase in the SFGS with MassCon symmetry of voluntary movement score (mean difference=−12.250, 95% CI −15.486 to −9.014; F_1=65.910; P<.001) and SFGS with MassCon composite score (mean difference=−0.196, 95% CI −0.231 to −0.161; F_1=143.499; P<.001) were observed (Figures 4A and 4B, respectively). No dropouts were registered in either treatment group, indicating strong patient engagement and good adherence to the intervention across both delivery modalities.

XSL•FO
RenderX

Table . Clinical assessments for each treatment condition (in-person and online) at baseline (T0) and the posttreatment time point (T1) and 2-way ANOVA results.

	In-person, mean (SD)		Online, mean (SD)		2-way ANOVA, *P* value		
	T0	T1	T0	T1	Time	Treatment	Time × treatment
SFGS[a]							
Resting symmetry score	11.25 (3.54)	5 (2.67)	13.13 (2.59)	6.88 (4.58)	*<.001* [b]	.22	>.99
Symmetry of voluntary movement score	46 (14.81)	63 (11.26)	45.5 (10.46)	63 (12.60)	*<.001*	.97	.89
Synkinesis score	1.38 (2.13)	1.25 (1.49)	1.38 (1.69)	0.63 (1.06)	.32	.66	.47
Composite score (%)	33.4 (15.1)	56.8 (11.6)	31 (11.8)	55.5 (15.4)	*<.001*	.78	.81
SFGS with MassCon[c]							
Symmetry of voluntary movement score	61.5 (12.64)	71.5 (8.12)	60.5 (9.42)	75 (10.2)	*<.001*	.80	.16
Composite score (%)	48.5 (13.6)	65.9 (9)	46.1 (10.2)	67.9 (12)	*<.001*	.97	.20
FDI[d]							
Physical function subscale	45 (18.51)	66.25 (12.17)	62.5 (23.45)	70 (19.46)	*<.001*	.27	*.002*
Social and well-being function subscale	52 (26.01)	64.5 (27.41)	64 (25.02)	73.5 (19.94)	*<.001*	.41	.45
BAI[e] total score (0-63)	20.13 (15.11)	13.12 (9.34)	13.5 (8.85)	9.75 (5.95)	*.002*	.33	.27

[a]SFGS: Sunnybrook Facial Grading System.

[b]Italicized values are significant.

[c]MassCon: bilateral masseter contraction.

[d]FDI: Facial Disability Index.

[e]BAI: Beck Anxiety Inventory.

XSL•FO
RenderX

Figure 2. Changes in (A) Facial Disability Index (FDI) physical function subscale, (B) FDI social and well-being function subscale, and (C) Beck Anxiety Inventory (BAI) total score. Data are given as mean and SD. The statistical significance refers to the comparison between in-person and online treatments. ***P*<.01; ****P*<.001 (2-way ANOVA); T0: baseline; T1: posttreatment time point.

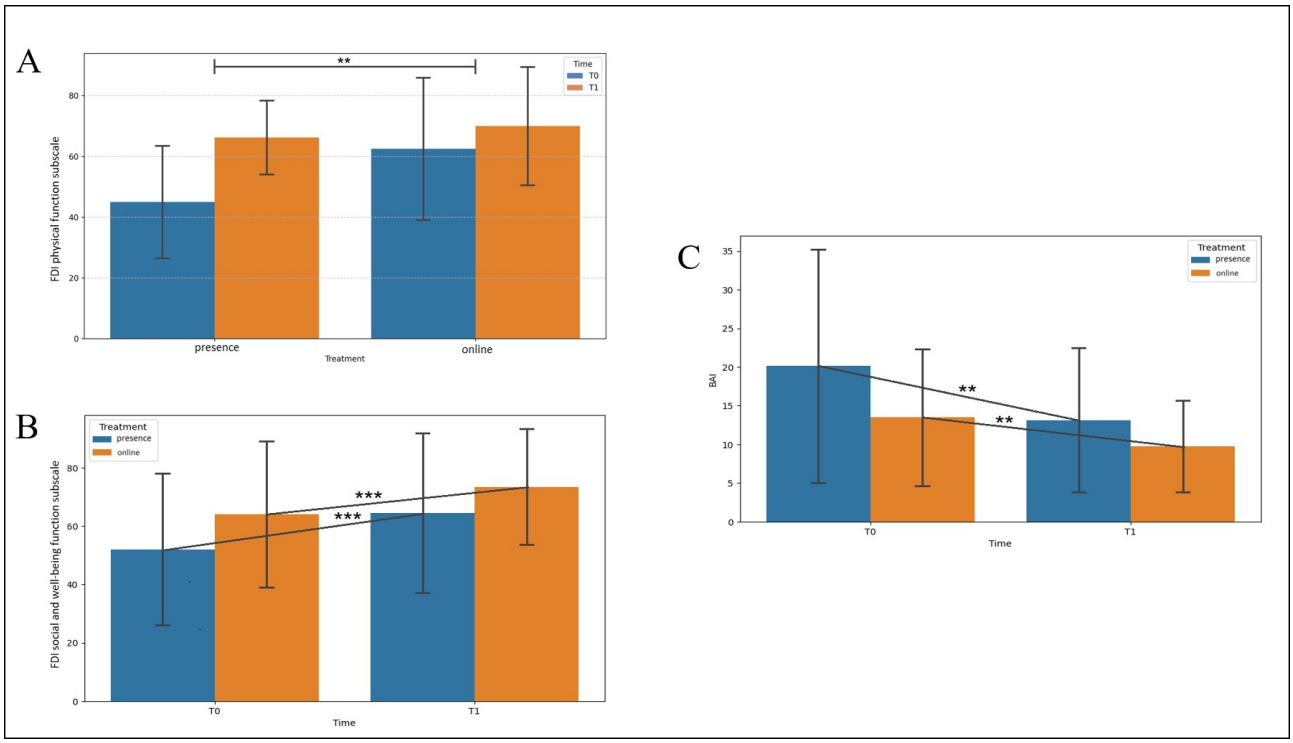
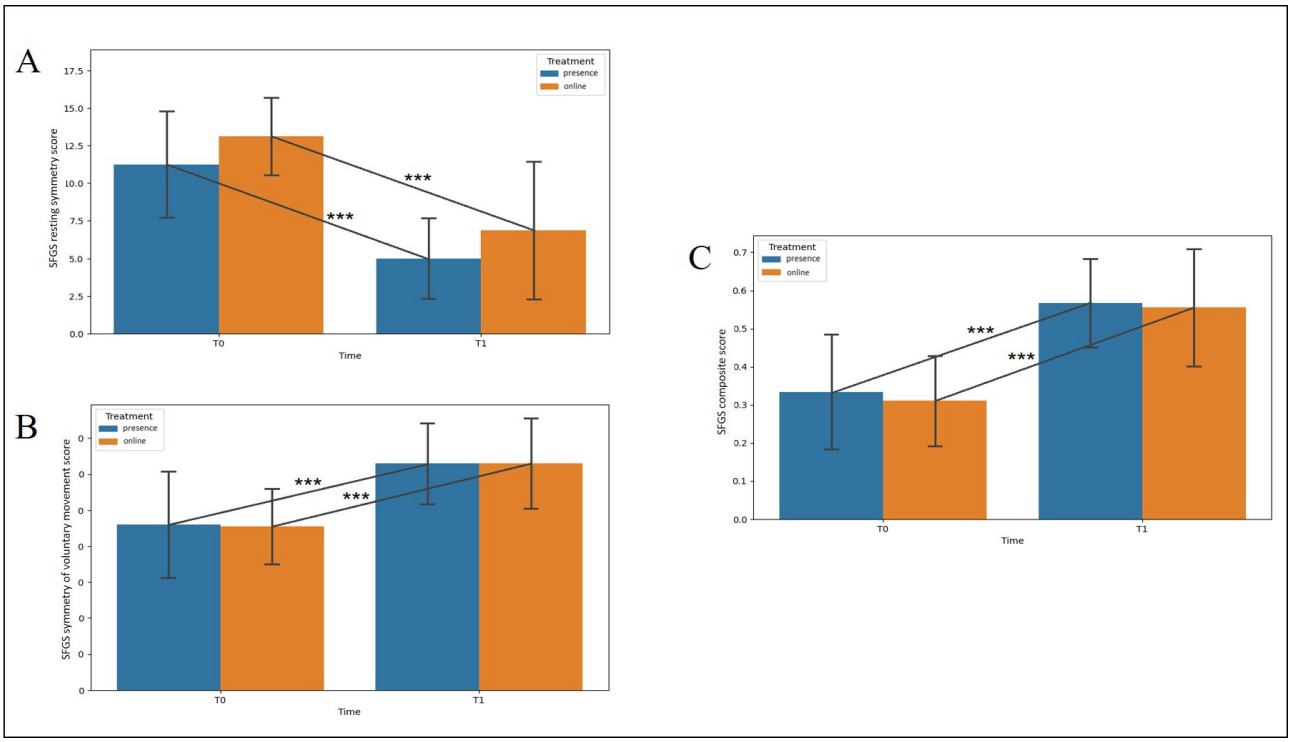
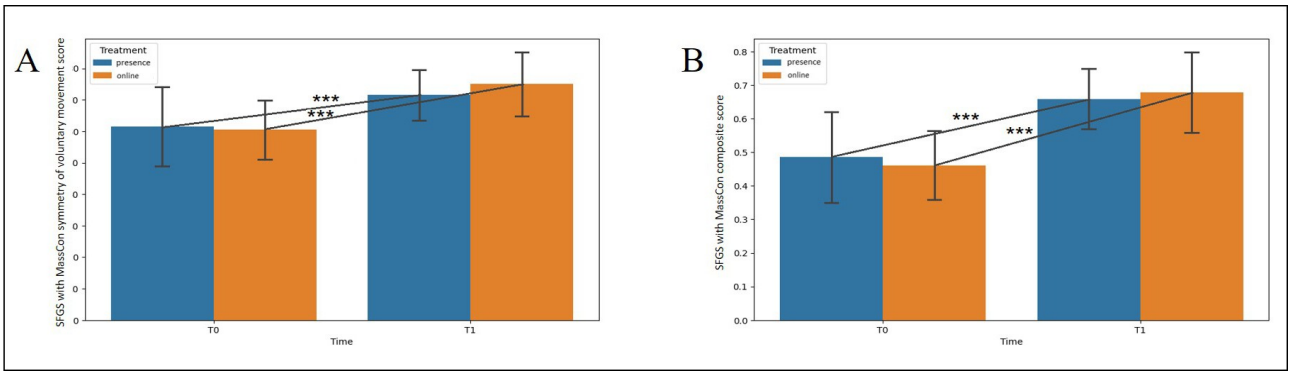


Figure 3. Changes in (A) Sunnybrook Facial Grading System (SFGS) resting symmetry score, (B) SFGS symmetry of voluntary movement score, and (C) SFGS composite score. Data are given as mean and SD. The statistical significance refers to the comparison between in-person and online treatments. ****P*<.001 (2-way ANOVA); T0: baseline; T1: posttreatment time point.



XSL•FO
RenderX

Figure 4. Changes in (A) Sunnybrook Facial Grading System (SFGS) with bilateral masseter contraction (MassCon) symmetry of voluntary movement score and (B) SFGS with MassCon composite score. Data are given as mean and SD. The statistical significance refers to the comparison between in-person and online treatments. ***P<.001 (2-way ANOVA); T0: baseline; T1: posttreatment time point.



Finally, correlation analysis revealed a moderate positive correlation between age and change in SFGS with MassCon symmetry of voluntary movement score (r=0.50, 95% CI 0.01-0.799; P=.05) and a strong positive correlation between the time from surgery to rehabilitation and change in SFGS composite score (r=0.94, 95% CI 0.832-0.979; P=.005; Table 3).

XSL•FO
RenderX

Table . Results of correlation analysis through Pearson correlation.

	Age		Onset to surgery		Surgery to rehabilitation		Onset to rehabilitation	
	Pearson *r*	*P* value	Pearson *r*	*P* value	Pearson *r*	*P* value	Pearson *r*	*P* value
ΔSFGS[a] resting symmetry score[b]	0.05	.87	0.44	.39	−0.75	.09	0.18	.73
ΔSFGS symmetry of voluntary movement score[c]	0.17	.52	−0.37	.47	0.8	.06	−0.1	.85
ΔSFGS synkinesis score	0.10	.70	−0.3	.57	0.046	.93	−0.28	.59
ΔSFGS composite score	0.1	.72	−0.43	.39	*0.94* [d]	*.005*	−0.11	.83
ΔSFGS with MassCon[e] symmetry of voluntary movement score	*0.50*	*.05*	−0.31	.54	−0.24	.65	−0.39	.44
ΔSFGS with MassCon composite score	0.34	.19	−0.54	.27	0.25	.64	−0.44	.38
ΔFDI[f] physical function subscale	0.43	.10	−0.04	.94	−0.36	.48	−0.16	.76
ΔFDI social and well-being function subscale	0.46	.07	−0.4	.44	−0.59	.22	−0.59	.22
ΔBAI[g] total score	−0.11	.70	0.32	.54	0.55	.26	0.49	.32

[a]SFGS: Sunnybrook Facial Grading System.

[b]Variation of outcome scores, calculated with the following formula: Δ=variation of outcome scores, as for the formula: Δ score=T1 score−T0 score.

[c]Posttreatment time point (T1) score − baseline (T0) score.

[d]Italicized values are significant.

[e]MassCon: bilateral masseter contraction.

[f]FDI: Facial Disability Index.

[g]BAI: Beck Anxiety Inventory.

Discussion

Principal Findings

Despite the pivotal importance of PT, the most effective PT approach after surgically treated peripheral FNP is still unclear, with a telerehabilitation-based emotional training protocol showing promising foundations given the specific characteristics of the facial impairments. In this exploratory pilot study, we confirmed the feasibility and potential clinical value of a telerehabilitation-based emotional training protocol in improving motor symptoms, facial disability, and anxiety in patients with peripheral FNP treated with triple innervation surgery also when compared with an in-person approach. However, these findings should be considered preliminary.

Emotional training is a novel therapeutic approach based on multiple phases that has shown preliminary promising results in the motor rehabilitation of FNP [30]. The first phase comprises a sensory-based intervention in which patients are asked to recognize the sensations induced by external stimuli on the lesioned hemiface (eg, intrinsic characteristics of objects or surfaces and passive movements performed by the physiotherapist) and perform facial movements—or imagine them, given that both motor imagery and actual movements activate the same neural circuits [48,49]. The rationale behind this first phase relies on the theories of motor control [55] that conceptualize the normal control of movement as a cognitive task [32] in which individuals integrate sensory inputs (feedback control), prior knowledge, and volitional intention (feedforward control) to select appropriate motor programs and regulate motor

XSL•FO
RenderX

output [56,57]. As reported by numerous studies, sensory stimulation reshapes cortical motor representations [58] given the interplay between motor (primary motor cortex, supplementary motor area, and premotor cortex) and somatosensory (primary and secondary somatosensory cortices and the sensory thalamus) regions [59,60]. In neurological motor rehabilitation, sensory-based interventions have been shown to enhance adaptive motor cortical plasticity and promote functional recovery [61-63]. Nevertheless, cognitive processes (eg, information perception and the elaboration, selection, and execution of motor plans) [31,64] play a critical role in mediating these mechanisms, supporting functional interactions between the body and the environment [65,66]. In addition, combining different types of sensory information stimulates the activity of associative brain areas that are important for learning, neuronal plasticity, and recovery [67]. Although evidence remains limited [64,68-71], promising findings suggest that sensory-based interventions can significantly improve the ability to perform activities of daily living and enhance quality of life [71] even for people with peripheral FNP [72,73].

The second phase aims to restore authentic facial expressions, emotional congruence, emotional symmetry, and spontaneity [74], which is crucial for optimal motor recovery [75]. At the neurophysiological level, facial movements in response to emotions are mediated by a widespread neural network involving various cortical regions that project to the facial nucleus, integrating motor planning, somatosensory feedback, and limbic system activity [76,77]. While corticofacial projections from the lateral frontal lobe contribute to the voluntary control of facial movements [24], the cingulate cortex plays a pivotal role in emotional expression through facial mimics, with projections from the amygdala providing the necessary emotional context [24]. For example, among the 2 different neural pathways controlling laughter [25], one, which involves primarily the frontal operculum [78,79] and the primary motor cortex, represents the neural basis of voluntary, nonemotional, conversational laughter; the other, which includes the pregenual anterior cingulate [80,81], ventral striatum and nucleus accumbens, amygdala, and hypothalamus [82], is involved in the production of emotional laughter, with the supplementary motor area likely modulating the production of both [27]. Notably, training voluntary movements to express different emotions helps reconnect the limbic-motor pathways [24], supporting the synergy between voluntary and emotional facial movements [24]. This is particularly important for facial nerve lesions, where voluntary movement and emotional expression are often dissociated [83]. The final phase of rehabilitation focuses on functional exercises designed to restore the ability to perform essential daily activities. These functional tasks, such as eating, speaking clearly, or demonstrating facial expressiveness during conversations, involve complex and coordinated facial muscle activity [84]—which is why these exercises are included as the final part of the treatment. By training patients in real-life contexts, they are engaged in meaningful activities that mirror their daily life. This approach has been shown to enhance neuroplasticity and functional recovery through task-specific repetition [85]—a principle known as “use-dependent cortical reorganization” [86]. In addition, focusing on real-world tasks aligns rehabilitation with the patient’s meaningful goals, improving motivation and adherence to the treatment protocol [87-89].

Taken together, our findings reinforce the conclusions of several systematic reviews on the feasibility of telerehabilitation in neurological conditions [33,90]. Recently, a randomized controlled trial investigated the effectiveness of video exercise–based telerehabilitation on motor and psychological symptoms in patients with peripheral FNP [35]. Patients in the telerehabilitation group not only found the treatment satisfactory and accessible but also demonstrated a significant improvement in FDI scores comparable to those receiving routine in-person care. Furthermore, both groups exhibited similar improvements in anxiety after treatment [35]. Despite the limitations of our study design, our findings are consistent with this evidence, suggesting the need to expand upon the value of telerehabilitation services for peripheral FNP, also considering that both patients and therapists generally have a positive attitude toward the integration of digital technologies [91,92]. Following the telerehabilitation-based emotional training, our patients reported improvements comparable to those of the face-to-face treatment in facial nerve function and facial symmetries (both at rest and during voluntary movement) even when considering the modified version of the SFGS (SFGS with MassCon). This is consistent with the only study assessing the effectiveness of emotional training in peripheral FNP, although the latter was performed in an in-person fashion [30]. Indeed, the authors reported that 50 patients who underwent face-to-face emotional training had improved House-Brackmann scale scores and symmetry of the face both at rest and during movements. In addition, in contrast to our study, only patients with iatrogenic peripheral FNP but not those who underwent surgery had gains in social function and subjectively perceived well-being. In our study, all patients experienced significant improvements in social and well-being function at the end of the treatment both in the in-person and online groups. This result was further confirmed by improved anxiety assessments. This is of particular importance given the substantial psychosocial burden that patients with facial impairments experience [93]. Nevertheless, the nonrandomized allocation and limited sample size warrant a cautious interpretation of these positive findings. Moreover, comparisons should be interpreted cautiously as the aforementioned study [30] was a retrospective analysis with no control group and the authors performed an in-person treatment protocol. However, in this study, self-reported physical function was greater at the end of the in-person treatment when compared to telerehabilitation treatment. At least for physical functioning, the therapeutic alliance and the therapist-patient relationship, which are strengthened by the direct, in-person interaction, and the emotional bond inherent to in-person rehabilitation may have had a positive effect, as previously suggested [33,94].

This study has some limitations. First, the small sample size may have underpowered the study and increased the risk of inflated effect estimates (eg, from correlations), whereas the use of subjective and operator-dependent assessments might have affected the results and interpretation. The nonrandomized allocation could have affected the internal validity of the study, whereas the absence of a follow-up assessment prevents conclusions regarding the durability of the observed

XSL•FO
RenderX

improvements. In addition, our neurophysiological hypotheses should be confirmed through neuroimaging studies, which were not feasible at this time. Second, although our patients were asked to maintain a stable pharmacological treatment throughout the study, the pharmacological therapies of the patients were not documented. The different biochemical effects induced by the medications could have influenced motor recovery, as previously suggested [95,96]. Similarly, we included surgically treated patients with peripheral FNP with multiple palsy etiologies (eg, neurinoma and trauma). The different underlying mechanisms of facial dysfunction might have affected the results and reduced the external validity of our findings.

Clinical Implications

The findings of this pilot study indicate that emotional training could be integrated into routine postoperative care for patients with peripheral FNP after triple innervation surgery, offering measurable benefits in facial symmetry, voluntary motor performance, social functioning, and anxiety reduction. The intervention's emphasis on both motor and psychosocial domains aligns well with the complex functional demands of patients undergoing facial reanimation, who often experience challenges that extend beyond motor impairment alone. Interestingly, the comparable improvements observed with the in-person protocol and telerehabilitation delivery suggest that emotional training can be implemented flexibly without compromising clinical outcomes, providing a viable option for individuals with mobility limitations, work constraints, or psychosocial discomfort associated with facial impairment. The structured nature of emotional training, centered on restoring selective motor control, improving emotional expression, and reducing maladaptive compensatory patterns, offers clinicians a reproducible framework that can be tailored to individual patient needs. Clinicians may consider telerehabilitation as a practical strategy to maintain continuity of care and reduce barriers to treatment access.

Conclusions

In conclusion, this exploratory pilot study provides preliminary evidence that telerehabilitation-based emotional training may have beneficial effects on motor, physical, and psychosocial functioning, as well as on anxiety, in surgically treated patients with peripheral FNP. Given the study's observational design and limited sample size, these findings should be interpreted cautiously and considered hypothesis generating rather than confirmatory. Further studies with optimized designs (eg, double-blind, parallel-group trials), larger sample sizes, and more objective and informative assessments (such as neurophysiological or neuroimaging evaluations) are warranted to confirm and extend these findings.

Acknowledgments

The authors thank all the enrolled participants who kindly gave their consent for the scientific use of their clinical data.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: MG, SC, RP
Data curation: MG, SC, JR, RP
Formal analysis: MG, SM, NVM
Investigation: SC, JR, FB
Methodology: MG, SC, FB
Supervision: RP, AP, AMP
Visualization: MG, SM, NVM
Writing—original draft: MG, JR, SM, AP, AMP
Writing—review and editing: MG, SC, JR, NVM, SM, RP, FB, AP, AMP
All authors read and approved the submitted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed, structured session plan with protocol phases, specific exercises, session organization, approximate durations, and progression criteria.
[PDF File, 110 KB - rehab_v13i1e79520_app1.pdf]

XSL•FO
RenderX

References

1. Psillas G, Antoniades E, Ieridou F, Constantinidis J. Facial nerve palsy in children: a retrospective study of 124 cases. J Paediatr Child Health 2019 Mar;55(3):299-304. [doi: 10.1111/jpc.14190] [Medline: 30144181]
2. Shafshak TS. The treatment of facial palsy from the point of view of physical and rehabilitation medicine. Eura Medicophys 2006 Mar;42(1):41-47. [Medline: 16565685]
3. Kleiss IJ, Hohman MH, Susarla SM, Marres HAM, Hadlock TA. Health-related quality of life in 794 patients with a peripheral facial palsy using the FaCE Scale: a retrospective cohort study. Clin Otolaryngol 2015 Dec;40(6):651-656. [doi: 10.1111/coa.12434] [Medline: 25858429]
4. Diels HJ, Combs D. Neuromuscular retraining for facial paralysis. Otolaryngol Clin North Am 1997 Oct;30(5):727-743. [Medline: 9295250]
5. Rotondo M, D'Avanzo R, Natale M, Conforti R, Pascale M, Scuotto A. Post-traumatic peripheral facial nerve palsy: surgical and neuroradiological consideration in five cases of delayed onset. Acta Neurochir (Wien) 2010 Oct;152(10):1705-1709. [doi: 10.1007/s00701-010-0747-x] [Medline: 20661598]
6. Finsterer J. Management of peripheral facial nerve palsy. Eur Arch Otorhinolaryngol 2008 Jul;265(7):743-752. [doi: 10.1007/s00405-008-0646-4] [Medline: 18368417]
7. Roob G, Fazekas F, Hartung HP. Peripheral facial palsy: etiology, diagnosis and treatment. Eur Neurol 1999 Jan;41(1):3-9. [doi: 10.1159/000007990] [Medline: 9885321]
8. Biglioli F, Colombo V, Pedrazzoli M, et al. Thoracodorsal nerve graft for reconstruction of facial nerve branching. J Craniomaxillofac Surg 2014 Jan;42(1):e8-e14. [doi: 10.1016/j.jcms.2013.03.001] [Medline: 23615388]
9. Paternostro-Sluga T, Herceg M, Frey M. Conservative treatment and rehabilitation in peripheral facial palsy [Article in German]. Handchir Mikrochir Plast Chir 2010 Apr;42(2):109-114. [doi: 10.1055/s-0029-1243215] [Medline: 20200817]
10. Ross B, Nedzelski JM, McLean JA. Efficacy of feedback training in long-standing facial nerve paresis. Laryngoscope 1991 Jul;101(7 Pt 1):744-750. [doi: 10.1288/00005537-199107000-00009] [Medline: 2062155]
11. Segal B, Zompa I, Danys I, et al. Symmetry and synkinesis during rehabilitation of unilateral facial paralysis. J Otolaryngol 1995 Jun;24(3):143-148. [Medline: 7674438]
12. Infante-Cossio P, Prats-Golczer VE, Lopez-Martos R, Montes-Latorre E, Exposito-Tirado JA, Gonzalez-Cardero E. Effectiveness of facial exercise therapy for facial nerve dysfunction after superficial parotidectomy: a randomized controlled trial. Clin Rehabil 2016 Nov;30(11):1097-1107. [doi: 10.1177/0269215515617309] [Medline: 26589401]
13. Wamkpah NS, Jeanpierre L, Lieu JE, Del Toro D, Simon LE, Chi JJ. Physical therapy for iatrogenic facial paralysis: a systematic review. JAMA Otolaryngol Head Neck Surg 2020 Nov 1;146(11):1065-1072. [doi: 10.1001/jamaoto.2020.3049] [Medline: 32970128]
14. Biglioli F, Allevi F, Rabbiosi D, et al. Triple innervation for re-animation of recent facial paralysis. J Craniomaxillofac Surg 2018 May;46(5):851-857. [doi: 10.1016/j.jcms.2018.02.014] [Medline: 29598895]
15. Biglioli F. Facial reanimations: part I--recent paralyses. Br J Oral Maxillofac Surg 2015 Dec;53(10):901-906. [doi: 10.1016/j.bjoms.2015.06.023] [Medline: 26188934]
16. Gibelli D, Tarabbia F, Restelli S, et al. Three-dimensional assessment of restored smiling mobility after reanimation of unilateral facial palsy by triple innervation technique. Int J Oral Maxillofac Surg 2020 Apr;49(4):536-542. [doi: 10.1016/j.ijom.2019.07.015] [Medline: 31447219]
17. Sforza C, Ulaj E, Gibelli DM, et al. Three-dimensional superimposition for patients with facial palsy: an innovative method for assessing the success of facial reanimation procedures. Br J Oral Maxillofac Surg 2018 Jan;56(1):3-7. [doi: 10.1016/j.bjoms.2017.11.015] [Medline: 29223635]
18. Barbara M, Monini S, Buffoni A, et al. Early rehabilitation of facial nerve deficit after acoustic neuroma surgery. Acta Otolaryngol 2003 Oct;123(8):932-935. [doi: 10.1080/00016480310000629] [Medline: 14606595]
19. Hussain SS, Winterburn SJ, Grace AR. Eutrophic electrical stimulation in long-standing facial palsy. In: Stennert ER, Kreutzberg GW, Michel O, Jungehülsing M, editors. The Facial Nerve: Springer; 1994. [doi: 10.1007/978-3-642-85090-5_40]
20. Choi BK, Goh RC, Chen PK, Chuang DC, Lo LJ, Chen YR. Facial nerve palsy after sagittal split ramus osteotomy of the mandible: mechanism and outcomes. J Oral Maxillofac Surg 2010 Jul;68(7):1615-1621. [doi: 10.1016/j.joms.2010.01.010] [Medline: 20434255]
21. Brach JS, VanSwearingen JM, Lenert J, Johnson PC. Facial neuromuscular retraining for oral synkinesis. Plast Reconstr Surg 1997 Jun;99(7):1922-1931. [doi: 10.1097/00006534-199706000-00017] [Medline: 9180715]
22. VanSwearingen JM, Brach JS. Changes in facial movement and synkinesis with facial neuromuscular reeducation. Plast Reconstr Surg 2003 Jun;111(7):2370-2375. [doi: 10.1097/01.PRS.0000061007.36637.88] [Medline: 12794483]
23. Pourmomeny AA, Asadi S. Management of synkinesis and asymmetry in facial nerve palsy: a review article. Iran J Otorhinolaryngol 2014 Oct;26(77):251-256. [Medline: 25320703]
24. Gothard KM. The amygdalo-motor pathways and the control of facial expressions. Front Neurosci 2014;8:43. [doi: 10.3389/fnins.2014.00043] [Medline: 24678289]
25. Gerbella M, Pinardi C, Di Cesare G, Rizzolatti G, Caruana F. Two neural networks for laughter: a tractography study. Cereb Cortex 2021 Jan 5;31(2):899-916. [doi: 10.1093/cercor/bhaa264] [Medline: 32969467]

XSL•FO
RenderX

26. Caruana F, Avanzini P, Gozzo F, Pelliccia V, Casaceli G, Rizzolatti G. A mirror mechanism for smiling in the anterior cingulate cortex. Emotion 2017 Mar;17(2):187-190. [doi: 10.1037/emo0000237] [Medline: 27854442]
27. Ferrari PF, Gerbella M, Coudé G, Rozzi S. Two different mirror neuron networks: the sensorimotor (hand) and limbic (face) pathways. Neuroscience 2017 Sep 1;358:300-315. [doi: 10.1016/j.neuroscience.2017.06.052] [Medline: 28687313]
28. Krane NA, Genther D, Weierich K, et al. Degree of self-reported facial impairment correlates with social impairment in individuals with facial paralysis and synkinesis. Facial Plast Surg Aesthet Med 2020;22(5):362-369. [doi: 10.1089/fpsam.2020.0082]
29. Norris JH, Longmire NM, Kilcoyne S, Johnson D, Fitzpatrick R, Klassen AF. Exploring patient experience of facial nerve palsy to inform the development of a PROM. Plast Reconstr Surg Glob Open 2019 Jan 9;7(1):e2072. [doi: 10.1097/GOX.0000000000002072] [Medline: 30859034]
30. Gervasoni F, LoMauro A, Ricci V, et al. Emotional training after facial nerve palsy: let's smile again!. NeuroRehabilitation 2025 Aug;57(1):111-121. [doi: 10.1177/10538135251335123] [Medline: 40432503]
31. Metzger JC, Lambercy O, Califfi A, Conti FM, Gassert R. Neurocognitive robot-assisted therapy of hand function. IEEE Trans Haptics 2014;7(2):140-149. [doi: 10.1109/TOH.2013.72] [Medline: 24968378]
32. Perfetti C. Rieducazione Motoria Dell'emiplegico: Libraio Ghedini Editore; 1979. URL: https://www.libraioghedini.it/neurologia/1149-rieducazione-motoria-dell-emiplegico.html [accessed 2026-04-06]
33. Suso-Martí L, La Touche R, Herranz-Gómez A, Angulo-Díaz-Parreño S, Paris-Alemany A, Cuenca-Martínez F. Effectiveness of telerehabilitation in physical therapist practice: an umbrella and mapping review with meta-meta-analysis. Phys Ther 2021 May 4;101(5):pzab075. [doi: 10.1093/ptj/pzab075] [Medline: 33611598]
34. de Sire A, Marotta N, Agostini F, et al. A telerehabilitation approach to chronic facial paralysis in the COVID-19 pandemic scenario: what role for electromyography assessment? J Pers Med 2022 Mar 19;12(3):497. [doi: 10.3390/jpm12030497] [Medline: 35330496]
35. Özden F, Golcuk Y, Tümtürk İ, Özkeskin M. The effects of telerehabilitation-based exercise therapy on motor and non-motor clinical outcomes in adults with facial palsy: a randomized controlled trial. Percept Mot Skills 2024 Dec;131(6):2182-2198. [doi: 10.1177/00315125241284677] [Medline: 39351643]
36. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS. Telerehabilitation: review of the state-of-the-art and areas of application. JMIR Rehabil Assist Technol 2017 Jul 21;4(2):e7. [doi: 10.2196/rehab.7511] [Medline: 28733271]
37. Maiorana NV, Aiello EN, Poletti B, et al. Relationship between reaction times and post-COVID-19 symptoms assessed by a web-based visual detection task. Healthcare (Basel) 2023 Jan 17;11(3):284. [doi: 10.3390/healthcare11030284] [Medline: 36766859]
38. Guidetti M, Averna A, Castellini G, et al. Physical activity during COVID-19 lockdown: data from an Italian survey. Healthcare (Basel) 2021 Apr 28;9(5):513. [doi: 10.3390/healthcare9050513] [Medline: 33925257]
39. Varelas EA, Gidumal S, Verma H, Vujovic D, Rosenberg JD, Gray M. Physical therapy for idiopathic facial paralysis: a systematic review. Am J Otolaryngol 2025;46(1):104511. [doi: 10.1016/j.amjoto.2024.104511] [Medline: 39647183]
40. André Texeira Iora M, Rodrigues Teixeira Netto M, Porto Cardoso C, et al. Effectiveness of hypoglossal-facial anastomosis in the rehabilitation of facial paralysis following vestibular schwannoma surgery: a systematic review. Cureus 2024 Apr;16(4):e57625. [doi: 10.7759/cureus.57625] [Medline: 38707182]
41. Buendia J, Loayza FR, Luis EO, Celorrio M, Pastor MA, Hontanilla B. Functional and anatomical basis for brain plasticity in facial palsy rehabilitation using the masseteric nerve. J Plast Reconstr Aesthet Surg 2016 Mar;69(3):417-426. [doi: 10.1016/j.bjps.2015.10.033] [Medline: 26683008]
42. Manktelow RT, Tomat LR, Zuker RM, Chang M. Smile reconstruction in adults with free muscle transfer innervated by the masseter motor nerve: effectiveness and cerebral adaptation. Plast Reconstr Surg 2006 Sep;118(4):885-899. [doi: 10.1097/01.prs.0000232195.20293.bd] [Medline: 16980848]
43. Hontanilla B, Cabello A. Spontaneity of smile after facial paralysis rehabilitation when using a non-facial donor nerve. J Craniomaxillofac Surg 2016 Sep;44(9):1305-1309. [doi: 10.1016/j.jcms.2016.06.031] [Medline: 27460946]
44. Pereira LM, Obara K, Dias JM, Menacho MO, Lavado EL, Cardoso JR. Facial exercise therapy for facial palsy: systematic review and meta-analysis. Clin Rehabil 2011 Jul;25(7):649-658. [doi: 10.1177/0269215510395634] [Medline: 21382865]
45. Ajimsha MS, Al-Mudahka NR, Al-Madzhar JA. Effectiveness of myofascial release: systematic review of randomized controlled trials. J Bodyw Mov Ther 2015 Jan;19(1):102-112. [doi: 10.1016/j.jbmt.2014.06.001] [Medline: 25603749]
46. Paolucci T, Cardarola A, Colonnelli P, et al. Give me a kiss! An integrative rehabilitative training program with motor imagery and mirror therapy for recovery of facial palsy. Eur J Phys Rehabil Med 2020 Feb;56(1):58-67. [doi: 10.23736/S1973-9087.19.05757-5] [Medline: 30916916]
47. Wilson CM, Ronan SL. Rehabilitation postfacial reanimation surgery after removal of acoustic neuroma: a case study. J Neurol Phys Ther 2010 Mar;34(1):41-49. [doi: 10.1097/NPT.0b013e3181cfc324] [Medline: 20212367]
48. Mulder T. Motor imagery and action observation: cognitive tools for rehabilitation. J Neural Transm (Vienna) 2007;114(10):1265-1278. [doi: 10.1007/s00702-007-0763-z] [Medline: 17579805]
49. Mustile M, Kourtis D, Edwards MG, Donaldson DI, Ietswaart M. Neural correlates of motor imagery and execution in real-world dynamic behavior: evidence for similarities and differences. Front Hum Neurosci 2024;18:1412307. [doi: 10.3389/fnhum.2024.1412307] [Medline: 38974480]

XSL•FO
RenderX

50. Osgood CE. Dimensionality of the semantic space for communication via facial expressions. Scand J Psychol 1966;7(1):1-30. [doi: 10.1111/j.1467-9450.1966.tb01334.x] [Medline: 5908205]
51. Pavese C, Tinelli C, Furini F, et al. Validation of the Italian version of the Sunnybrook Facial Grading System. Neurol Sci 2013 Apr;34(4):457-463. [doi: 10.1007/s10072-012-1025-x] [Medline: 22487815]
52. Pavese C, Cecini M, Camerino N, et al. Functional and social limitations after facial palsy: expanded and independent validation of the Italian version of the facial disability index. Phys Ther 2014 Sep;94(9):1327-1336. [doi: 10.2522/ptj.20130254] [Medline: 24786937]
53. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]
54. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. Ann Card Anaesth 2019;22(1):67-72. [doi: 10.4103/aca.ACA_157_18] [Medline: 30648682]
55. Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. Science 1995 Sep 29;269(5232):1880-1882. [doi: 10.1126/science.7569931] [Medline: 7569931]
56. Lee SA, Lee HS. Practical and methodological new cognitive exercise therapy. J Rehabil Res 2005;9:142-164.
57. Newell KM. Motor skill acquisition. Annu Rev Psychol 1991;42:213-237. [doi: 10.1146/annurev.ps.42.020191.001241] [Medline: 2018394]
58. Sanes JN, Suner S, Donoghue JP. Dynamic organization of primary motor cortex output to target muscles in adult rats. I. Long-term patterns of reorganization following motor or mixed peripheral nerve lesions. Exp Brain Res 1990;79(3):479-491. [doi: 10.1007/BF00229318] [Medline: 2340868]
59. Romo R, Ruiz S, Crespo P, Zainos A, Merchant H. Representation of tactile signals in primate supplementary motor area. J Neurophysiol 1993 Dec;70(6):2690-2694. [doi: 10.1152/jn.1993.70.6.2690] [Medline: 8120609]
60. Nudo RJ, Milliken GW, Jenkins WM, Merzenich MM. Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. J Neurosci 1996 Jan 15;16(2):785-807. [doi: 10.1523/JNEUROSCI.16-02-00785.1996] [Medline: 8551360]
61. Serrada I, Hordacre B, Hillier SL. Does sensory retraining improve sensation and sensorimotor function following stroke: a systematic review and meta-analysis. Front Neurosci 2019;13:402. [doi: 10.3389/fnins.2019.00402] [Medline: 31114472]
62. Carrico C, Chelette KC 2nd, Westgate PM, et al. Nerve stimulation enhances task-oriented training in chronic, severe motor deficit after stroke: a randomized trial. Stroke 2016 Jul;47(7):1879-1884. [doi: 10.1161/STROKEAHA.116.012671] [Medline: 27188405]
63. Castel-Lacanal E, Marque P, Tardy J, et al. Induction of cortical plastic changes in wrist muscles by paired associative stimulation in the recovery phase of stroke patients. Neurorehabil Neural Repair 2009 May;23(4):366-372. [doi: 10.1177/1545968308322841] [Medline: 19060132]
64. Chanubol R, Wongphaet P, Chavanich N, et al. A randomized controlled trial of Cognitive Sensory Motor Training Therapy on the recovery of arm function in acute stroke patients. Clin Rehabil 2012 Dec;26(12):1096-1104. [doi: 10.1177/0269215512444631] [Medline: 22649162]
65. McEwen SE, Huijbregts MP, Ryan JD, Polatajko HJ. Cognitive strategy use to enhance motor skill acquisition post-stroke: a critical review. Brain Inj 2009 Apr;23(4):263-277. [doi: 10.1080/02699050902788493] [Medline: 19277922]
66. Turville M, Carey LM, Matyas TA, Blennerhassett J. Change in functional arm use is associated with somatosensory skills after sensory retraining poststroke. Am J Occup Ther 2017;71(3):7103190070p1-7103190070p9. [doi: 10.5014/ajot.2017.024950] [Medline: 28422633]
67. Van de Winckel A, Wenderoth N, De Weerdt W, et al. Frontoparietal involvement in passively guided shape and length discrimination: a comparison between subcortical stroke patients and healthy controls. Exp Brain Res 2012 Jul;220(2):179-189. [doi: 10.1007/s00221-012-3128-2] [Medline: 22648204]
68. Sallés L, Martín-Casas P, Gironès X, Durà MJ, Lafuente JV, Perfetti C. A neurocognitive approach for recovering upper extremity movement following subacute stroke: a randomized controlled pilot study. J Phys Ther Sci 2017 Apr;29(4):665-672. [doi: 10.1589/jpts.29.665] [Medline: 28533607]
69. Morreale M, Marchione P, Pili A, et al. Early versus delayed rehabilitation treatment in hemiplegic patients with ischemic stroke: proprioceptive or cognitive approach? Eur J Phys Rehabil Med 2016 Feb;52(1):81-89. [Medline: 26220327]
70. Carey L, Macdonell R, Matyas TA. SENSe: Study of the Effectiveness of Neurorehabilitation on Sensation: a randomized controlled trial. Neurorehabil Neural Repair 2011 May;25(4):304-313. [doi: 10.1177/1545968310397705] [Medline: 21350049]
71. Lee S, Bae S, Jeon D, Kim KY. The effects of cognitive exercise therapy on chronic stroke patients' upper limb functions, activities of daily living and quality of life. J Phys Ther Sci 2015;27(9):2787-2791. [doi: 10.1589/jpts.27.2787] [Medline: 26504294]
72. Calabrese M, Ciriello M, Calabrese M. The role of rehabilitation in peripheral paralysis of the facial nerve: a case report of a patient treated with a neuro-cognitive approach. Glob Clin Case Rep 2020 May 29;2(2):1-9. [doi: 10.31487/j.GCCR.2020.01.02]
73. Cavallaro F, Portaro S, Pintaudi T, Ceccio M, Alito A. Remote cognitive therapeutic exercise in facial nerve palsy rehabilitation: pandemic tips and tricks. Innov Clin Neurosci 2023;20(1-3):10-12. [Medline: 37122569]

XSL•FO
RenderX

74. Bress KS, Cascio CJ. Sensorimotor regulation of facial expression - an untouched frontier. Neurosci Biobehav Rev 2024 Jul;162:105684. [doi: 10.1016/j.neubiorev.2024.105684] [Medline: 38710425]
75. Coulson SE, O'dwyer NJ, Adams RD, Croxson GR. Expression of emotion and quality of life after facial nerve paralysis. Otol Neurotol 2004 Nov;25(6):1014-1019. [doi: 10.1097/00129492-200411000-00026] [Medline: 15547436]
76. Müri RM. Cortical control of facial expression. J Comp Neurol 2016 Jun 1;524(8):1578-1585. [doi: 10.1002/cne.23908] [Medline: 26418049]
77. Maiorana N, Guidetti M, Dini M, Priori A, Ferrucci R. Cerebellar tDCS as therapy for cerebellar ataxias. Cerebellum 2022 Oct;21(5):755-761. [doi: 10.1007/s12311-021-01357-1] [Medline: 35060077]
78. Wattendorf E, Westermann B, Fiedler K, Kaza E, Lotze M, Celio MR. Exploration of the neural correlates of ticklish laughter by functional magnetic resonance imaging. Cereb Cortex 2013 Jun;23(6):1280-1289. [doi: 10.1093/cercor/bhs094] [Medline: 22508768]
79. Salas CE, Castro O, Yuen KS, Radovic D, d'Avossa G, Turnbull OH. "Just can't hide it": a behavioral and lesion study on emotional response modulation after right prefrontal damage. Soc Cogn Affect Neurosci 2016 Oct;11(10):1528-1540. [doi: 10.1093/scan/nsw075] [Medline: 27317928]
80. Caruana F, Avanzini P, Gozzo F, Francione S, Cardinale F, Rizzolatti G. Mirth and laughter elicited by electrical stimulation of the human anterior cingulate cortex. Cortex 2015 Oct;71:323-331. [doi: 10.1016/j.cortex.2015.07.024] [Medline: 26291664]
81. Caruana F, Gerbella M, Avanzini P, et al. Motor and emotional behaviours elicited by electrical stimulation of the human cingulate cortex. Brain 2018 Oct 1;141(10):3035-3051. [doi: 10.1093/brain/awy219] [Medline: 30107501]
82. Gibson WS, Cho S, Abulseoud OA, et al. The impact of mirth-inducing ventral striatal deep brain stimulation on functional and effective connectivity. Cereb Cortex 2017 Mar 1;27(3):2183-2194. [doi: 10.1093/cercor/bhw074] [Medline: 27001680]
83. Bradbury ET, Simons W, Sanders R. Psychological and social factors in reconstructive surgery for hemi-facial palsy. J Plast Reconstr Aesthet Surg 2006;59(3):272-278. [doi: 10.1016/j.bjps.2005.09.003] [Medline: 16673541]
84. Adamov L, Petrović B, Milić L, et al. Comparative analysis of electrical signals in facial expression muscles. Biomed Eng Online 2025 Feb 12;24(1):17. [doi: 10.1186/s12938-025-01350-3] [Medline: 39939995]
85. Hubbard IJ, Parsons MW, Neilson C, Carey LM. Task-specific training: evidence for and translation to clinical practice. Occup Ther Int 2009;16(3-4):175-189. [doi: 10.1002/oti.275] [Medline: 19504501]
86. Li Y, Hollis ER2. The role of motor network reorganization during rehabilitation. Neural Regen Res 2017 May;12(5):745-746. [doi: 10.4103/1673-5374.206641] [Medline: 28616027]
87. Littooij E, Doodeman S, Holla J, et al. Setting meaningful goals in rehabilitation: a qualitative study on the experiences of clients and clinicians in working with a practical tool. Clin Rehabil 2022 Mar;36(3):415-428. [doi: 10.1177/02692155211046463] [Medline: 34730459]
88. Crawford L, Maxwell J, Colquhoun H, et al. Facilitators and barriers to patient-centred goal-setting in rehabilitation: a scoping review. Clin Rehabil 2022 Dec;36(12):1694-1704. [doi: 10.1177/02692155221121006] [Medline: 36017567]
89. Kingston K, Jenkins D, Kingston G. Promoting adherence to rehabilitation through supporting patient well-being: a self-determination perspective. In: Zenko Z, Jones L, editors. Essentials of Exercise and Sport Psychology: An Open Access Textbook: Society for Transparency, Openness, and Replication in Kinesiology; 2021:759-782. [doi: 10.51224/B1032]
90. Federico S, Cacciante L, Cieślik B, et al. Telerehabilitation for neurological motor impairment: a systematic review and meta-analysis on quality of life, satisfaction, and acceptance in stroke, multiple sclerosis, and Parkinson's disease. J Clin Med 2024 Jan 4;13(1):299. [doi: 10.3390/jcm13010299] [Medline: 38202306]
91. Szczepura A, Holliday N, Neville C, et al. Raising the digital profile of facial palsy: national surveys of patients' and clinicians' experiences of changing UK treatment pathways and views on the future role of digital technology. J Med Internet Res 2020 Oct 5;22(10):e20406. [doi: 10.2196/20406] [Medline: 32763890]
92. Tan JR, Coulson S, Keep M. Face-to-face versus video assessment of facial paralysis: implications for telemedicine. J Med Internet Res 2019 Apr 12;21(4):e11109. [doi: 10.2196/11109] [Medline: 30977734]
93. Hotton M, Huggons E, Hamlet C, et al. The psychosocial impact of facial palsy: a systematic review. Br J Health Psychol 2020 Sep;25(3):695-727. [doi: 10.1111/bjhp.12440] [Medline: 32538540]
94. Hall AM, Ferreira PH, Maher CG, Latimer J, Ferreira ML. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. Phys Ther 2010 Aug;90(8):1099-1110. [doi: 10.2522/ptj.20090245] [Medline: 20576715]
95. Khan AJ, Szczepura A, Palmer S, et al. Physical therapy for facial nerve paralysis (Bell's palsy): an updated and extended systematic review of the evidence for facial exercise therapy. Clin Rehabil 2022 Nov;36(11):1424-1449. [doi: 10.1177/02692155221110727] [Medline: 35787015]
96. Shi J, Lu D, Chen H, et al. Efficacy and safety of pharmacological and physical therapies for Bell's palsy: a Bayesian network meta-analysis. Front Neurol 2022;13:868121. [doi: 10.3389/fneur.2022.868121] [Medline: 35528739]

Abbreviations

BAI: Beck Anxiety Inventory

XSL•FO
RenderX

CONSORT: Consolidated Standards of Reporting Trials
FDI: Facial Disability Index
FNP: facial nerve palsy
PT: physiotherapy
SFGS: Sunnybrook Facial Grading System
SFGS with MassCon: Sunnybrook Facial Grading System with bilateral masseter contraction

Edited by S Munce; submitted 23.Jun.2025; peer-reviewed by A Baroni, A Edalatpour, C Cordani, F Ozden, F Calvaruso; revised version received 10.Dec.2025; accepted 19.Dec.2025; published 27.Apr.2026.

Please cite as:
Guidetti M, Cupello S, Reali J, Maiorana NV, Marceglia S, Pagani R, Biglioli F, Priori A, Previtera AM
Emotional Training via Telerehabilitation After Surgical Treatment for Facial Palsy: Prospective, Assessor-Blinded, 2-Arm Pilot Cohort Study
JMIR Rehabil Assist Technol 2026;13:e79520
URL: https://rehab.jmir.org/2026/1/e79520
doi:10.2196/79520

© Matteo Guidetti, Silvia Cupello, Jacopo Reali, Natale Vincenzo Maiorana, Sara Marceglia, Rossella Pagani, Federico Biglioli, Alberto Priori, Antonino Michele Previtera. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Goal Setting During Telerehabilitation in Japan: Qualitative Study of Occupational Therapists' and Physical Therapists' Experiences Using Reflexive Thematic Analysis

Yuho Okita[1,2], BExSci(Hons), MOT; Kounosuke Tomori[3], PhD; Minoru Okita[2], PhD

[1]School of Health Sciences, Swinburne University of Technology, John Street, Hawthorn, Victoria, Australia

[2]Department of Physical Therapy Science, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

[3]School of Health Sciences, Department of Rehabilitation, Major of Occupational Therapy, Tokyo University of Technology, Tokyo, Japan

Corresponding Author:
Yuho Okita, BExSci(Hons), MOT
School of Health Sciences, Swinburne University of Technology, John Street, Hawthorn, Victoria, Australia

Abstract

Background: Telerehabilitation expanded rapidly during the COVID-19 pandemic and is increasingly used alongside, or as an alternative to, in-person rehabilitation. Goal setting is a core component of client-centered rehabilitation; however, how therapists conduct goal setting remotely remains poorly understood, particularly in East Asian contexts. To our knowledge, no qualitative study has specifically examined remote goal setting in telerehabilitation in Japan, where remote delivery was not yet routine when these data were collected.

Objective: This qualitative study aimed to explore how Japanese occupational therapists and physical therapists experienced negotiating and adapting rehabilitation goals during telerehabilitation, including how remote delivery shaped assessment, collaboration, and the purposes assigned to goals.

Methods: Individual semistructured interviews (18 - 51 minutes; mean 32 minutes) were conducted via Zoom (Zoom Communications, Inc) in Japanese with 9 therapists (7 occupational therapists and 2 physical therapists; 7 male and 2 female; aged 29 - 43 years) who had direct experience of remote goal setting. Participants were recruited by purposive and snowball sampling from 9 different workplaces across Japan, including metropolitan areas such as Tokyo and regional areas such as Hokkaido and Nagasaki, and worked in hospital, private-practice, community-based, and academic settings. Data were analyzed using reflexive thematic analysis within a contextualist epistemology, following Braun and Clarke. Trustworthiness was considered using Nowell and colleagues' criteria, and reporting followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.

Results: Four themes were constructed: (1) losing touch: uncertainty in remote assessment and goal setting (9/9 participants); (2) from direct intervention to coaching-oriented goal setting (8/9 participants); (3) at home, at ease: needs that surface in clients' own environments (6/9 participants); and (4) when continuity becomes the goal (6/9 participants). Therapists described remote goal setting as shaped by both constraint and possibility. Partial bodily, nonverbal, and contextual information reduced their confidence in assessment and goal negotiation, and they responded through preparation, visual resources, self-monitoring tools, and collaboration with clients and families. Sessions conducted in clients' own environments appeared to support comfort and sometimes made everyday needs more visible. When in-person care was unavailable or insufficient, therapists sometimes broadened goals from functional improvement alone to include continuity, ongoing monitoring, and prevention of disengagement.

Conclusions: In these therapists' accounts, remote goal setting was not simply a diminished version of in-person practice but a distinct clinical process requiring deliberate preparation, collaborative communication, and management of uncertainty. Hybrid pathways may combine in-person assessment, when hands-on examination is needed, with continued remote negotiation, monitoring, and revision of goals. Further research should incorporate clients' and families' perspectives and evaluate hybrid models of goal setting across diverse rehabilitation populations.

(JMIR Rehabil Assist Technol 2026;13:e101355) doi:10.2196/101355

KEYWORDS

telerehabilitation; goal setting; reflexive thematic analysis; occupational therapy; physical therapy

XSL•FO
RenderX

Introduction

Telerehabilitation and Remote Goal Setting

Telerehabilitation, defined as the delivery of rehabilitation services through information and communication technologies, such as videoconferencing, mobile apps, sensors, and telephone-based communication, has expanded rapidly in recent years [1,2]. This growth was particularly accelerated by the COVID-19 pandemic, which increased the uptake of remote health care across rehabilitation settings [3]. This expansion has created new opportunities to improve access to care and support continuity of rehabilitation services, particularly for people who face barriers to attending in-person appointments [4]. In addition to improving accessibility, telerehabilitation can offer other benefits, including greater convenience, reduced travel-related costs, improved adherence, and enhanced continuity of care [5]. Recent evidence indicates that telerehabilitation may have a positive role in supporting physical activity and functional outcomes among older adults. For example, a systematic review and meta-analysis of telerehabilitation-delivered exercise interventions for older adults with frailty, cognitive impairment, or mobility disability reported improvements in mobility, muscle strength, and balance [6]. Similarly, a recent meta-analysis of digital interventions for frail older adults found reductions in frailty scores and improvements in grip strength and health-related quality of life [7]. This evidence suggests that telerehabilitation may be a useful alternative or complement to in-person rehabilitation for diverse populations, including frail older adults and others with limited mobility [1,8]. However, its implementation remains constrained by several challenges, including limited digital literacy, difficulties in conducting physical assessments remotely, and insufficient training opportunities for clinicians [3]. Because telerehabilitation was initially adopted rapidly in response to urgent service needs, evidence regarding how to deliver client-centered rehabilitation in remote contexts remains limited [3].

Among the core processes of client-centered rehabilitation, reaching an agreement on goals and tasks is a central element of the therapeutic alliance that requires complex adaptations during the transition from in-person to remote service delivery [9]. Both client-centeredness and the therapeutic alliance are considered beneficial for rehabilitation: a stronger therapist-patient alliance has been associated with better outcomes in physical rehabilitation settings, including greater treatment adherence and satisfaction, improved physical function, and reduced depressive symptoms [10], and goal setting aligned with clients' own priorities may improve engagement, motivation, and psychosocial outcomes such as self-efficacy [11]. Goal setting is widely regarded as essential to client-centered practice [11]. It helps establish priorities, guide intervention, and align rehabilitation efforts with what matters most to clients in their everyday lives [12,13]. In principle, rehabilitation goal setting should be client-centered and involve shared decision-making; however, previous research suggests that this is not always achieved consistently in routine practice [13,14]. When therapists and clients meet in person, goal setting is typically informed by clinical observation, physical assessment, and nonverbal communication [15], and is supported by embodied interaction where meaning and goals are often co-constructed and confirmed through physical guidance and felt bodily changes [16]. In telerehabilitation, these elements are constrained, as physical distance and technology alter nonverbal communication and eliminate hands-on interaction. Although technology may support goal setting through functions such as monitoring, feedback, and communication, this does not in itself explain how goal setting is actually carried out in remote rehabilitation encounters [12,17]. A recent scoping review identified a range of goal-setting tools used in adult rehabilitation, yet highlighted limited evidence regarding their design or validation for remote use [12]. Another scoping review focusing specifically on telerehabilitation and goal setting identified 68 studies, but none were conducted in East Asian countries, including Japan [17]. Goal setting and shared decision-making within remotely delivered rehabilitation have begun to attract attention elsewhere, for example, among older adults starting mobile health cardiac rehabilitation in the United States [18] and in a trial of decision aid-supported shared decision-making within hybrid in-person and virtual pulmonary rehabilitation in China [19]; however, no qualitative study has examined therapists' experiences of goal setting in telerehabilitation in East Asia, including Japan. Examining how goal setting is experienced in telerehabilitation is therefore important, particularly in settings where it is not yet routine, as the challenges and adaptations of early implementation can become less visible once remote practice is normalized and embedded within established workflows.

Japanese Rehabilitation Context

In Japan, goal setting in rehabilitation continues to be strongly influenced by traditional medical-model and clinician-led approaches [20]. This has prompted ongoing discussion about how goal-setting practices can become more collaborative and client-centered [14]. A cross-sectional study involving clients and occupational therapists in Japan found that, although both groups reported high levels of client engagement in the goal-setting process, nearly 80% of goal statements were mismatched between occupational therapists and their clients [13]. This finding suggests that even when both parties perceive goal setting as collaborative, their understandings of the goals may not necessarily be aligned, highlighting the need for higher-quality goal-setting practices [13]. This may also reflect culturally specific factors, as rehabilitation in Japan still tends to focus predominantly on functional training rather than activity- and participation-level goals aligned with what is personally meaningful to the individual [21]. Consistent with this, a recent qualitative study on goal setting in home rehabilitation in Japan highlighted the need for more explicit explanation and more collaborative approaches tailored to the Japanese context [21].

Japan's rapidly aging population and geographic challenges point to the potential value of telerehabilitation for improving access to rehabilitation, particularly for people with mobility limitations, including frail older adults [1,22]. Regional disparities in rehabilitation provision continue to create uneven access across the country [23-25], and in this context,

XSL•FO
RenderX

telerehabilitation may be particularly relevant. Recent studies have reported improvements in balance, gait, and lower-limb strength following remotely delivered exercise programs for pre-frail and frail older adults [26], and a pilot study in Japan found improvements in health-related quality of life after a multiuser telerehabilitation exercise intervention for older adults with frailty [27]. However, the uptake of telerehabilitation in Japan remains limited compared with some other countries, partly because of regulatory and reimbursement constraints, and remote rehabilitation may represent a particularly significant shift in practice for Japanese therapists [20,28,29]. At the time of this study, telerehabilitation in Japan was not yet routine practice: it was typically introduced when in-person care was difficult or impossible, for example, during the COVID-19 pandemic, after hospital discharge, or in geographically underserved areas, and it was largely delivered outside standard reimbursement schemes [28,29]. The experiences examined in this study should therefore be understood against this backdrop of early, often improvised, adoption.

Study Aim

Considering these contextual factors, it is important to understand how therapists in Japan experience the delivery of telerehabilitation, including the goal-setting process. Such knowledge may help identify research priorities, inform clinical training, and support policy discussions on reimbursement and service delivery. Because telerehabilitation was not yet routine in Japan when these data were collected, therapists' accounts of this period can document challenges and adaptations that may become less visible once remote practice is normalized. Therefore, this qualitative study aimed to explore how Japanese occupational therapists and physical therapists experienced negotiating and adapting rehabilitation goals when delivering telerehabilitation across a range of clinical settings in Japan, including how remote delivery shaped assessment, collaboration, and the purposes assigned to goals. Because goal setting is embedded within the broader flow of remote practice, therapists' wider experiences of telerehabilitation were also explored to the extent that they shaped how goals were identified, negotiated, and followed up. To obtain a broad understanding that can inform future research, participants were not limited by practice setting or target population.

Methods

Study Design

This study used a qualitative descriptive design informed by reflexive thematic analysis (RTA), following the 6-phase approach articulated by Braun and Clarke [30,31]. Consistent with RTA, the study was positioned within a contextualist epistemology, which acknowledges that participants' accounts reflect situated experience shaped by cultural, institutional, and interpersonal contexts, and treats the researcher as the analytic instrument through which meaning is constructed rather than discovered [30-32]. Reporting was guided by the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Checklist 1) [33]. A reflexive thematic approach was selected because the study aimed to explore the experiences, perceptions, and meanings that therapists ascribed to remote goal setting in rehabilitation practice.

Ethical Considerations

Ethical approval for this study was obtained from the Ethics Committee of the Graduate School of Biomedical Sciences (Health Sciences), Nagasaki University (approval number: 22041401). All participants provided written informed consent before participation: each participant received a written information sheet by email as part of the study invitation (see Participants and Recruitment), and signed consent forms were returned by post, with the original forms stored securely at the second author's institution. Participants were informed that participation was entirely voluntary and that consent could be withdrawn at any time without disadvantage. Participants did not receive financial or other compensation for taking part in this study. Interviews were conducted in Japanese and transcribed verbatim. Participant anonymity was protected through the use of identification codes, and all direct identifiers were removed from transcripts. Video recordings and study data were stored securely and will be retained for 5 years in accordance with institutional policy.

Research Team and Reflexivity

The research team comprised 3 authors. The first author (YO) is a male occupational therapist holding BExSc(Hons) and MOT (Master of Occupational Therapy) qualifications at the time of the study, with clinical experience in Australia and formal training in qualitative research through doctoral study; YO conducted all interviews and led the analysis. YO had a research interest in client-centered practice and telerehabilitation prior to the study and approached telerehabilitation as promising but underevidenced in Japanese practice. The second author (KT) is an occupational therapist and university-based researcher with expertise in collaborative goal-setting tools and client-centered practice in Japan; KT contributed to study design, interview guide development, and analytic oversight. The third author (MO) is a physical therapist and university-based researcher with expertise in rehabilitation science in Japan; MO contributed to study design, participant recruitment, and analytic supervision. None of the authors had a prior clinical relationship with the participants at the time of interview; however, several participants were known to the authors through prior professional contact or professional-network introductions, and this was disclosed and discussed in team meetings as part of reflexivity. An audit trail of the sequential analytic process was maintained in a separate document, documenting initial analysis, integration of initial analysis, initial theme generation, and final theme development with illustrative quotes [34]. These materials were shared with all authors for review and discussion [34,35]. Consistent with RTA, interpretation was recognized as shaped by researcher subjectivity and not expected to be identical across researchers; team discussions were used to challenge interpretations, refine thematic boundaries, and strengthen the coherence, credibility, and acceptability of the final thematic structure [35,36].

XSL•FO
RenderX

Participants and Recruitment

Participants were recruited using purposive and snowball sampling to identify therapists with direct experience of remote goal setting in rehabilitation practice [37,38]. For the purposes of this study, remote goal setting was defined broadly to include remote discussions of clients' needs, priorities, and goals that informed rehabilitation planning. Licensed physical therapists and occupational therapists practicing in Japan who had experience with remote goal setting were eligible to participate. No additional exclusion criteria were applied.

Potential participants were identified through the professional networks of the research team and through participant referral. Each potential participant was contacted individually by email with an explanation of the study purpose and an invitation to take part. Each participant worked at a different organization; no two participants were employed at the same workplace at the time of the interview. This was a deliberate consequence of the sampling strategy and reduces the likelihood that the themes constructed reflect a single institutional culture. Participants were drawn from both metropolitan areas (eg, Tokyo) and regional areas of Japan (including Hokkaido and Nagasaki) and worked across a variety of settings, including hospital-based services, private practices, and community-based services. Information on participants' clinical setting and the geographic location of their practice was identified from interview content during the familiarization phase of analysis and recorded in the audit trail; this level of demographic information was not captured in the structured online questionnaire. For one participant (Participant 8), the region of practice could not be confirmed from these sources and is therefore not reported in Table 1.

Table . Participant characteristics (N=9)[a].

ID	Profession; sex; age (years)	Experience, years (remote)	Remote practice context	Region
P1	OT; male; 34	12 (<1)	Adults (19-65 years); hospital; neuromuscular disorders	Kanto
P2	OT; male; 38	16 (1)	Children (0 - 18 years); community; neurodevelopmental disorders	Hokkaido
P3	PT; male; 29	8 (6)	Adults (19-65 years); private practice; musculoskeletal conditions	Kanto
P4	PT; male; 42	21 (1)	Older adults (65+ years); hospital; musculoskeletal conditions	Kyushu
P5	OT; male; 39	17 (7[b])	Adults (19-65 years); community; psychosocial disabilities	Kansai
P6	OT; male; 37	14 (3)	Children (0 - 18 years); academic; neurodevelopmental disorders	Chubu
P7	OT; female; 41	19 (1)	Older adults (65+ years); hospital; musculoskeletal and neuromuscular conditions	Hokkaido
P8	OT; male; 43	19 (1)	Children (0 - 18 years); community; musculoskeletal and neuromuscular conditions	Not reported[c]
P9	OT; female; 31	12 (1)	Children (0 - 18 years); community; neurodevelopmental disorders	Kanto

[a]OT: occupational therapist; PT: physical therapist. Experience is reported as total years of therapist experience, with years of remote service experience in parentheses. Region indicates the broad geographic region of Japan in which the participant practiced, reported at this level to protect participant anonymity: Kanto (eastern Honshu, including the Greater Tokyo area), Chubu (central Honshu), Kansai (western Honshu, including Osaka and Kyoto), Kyushu (the southwestern main island, including Nagasaki), and Hokkaido (the northernmost main island).

[b]Approximately 7 years, with infrequent use.

[c]The region of practice for this participant could not be confirmed from the interview or questionnaire data and is therefore not reported.

Sample size was guided by the concept of information power rather than by a predetermined target or data saturation [39]. Applying Malterud and colleagues' [39] five criteria, we judged that the sample could support the study aim on the basis of (1) a narrow and focused aim (therapists' experiences of remote goal setting); (2) dense specificity of participants (licensed, experienced therapists with direct remote goal-setting experience, drawn from 9 workplaces across metropolitan and regional Japan and a variety of clinical settings); (3) an established conceptual grounding in client-centered goal setting;

XSL•FO
RenderX

(4) focused semistructured interviews conducted by a trained qualitative interviewer using a flexibly applied topic guide (see Data Collection); and (5) a cross-case, theme-oriented analytic strategy. Information power offers a structured basis for judgment rather than a guarantee of completeness, and the uncertainty inherent in this judgment is acknowledged in the Strengths and Limitations section.

Participant Characteristics

Nine therapists participated in the study, including 7 occupational therapists and 2 physical therapists. Of these, 7 were male and 2 were female. Participant ages ranged from 29 to 43 years, and years of therapist experience ranged from 8 to 21 years. Experience with remote service delivery ranged from less than 1 year to approximately 7 years. Participants had experience providing remote services to children, adults, and older adults across a range of clinical areas, including neurodevelopmental disorders in children, musculoskeletal disorders or injuries, neuromuscular disorders, and other conditions. Detailed participant characteristics are presented in Table 1.

Participants delivered remote services mainly by videoconference. Consistent with the early-adoption context described in the Introduction, they had typically introduced remote delivery when in-person care was difficult or impossible, for example, during pandemic-related service disruption, after hospital discharge, or for clients in geographically underserved areas. The frequency and duration of remote sessions were not collected systematically in the questionnaire; the following patterns were identified from interview content during the familiarization phase. Remote service formats were heterogeneous: they ranged from brief, repeated contacts of about 10-30 minutes to sessions of about an hour, delivered at rhythms from weekly to monthly, and several participants combined scheduled video sessions with asynchronous contact between sessions, such as email, messaging applications, activity diaries, and videos recorded and submitted by family members. Examples included weekly 30-40-minute videoconference sessions continuing a hospital program after discharge; a 3-month package of 5 video consultations (an initial 40-minute goal-setting session followed by 20-minute follow-ups) supplemented by unrestricted text chat; and regular 15‐30-minute video calls that complemented, rather than replaced, in-person visits. Because no standardized comparison with usual in-person schedules was collected, the study does not infer that remote delivery itself shortened sessions or changed their frequency.

Data Collection

Semistructured interviews were conducted individually via Zoom videoconference between September and October 2022. Interviews lasted between 18 and 51 minutes, with a mean duration of 32 minutes. Zoom was selected for its ease of use, cost-effectiveness, data management capabilities, and security features, which have been positively appraised in prior qualitative research [40]; known limitations of videoconferencing in qualitative research (eg, technical issues and partial loss of nonverbal cues) were considered acceptable given the nature of the study topic. The interview guide (Multimedia Appendix 1) was developed iteratively by the research team and included open-ended questions about therapists' experiences before, during, and after remote goal-setting encounters, including perceived advantages and barriers, adaptation strategies, and impacts on the therapeutic relationship. The guide was used flexibly rather than as a fixed script: questions served as points of departure for a conversation around each topic, the order of questions varied with the flow of each interview, and follow-up probes were used to invite elaboration, while the interviewer ensured that all topic areas were covered in every interview. The interview guide was not formally pilot tested; however, it was reviewed and refined through team discussion among the coauthors, who have established expertise in qualitative research methods and in rehabilitation goal setting. All interviews were video-recorded with participants' permission. No one other than the interviewer and the participant was present during the interviews, and no repeat interviews were conducted; each participant was interviewed once. Formal field notes were not maintained during or immediately after data collection. Instead, the interviewer (YO) reviewed each video recording in full against the corresponding transcript on multiple occasions during the familiarization phase and recorded reflective observations and interpretive notes in the audit trail. For example, notes made during theme development recorded that therapists repeatedly described relying on partial or secondhand information (eg, answers relayed only as "probably..."), and these notes informed the interpretation, developed in theme 1, that remote delivery reduced therapists' confidence in their clinical judgments.

Transcripts were produced in Japanese using AI-assisted first-pass transcription (Notta), after which YO manually reviewed each recording in full and corrected any errors against the audio and video. Analysis was conducted on the original Japanese transcripts to preserve contextual and cultural meaning. Illustrative quotations presented in this manuscript were translated from Japanese into English with the assistance of generative AI tools as a first-pass aid (see the Acknowledgments section), and were subsequently reviewed and corrected by the corresponding author (YO), a native Japanese speaker fluent in English, against the original Japanese audio and transcripts. Where culturally specific expressions had no direct English equivalent, the translation prioritized conceptual meaning over literal rendering, with the source Japanese preserved in the audit trail for transparency. Demographic and professional background information was collected through a brief online questionnaire covering profession, age, years of therapist experience, years of remote service experience, client age groups served, and broad diagnostic areas served remotely.

Data Analysis

Data were analyzed using RTA as articulated by Braun and Clarke [30,31]. The analysis was primarily inductive and data-driven, and no predetermined coding framework was applied. Coding stayed close to the explicit, semantic content of participants' accounts; more interpretive analysis was introduced in the later phases, for example, when accounts of hesitancy and reduced confidence were interpreted together as a disruption of professional certainty (theme 1). The reflective observations and interpretive notes described above contributed

XSL•FO
RenderX

to this interpretive work. YO led all analytic phases. In phase 1 (familiarization), YO read each transcript repeatedly while checking it against the interview recordings. In phase 2 (initial coding), YO generated codes across the dataset, primarily at the semantic level. In phase 3 (candidate theme generation), codes were grouped to identify broader patterns of meaning and construct candidate themes. In phase 4 (theme review), candidate themes were reviewed against both the coded extracts and the full dataset for internal coherence and external distinctiveness. In phase 5 (theme definition and naming), themes were refined and named to capture the central organizing concept of each. In phase 6 (report production), the analytic narrative was developed by selecting illustrative quotations and relating the final themes back to the study aim [41]. Analytic materials were shared with all authors throughout the process; discussions were used to challenge interpretations and refine thematic boundaries rather than to establish coder agreement, consistent with RTA. Data were managed using Microsoft Excel and Microsoft Word; no dedicated qualitative data analysis software was used, as the dataset comprised 9 transcripts and the audit trail was maintained in a shared spreadsheet that all authors could access directly, preserving the traceability of coding decisions.

Rigor and Trustworthiness

Trustworthiness was addressed through the criteria of credibility, transferability, dependability, and confirmability, following Nowell and colleagues' [42] application of these criteria to thematic analysis. Credibility refers to the confidence that can be placed in the fit between participants' accounts and the researchers' representation of them. Credibility was strengthened through prolonged engagement with the data, including repeated reading of transcripts alongside audio recordings, collaborative analytic discussions among the 3 authors, attention to divergent and less common cases, and retention of the original Japanese data alongside English translations [33,35]. Transferability refers to the extent to which findings may be applicable in other settings, which readers can judge only if the study context is described in sufficient detail. Transferability was supported by detailed reporting of the study setting, participant characteristics, and analytic decisions, as presented in Table 1 and the Methods section [33]. Dependability refers to whether the research process is logical, traceable, and clearly documented. Dependability was enhanced through an audit trail documenting analytic decisions across the 6 phases of analysis, maintained in a shared spreadsheet [34]. Confirmability refers to the demonstration that interpretations are clearly derived from the data rather than from researcher predisposition alone. Confirmability was supported by this audit trail, team review of the links between data extracts and final themes, and reflexive consideration of the research team's positioning, as described in the Research Team and Reflexivity section [34,35]. Participant transcript review, or member checking, was not undertaken. This decision was consistent with the study's contextualist stance, which treats meaning as co-constructed through the analytic process rather than as something to be verified retrospectively by participants [31,35].

Results

Overview

Four themes were constructed that described therapists' experiences of remote goal setting in rehabilitation: (1) losing touch: uncertainty in remote assessment and goal setting; (2) from direct intervention to coaching-oriented goal setting; (3) at home, at ease: needs that surface in clients' own environments; and (4) when continuity becomes the goal. Each theme is described in turn below, and the relationships among the themes are then summarized. Table 2 summarizes each theme and the number of participants whose accounts contributed to it. Within-theme variation was retained in the analysis through close attention to less prominent perspectives and is reflected in the selection of quotations across participants; no minor themes or divergent cases inconsistent with the 4 main themes were identified that warranted construction as a separate theme.

Table . Themes, central organizing concepts, and participant contributions (N=9)[a].

Theme	Central organizing concept	Participants contributing (n/9)
1. Losing touch: uncertainty in remote assessment and goal setting	Loss of embodied and contextual information disrupts clinical certainty.	9/9
2. From direct intervention to coaching-oriented goal setting	Remote constraints prompt a shift from provider to facilitator/coach.	8/9
3. At home, at ease: needs that surface in clients' own environments	Home-based interaction puts clients at ease and discloses context and needs less visibility in clinic settings.	6/9
4. When continuity becomes the goal	Goals are reoriented toward continuity through ongoing monitoring, structured follow-up, and prevention of decline.	6/9

[a]Counts are based on at least one substantive contribution to the theme; counts do not imply quantification of prevalence.

Theme 1: Losing Touch: Uncertainty in Remote Assessment and Goal Setting

The first theme captured therapists' experiences of uncertainty in assessment and goal setting when working remotely. All 9 participants described how the inability to physically touch clients and the reduced access to nonverbal and contextual information limited their confidence in understanding clients' conditions, needs, and priorities. In this study, "reduced visibility" did not mean that clients' homes were invisible;

XSL•FO
RenderX

rather, what therapists could see was restricted to the fragment of the environment and the body captured within the camera frame, shaped by camera angle, device placement, connection quality, and what the client or another person was able or willing to show. Unlike an in-person encounter, in which therapists could walk around, change their vantage point, and observe the whole body and setting, remote observation was partial and fixed. Therapists highlighted limitations in physical assessment and intervention, particularly in situations in which whole-body observation or hands-on assessment would normally inform clinical reasoning. One therapist stated, "It is difficult to understand everyday life scenes remotely...it is hard to picture daily life, or to imagine the actual situation just by listening" [Occupational therapist, participant 9]. Another reflected on the limits of remote postural assessment, explaining that a client's sitting posture might look acceptable during a brief observation on screen but deteriorate over a longer period, something that could be detected only through sustained in-person observation: "With postural adjustment, there are aspects you only understand if you take time. It may seem to be going well in the moment, but posture can collapse over time, so you really need to observe it in person to know" [Occupational therapist, participant 8].

Therapists mentioned difficulty interpreting pauses, atmosphere, and other nonverbal features of interaction that would typically shape in-person communication. In addition, some felt constrained by having to depend on information mediated through third parties. In the Japanese service contexts described by participants, these third parties included, for older clients, care managers (care coordinators who arrange services under Japan's long-term care insurance system) and visiting care staff, and, for children, parents and teachers who acted as the child's spokespersons in remote consultations. Information about the client's wishes then reached the therapist secondhand, already filtered through another person's interpretation. This limited their ability to directly understand clients' perspectives and the reasons underlying their preferences and behaviors. One therapist, who was supporting an older client with Parkinson disease, received information about the client's wishes mainly through the care manager and care staff involved in the client's daily care. She found that these relayed answers were often vague, which made it difficult to identify acceptable alternatives to the client's stated wish. She explained, "I want to know why they want to do it that way, their particular preference, or their true feelings. Without that...it would really have been better if I could have directly asked the person themselves" [Occupational therapist, participant 7].

These accounts described more than practical inconvenience. Without touch and whole-body observation, assessments that would ordinarily feel routine became tentative, therapists made proposals by inference from secondhand information, and one therapist noted that whether a suggestion had worked often became clear only at the following session. In their accounts, information that therapists usually gather through vision, touch, and the atmosphere of an encounter had supported the confidence of their goal setting. Remote goal setting therefore involved interpretive uncertainty: a reduction in therapists' professional certainty that they worked to manage through the adaptations described in the following themes.

Theme 2: From Direct Intervention to Coaching-Oriented Goal Setting

The second theme, reported by 8 of the 9 participants, reflected a shift in therapists' roles from directly intervening in clients' bodies and activities to supporting goal setting through coaching. In this theme, "coaching-oriented" refers to the interactional role therapists adopted: facilitating clients' reflection, self-monitoring, and problem-solving rather than delivering hands-on intervention. It is related to, but not the same as, client-centeredness (see Introduction): client-centeredness concerns whose priorities guide care, whereas coaching describes how therapists worked with clients when hands-on guidance was unavailable. For these therapists, coaching became a principal means of conducting client-centered goal setting remotely. Therapists described compensating for the limitations of remote care by gathering information in advance, using visual materials, and encouraging clients and family members to take a more active role in implementing strategies in daily life. Rather than positioning themselves primarily as direct providers of intervention, therapists described working more as partners who facilitated reflection, supported self-monitoring, and collaboratively developed practical strategies with clients and families.

Several therapists reported using self-monitoring tools, including pedometers and check sheets, to monitor clients' physical activity levels and daily routines between sessions. Remote sessions gave therapists only a brief and delayed view of clients' daily lives, and self-monitoring tools extended assessment and goal setting beyond the session itself: they supported clients' reflection on their activity patterns and gave therapists more structured information about clients' everyday activities, which informed subsequent goal-setting discussions. For example, one hospital-based therapist reviewed a weekly step count from a loaned pedometer with each client and negotiated a graded target for the following week. Notably, several therapists described these strategies not merely as substitutes for in-person methods but as improvements they valued in their own right: one explained that in conventional practice his advice would have ended in the session, whereas the online tools allowed him to follow clients' behavior change between sessions, and another described slide-supported goal setting as something good "in its own way" that was unique to remote delivery.

Others used visual materials, including slide summaries and submitted videos, to overcome what they described as the limitations of verbal communication alone. One therapist, who had noticed in earlier group-based remote sessions that sharing slides on screen helped conversations run smoothly, applied this to individual goal setting with a young school-aged client: rating cards from a goal-setting tool were embedded in the slides, and the client's own words were typed into the slides during the session and summarized visually, so that therapist and client were, in effect, looking at a shared record of the conversation as it developed: "I input the client's narrative directly into the slides and summarised it for them, and by doing that, the conversation became smoother" [Occupational therapist, participant 6]. Physical activities such as skipping rope and using a horizontal bar could not be practiced live through the screen, so the same therapist treated them as a home

XSL•FO
RenderX

program: the online sessions were used to agree on practical strategies the child would try at home, the parent recorded videos of the practice and sent them for monitoring, and the child reported back at the next session on which strategies had worked: "When practising tasks at home on a task-based approach, we discussed online what kinds of strategies could be used" [Occupational therapist, participant 6].

Family members and other supporters occupied a distinctive position in these remote encounters. Particularly for children, a parent or teacher was often physically present with the client while the therapist was not: they operated the device, recorded and sent videos, and carried out with the client the activities that the therapist would ordinarily have demonstrated in person. Therapists therefore had to put into words, and coach others through, actions they would normally have shown directly; as one therapist described, instead of demonstrating a technique and gradually fading support, she had to verbalize it for the mother to understand and carry out, which she found demanding. These arrangements created new demands, but they also distributed ownership of implementation across the therapist, the client, and the people around them. These adaptations nevertheless appeared to support a more collaborative form of goal setting, in which clients and families were encouraged to identify and monitor problems in their own daily lives. One therapist supported a middle-aged client who had returned to work following a neuromuscular condition and who completed an occupational balance check sheet several times a week and emailed it to him between weekly video sessions; the sheet gave the client herself a structured way of noticing and reporting what mattered: "Because they filled in the occupational balance check sheet, they always told me or reported what they, as the person concerned, thought was a problem that week" [Occupational therapist, participant 1].

Theme 3: At Home, at Ease: Needs That Surface in Clients' Own Environments

The third theme, raised by 6 of the 9 participants, described how remote sessions conducted in clients' own living environments changed the goal-setting encounter in 2 intertwined ways: clients often appeared more at ease, and their everyday needs and circumstances became more visible to therapists. Therapists suggested that remote rehabilitation was not only a way of overcoming physical distance, but also a way of shifting the interaction from the clinician's setting to the client's own space. This appeared to allow clients to participate in a more relaxed and natural manner and enabled therapists to observe aspects of clients' everyday environments that were not readily available in clinic-based encounters. Therapists described how being at home could reduce tension and make it easier for clients to disclose concerns and preferences. One therapist, whose clients often joined video consultations while working from home, noted, "Because the user could participate in a relaxed state, I felt as a therapist that it was easier to draw out their needs in a more natural way" [Physical therapist, participant 3]. Everyday remote contact could also create natural openings for goal-related conversation. Another therapist, who kept in regular video contact with a community-dwelling client, described how casual exchanges about the previous evening became shared topics that could be raised naturally in later sessions, weaving needs identification into ordinary conversation rather than formal questioning [Occupational therapist, participant 5].

Therapists also reported that remote sessions allowed them to directly observe everyday environments, such as desk setups, room layouts, and family situations, which supported more concrete assessment and goal setting. As one participant stated, "You can also see the desk environment here, and even the family situation is visible, so based on that, we can immediately check things and assess them" [Physical therapist, participant 3]. These accounts suggest that remote care sometimes enabled therapists to identify needs and priorities that may have been less visible in institutional settings. This visibility was of a different kind from the reduced visibility described in theme 1: relative to hands-on, whole-body assessment, remote observation remained partial, but relative to clinic-based encounters, in which the home is available only through the client's description, the video call brought therapists directly into contact with the client's everyday world.

Theme 4: When Continuity Becomes the Goal

The fourth theme, articulated most explicitly by 6 of the 9 participants, concerned a shift in how therapists conceptualized the goals of rehabilitation in remote contexts. Faced with disrupted service access during the pandemic and with ongoing geographical or systemic barriers, therapists described reorienting goals away from functional improvement alone and toward maintaining continuity, ongoing monitoring of well-being, functional status, and daily engagement. Remote rehabilitation was often positioned as a means of sustaining structured follow-up and preventing decline or disengagement from services, rather than as a direct substitute for intensive physical intervention. In some cases, this monitoring function also carried an important relational dimension, as regular remote contact provided psychological reassurance and helped prevent social isolation. Some emphasized the importance of continuity across service transitions and the potential for remote follow-up to support self-management, including pain management. One hospital-based therapist, whose older clients returned to rural areas with limited transport after discharge, continued the hospital's activity program through weekly video sessions and described the rationale for this follow-up: "By continuing the same kind of service remotely even after discharge, [we wondered] whether the onset of chronic pain could be prevented" [Physical therapist, participant 4]. He also noted that clients reported feeling reassured that support did not simply stop at discharge.

Others described the therapeutic value of remaining connected to clients over time. One therapist made regular evening video calls to a client with a psychosocial disability who had recently begun living alone and who was at risk of self-harm driven by loneliness and self-blame. He deliberately avoided making goals explicit with this client, because putting them into words risked drawing the client's attention to questions of life and death; instead, the everyday contact itself carried the therapeutic purpose, oriented toward the client feeling that their life had worth and that tomorrow was worth reaching: "Rather than putting it into words or things like that, what I wanted them to

XSL•FO
RenderX

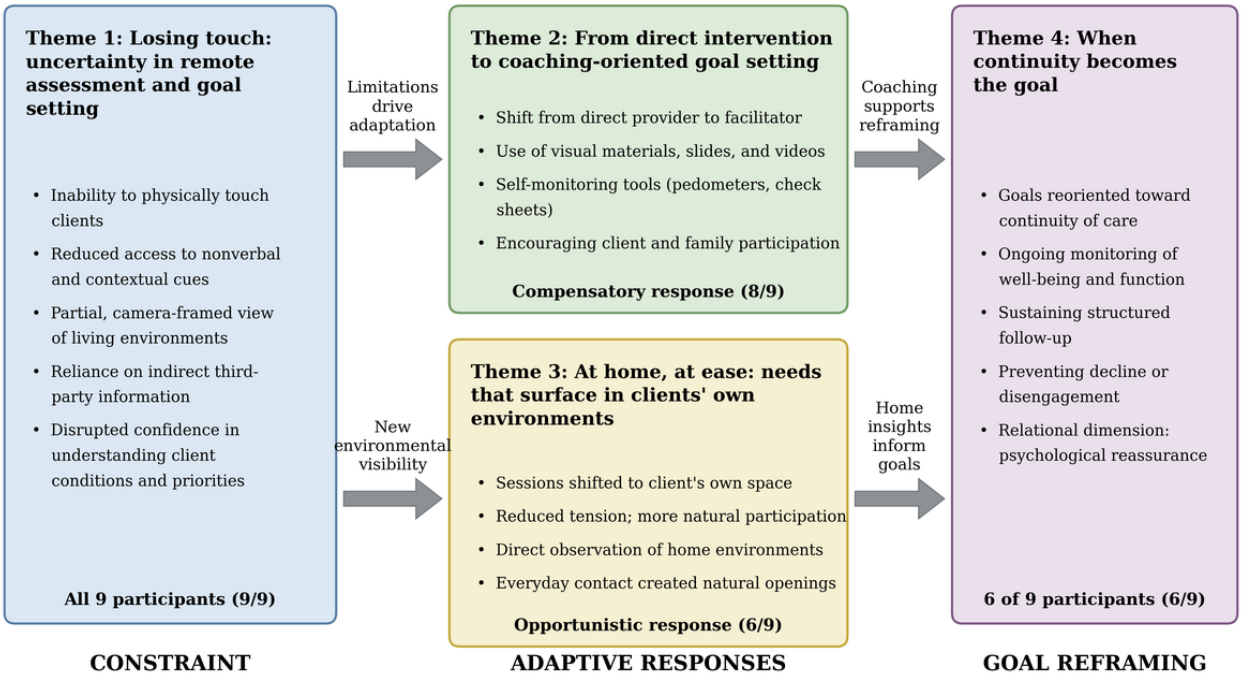
feel was that it is okay to be alive. That became our goal, or rather our purpose" [Occupational therapist, participant 5]. A further participant, a pediatric therapist practicing in a region with many underserved communities who had heard of families relocating simply to access developmental services, emphasized equity of access, stating, "I wish people could receive services no matter where they live...I want to make sure that is guaranteed for them" [Occupational therapist, participant 2].

This reframing did not appear in therapists' accounts as an abandonment of functional goals. Therapists continued to pursue functional gains where feasible through coached self-management (theme 2), while the immediate purpose of the remote encounter became sustaining engagement, monitoring change, and preserving the therapeutic relationship until more intensive intervention was possible or necessary. Continuity was described as an interim, complementary, or sometimes intrinsically meaningful goal, consistent with a client-centered orientation. The accounts also marked a clinical boundary: goals requiring hands-on assessment or intervention could not be fully pursued remotely and required a hybrid or later in-person response. These accounts indicate that, in remote contexts, therapists sometimes redefined the purpose of goal setting to include ongoing monitoring, sustained support, and continuity of care.

Relationships Among Themes

We interpreted the 4 themes as interconnected rather than independent. In our reading of the data, the constraints described in theme 1 formed the background against which 2 adaptive responses were described: a compensatory shift toward coaching-oriented goal setting (theme 2) and a recognition of needs made visible in clients' home environments (theme 3). These responses, in turn, informed a broader reframing of the purpose of goal setting toward continuity and ongoing monitoring (theme 4). Figure 1 presents this interpretation of the relationships among the themes; it summarizes our analytic account rather than an observed temporal or causal sequence.

Figure 1. The authors' interpretation of the relationships among the four themes: constraint (theme 1), adaptive responses (themes 2 and 3), and goal reframing (theme 4).



Discussion

Principal Findings

In these therapists' accounts, remote goal setting changed both the information available for clinical judgment and the purposes assigned to rehabilitation goals; it was described as a distinct form of practice shaped by constraint and possibility rather than as a diminished copy of in-person goal setting. This study examined the goal-setting experiences of Japanese therapists who practiced telerehabilitation when remote service delivery was not yet routine, frequently under pandemic-related pressure. As summarized in Figure 1, the constraints described in theme 1 were met with 2 adaptive responses, a shift toward coaching-oriented practice (theme 2) and attention to needs visible in clients' home environments (theme 3), and these together informed a reframing of goals toward continuity and ongoing monitoring (theme 4). Reduced visibility and the inability to touch introduced uncertainty into assessment and interpretation, yet remote care also appeared to create opportunities for collaborative and contextually grounded goal setting. Our analysis suggests that the conduct of remote goal setting may depend on how telerehabilitation encounters are prepared, structured, and supported, and not on access to technology alone; this interpretation, derived from the present data, is consistent with the broader telehealth implementation literature [43,44]. This is relevant given the growing recognition of telerehabilitation as a means of supporting physical activity and preventing functional decline among populations with

XSL•FO
RenderX

limited access to conventional rehabilitation, including frail older adults [6,7].

Comparison With Prior Work

The first theme, losing touch, is broadly consistent with previous literature showing that telerehabilitation can limit physical assessment and reduce access to nonverbal and contextual cues [45,46]. However, this study extends that literature by suggesting that these constraints may be experienced not merely as practical barriers but as a disruption to therapists' professional certainty, as described in the final paragraph of theme 1. In other words, remote care appeared to introduce interpretive uncertainty, defined as uncertainty in making clinical sense of partial or mediated information, by reducing the embodied and contextual information on which therapists often rely, sometimes implicitly, to make sense of clients' needs, priorities, and readiness for goal setting [16,47]. In this sample, the challenge of remote goal setting was therefore not simply that it is "harder through a screen," but that it could alter the foundations of clinical reasoning and confidence in goal negotiation.

The second theme, from direct intervention to coaching-oriented goal setting, also aligns with prior work suggesting that telerehabilitation can shift therapists' roles from hands-on intervention toward facilitation, coaching, and collaborative problem-solving [48,49]. In the present study, therapists described using advanced information gathering, visual tools, and self-monitoring resources to compensate for the limits of remote interaction. In particular, pedometers and activity check sheets were used to monitor clients' physical activity between sessions. This aligns with recent evidence suggesting that wearable activity trackers and structured self-monitoring tools may contribute to improvements in physical function among frail older adults, although their effects on overall daily physical activity levels remain inconclusive [50,51]. These adaptations did not appear to function only as substitutes for in-person practice. Rather, they sometimes created a more co-constructed form of goal setting by encouraging clients and families to take a more active role in identifying priorities, monitoring difficulties, and implementing strategies in daily life. This connects directly to the calls, outlined in the Introduction, for more collaborative and client-centered goal setting in Japanese rehabilitation [13,14,21]: coaching, in these accounts, was the interactional form through which client-centered principles were enacted when hands-on guidance was unavailable. Coaching should not, however, be equated with client-centeredness: remote delivery does not automatically make practice client-centered, which continues to depend on clients genuinely owning and negotiating their goals. When adequately supported, the constraints of remote delivery may create opportunities for more participatory and, thereby, more client-centered, goal-setting practices.

The third theme, at home, at ease, highlights a potential advantage of remote care. Previous studies have suggested that, by situating rehabilitation within the client's everyday environment, telerehabilitation can make assessment and goal setting more closely reflect clients' actual daily lives [52,53]. The present findings add that the value of home-based interaction may lie in moving the encounter from the clinician's setting to the client's own space, as well as in observing the physical environment. Therapists perceived that clients were often more relaxed at home, and this appeared to make it easier for some clients to express concerns, preferences, and personally meaningful goals. In this sample, remote goal setting sometimes appeared to make everyday needs more visible both because parts of the home environment could be observed directly and because clients seemed more comfortable in their own surroundings [54].

The fourth theme, when continuity becomes the goal, suggests that the goals of rehabilitation may be repositioned in remote contexts. Rather than focusing solely on functional improvement, therapists described reorienting goals toward maintaining continuity through ongoing monitoring of well-being, functional status, and daily engagement. Remote rehabilitation was often positioned as a means of sustaining structured follow-up and preventing decline or disengagement from services. In some cases, this monitoring function also carried an important relational dimension, as regular remote contact provided psychological reassurance and helped prevent social isolation. This should not be interpreted as a retreat from rehabilitation goals, but rather as a clinically meaningful repositioning of goals under specific service constraints. Such a shift may be particularly relevant in chronic-condition management and longer-term community-based rehabilitation, where self-management, continuity, and therapeutic connection are often central to meaningful care [55,56]. In frailty management, this may be especially important, as sustained physical activity and ongoing monitoring are needed to help prevent further decline. Remote goal setting and digital tracking may provide a practical way to maintain therapeutic engagement and support regular activity beyond scheduled rehabilitation sessions [27,50,51].

Implications for Telerehabilitation Design, Implementation, and Practice

These findings have several implications for the design, implementation, and delivery of telerehabilitation. Remote delivery of rehabilitation, including its goal-setting component, should not be assumed to be suitable for every client, clinical purpose, or stage of care. Rather, decisions about when goal setting is conducted remotely, in person, or in combination need to consider client characteristics, family support, communication needs, home and service environments, and the type of clinical judgment required. The findings also suggest that remote goal setting benefits from structured preparation. Collection of videos, check sheets, or contextual information may help compensate for therapists' reduced access to embodied and situational cues; this implication is consistent with prior work on telehealth-based goal setting and goal management [57]. In this process, families and support persons may need to be positioned not simply as assistants, but as active collaborators whose roles and whose influence on the client's own voice are explicitly negotiated. Visual tools, self-monitoring resources, and shared documentation may also support shared understanding and continuity between sessions. For therapists, this suggests a need for telerehabilitation-specific training that extends beyond technical competence to include rapport-building, verbal coaching, collaborative goal negotiation,

XSL•FO
RenderX

and managing interpretive uncertainty; our findings align here with previous work on remote consultation skills [58].

The findings also support hybrid models in which remote goal setting is used selectively alongside in-person assessment or intervention. In-person review may remain important when hands-on assessment, detailed observation, or risk management is required: initial goal setting and physical assessment might take place in person, with goals then reviewed, monitored, and renegotiated remotely, and needs surfacing during remote sessions might prompt targeted in-person review. Because goal negotiation is an ongoing process embedded in the therapeutic alliance rather than a one-off event, hybrid delivery does not remove the need for remote goal-setting skills; it distributes goal-setting work across delivery modes according to clinical purpose. Organizations may therefore benefit from clear hybrid pathways that identify when remote goal setting is appropriate, what preparation is needed, and when in-person review should be prioritized.

Strengths and Limitations

A strength of this study is that it qualitatively examined remote goal setting from the perspective of rehabilitation therapists practicing in Japan, a context that has received limited attention in the telerehabilitation literature. Participants were recruited from 9 different workplaces across metropolitan and regional areas of Japan, including hospital-based, private-practice, and community-based services. This reduces the likelihood that the patterns identified reflect the practices or culture of a single institution and supports modest transferability to other Japanese rehabilitation settings. Another strength is that the study captured therapists' experiences across the phases before, during, and after remote goal-setting encounters. This enabled a broader understanding of how preparation, interaction, and follow-up shaped remote goal setting in practice. The analysis also identified not only barriers, but also potential advantages of remote goal setting, generating practical implications for service design and implementation.

The study also has limitations. The sample was small, included more occupational therapists (n=7) than physical therapists (n=2), and included more men (n=7) than women (n=2); the perspectives of physical therapists and of female therapists are therefore less fully represented, and the findings should not be read as a balanced account of Japanese rehabilitation practice. Although information power guided our judgment of the dataset, therapists who declined or were not reached may have experienced remote goal setting differently, and a different sample might have led to additional or different themes. The analysis is therefore presented as a contextually situated account of these 9 therapists' experiences rather than an exhaustive description of remote goal setting in Japan. The findings are based on retrospective self-report rather than observation of practice. Session frequency and appointment duration were not collected systematically, which limits comparison with in-person service schedules. Participants represented selected clinical areas and all had some telerehabilitation experience, which may limit transferability to other contexts or to therapists without such experience. The study did not include clients' or families' perspectives and therefore captures only one side of the goal-setting process. Microsoft Excel and Word provided a transparent audit trail but fewer automated retrieval and version-control functions than dedicated qualitative data analysis software. Finally, because the data were collected during the COVID-19 pandemic, some findings may reflect pandemic-related service disruption as well as more enduring features of telerehabilitation.

Future Research

Future research should examine remote goal setting from the perspectives of clients and families to better understand how shared understanding, comfort, participation, and collaboration are experienced across stakeholders. Further research is also needed on tools and supports designed specifically for remote goal setting, such as visual prompts, structured preparation tools, and shared documentation resources. Future studies should also evaluate hybrid models of goal setting that combine remote and in-person elements according to clinical purpose, client preference, and service context. Comparative research across health care systems and cultural contexts may help clarify how institutional, regulatory, and social factors shape remote goal setting in practice. Finally, while some implications of the present findings may be relevant to chronic-condition management and long-term community-based rehabilitation, further research is needed to examine whether similar patterns are observed across other rehabilitation populations and service models.

Conclusions

This qualitative study showed that therapists experienced remote goal setting as both constraining and enabling. Reduced visibility and the inability to touch created uncertainty in assessment and interpretation, yet remote encounters also supported new forms of collaboration, greater visibility of clients' everyday contexts, and a reorientation of goals toward continuity and ongoing monitoring. Remote goal setting should therefore not be viewed simply as a lesser version of in-person rehabilitation. Rather, it represents a distinct form of practice that requires its own clinical reasoning, communication strategies, and implementation supports.

Acknowledgments

The authors thank all participating therapists for their time and candor, and the colleagues in their professional networks who supported recruitment.

AI-assisted transcription tools (Notta) were used for the first-pass transcription of interview audio, after which all transcripts were reviewed and corrected manually by the first author against the original recordings. After the four themes had been constructed by the authors from the Japanese transcripts, generative AI tools (ChatGPT, OpenAI; Claude, Anthropic) were used in three

XSL•FO
RenderX

further capacities: (1) as a first-pass aid for translating selected illustrative quotations from Japanese to English, (2) for grammar and stylistic editing of the manuscript, and as a sounding board for clarifying the authors' existing arguments, and (3) to draft an initial version of Figure 1. Generative AI was not used to generate codes or to construct themes; analytic decisions and final interpretations are the authors' own. All AI-assisted outputs—translations, edited text, and figure drafts—were reviewed, verified, and edited by the corresponding author, a native Japanese speaker fluent in English, before inclusion in the manuscript.

Funding

This work was supported by JSPS KAKENHI (grant JP23K18374).

Data Availability

The qualitative data (interview transcripts) generated during this study are not publicly available due to privacy and ethical restrictions related to the consent provided by participants. Anonymized data summaries may be available from the corresponding author upon reasonable request.

Authors' Contributions

YO conceptualized and designed the study, collected and analyzed the data, and drafted the manuscript. KT contributed to study design, interview guide development, analytic supervision, and critical revision of the manuscript. MO contributed to study design, participant recruitment, analytic supervision, and critical revision of the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide (translated from Japanese to English).
[DOCX File, 19 KB - rehab_v13i1e101355_app1.docx]

Checklist 1
COREQ checklist.
[PDF File, 89 KB - rehab_v13i1e101355_app2.pdf]

References

1. Gamble CJ, van Haastregt J, van Dam van Isselt EF, Zwakhalen S, Schols J. Effectiveness of guided telerehabilitation on functional performance in community-dwelling older adults: a systematic review. Clin Rehabil 2024 Apr;38(4):457-477. [doi: 10.1177/02692155231217411] [Medline: 38013415]
2. Baroni MP, Jacob MFA, Rios WR, et al. The state of the art in telerehabilitation for musculoskeletal conditions. Arch Physiother 2023 Jan 4;13(1). [doi: 10.1186/s40945-022-00155-0] [Medline: 36597130]
3. Lee AC, Deutsch JE, Holdsworth L, et al. Telerehabilitation in physical therapist practice: a clinical practice guideline from the American Physical Therapy Association. Phys Ther 2024 May 1;104(5):zae045. [doi: 10.1093/ptj/pzae045] [Medline: 38513257]
4. Nicolas B, Leblong E, Fraudet B, Gallien P, Piette P. Telerehabilitation solutions in patient pathways: an overview of systematic reviews. Digit Health 2024;10. [doi: 10.1177/20552076241294110] [Medline: 39502480]
5. Rabanifar N, Hoseini MA, Abdi K. Exploring barriers to implementing telerehabilitation from experiences of managers, policymakers, and providers of rehabilitation services in Iran: a qualitative study. Med J Islam Repub Iran 2022;36:157. [doi: 10.47176/mjiri.36.157] [Medline: 36660003]
6. Dawson R, Oliveira JS, Kwok WS, et al. Exercise interventions delivered through telehealth to improve physical functioning for older adults with frailty, cognitive, or mobility disability: a systematic review and meta-analysis. Telemed J E Health 2024 Apr;30(4):940-950. [doi: 10.1089/tmj.2023.0177] [Medline: 37975811]
7. Liang MY, Cui JR, Fan X, Zhang JJ, Liu XL, Liu DG. The effect of digital health interventions in older adults with frailty: a systematic review and meta-analysis. Int J Nurs Stud Adv 2026 Jun;10:100470. [doi: 10.1016/j.ijnsa.2025.100470] [Medline: 41551860]
8. Wicks M, Dennett AM, Peiris CL. Physiotherapist-led, exercise-based telerehabilitation for older adults improves patient and health service outcomes: a systematic review and meta-analysis. Age Ageing 2023 Nov 2;52(11):afad207. [doi: 10.1093/ageing/afad207] [Medline: 37979183]
9. Seebacher B, Geimer C, Neu J, Schwarz M, Diermayr G. Identifying central elements of the therapeutic alliance in the setting of telerehabilitation: a qualitative study. PLOS ONE 2024;19(3):e0299909. [doi: 10.1371/journal.pone.0299909] [Medline: 38457374]

XSL•FO
RenderX

10. Hall AM, Ferreira PH, Maher CG, Latimer J, Ferreira ML. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. Phys Ther 2010 Aug;90(8):1099-1110. [doi: 10.2522/ptj.20090245] [Medline: 20576715]
11. Levack WMM, Weatherall M, Hay-Smith EJC, Dean SG, McPherson K, Siegert RJ. Goal setting and strategies to enhance goal pursuit for adults with acquired disability participating in rehabilitation. Cochrane Database Syst Rev 2015 Jul 20;2015(7):CD009727. [doi: 10.1002/14651858.CD009727.pub2] [Medline: 26189709]
12. Okita Y, Kawaguchi Y, Inoue Y, et al. Characteristics of goal-setting tools in adult rehabilitation: a scoping review. Clin Rehabil 2024 Feb;38(2):234-250. [doi: 10.1177/02692155231197383] [Medline: 37644842]
13. Saito Y, Tomori K, Sawada T, et al. Determining whether occupational therapy goals match between pairs of occupational therapists and their clients: a cross-sectional study. Disabil Rehabil 2021 Mar;43(6):828-833. [doi: 10.1080/09638288.2019.1643417] [Medline: 31352840]
14. Ogawa T, Fujimoto S, Omon K, Ishigaki T, Morioka S. Shared decision-making in physiotherapy: a cross-sectional study of patient involvement factors and issues in Japan. BMC Med Inform Decis Mak 2023 Jul 24;23(1):135. [doi: 10.1186/s12911-023-02208-1] [Medline: 37488562]
15. Sivertsen M, De Jaegher H, Arntzen EC, Alstadhaug KB, Normann B. Embodiment, tailoring, and trust are important for co-construction of meaning in physiotherapy after stroke: a qualitative study. Physiother Res Int 2022 Jul;27(3):e1948. [doi: 10.1002/pri.1948] [Medline: 35306716]
16. García E, Di Paolo EA, De Jaegher H. Embodiment in online psychotherapy: a qualitative study. Psychol Psychother 2022 Mar;95(1):191-211. [doi: 10.1111/papt.12359] [Medline: 34390129]
17. Okita Y, Kendal E, Tomori K, McDonald R. Telehealth and goal-setting: a scoping review of mobile phones and computing devices. Disabil Rehabil Assist Technol 2026 Feb 21:1-15. [doi: 10.1080/17483107.2026.2632365] [Medline: 41722035]
18. Shwayder E, Dodson JA, Tellez K, et al. Goal setting among older adults starting mobile health cardiac rehabilitation in the RESILIENT trial. J Am Geriatr Soc 2024 Jul;72(7):2157-2166. [doi: 10.1111/jgs.18868] [Medline: 38450759]
19. Jiang Y, Nuerdawulieti B, Chen Z, et al. Effectiveness of patient decision aid supported shared decision-making intervention in in-person and virtual hybrid pulmonary rehabilitation in older adults with chronic obstructive pulmonary disease: a pilot randomized controlled trial. J Telemed Telecare 2024 Dec;30(10):1532-1542. [doi: 10.1177/1357633X231156631] [Medline: 36919365]
20. Asai A, Okita T, Bito S. Discussions on present Japanese psychocultural-social tendencies as obstacles to clinical shared decision-making in Japan. Asian Bioeth Rev 2022 Apr;14(2):133-150. [doi: 10.1007/s41649-021-00201-2] [Medline: 35069883]
21. Saito Y, Tomori K, Sawada T, Ohno K. Longitudinal process of setting and achieving activity- and participation-level goals in home rehabilitation in Japan: a qualitative study using trajectory equifinality modeling. Int J Environ Res Public Health 2023 May 8;20(9):5746. [doi: 10.3390/ijerph20095746] [Medline: 37174263]
22. Fujii R, Tateishi T, Tanaka S. A pilot trial of telerehabilitation for chronic stroke survivors: a case-series study of three individuals. Prog Rehabil Med 2024;9:20240033. [doi: 10.2490/prm.20240033] [Medline: 39372191]
23. Yamaguchi K, Nakanishi Y, Tangcharoensathien V, et al. Rehabilitation services and related health databases, Japan. Bull World Health Organ 2022 Nov 1;100(11):699-708. [doi: 10.2471/BLT.22.288174] [Medline: 36324547]
24. Morii Y, Abiko K, Ishikawa T, Fujiwara K, Konomura K, Ogasawara K. Regional disparity in the provision of rehabilitation services using the open data from the Japanese national claims database: an ecological study. BMJ Open 2023 Jun 20;13(6):e071670. [doi: 10.1136/bmjopen-2023-071670] [Medline: 37339845]
25. Futohashi T. Regional differences in the provision of home-visit rehabilitation services in Japan. J Phys Ther Sci 2025 Jan;37(1):34-38. [doi: 10.1589/jpts.37.34] [Medline: 39764297]
26. Kannan L, Sahu U, Subramaniam S, et al. Gaming-based tele-exercise program to improve physical function in frail older adults: feasibility randomized controlled trial. J Med Internet Res 2024 Nov 27;26:e56810. [doi: 10.2196/56810] [Medline: 39602215]
27. Yamada N, Sato I, Kinoshita S, et al. Multiuser exercise-based telerehabilitation intervention for older adults with frailty: a pilot study. NeuroSci 2026 Jan 13;7(1):11. [doi: 10.3390/neurosci7010011] [Medline: 41562836]
28. Kinoshita S, Kishimoto T. Current status and challenges of the dissemination of telemedicine in Japan after the start of the COVID-19 pandemic. Telemed J E Health 2022 Aug;28(8):1220-1224. [doi: 10.1089/tmj.2021.0336] [Medline: 34918952]
29. Kaihara T, Scherrenberg M, Falter M, et al. Cardiac telerehabilitation – a solution for cardiovascular care in Japan. Circ Rep 2021 Dec 10;3(12):733-736. [doi: 10.1253/circrep.CR-21-0126] [Medline: 34950799]
30. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
31. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
32. Dodgson JE. Reflexivity in qualitative research. J Hum Lact 2019 May;35(2):220-222. [doi: 10.1177/0890334419830990] [Medline: 30849272]

XSL•FO
RenderX

33. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
34. Carcary M. The research audit trail: methodological guidance for application in practice. Electron J Bus Res Methods 2021;18(2):166-177. [doi: 10.34190/JBRM.18.2.008]
35. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. Med Teach 2023 Mar 4;45(3):241-251. [doi: 10.1080/0142159X.2022.2057287]
36. Berger R. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. Qual Res 2015 Apr;15(2):219-234. [doi: 10.1177/1468794112468475]
37. Ahmad M, Wilkins S. Purposive sampling in qualitative research: a framework for the entire journey. Qual Quant 2025 Apr;59(2):1461-1479. [doi: 10.1007/s11135-024-02022-5]
38. Naderifar M, Goli H, Ghaljaie F. Snowball sampling: a purposeful method of sampling in qualitative research. Strides Dev Med Educ 2017;14(3):e67670. [doi: 10.5812/sdme.67670]
39. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
40. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. Int J Qual Methods 2019 Jan 1;18. [doi: 10.1177/1609406919874596]
41. Ahmed SK, Mohammed RA, Nashwan AJ, et al. Using thematic analysis in qualitative research. J Med Surg Public Health 2025 Aug;6:100198. [doi: 10.1016/j.glmedi.2025.100198]
42. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1). [doi: 10.1177/1609406917733847]
43. Cottrell MA, Russell TG. Telehealth for musculoskeletal physiotherapy. Musculoskelet Sci Pract 2020 Aug;48:102193. [doi: 10.1016/j.msksp.2020.102193] [Medline: 32560876]
44. Seron P, Oliveros MJ, Gutierrez-Arias R, et al. Effectiveness of telerehabilitation in physical therapy: a rapid overview. Phys Ther 2021 Jun 1;101(6):zab053. [doi: 10.1093/ptj/pzab053]
45. Tenforde AS, Borgstrom H, Polich G, et al. Outpatient physical, occupational, and speech therapy synchronous telemedicine: a survey study of patient satisfaction with virtual visits during the COVID-19 pandemic. Am J Phys Med Rehabil 2020 Nov;99(11):977-981. [doi: 10.1097/PHM.0000000000001571] [Medline: 32804713]
46. Lawford BJ, Bennell KL, Hinman RS. Consumer perceptions of and willingness to use remotely delivered service models for exercise management of knee and hip osteoarthritis: a cross‐sectional survey. Arthritis Care Res (Hoboken) 2017 May;69(5):667-676. [doi: 10.1002/acr.23122]
47. Scott IA, Doust JA, Keijzers GB, Wallis KA. Coping with uncertainty in clinical practice: a narrative review. Med J Aust 2023 May 15;218(9):418-425. [doi: 10.5694/mja2.51925] [Medline: 37087692]
48. Stark AL, Krayter S, Dockweiler C. Competencies required by patients and health professionals regarding telerehabilitation: a scoping review. Digit Health 2023;9. [doi: 10.1177/20552076231218841] [Medline: 38107985]
49. Hawley-Hague H, Tacconi C, Mellone S, et al. Smartphone apps to support falls rehabilitation exercise: app development and usability and acceptability study. JMIR Mhealth Uhealth 2020;8(9):e15460. [doi: 10.2196/15460]
50. Han HW, Park SW, Kim DY, et al. E-health interventions for older adults with frailty: a systematic review. Ann Rehabil Med 2023 Oct;47(5):348-357. [doi: 10.5535/arm.23090]
51. Liu JYW, Kwan RYC, Yin YH, Lee PH, Siu JYM, Bai X. Enhancing the physical activity levels of frail older adults with a wearable activity tracker-based exercise intervention: a pilot cluster randomized controlled trial. Int J Environ Res Public Health 2021 Sep 30;18(19):10344. [doi: 10.3390/ijerph181910344] [Medline: 34639644]
52. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS. Telerehabilitation: review of the state-of-the-art and areas of application. JMIR Rehabil Assist Technol 2017 Jul 21;4(2):e7. [doi: 10.2196/rehab.7511] [Medline: 28733271]
53. Dinesen B, Nonnecke B, Lindeman D, et al. Personalized telehealth in the future: a global research agenda. J Med Internet Res 2016;18(3):e53. [doi: 10.2196/jmir.5257]
54. Hale-Gallardo JL, Kreider CM, Jia H, et al. Telerehabilitation for rural veterans: a qualitative assessment of barriers and facilitators to implementation. J Multidiscip Healthc 2020;13:559-570. [doi: 10.2147/JMDH.S247267]
55. Slightam C, Gregory AJ, Hu J, et al. Patient perceptions of video visits using Veterans Affairs telehealth tablets: survey study. J Med Internet Res 2020 Apr 15;22(4):e15682. [doi: 10.2196/15682] [Medline: 32293573]
56. Sood S, Mbarika V, Jugoo S, et al. What is telemedicine? A collection of 104 peer-reviewed perspectives and theoretical underpinnings. Telemed J E Health 2007 Oct;13(5):573-590. [doi: 10.1089/tmj.2006.0073] [Medline: 17999619]
57. Kang E, Chen J, Foster ER. Feasibility, client engagement, and person-centeredness of telehealth goal setting and goal management intervention. OTJR Occup Ther J Res 2023 Jul;43(3):408-416. [doi: 10.1177/15394492231172930]
58. Tuijt R, Rait G, Frost R, Wilcock J, Manthorpe J, Walters K. Remote primary care consultations for people living with dementia during the COVID-19 pandemic: experiences of people living with dementia and their carers. Br J Gen Pract 2021 Aug;71(709):e574-e582. [doi: 10.3399/BJGP.2020.1094]

XSL•FO
RenderX

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
MOT: Master of Occupational Therapy
RTA: reflexive thematic analysis

Edited by S Munce; submitted 14.May.2026; peer-reviewed by M Wilhelmsen, M Holter; revised version received 20.Jul.2026; accepted 22.Jul.2026; published 21.Aug.2026.

Please cite as:
Okita Y, Tomori K, Okita M
Goal Setting During Telerehabilitation in Japan: Qualitative Study of Occupational Therapists' and Physical Therapists' Experiences Using Reflexive Thematic Analysis
JMIR Rehabil Assist Technol 2026;13:e101355
URL: https://rehab.jmir.org/2026/1/e101355
doi:10.2196/101355

© Yuho Okita, Kounosuke Tomori, Minoru Okita. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 21.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Acceptance of Remote Care Assist, a Telecare System for Home Care Among Care and Nursing Staff: Cross-Sectional Pilot Study

Friedrich Ebner[1], PhD; Ulrike Schneider[2], Prof Dr; Cornelia Schneider[3], Dr rer; Birgit Trukeschitz[1], Dr rer

[1]Research Institute for Economics of Aging, WU Vienna University of Economics and Business, Welthandelsplatz 1, Building D5, Vienna, Austria

[2]Department of Socioeconomics, Institute for Social Policy, Research Institute for Economics of Aging, WU Vienna University of Economics and Business, Vienna, Austria

[3]Department of Information Technologies and Digitalisation, Salzburg University of Applied Sciences, Salzburg, Austria

Corresponding Author:
Friedrich Ebner, PhD
Research Institute for Economics of Aging, WU Vienna University of Economics and Business, Welthandelsplatz 1, Building D5, Vienna, Austria

Abstract

Background: Demographic and epidemiological changes are increasing pressure on health and long-term care systems, underscoring the need for digital innovations. Remote Care Assist is a digital system that enables home care staff to connect with care experts for exchange and support via real-time video calls. Although technology acceptance is crucial for successful implementation, little is known about how care staff's expected benefits for care recipients influence acceptance in professional home care.

Objective: This study examined predictors of user acceptance of the Remote Care Assist among home care staff, with a particular focus on the role of staff's expectations of benefits for home care service users.

Methods: Technology acceptance data were collected from staff in home care organizations in Austria and Luxembourg. Among 337 survey respondents, 139 participants who reported using Remote Care Assist at least once per month over a period of 5-6.5 months were included in the acceptance analysis (45 care experts and 94 on-site care staff). Partial least squares structural equation modeling was used to test a contextualized technology acceptance model.

Results: Technology acceptance was measured by "Behavioral Intention to Use" the Remote Care Assist. "Behavioral Intention to Use" was positively associated with "Expected Benefit for Home Care Service Users" (EBC; β=0.506, 95% CI 0.364 to 0.658; *P*<.001), "Perceived Usefulness (PU)" for care staff (β=0.314, 95% CI 0.151 to 0.460; *P*<.001), and "Perceived Ease of Use" (PEOU; β=0.130, 95% CI 0.038 to 0.231; *P*=.01). "EBC" (β=0.415, 95% CI 0.276 to 0.537; *P*<.001), "Perceived Efficiency" (β=0.396, 95% CI 0.267 to 0.531; *P*<.001), and "PEOU" (β=0.170, 95% CI 0.083 to 0.266; *P*=.001) were positively associated with "PU" for care staff. "PU" also positively mediated the associations of "EBC" (β=0.130, 95% CI 0.061 to 0.194; *P*=.001) and "PEOU" (β=0.053, 95% CI 0.017 to 0.101; *P*=.02) with "Behavioral Intention to Use." "Reliable Functionality" was not significantly associated with "PU."

Conclusions: This study suggests that the technology acceptance of a digital system for enhancing professional exchange between different staff groups in home care is shaped not only by established predictors of acceptance, such as PU and PEOU, but also by a currently neglected predictor, namely care staff's expectations that the technology will benefit home care service users, which plays an important role in technology acceptance. In addition to usability and workflow support, successful implementation strategies for digital technologies should clearly communicate the technology's potential benefits for care staff, care service users, and the broader care ecosystem.

(JMIR Rehabil Assist Technol 2026;13:e80514) doi:10.2196/80514

KEYWORDS

home care services; nurse-patient relation; nursing staff; remote consultation; user acceptance; videoconferencing

XSL•FO
RenderX

Introduction

Background

The rapid proliferation of digital communication technologies has transformed a number of sectors, with the COVID-19 pandemic accelerating this trend [1]. Although the health care sector is no exception, it has been slower to adopt digital solutions compared to other sectors, such as banking and general insurance [2]. Electronic documentation, for example, has been rolled out later and more slowly [3]. Even within the health care sector, the adoption of digital technologies varies greatly, with long-term care lagging further behind than acute care. In particular, implementing new digital apps in home care settings poses challenges, such as limited time and financial resources [4,5], concerns regarding data security and privacy [6], and resistance or lack of acceptance among care staff. These barriers underscore the importance of designing and evaluating telecare technologies that fit the needs of home care practice.

Remote Care Assist

Against this background, research and development projects cofunded by national and European funding agencies aim to develop tailored digital solutions for home care settings. One of these projects, titled "Care about Care," developed and tested a communication system to support care workers in the homes of people receiving care. This so-called "Remote Care Assist" system enabled highly qualified and more experienced care staff (henceforth "care experts") to provide support and advice via video calls [7]. To this end, Remote Care Assist consisted of 2 components. The first component was a "remote support app" on care and nursing staff's Android and iOS company phones to enable video streaming between members of the care team. In addition, care experts also had access to a web app called "the expert center" to receive and make such calls in their offices using laptops or stationary computers. The system enabled care experts to see what their colleagues see in their clients' homes. Different tools supported them in interactive tasks, such as marking and measuring specific wound regions on the screen. In addition, care experts could save screenshots as part of the care documentation. This digital technology thus aimed to facilitate communication and decisions on the appropriate management of wounds, medication, and other care and nursing issues.

Technology Acceptance

For technologies such as Remote Care Assist, successful implementation depends not only on technical performance, but also on end-user acceptance. Acceptance research aims to identify predictors that positively or negatively influence end users' behavioral intention to use a product or service [8]. Because acceptance determinants can vary across contexts, technology acceptance models (TAMs) should be validated for specific contexts, technologies, and end-user groups [9,10]. Established models such as the TAM [11] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [12] are regarded as the gold standard in acceptance research across various apps and industries, including the health care sector [9].

Analytical Model and Hypotheses

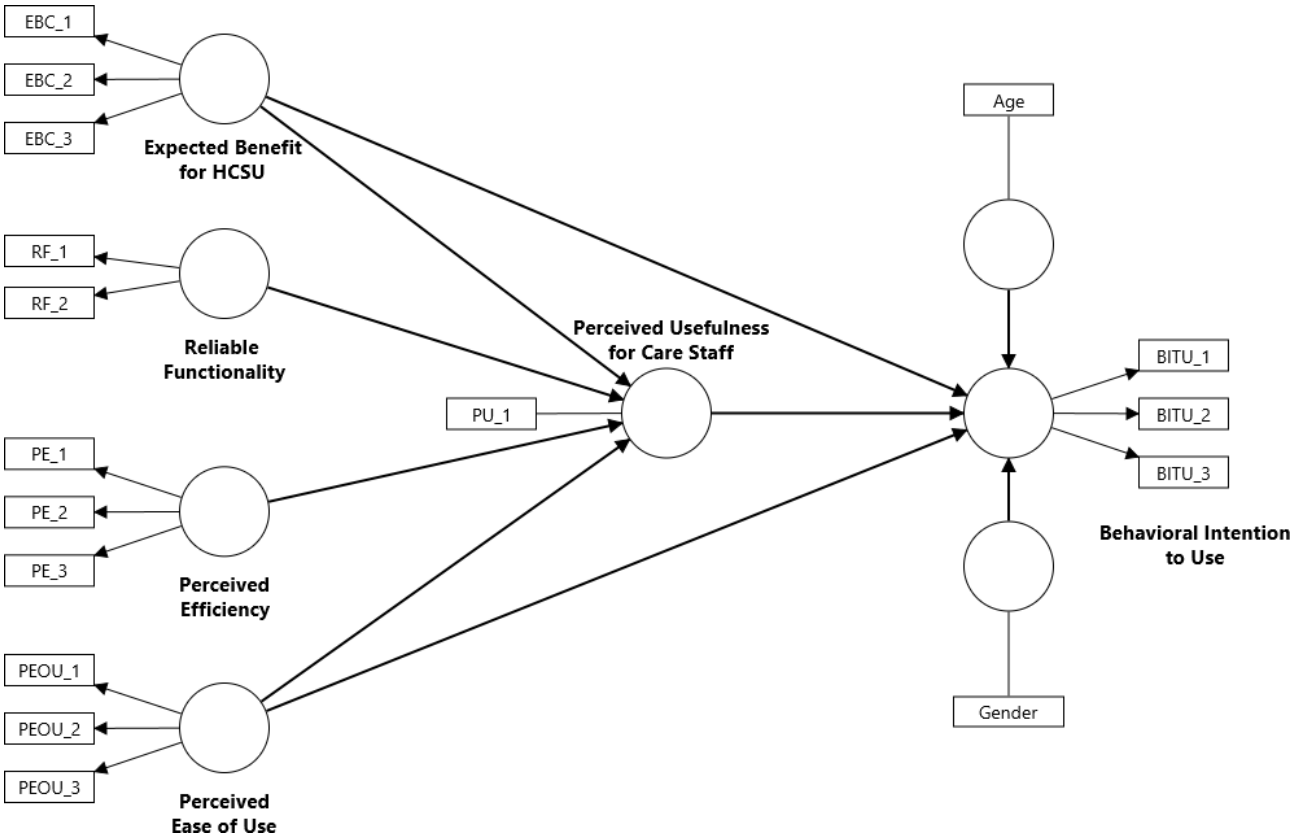
TAMs have been widely adopted and further developed in previous studies [13-20]. Building on this body of research, we propose a contextualized model (Figure 1) in which care staff's "Behavioral Intention to Use" Remote Care Assist is associated with "Perceived Usefulness (PU)" and "Perceived Ease of Use (PEOU)." We extend the model by incorporating "Expected Benefit for Home Care Service Users (EBC)" and by modeling "Perceived Efficiency (PE)" and "Reliable Functionality (RF)" as antecedents of "PU."

"PU" has emerged as one of the most influential factors in acceptance research [10,21-23]. The predictor is described as the degree to which an individual believes that using the system will help him or her to attain gains in job performance [12]. We hypothesized a direct association of "PU" for care staff with "Behavioral Intention to Use" Remote Care Assist. Hypothesis 1 was that "PU" for care staff is positively associated with the "Behavioral Intention to Use" Remote Care Assist among care staff.

"EBC" reflects the degree to which care staff believe that home care service users will benefit from the use of the system. We hypothesized a dual motivation for technology acceptance in long-term care settings, which consists of both a direct association of "EBC" with "Behavioral Intention to Use" Remote Care Assist and an indirect mediated relationship through "PU" for care staff. The mediation suggests that care staff who expect digital care technology to be beneficial for their care recipients are more likely to perceive it as useful, which, in turn, enhances their future "Behavioral Intention to Use" the Remote Care Assist. Hypothesis 2a was that "EBC" is positively associated with "Behavioral Intention to Use" Remote Care Assist among care staff, Hypothesis 2b was that "EBC" is positively associated with "PU" for care staff, and Hypothesis 2c was that the relationship between "EBC" and "Behavioral Intention to Use" Remote Care Assist among care staff is positively mediated by "PU" among care staff. While "PU" in acceptance research [12] typically emphasizes efficiency improvement, our study proposed "PU" among care staff as an attitude, with "PE" and "RF" as direct antecedents. "PE" is defined as the degree to which the system will increase the productivity of care staff. "RF" is related to "Output Quality" [8,20]. However, because telecare technologies are not primarily output-oriented, we defined "RF" as the degree to which the system is considered to work consistently and error-free in everyday professional use. Hypothesis 3 was that "PE" is positively associated with "PU" among care staff, and Hypothesis 4 was that "RF" is positively associated with "PU" among care staff.

XSL•FO
RenderX

Figure 1. Proposed technology acceptance model. EBC_1-3: measurement items for "Expected Benefit for HCSU," RF_1-2: measurement items for "Reliable Functionality," PE_1-3: measurement items for "Perceived Efficiency," PEOU_1-3: measurement items for "Perceived Ease of Use," PU_1: measurement item for "Perceived Usefulness for Care Staff," and BITU_1-3: measurement items for "Behavioral Intention to Use." HCSU: home care service user.



Consistent with previous studies [24,25], "PEOU" often emerges as an important predictor of "Behavioral Intention to Use," proposing that end users who perceive information technology as easy to use are more likely to accept it for further use. Therefore, "PEOU" is defined as the degree of ease associated with consumers' use of technology [19]. We hypothesized a direct association of "PEOU" with the "Behavioral Intention to Use" Remote Care Assist and an indirect mediated relationship through "PU" for care staff. The hypothesized mediation suggests that care staff who perceive telecare technology as user-friendly are more likely to perceive it as useful, which in turn enhances their future "Behavioral Intention to Use" Remote Care Assist. Hypothesis 5a was that "PEOU" is positively associated with "Behavioral Intention to Use" Remote Care Assist among care staff, Hypothesis 5b was that "PEOU" is positively associated with "PU" for care staff, and Hypothesis 5c was that the relationship between the "PEOU" and "Behavioral Intention to Use" Remote Care Assist among care staff is positively mediated by "PU."

"Behavioral Intention to Use" Remote Care Assist is defined as the care staff's subjective probability of using the system in the future [26]. In accordance with previous acceptance studies [13-15,27,28], this construct was designated as the principal dependent variable of this study. This choice was justified by the fact that Remote Care Assist was examined in a pilot context and was therefore not available on the open market at the time of the study.

Objective

The objective of this study was to test a contextualized TAM for the "Remote Care Assist" system in professional home care. Specifically, we examined the established associations of "PU" for care staff and "PEOU" with "Behavioral Intention to Use" Remote Care Assist. Additionally, we evaluated "PE" and "RF" as antecedents of "PU." In professional care contexts, however, adoption decisions may also reflect considerations beyond the individual user's own efficiency, including whether a technology is expected to benefit care recipients. While prior research has examined "Perceived Customer Usefulness" in social work contexts [29], little is known about how care staff's expectations of benefits for home care service users influence their intention to use telecare in routine professional practice. This study addresses this gap by extending established acceptance predictors with the EBC and by exploring system- and workflow-related antecedents of the behavioral intention to use a newly developed telecare application.

Methods

Study Design and Participants

The Remote Care Assist was tested in 2 home care organizations in Austria and Luxembourg [30]. Study participants were recruited by the participating care organizations using a 2-stage sampling procedure. First, all care units within each organization were screened for eligibility. Care units were eligible if they did not have major staffing shortages and did not participate in

XSL•FO
RenderX

pretests of the technology. Second, all employees of eligible care units were considered eligible for participation. Participants were classified according to their professional role as care experts (registered nurses or the unit's care managers) or care and nursing staff (nursing assistants and home helpers). All participants were trained in the use of the system (mobile app and expert center) using a peer-to-peer training approach. For more details, refer to the study [30].

Approximately half of the eligible care units were randomly assigned to the test group. Care workers in test-group care units received access to the Remote Care Assist app on their company smartphones. Care experts had access to the web-based care expert center within their assigned units. They did not provide support across units. The acceptance analysis focuses on exposed users rather than on comparisons between randomized groups. Only participants who reported using Remote Care Assist at least once per month were included in the acceptance analyses. To ensure transparency and complete reporting, we followed the CHERRIES (Checklist for Reporting Results of Internet E-Surveys checklist [Checklist 1]).

Sample Size

Sample size considerations were guided by the 10-times rule for partial least squares structural equation modeling (PLS-SEM) based on the maximum number of direct incoming paths to any endogenous construct [31]. In our structural model, "Behavioral Intention to Use" was the endogenous construct with the largest number of direct predictors: "EBC," "PU" for care staff, "PEOU," and the control variables age and gender. Therefore, we determined the minimum required sample size to be 50 valid responses. To assess the adequacy of the achieved sample after the survey, we additionally conducted a PLS-SEM power analysis in line with established recommendations [32].

Data Collection

The acceptance survey was conducted as a special topic survey and as part of the second wave of the study's data collection. The original questionnaire was developed in German and translated into French for French-speaking participants in Luxembourg. Before administration, both language versions were pretested in Austria and Luxembourg. Eligible participants were informed about the project by their colleagues participating in the project consortium. The survey was administered to all eligible care and nursing staff (care workers and nursing assistants) and care experts (registered nurses, wound managers, and care managers). The link provided in a text message directed them to a language-specific LimeSurvey (LimeSurvey GmbH) questionnaire. It comprised modules with questions on potential outcomes of the technology, usability, usage behavior, and technology acceptance. The survey included 65 questions for care staff and 74 questions for care experts. Responses were recorded automatically. Participants could discontinue the survey and resume it at any time using their individual token. Trial periods lasted from March to September 2023, with survey data collection from September 13 to October 15, 2023, in Austria and from May to October 2023, with survey data collection from September 14 to October 30, 2023, in Luxembourg. Differences in the data collection period resulted from a delay in ethical approval in Luxembourg.

Measures

The technology acceptance section of the online survey included 15 items adapted from established acceptance research [11,12] and tailored to the Remote Care Assist context. Construct definitions, item wordings, and variable names were adapted from prior validated instruments, particularly the UTAUT [12], and tailored to the technology used in home care settings (Multimedia Appendix 1).

The outcome variable, "Behavioral Intention to Use" Remote Care Assist, reflected respondents' willingness to use the system in the future and represented technology acceptance. It was assessed using 3 items (BITU_1 to BITU_3). "PU" for care staff, specified as a mediator, captured the perceived benefits for care and nursing staff and was measured using a single global item (PU_1). This was in line with prior research using parsimonious measures [33,34]. The remaining predictors were operationalized as follows: EBC_1 to EBC_3, PE_1 to PE_3, RF_1 to RF_2, and PEOU_1 to PEOU_3. All acceptance items were rated on a 6-point Likert-type agreement scale ranging from 1 ("strongly disagree") to 6 ("strongly agree"). An even-numbered response format was used to reduce midpoint responding, as midpoint selections may indicate forced neutrality [35], serve as a reference point [36], or fail to reflect respondents' true opinions [37].

In addition, consistent with previous studies on technology acceptance [12,25,38,39], age and gender were included as covariates to control for potential differences related to the "Behavioral Intention to Use" the Remote Care Assist. Age was dichotomized using two groups: 20-45 years (lower-age group=0) and 46-60 years (higher-age group=1). Gender was coded as women=0 and men=1.

Statistical Analysis

Descriptive statistics were conducted using R (R Foundation for Statistical Computing). Technology acceptance data were analyzed using SmartPLS (version 4; SmartPLS GmbH). The proposed model was analyzed using PLS-SEM, enabling simultaneous assessment of the measurement model and testing of the hypothesized structural relationships. This approach is particularly suited for exploratory research with small sample sizes [40]. It is a nonparametric method and therefore does not require any distributional assumptions. We used statistical bootstrapping to provide more accurate SEs and CIs and thus reliable parameter estimation [31].

First, we evaluated the measurement models by examining their reliability and validity, with a focus on internal consistency reliability, indicator reliability, convergent validity, and discriminant validity [40]. Internal consistency was evaluated by Cronbach alpha and composite reliability (RhoA and RhoC [41]). Indicator reliability was evaluated based on indicator loadings associated with the underlying constructs. Convergent validity was evaluated using the average variance extracted. Discriminant validity was assessed using the Heterotrait-Monotrait ratio of correlations (HTMT).

After confirming the validity and reliability of the measurement models, we evaluated the hypothesized structural model [40]. This included assessing collinearity among predictor constructs

XSL•FO
RenderX

using variance inflation factors (VIFs), estimating the magnitude and significance of the structural path coefficients, and examining the model's explanatory and predictive power. We applied the PLS-SEM bootstrapping procedure using 10,000 subsamples and assessed directional hypotheses using one-tailed tests at the 5% level. In addition, we report bootstrap CIs. Further, we conducted a mediation analysis to test the hypothesized direct and indirect mediations.

Additionally, we performed an importance-performance map analysis [42] to visualize and classify predictors into 4 quadrants according to their relative importance and performance on "Behavioral intention to Use" Remote Care Assist. The importance threshold was defined as the mean total effect across the key constructs, and the performance threshold was defined as the grand mean of the corresponding construct mean scores. Based on these thresholds, constructs were assigned to 1 of 4 quadrants: low importance and low performance, low importance and high performance, high importance and low performance, and high importance and high performance. The positions of each predictor in the matrix were visualized to identify potential intervention priorities. Highly important constructs were considered priorities, particularly when their performance was low, indicating the greatest room for improvement. In contrast, constructs characterized by low importance, but high performance were considered potential areas from which resources could be reallocated to higher-priority constructs.

Finally, we evaluated the model's in-sample explanatory power using the adjusted coefficient of determination (adjusted R^2). We then assessed its out-of-sample predictive power using the PLSpredict procedure, defined as the model's ability to generate accurate predictions for holdout-sample observations not used in model estimation [43]. Specifically, we applied a 10-fold cross-validation procedure with 10 repetitions [44]. We further compared the root mean square error (RMSE) and mean absolute error (MAE) of the PLS-SEM model with those of the naive linear model benchmark.

Ethical Considerations

The study design was approved by the Ethics Committee of the University of Applied Science Wiener Neustadt (September 2021), the Comité national d'éthique de recherche in Luxembourg (202209/07 version 4.0), and the Ministère de la Santé in Luxembourg (840xe52b7). All participants received written information regarding the purpose, potential benefits, and risks of the study and were assured that all data would remain confidential and processed in line with the General Data Protection Regulation. All participants provided electronic consent based on the information provided and voluntarily agreed to participate in the study. All participants had the option to decline participation or withdraw from the study at any time. Participation was voluntary. No monetary or nonmonetary incentives were provided. However, survey participation was allowed during work hours. Participant information was pseudonymized, and access to the data was restricted to the research team.

Results

Sample Characteristics

In total, 337 care and nursing staff completed the online survey, including 55 care experts. Of these, 139 (41%) reported using Remote Care Assist at least once per month and were therefore included in the technology acceptance analysis. Among these 139 eligible participants, 45 (32%) were care experts (28 from Austria and 17 from Luxembourg) and 94 (68%) were home care staff (75 from Austria and 19 from Luxembourg).

The sample consisted predominantly of women: 123 (89%) were women, 14 (10%) were men, and 2 (1%) did not disclose their gender. The sample reflected the gender composition of the care workforce in typical care settings [45]. About half of the participants (n=75, 54%) were aged 20-45 years, 62 (45%) were aged 46 years or older, and 2 (1%) did not disclose their age.

Behavioral Intention to Use Remote Care Assist

Overall, 113 (81%) participants reported a positive "Behavioral Intention to Use" Remote Care Assist (Likert scores 4 - 6: somewhat agree, mostly agree, and strongly agree). Of these, 44 (32%) selected "somewhat agree," 34 (25%) "mostly agree," and 35 (25%) "strongly agree." In contrast, 26 (19%) participants reported a negative "Behavioral Intention to Use" Remote Care Assist (Likert scores 1 - 3: strongly disagree, mostly disagree, and somewhat disagree). Of these, 6 (4%) selected "strongly disagree," 5 (4%) "mostly disagree," and 15 (11%) "somewhat disagree." The odds of agreement vs disagreement were 4.35:1.

Measurement Model Evaluation

The measurement model was evaluated for indicator reliability, internal consistency reliability, convergent validity, and discriminant validity. We first evaluated the initial measurement model. Due to high internal consistency estimates (Cronbach α and composite reliability values >0.95) and concerns regarding discriminant validity, particularly between "EBC" and "Behavioral Intention to Use" Remote Care Assist (HTMT ratio=0.884), we reassessed redundancy and content validity, resulting in the removal of 4 redundant items: BITU_2, BITU_3, PE_3, and EBC_2. Consistent with prior studies [46-51] and following parsimonious, pilot-oriented item reduction, "Behavioral Intention to Use" was measured as a single-item construct. For a more detailed description of the variable selection, refer to Multimedia Appendix 1.

In the final measurement model, all multi-item constructs showed strong internal consistency, with Cronbach α values ranging from 0.903 to 0.941. All indicator loadings exceeded the recommended threshold of 0.708 and were statistically significant, supporting indicator reliability. Average variance extracted values were above 0.5 for all multi-item constructs, indicating adequate convergent validity. Discriminant validity was also supported, as all HTMT values (Multimedia Appendix 2) were below the conservative threshold of 0.85 [52] and the bootstrapped HTMT upper 95% CI did not include 1 [31]. Thus, all measurement models were found to be reliable and valid.

XSL•FO
RenderX

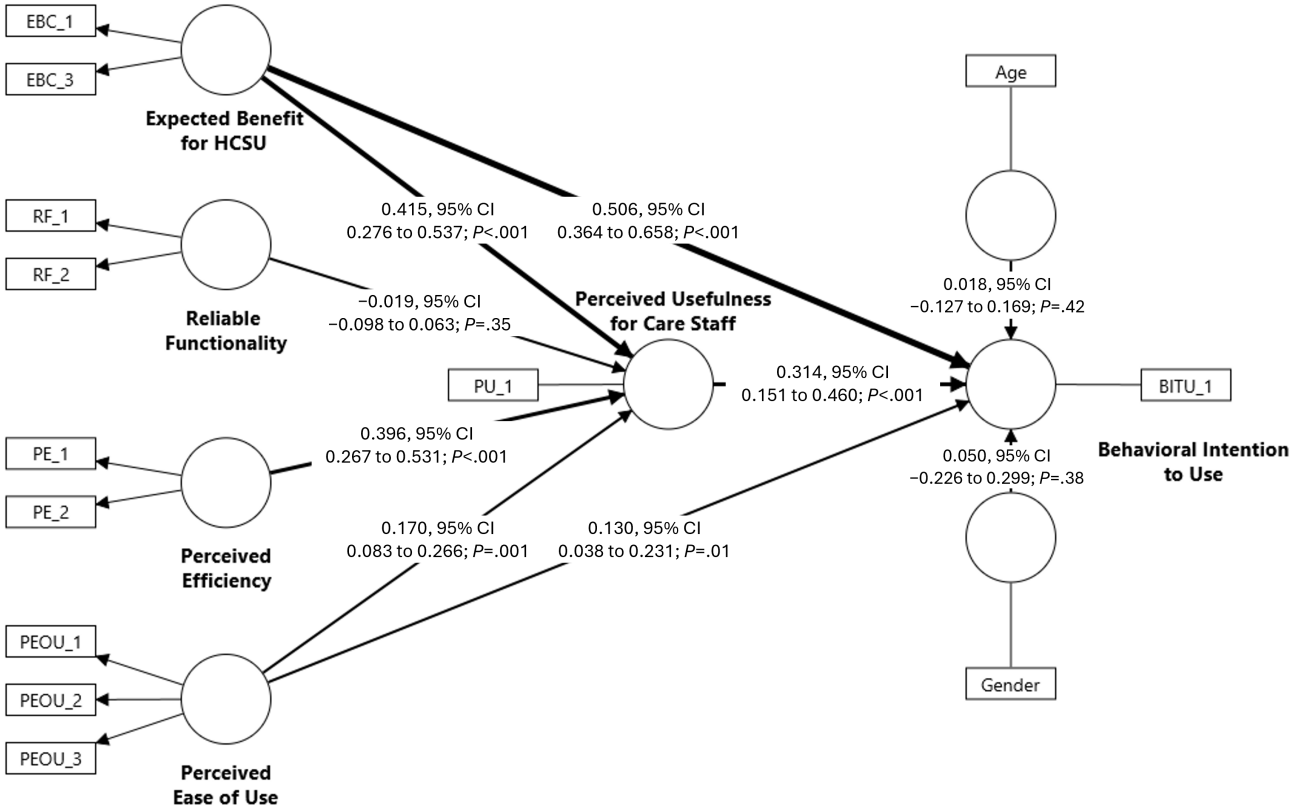
Multimedia Appendix 3 summarizes the evaluation results for both the proposed model and the final model.

Structural Model Evaluation and Hypothesis Testing

Based on the achieved sample size of 139 participants, the PLS-SEM power analysis (80% power, α=.05) indicated a minimum detectable effect size of 0.211 [32]. VIFs indicated no problematic multicollinearity. All VIFs were below the commonly used threshold of 3, except for the path from "PU" for care staff to "Behavioral Intention to Use" Remote Care Assist (VIF=3.060), which remained below the more liberal threshold of 5 [53] (Multimedia Appendix 4).

Figure 2 presents the structural model results. "Behavioral Intention to Use" the Remote Care Assist was positively associated with "EBC" (β=0.506, 95% CI 0.364 to 0.658; *P*<.001), "PU" for care staff (β=0.314, 95% CI 0.151 to 0.460; *P*<.001), and "PEOU" (β=0.130, 95% CI 0.038 to 0.231; *P*=.01), which supported Hypotheses 1, 2a, and 5a. Among all predictors, "EBC" showed the strongest direct association with "Behavioral Intention to Use."

Figure 2. Structural equation modeling results. EBC_1-3: measurement items for "Expected Benefit for HCSU," RF_1-2: measurement items for "Reliable Functionality," PE_1-2: measurement items for "Perceived Efficiency," "PEOU_1-3: measurement items for "Perceived Ease of Use," PU_1: measurement item for "Perceived Usefulness for Care Staff," and BITU_1: measurement item for "Behavioral Intention to Use." HCSU: home care service user.



According to our expectations, "PU" of the Remote Care Assist for care staff was positively associated with "EBC" (β=0.415, 95% CI 0.276 to 0.537; *P*<.001), "PE" (β=0.396, 95% CI 0.267 to 0.531; *P*<.001), and "PEOU" (β=0.170, 95% CI 0.083 to 0.266; *P*=.001), which confirmed Hypotheses 2b, 3, and 5b. In contrast, we found no evidence for Hypothesis 4, as "RF" was not significantly associated with "PU" for care staff (β=−0.019, 95% CI −0.098 to 0.063; *P*=.35).

In addition to the direct relationship between "EBC" and "Behavioral Intention to Use," the association between "EBC" and "Behavioral Intention to Use" Remote Care Assist among care staff was also positively mediated by "PU" for care staff (β=0.130, 95% CI 0.061 to 0.194; *P*=.001), supporting Hypothesis 2c. These findings support the hypothesized dual role of "EBC" in shaping care staff's "Behavioral Intention to Use" Remote Care Assist. On the one hand, care and nursing staff appeared to be guided by a sense of responsibility toward care recipients, as their expectation that Remote Care Assist would benefit home care service users positively contributed to technology acceptance. On the other hand, higher expected benefits for home care service users were associated with greater PU among care and nursing staff, which in turn was positively associated with technology acceptance.

In addition, the association between "PEOU" and "Behavioral Intention to Use" Remote Care Assist among care staff was positively mediated by "PU" (β=0.053, 95% CI 0.017 to 0.101; *P*=.02), supporting Hypothesis 5c. The result confirms that easy-to-use apps are perceived as more useful, which in turn increases the acceptance of the technology. Because both the direct and indirect associations were statistically significant and pointed in the same direction, both mediation patterns can be interpreted as complementary mediation [33]. Overall, the results supported Hypotheses 1, 2a, 2b, 2c, 3, 5a, 5b, and 5c, whereas Hypothesis 4 was not supported (Table 1).

Finally, age (β=0.018, 95% CI −0.127 to 0.169; *P*=.42) and gender (β=0.050, 95% CI −0.226 to 0.299; *P*=.38) as control

XSL•FO
RenderX

variables showed no statistically significant associations with "Behavioral Intention to Use" Remote Care Assist.

Table . Evaluation of path coefficients from the structural equation model and results from hypothesis testing.

Hypothesis	Path	Estimate β (95% CI)	SE	*P* value	Supported
H1	PU[a] → BITU[b]	0.314 (0.151 to 0.460)	0.094	<.001	Yes
H2a	EBC[c] → BITU	0.506 (0.364 to 0.658)	0.089	<.001	Yes
H2b	EBC → PU	0.415 (0.276 to 0.537)	0.079	<.001	Yes
H2c	EBC → PU → BITU	0.130 (0.061 to 0.194)	0.041	.001	Yes
H3	PE[d] → PU	0.396 (0.267 to 0.531)	0.080	<.001	Yes
H4	RF[e] → PU	–0.019 (–0.098 to 0.063)	0.038	.35	No
H5a	PEOU[f] → BITU	0.130 (0.038 to 0.231)	0.059	.01	Yes
H5b	PEOU → PU	0.170 (0.083 to 0.266)	0.056	.001	Yes
H5c	PEOU → PU → BITU	0.053 (0.017 to 0.101)	0.026	.02	Yes

[a]PU: Perceived Usefulness.

[b]BITU: Behavioral Intention to Use.

[c]EBC: Expected Benefit for Home Care Service Users.

[d]PE: Perceived Efficiency.

[e]RF: Reliable Functionality.

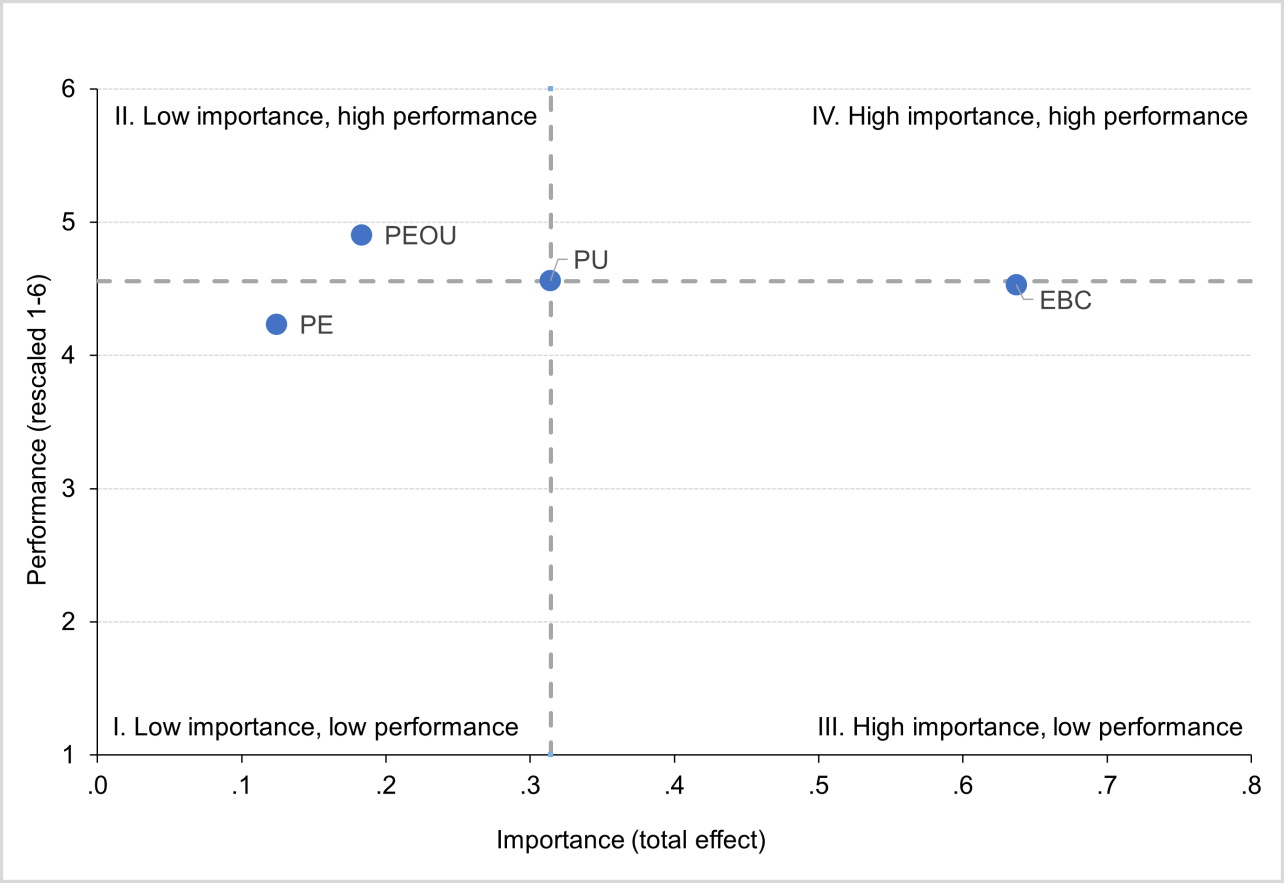
[f]PEOU: Perceived Ease of Use.

Importance-Performance Map Analysis

Figure 3 shows the importance-performance map analysis for the predictors of "Behavioral Intention to Use." The mean importance threshold, defined as the average total effect across the key constructs, was 0.315. The mean performance threshold, defined as the grand mean of the latent variable scores, was 4.557. The predictors differed in both importance and performance. "EBC" (importance=0.637; performance=4.531) was positioned in the high-importance/low-performance quadrant and emerged as the strongest predictor overall. "PU" (importance=0.314; performance=4.561) and "PEOU" (importance=0.183; performance=4.904) were in the low-importance/high-performance quadrant, indicating these constructs as secondary targets for further optimization. In contrast, "PE" (importance=0.124; performance=4.233) fell within the low-importance/low-performance quadrant and was the least influential predictor.

XSL•FO
RenderX

Figure 3. Importance-performance map analysis (IPMA) for the predictors of "Behavioral Intention to Use" Remote Care Assist. EBC: Expected Benefit for Home Care Service Users; PE: Perceived Efficiency; PEOU: Perceived Ease of Use; and PU: Perceived Usefulness.



Predictive Performance Assessment

The model showed strong explanatory power. Adjusted R^2 values were 0.724 for "PU" for care staff and 0.726 for "Behavioral Intention to Use" Remote Care Assist. These results were slightly above the coefficient of determination compared to the UTAUT [12]. Out-of-sample predictive performance was also supported [43]. The $Q^2_{predict}$ values were positive for both constructs ("PU": 0.716 and "Behavioral Intention to Use": 0.687), indicating that the PLS-SEM model outperformed the naive indicator-average benchmark [31]. In addition, prediction errors for the PLS-SEM model were lower than those for the naive linear benchmark model. For "PU" and "Behavioral Intention to Use," the PLS-SEM model yielded lower RMSE (ie, 0.664 and 0.736, respectively) than the benchmark model (RMSE=0.671 and RMSE=0.773). The same pattern was observed for MAE=0.486 and MAE=0.507 vs MAE=0.489 and MAE=0.545, respectively. The lower RMSE and MAE values in the PLS-SEM analysis demonstrated that our model provided more accurate predictions than the naive linear benchmark model, confirming its high predictive power. In summary, the final model demonstrated excellent in-sample predictive power and high out-of-sample predictive power, enabling forecasts of the "Behavioral Intention to Use" Remote Care Assist among care staff.

Discussion

Principal Findings

This study tested a contextualized TAM for a newly developed digital communication tool for home care and nursing staff that enabled support and advice via video calls, called Remote Care Assist. Technology acceptance was measured by the "Behavioral Intention to Use" the system among care staff who had already used the system for several months in practice. The most important finding was that a new aspect of technology acceptance, the "EBC," was the strongest predictor of "Behavioral Intention to Use" the Remote Care Assist. This association was observed both directly and indirectly through "PU" for care staff. These findings suggest that technology acceptance in professional home care is shaped not only by care staff's expectations of personal or work-related utility, but also by their expectations of the technology's benefits for home care service users.

The results extend established acceptance theory, building on rational choice [11], by highlighting a relational dimension of technology acceptance in care work. We consider the social influence of home care service users not as a kind of external pressure from other people who want to persuade care staff to perform a behavior, but as an internal perception of care staff that promotes technology acceptance in long-term care settings. In other words, care staff may be more willing to adopt such digital technologies when they believe that they will benefit people receiving care. In addition, this perceived benefit for

XSL•FO
RenderX

home care service users may also strengthen their own perception that the technology is useful in practice. Future research could build on these findings and extend them to focus on whether care staff's emotional attachment to their home care service users or their own standards of professionalism explain why they consider the interests of those in need of care when deciding to accept or reject a particular technology.

Our analysis revealed "PU" as a central and multifaceted predictor and mediator of user acceptance, as supported by previous research [27]. "PU" for care staff was positively associated with care staff's intention to use Remote Care Assist. "PU" for care staff mediated the associations of both "PEOU" and "EBC" with "Behavioral Intention to Use" the Remote Care Assist.

In addition, the study reaffirmed the relevance of "PEOU" as a positive predictor of "Behavioral Intention to Use" [54,55]. "PEOU" was associated with acceptance both directly and indirectly via "PU," consistent with established acceptance theory [8,20,38]. These findings indicate that user-friendly digital technologies are more likely to be perceived as useful because they simplify tasks and workflows, which in turn may increase care staff's willingness to integrate them into daily practice.

Our study confirmed "PE" as an antecedent of "PU" for care staff. The predictor encompasses factors related to improving workflow, speed, and work convenience. By contrast, we found no statistically significant association between "RF" and "PU" for care staff. This nonsignificant result aligns with similar findings in the literature, where the predictor "insecurity," referring to distrust in telehealth systems' ability to function well and properly, also did not significantly affect the outcome [56]. A possible explanation for this result is that the Remote Care Assist technology was tested in an exploratory pilot study. Participants in the trial may have been more tolerant of errors and technological issues. However, once Remote Care Assist technology is fully integrated into routine home care practice, users may be less forgiving of technical problems, which could lead to different acceptance outcomes along the implementation process.

Consistent with previous studies [57,58] and in contrast to others who found age [59] and gender [60] to moderate user acceptance, our findings show that the covariates age [55,60] and gender were not statistically significantly associated with the "Behavioral Intention to Use" Remote Care Assist.

Research on technology acceptance provides valuable insights into the potential predictors of care staff's future intention to use Remote Care Assist or similar digital technologies in care settings. These insights are crucial for management decision-making regarding whether to promote and support technology within a care organization or discontinue its use if it is not well received by end users. For instance, if a predictor of technology acceptance demonstrates low performance but high importance, management should focus on targeted improvements to this predictor. Conversely, if a predictor appears both underperforming and of low importance, it may be appropriate to neglect it and reallocate resources to more promising ones.

Limitations

Several limitations should be considered when interpreting these findings. First, this was an exploratory pilot study, and the analysis was restricted to participants who reported using the Remote Care Assist system at least once per month. As a result, nonusers and less frequent users were not represented, limiting the generalizability of the findings. Second, the role of the "EBC" was examined for the first time in long-term care settings. Although the findings indicate that this construct may be relevant, further research is needed to establish its conceptual and psychometric validity and to clarify its role in TAMs in care settings. Third, the small sample size did not allow for more detailed multigroup analyses by country or professional role, including direct comparisons between Austria and Luxembourg. Moreover, the cross-sectional design precludes causal inference, and the observed associations should not be interpreted as evidence of temporal or causal relationships. Fourth, the analyses were conducted at the individual level and did not account for potential clustering within care units or organizations. If responses were correlated within units, SEs may have been underestimated, leading to overly narrow CIs and optimistic *P* values. Fifth, "PU" and "Behavioral Intention to Use" were modeled as single-item constructs. Although this may be acceptable in parsimonious or pilot research, single-item measures may increase measurement error and may not fully capture the underlying constructs. Sixth, because predictors and outcomes were collected from the same respondents in a single survey, common method bias cannot be excluded. Future research should include nonusers, use larger samples, further validate the newly introduced construct, and, where feasible, apply longitudinal, multi-item, and multilevel designs.

Conclusion

Our findings indicate that care organizations intending to implement digital technologies should prioritize utility for both groups: care and nursing staff and their home care service users. Three factors were positively associated with acceptance of telecare technologies: PU, PEOU, and EBC. In their adoption decisions, care staff appear to consider both their own work-related benefit and the expected benefits for the people receiving care. For applications with a clear EBC, this benefit should be explicitly communicated and clearly justified during the implementation process. In contrast, reliability in the sense of flawless functionality of Remote Care Assist was not statistically significantly associated with user acceptance during the pilot implementation phase. Overall, understanding the antecedents of user acceptance among decision-makers can improve the implementation process and the user acceptance of digital applications in long-term care settings, thereby contributing to a better understanding of crucial factors for a successful digitalization of this field.

XSL•FO
RenderX

Acknowledgments

The authors thank all care staff and care experts who participated in this study. The authors also acknowledge Lara Arth and David Schermann, both affiliated with the WU Research Institute of Aging at the time of data collection, for their support throughout the project. Artificial intelligence–assisted language editing was used to improve the grammar and phrasing of prewritten text. All suggestions were reviewed by the authors, who remain fully responsible for the final content.

Funding

The project Care about Care–Active and Assisted Living (AAL) Joint Program (grant AAL-2020- 7-144-CP) received funding from the AAL Program, cofunded by the European Commission, national funding authorities of Austria, the Vienna Social Fund (FSW), Belgium, Luxembourg, and Switzerland, as well as individual project partners. The WU Research Institute of Aging is cofunded by the Vienna Social Fund (FSW) with funds from the City of Vienna (Austria). Collaborating project partners were the University of Applied Sciences Wiener Neustadt (lead partner), MOCCA GmbH, Eichenberger-Szenografie, Vienna University of Economics and Business (co-lead), Hilfswerk Niederösterreich, Stëftung Hëllef Doheem, and Distrac Group. The project ran from June 2021 until November 2023.

Data Availability

The datasets generated or analyzed during this study are not publicly available because of data protection and privacy regulations. They are available from the corresponding author on reasonable request, subject to applicable legal and ethical restrictions.

Authors' Contributions

FE contributed to conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, and writing of the original draft. CS contributed to funding acquisition, software, resources, supervision, writing—review and editing, and project administration. US contributed to writing—review and editing, funding acquisition, and project administration. BT contributed to conceptualization, funding acquisition, investigation, methodology, project administration, supervision, resources, and writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Construct definitions and item wording (English, German, and French).
[DOCX File, 31 KB - rehab_v13i1e80514_app1.docx]

Multimedia Appendix 2
Heterotrait-Monotrait ratio of correlations (HTMT) matrix.
[DOCX File, 32 KB - rehab_v13i1e80514_app2.docx]

Multimedia Appendix 3
Evaluation results of the proposed and final measurement models.
[DOCX File, 37 KB - rehab_v13i1e80514_app3.docx]

Multimedia Appendix 4
Variance inflation factor (VIF) statistics.
[DOCX File, 32 KB - rehab_v13i1e80514_app4.docx]

Checklist 1
CHERRIES checklist.
[PDF File, 251 KB - rehab_v13i1e80514_app5.pdf]

References

1. Döhring B, Hristov A, Maier C, Roeger W, Thum-Thysen A. COVID-19 acceleration in digitalisation, aggregate productivity growth and the functional income distribution. Int Econ Econ Policy 2021;18(3):571-604. [doi: 10.1007/s10368-021-00511-8] [Medline: 40477811]
2. Niedermann F, Kellner K, Friesdorf M, Levin R. Rewiring healthcare payers: a guide to digital and AI transformation. McKinsey & Company. 2025. URL: https://www.mckinsey.com/industries/healthcare/our-insights/rewiring-healthcare-payers-a-guide-to-digital-and-ai-transformation [accessed 2026-05-23]

XSL•FO
RenderX

3. Duval Jensen J, Ledderer L, Kolbæk R, Beedholm K. Fragmented care trajectories in municipal healthcare: local sensemaking of digital documentation. Digit Health 2023;9:20552076231180521. [doi: 10.1177/20552076231180521] [Medline: 37312959]
4. Arikan F, Kara H, Erdogan E, Ulker F. Barriers to adoption of electronic health record systems from the perspective of nurses: a cross-sectional study. Comput Inform Nurs 2021 Nov 15;40(4):236-243. [doi: 10.1097/CIN.0000000000000848] [Medline: 34812779]
5. Schilgen B, Handtke O, Nienhaus A, Mösko M. Work-related barriers and resources of migrant and autochthonous homecare nurses in Germany: a qualitative comparative study. Appl Nurs Res 2019 Apr;46:57-66. [doi: 10.1016/j.apnr.2019.02.008] [Medline: 30853077]
6. Galavi Z, Montazeri M, Ahmadian L. Barriers and challenges of using health information technology in home care: a systematic review. Int J Health Plann Manage 2022 Sep;37(5):2542-2568. [doi: 10.1002/hpm.3492] [Medline: 35643906]
7. Schneider C, Rameder P, Kolmann P, Trukeschitz B. Remote assistance for home care workers: concept and technical implementation at a glance. Stud Health Technol Inform 2023 May 2;301:39-47. [doi: 10.3233/SHTI230009] [Medline: 37172150]
8. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decision Sciences 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
9. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
10. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. Appl Sci (Basel) 2021;11(22):10537. [doi: 10.3390/app112210537]
11. Davis FD. A technology acceptance model for empirically testing new end- user information systems: theory and results. : Massachusetts Institute of Technology; 1985 URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2026-05-23]
12. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
13. Kim JH, Kim J, Youn BY. Using a technology acceptance model to explore the intention to use digital health technologies among people with disabilities: cross-sectional survey study. J Med Internet Res 2025 Nov 20;27:e79595. [doi: 10.2196/79595] [Medline: 41264861]
14. Matthes J, Reinhardt A, Hodzic S, et al. Predicting the intention to use generative artificial intelligence for health information: comparative survey study. J Med Internet Res 2026 Jan 28;28:e75648. [doi: 10.2196/75648] [Medline: 41604597]
15. Mirabootalebi N, Holl F, Meidani Z, et al. Determinants of nurses' continuance intention to use mobile health apps in clinical nursing practice: structural equation modeling to extend the expectation-confirmation model. JMIR Nurs 2025 Oct 31;8:e68048. [doi: 10.2196/68048] [Medline: 41172281]
16. von Winckelmann C, Vanherle R, Schreurs L, et al. Understanding the behavioral determinants of first responder app adoption by integrating perspectives from the unified theory of acceptance and use of technology and health belief model: cross-sectional survey. JMIR Hum Factors 2025 Sep 9;12:e69934. [doi: 10.2196/69934] [Medline: 40924999]
17. Chen CH, Lee WI. Exploring nurses' behavioural intention to adopt AI technology: the perspectives of social influence, perceived job stress and human-machine trust. J Adv Nurs 2025 Jul;81(7):3739-3752. [doi: 10.1111/jan.16495] [Medline: 39340769]
18. Sheikhtaheri A, Taheri Moghadam S, Dehnad A, Tatarpoor P. Factors influencing nurses' acceptance of patient safety reporting systems based on the unified theory of acceptance and use of technology (UTAUT). Informatics in Medicine Unlocked 2024;49:101554. [doi: 10.1016/j.imu.2024.101554]
19. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology1. MIS Q 2012 Mar 1;36(1):157-178. [doi: 10.2307/41410412]
20. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
21. Khan MI, Saleh MA, Quazi A. Social media adoption by health professionals: a TAM-based study. Informatics (MDPI) 2021;8(1):6. [doi: 10.3390/informatics8010006]
22. Asua J, Orruño E, Reviriego E, Gagnon MP. Healthcare professional acceptance of telemonitoring for chronic care patients in primary care. BMC Med Inform Decis Mak 2012 Nov 30;12:139. [doi: 10.1186/1472-6947-12-139] [Medline: 23194420]
23. Gagnon MP, Orruño E, Asua J, Abdeljelil AB, Emparanza J. Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system. Telemed J E Health 2012;18(1):54-59. [doi: 10.1089/tmj.2011.0066] [Medline: 22082108]
24. Schlottke JA, Alvarez-Risco A, Bertiche EL, et al. Acceptance factors of telemedicine in times of COVID-19: case Argentina. PHAR 2024;71:1-10. [doi: 10.3897/pharmacia.71.e113509]
25. Alviani R, Purwandari B, Eitiveni I, Purwaningsih M. Factors affecting adoption of telemedicine for virtual healthcare services in Indonesia. J Inf Syst Eng Bus Intell 2023;9(1):47-69. [doi: 10.20473/jisebi.9.1.47-69]
26. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research: Addison-Wesley; 1975.

XSL•FO
RenderX

27. Rouidi M, Elouadi AE, Hamdoune A, Choujtani K, Chati A. TAM-UTAUT and the acceptance of remote healthcare technologies by healthcare professionals: a systematic review. Inform Med Unlocked 2022;32:101008. [doi: 10.1016/j.imu.2022.101008]
28. Pare G, Raymond L, Etindele Sosso FA. Nurses' intention to integrate AI into their practice: survey study in Canada. JMIR Nurs 2025 Sep 5;8:e76795. [doi: 10.2196/76795] [Medline: 40911813]
29. Zhang W, Gutierrez O. Information technology acceptance in the social services sector context: an exploration. Social Work (Stellenbosch) 2007 Jul 1;52(3):221-231. [doi: 10.1093/sw/52.3.221]
30. Trukeschitz B, Arth L. The C^C remote care assist system trial in Austria: results of the second wave of data collection – care/nursing staff: deliverable 19 of the "Care about Care" Project (part: field trial, RCA, care/nursing staff AUT). : WU Vienna University of Economics and Business; 2023 URL: https://research.wu.ac.at/de/publications/the-cc-remote-care-assist-system-trial-in-austria-results-of-the-/ [accessed 2026-05-23]
31. Hair JF J, Hult GTM, Ringle CM, Sarstedt M, Danks NP, Ray S. An introduction to structural equation modeling. In: Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: Springer; 2021:1-29. [doi: 10.1007/978-3-030-80519-7]
32. Ansani A, Rinallo E. PLS-SEM-power. ShinyApps. 2025. URL: https://aleansani.shinyapps.io/pls-sem-power/ [accessed 2026-05-23]
33. Stefanelli V, Manta F. Digital financial services and open banking innovation: are banks becoming 'invisible'? Glob Bus Rev 2023. [doi: 10.1177/09721509231151491]
34. Lewis JR, Sauro J. Effect of perceived ease of use and usefulness on UX and behavioral outcomes. Int J Hum Comput Interact 2024 Oct 17;40(20):6676-6683. [doi: 10.1080/10447318.2023.2260164]
35. Saris WE, Gallhofer IN. Design, Evaluation, and Analysis of Questionnaires for Survey Research, 2nd edition: John Wiley & Sons; 2014. [doi: 10.1002/9781118634646]
36. Tourangeau R, Couper MP, Conrad F. Color, labels, and interpretive heuristics for response scales. Public Opin Q 2007;71(1):91-112. [doi: 10.1093/poq/nfl046]
37. Sturgis P, Roberts C, Smith P. Middle alternatives revisited: how the neither/nor response acts as a way of saying "I don't know"? Sociol Methods Res 2014;43(1):15-38. [doi: 10.1177/0049124112452527]
38. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
39. Wynn M, Garwood-Cross L, Vasilica C, Griffiths M, Heaslip V, Phillips N. Digitizing nursing: a theoretical and holistic exploration to understand the adoption and use of digital technologies by nurses. J Adv Nurs 2023 Oct;79(10):3737-3747. [doi: 10.1111/jan.15810] [Medline: 37530425]
40. Hair JFJ, Hult GTM, Ringle CM, Sarstedt M, Richter NF, Hauff S. Partial Least Squares Strukturgleichungsmodellierung (PLS-SEM): Eine anwendungsorientierte Einführung: Vahlen Verlag; 2017.
41. Dijkstra TK, Henseler J. Consistent partial least squares path modeling1. MIS Q 2015 Jun 1;39(2):297-316. [doi: 10.25300/MISQ/2015/39.2.02]
42. Hair JFJ, Sarstedt M, Ringle CM, Gudergan SP. Advanced Issues in Partial Least Squares Structural Equation Modeling, 2nd edition: SAGE Publications; 2024.
43. Shmueli G, Ray S, Velasquez Estrada JM, Chatla SB. The elephant in the room: predictive performance of PLS models. J Bus Res 2016 Oct;69(10):4552-4564. [doi: 10.1016/j.jbusres.2016.03.049]
44. Shmueli G, Sarstedt M, Hair JF, et al. Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. EJM 2019 Nov 11;53(11):2322-2347. [doi: 10.1108/EJM-02-2019-0189]
45. Nursing workforce grows, but inequities threaten global health goals. World Health Organization. 2025. URL: https://www.who.int/news/item/12-05-2025-nursing-workforce-grows--but-inequities-threaten-global-health-goals? [accessed 2026-05-23]
46. Rohles B, Backes S, Fischbach A, Amadieu F, Koenig V. Creating positive learning experiences with technology: a field study on the effects of user experience for digital concept mapping. Heliyon 2022 Apr;8(4):e09246. [doi: 10.1016/j.heliyon.2022.e09246] [Medline: 35497031]
47. Vitari C, Ologeanu-Taddei R. The intention to use an electronic health record and its antecedents among three different categories of clinical staff. BMC Health Serv Res 2018 Mar 21;18(1):194. [doi: 10.1186/s12913-018-3022-0] [Medline: 29562942]
48. Kwon SH, Kim YJ, Park SH. Factors influencing home care nurses' behavioral intention to use digital healthcare. Asian Nurs Res (Korean Soc Nurs Sci) 2026 Feb;20(1):65-75. [doi: 10.1016/j.anr.2025.10.003] [Medline: 41192591]
49. Lee J, Kwon S, Ko JY, et al. Caregiver survey-based perspectives on digital therapeutics for children with delayed language development. Healthcare (Basel) 2025;13(24):3290. [doi: 10.3390/healthcare13243290]
50. Zeleke EA, Stephens JH, Gesesew HA, Gello BM, Worsa KT, Ziersch A. A community-based study of intention to use HIV self-testing among young people in urban areas of southern Ethiopia. Sci Rep 2024 Nov 3;14(1):26519. [doi: 10.1038/s41598-024-77728-5] [Medline: 39489831]
51. Fishman J, Lushin V, Mandell DS. Predicting implementation: comparing validated measures of intention and assessing the role of motivation when designing behavioral interventions. Implement Sci Commun 2020;1(81):81. [doi: 10.1186/s43058-020-00050-4] [Medline: 33005900]

XSL•FO
RenderX

52. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J of the Acad Mark Sci 2015 Jan;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]
53. Hair JFJ, Hult GTM, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 2nd edition: SAGE Publications; 2017.
54. Molfenter T, Roget N, Chaple M, et al. Use of telehealth in substance use disorder services during and after COVID-19: online survey study. JMIR Ment Health 2021 Feb 8;8(2):e25835. [doi: 10.2196/25835] [Medline: 33481760]
55. Dünnebeil S, Sunyaev A, Blohm I, Leimeister JM, Krcmar H. Determinants of physicians' technology acceptance for e-health in ambulatory care. Int J Med Inform 2012 Nov;81(11):746-760. [doi: 10.1016/j.ijmedinf.2012.02.002] [Medline: 22397989]
56. Li X, Zhou Y, Liu Y, Wang X, Yuen KF. Psychological antecedents of telehealth acceptance: a technology readiness perspective. Int J Disaster Risk Reduct 2023 Jun 1;91:103688. [doi: 10.1016/j.ijdrr.2023.103688] [Medline: 37089615]
57. Kim J, DelliFraine JL, Dansky KH, McCleary KJ. Physicians' acceptance of telemedicine technology: an empirical test of competing theories. IJISCM 2010;4(3):210. [doi: 10.1504/IJISCM.2010.033076]
58. Saigí-Rubió F, Torrent-Sellens J, Jiménez-Zarco A. Drivers of telemedicine use: comparative evidence from samples of Spanish, Colombian and Bolivian physicians. Implement Sci 2014 Oct 8;9(128):128. [doi: 10.1186/s13012-014-0128-6] [Medline: 25293651]
59. Segrelles-Calvo G, López-Padilla D, Chiner E, Fernández-Fabrellas E, de Granda-Orive JI. Acceptance of telemedicine among respiratory healthcare professionals. European Research in Telemedicine / La Recherche Européenne en Télémédecine 2017 Nov;6(3-4):147-155. [doi: 10.1016/j.eurtel.2017.10.002]
60. Garavand A, Samadbeik M, Nadri H, Rahimi B, Asadi H. Effective factors in adoption of mobile health applications between medical sciences students using the UTAUT model. Methods Inf Med 2019 Nov;58(4-05):131-139. [doi: 10.1055/s-0040-1701607] [Medline: 32170717]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
EBC: Expected Benefit for Home Care Service Users
HCSU: home care service users
HTMT: Heterotrait-Monotrait ratio of correlations
IMP: construct importance
MAE: mean absolute error
PE: Perceived Efficiency
PEOU: Perceived Ease of Use
PERF: construct performance
PLS-SEM: partial least squares structural equation modeling
PU: Perceived Usefulness
RF: Reliable Functionality
RMSE: root mean square error
TAM: Technology Acceptance Model
TE: construct total effect
UTAUT: Unified Theory of Acceptance and Use of Technology
VIF: variance inflation factor

Edited by S Munce; submitted 14.Jul.2025; peer-reviewed by J Qu; revised version received 08.Mar.2026; accepted 08.Apr.2026; published 03.Jun.2026.

Please cite as:
Ebner F, Schneider U, Schneider C, Trukeschitz B
User Acceptance of Remote Care Assist, a Telecare System for Home Care Among Care and Nursing Staff: Cross-Sectional Pilot Study
JMIR Rehabil Assist Technol 2026;13:e80514
URL: https://rehab.jmir.org/2026/1/e80514
doi:10.2196/80514

© Friedrich Ebner, Ulrike Schneider, Cornelia Schneider, Birgit Trukeschitz. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 3.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Iterative Development and Usability Evaluation of an Open-Source Rehabilitation Follow-Up Platform: User-Centered Development Study

Noora Emilia Angelva[1,2], MSc; Franziska Gurschler[2,3], MSc; Michael Single[1], PhD; Valerie Frischknecht[2,3], MSc; Dominik Näf[4], BSc; Noel Kampus[4], BA; Eva Cioffi[2,3], MSc; Thimo Marcin[3], PhD; Branislav Savic[5], PhD; Matthias Wilhelm[3], Prof Dr med; Anke Scheel-Sailer[3], PD Dr med; Anna Lisa Martin-Niedecken[4], Prof Dr; Urs Mosimann[1], Prof Dr med; Tobias Nef[1,5], Prof Dr

[1]Gerontechnology and Rehabilitation Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Freiburgstrasse 3, Bern, Switzerland

[2]Graduate School for Health Sciences, University of Bern, Bern, Switzerland

[3]Centre for Rehabilitation and Sports Medicine, Inselspital Bern, University Hospital Bern, Bern, Switzerland

[4]Digital Health Design Living Lab, Institute for Design Research, Zurich University of the Arts, Zurich, Switzerland

[5]Department of Neurology, Inselspital Bern, University Hospital of Bern, Bern, Switzerland

Corresponding Author:
Noora Emilia Angelva, MSc
Gerontechnology and Rehabilitation Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Freiburgstrasse 3, Bern, Switzerland

Abstract

Background: Rehabilitation frequently extends beyond supervised clinical settings, creating a demand for scalable digital tools to support structured follow-up. We propose a novel open-source telerehabilitation platform.

Objective: We describe the iterative, user-centered development of an open-source rehabilitation follow-up platform and report findings from an initial formative usability evaluation with patients, health care professionals, and additional end users.

Methods: The platform was developed through repeated cycles of stakeholder engagement, prototyping, and evaluation within the rehabilitation department of Insel Group, University of Bern. A formative usability evaluation was conducted with 24 participants (8 health care professionals, 9 patients, and 7 additional participants; mean age 41.1, SD 12.9 y) using a task-based protocol. Perceived usability was assessed with the System Usability Scale (scored 0 - 100) and the Post-Study System Usability Questionnaire (scored 1 - 7; higher is better). Open-ended written responses were processed by one researcher using rapid content analysis.

Results: A web-based telerehabilitation platform enables therapists to manage patient caseloads, assign and track personalized rehabilitation interventions, and monitor patient progress through Fitbit activity data and rehabilitation plan adherence. It supports multiclinic workflows with role-based access, REDCap patient importing (Vanderbilt University), automated background processing via Celery, and multilingual content. Its mean System Usability Scale score was 82.8 (SD 14.1), placing the platform in the “excellent” usability range. The mean Post-Study System Usability Questionnaire overall score was 6.02 (SD 0.77), with all 3 subscales scoring above 5.7. Most participants (20/24, 83%) rated the task set as easy or very easy; no participant rated any task as difficult. Among health care professionals, 7 out of 8 (88%) reported an intention to use the platform in practice. Qualitative themes highlighted the value of a rehabilitation plan overview, structured intervention scheduling, and progress tracking.

Conclusions: The platform, released as open-source (Massachusetts Institute of Technology [MIT] License), demonstrated promising perceived usability across both patient and clinician user groups in a formative evaluation. These findings provide preliminary support for the platform’s perceived usability and justify further evaluation in real-world rehabilitation settings. Future work will evaluate deployment in real-world outpatient settings and assess clinical effectiveness.

(JMIR Rehabil Assist Technol 2026;13:e100614) doi:10.2196/100614

KEYWORDS

telerehabilitation; digital health; open-source software; usability evaluation; rehabilitation platform; patient self-management; user-centered design

XSL•FO
RenderX

Introduction

The aging population is associated with an increase in chronic diseases, and population growth is placing pressure on health care systems [1]. Globally, the need for rehabilitation is increasing and remains substantially unmet, driven in part by population aging, chronic disease, and disability (Checklist 1)[2,3]. Rehabilitation often continues beyond the supervised clinical setting: to maintain functional gains and prevent complications, patients require structured exercise programs, tailored health information, and regular follow-up support after discharge. The resulting mismatch between limited clinical resources and the growing demand for long-term rehabilitation management underscores the need for scalable digital tools that bridge the gap between clinical discharge and sustained recovery at home.

Telerehabilitation has emerged as a promising approach for such tools, using telecommunications technologies to deliver rehabilitation services across geographic distances, supporting patient self-management and reducing travel barriers, particularly in rural or underserved areas [4-6]. Internet-based approaches have also demonstrated potential for cost savings by substituting for services that would otherwise be inaccessible [7]. Evidence from multiple systematic reviews indicates that telerehabilitation can achieve outcomes comparable to in-person rehabilitation across several rehabilitation contexts, including physical therapy broadly, Parkinson disease, and musculoskeletal disorders, and may be more effective than no rehabilitation in settings where face-to-face care is not feasible [5-8].

Well-designed telerehabilitation platforms can offer several advantages for clinical practice: they may enable clinicians to create and schedule individualized rehabilitation plans remotely, allow patients to access structured exercise instructions and educational materials at home, and provide longitudinal data on adherence and health parameters that can inform follow-up decisions [6,8,9]. By integrating wearable device data, patient-reported outcomes, and clinician-defined intervention schedules into a unified workflow, such platforms can support both clinical oversight and patient engagement, which are important for sustained rehabilitation outcomes [9,10].

However, platforms capable of delivering these benefits remain incompletely integrated into routine clinical use. A key barrier in digital health is vendor lock-in: proprietary solutions may create dependency on a single vendor's technology, pricing, and product roadmap, limiting flexibility, portability, interoperability, and reproducibility in research [11]. Reviews of open-source clinical software and open-source electronic health record systems suggest that open-source development can enable local customization, support long-term sustainability, and foster collaborative development. However, adoption remains limited, and reporting is inconsistent [12,13]. More work is needed to translate the open-source paradigm into practical, clinically validated rehabilitation platforms.

To address these gaps, we developed an open-source rehabilitation follow-up platform guided by 3 principles: iterative, user-centered development; open-source transparency and reproducibility; and adaptability across diverse institutional and clinical contexts. The platform supports shared rehabilitation planning and patient self-management after discharge by integrating existing evidence-based rehabilitation content, structuring it into individualized plans, and providing clinicians and patients with a unified workflow for scheduling, monitoring, and feedback. It was developed at the Center for Rehabilitation & Sports Medicine (Inselspital and Berner Reha Zentrum), Bern University Hospital, using an approach that embedded clinical stakeholder input into each development cycle. Existing open-source platforms address related but distinct needs: OpenTera [14] provides a microservice architecture for rapid prototyping of telehealth and robotic applications, primarily in long-term care settings, but it does not focus on structured rehabilitation planning workflows. OpenTeleRehab [15], developed by Humanity & Inclusion and recognized as a Digital Public Good, offers treatment plan templates, an exercise library, and therapist-patient communication features, but it is designed primarily for deployment in low- and middle-income countries with limited digital infrastructure, including offline and SMS text messaging functionality. The proposed platform differs in its focus on integration with existing clinical workflows in high-resource hospital settings, its containerized architecture designed for institutional self-hosting under local data governance frameworks, its integration of wearable device data into structured rehabilitation plans, and its formal usability evaluation using validated instruments. Taken together, these distinctions define the space the platform is designed to fill: a formally evaluated, workflow-oriented, open-source platform for structured rehabilitation follow-up in high-resource hospital settings.

The platform's core features were prioritized to address recognized barriers to rehabilitation follow-up, rather than solely for technical reasons. Limited adherence and engagement after discharge motivated structured scheduling, daily reminders, and visible progress feedback; discontinuity between inpatient and outpatient care motivated shared rehabilitation plans, an intervention library, and a unified clinician-patient workflow; and clinician workload motivated reusable plan templates, role-based caseload management, and consolidated patient overviews. Table 1 summarizes how each platform component is intended to address these barriers.

This paper makes 2 contributions: first, we describe the design and implementation of the platform as an open-source, modular web application; and second, we report findings from an initial formative usability evaluation conducted with patients, health care professionals, and additional end users within the rehabilitation department of the University Hospital of Bern [16].

XSL•FO
RenderX

Table . Conceptual mapping of barriers to rehabilitation follow-up and the platform components intended to address them.

Barrier to follow-up	Platform component intended to address it
Adherence or engagement	Scheduling, daily browser reminders, progress tracking, and "tiny-wins" feedback
Continuity of care	Shared rehabilitation plans, intervention library, and unified clinician-patient workflow
Clinician workload	Reusable plan templates, role-based caseload management, consolidated patient overview, and data export
Patient self-management	Patient dashboard, structured exercise and education content, and wearable-derived activity data

Methods

Design Approach and Iterative Development

We adopted an iterative, user-centered development approach informed by the IDEAS (integrate, design, assess, and share) framework [17]. Rather than following a linear development process, the platform was refined through repeated cycles of stakeholder input, rapid prototyping, implementation, and formative evaluation. In each cycle, feedback from patients, clinicians, and the research team was translated into concrete design requirements, which were implemented in the prototype and subsequently reviewed to identify usability issues, missing features, and opportunities for improvement. This process allowed the platform to evolve progressively in response to user needs and clinical workflow considerations.

Throughout these cycles, an interprofessional team (cardiology and sports medicine, rehabilitation medicine, neurology and neuropsychology, biomedical engineering and gerontechnology, exercise and rehabilitation science, and design research) met regularly to steer development and was supported by 3 structured co-design workshops. In the first workshop (8 team members; August 2024), candidate features were generated through brain-writing and ranked on a relevance-by-feasibility matrix. The priorities carried into development included a structured intervention library with expected benefits, goal-based intervention suggestions, progress tracking with motivating ("tiny wins") feedback, and expert-approved constraints. These priorities shaped the rehabilitation-plan structure and intervention-metadata workflows. A second workshop with 3 patient representatives (October 2024) examined postdischarge needs: participants highlighted the loss of structure, reduced clinical support, and emotional strain, and called for a simple, accessible design with diagnosis-specific guidance and motivating feedback. When shown the team's prioritized ideas, they endorsed them while cautioning that automated suggestions should complement rather than replace human feedback. A third workshop with patient representatives and clinicians (March 2025) compared 2 interaction concepts: a guided flowchart and a structured plan. Participants judged both useful at different recovery stages and stressed the importance of ease of use, web and mobile access, and data protection. These inputs underpin the barrier-to-feature mapping in Table 1 and the prototype evaluated here.

Stakeholder involvement was conducted in hospital clinics within the rehabilitation department of the University Hospital of Bern, engaging patients, clinicians, and researchers to identify needs and priorities for rehabilitation follow-up and digital support. Outputs from these activities informed core platform workflows (eg, rehabilitation plan structure, intervention metadata, and feedback capture) and prioritized features for successive development cycles.

Development proceeded through repeated cycles of (1) requirements elicitation and prioritization, (2) low- and high-fidelity prototyping, (3) implementation of prioritized features, and (4) formative evaluation and refinement. Users participated at each stage, as summarized in Table 2. The formal usability evaluation reported in this paper corresponds to one of these stages and was followed by an expert user experience (UX) review whose recommendations partly informed the current release.

XSL•FO
RenderX

Table . Iterative development phases, user involvement, and outputs[a].

Phase	User involvement	Output
Requirements/needs elicitation	Patients, clinicians, and researchers	Core workflows: plan structure, intervention metadata, and feedback capture
Prototyping (low/high fidelity)	Stakeholder feedback on prototypes	Prioritized feature set and role-based interface designs
Formative usability evaluation (this study)	8 health care professionals, 9 patients, and 7 additional participants	SUS[b]/PSSUQ[c] scores, task completion, and qualitative themes
Expert UX[d] review	UX experts	Refinement priorities (landing page, navigation, dashboards, and motivation)
Postevaluation refinement	Implemented from review	For example, browser-based daily reminders (service worker) added to the current release

[a]The formative usability evaluation reported here represents one stage within this process.

[b]SUS: System Usability Scale.

[c]PSSUQ: Post-Study System Usability Questionnaire.

[d] UX: user experience.

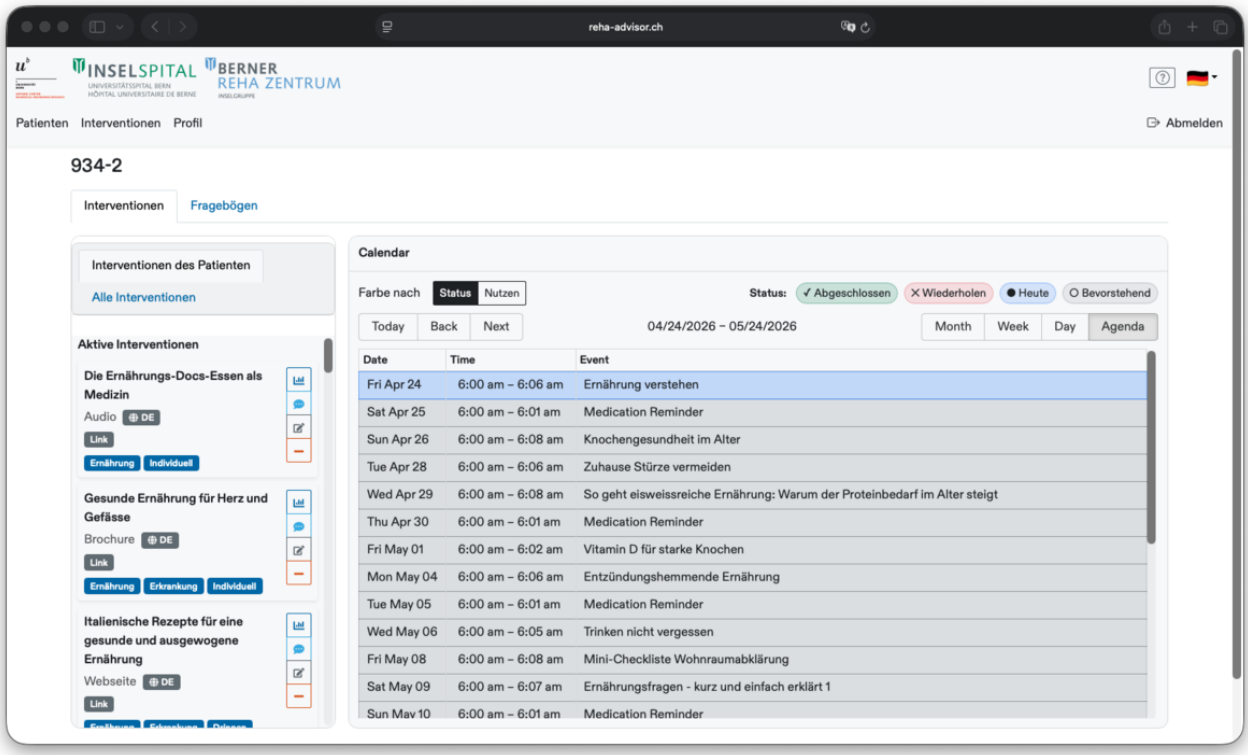
Platform Overview

The proposed platform is a modular web application designed for 2 primary user roles: patients and health care professionals. It provides role-based interfaces: a clinician-facing web portal to manage patients and rehabilitation plans and a patient-facing interface to view assigned interventions, respond to questionnaires, and review relevant health parameters (eg, activity or symptom tracking). Both interfaces are delivered as Progressive Web Apps, allowing them to be installed on mobile and desktop devices without requiring distribution through an app store. For patients, it includes a dashboard for quick status review and a calendar-based view of scheduled rehabilitation activities.

Core functional modules include (1) user and role management; (2) patient administration; (3) an intervention library that supports searching, filtering, and creation of clinic-specific interventions; (4) rehabilitation plan templates and scheduling with calendar-based views; (5) patient feedback capture via questionnaires and postintervention feedback (including optional video feedback with time-limited storage); (6) data export for clinical documentation and research (eg, as PDFs or CSV files); and (7) daily notification reminders to support adherence, configurable through a dedicated settings interface. Figures 1 and 2 illustrate the end-to-end workflow from clinician planning to patient completion and feedback capture.

XSL•FO
RenderX

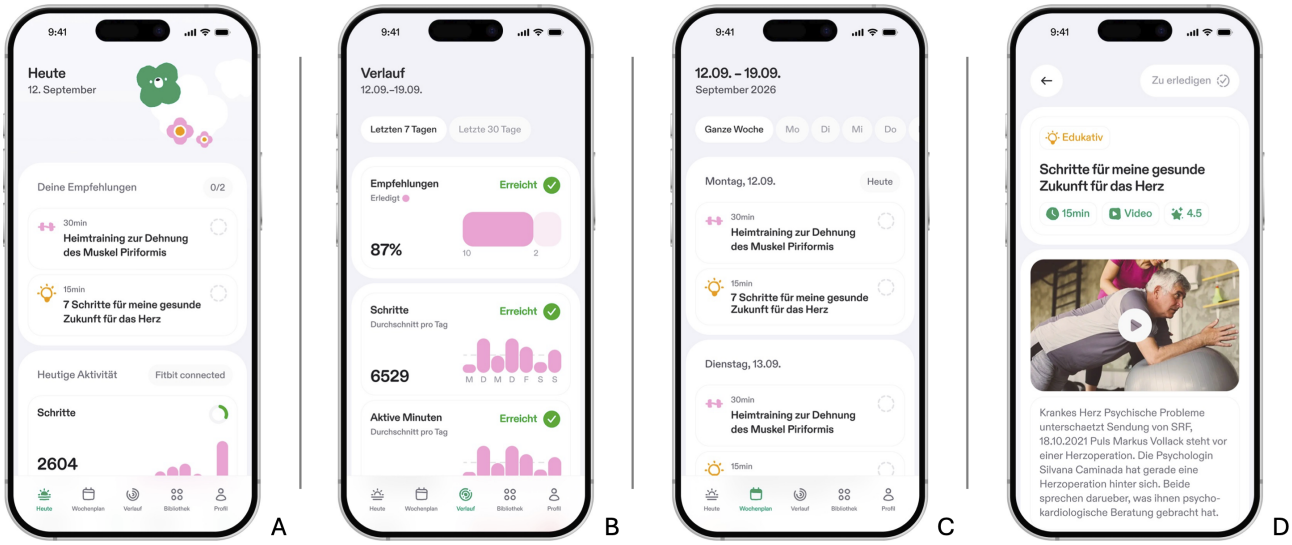
Figure 1. Clinical interface for intervention scheduling and monitoring. Calendar views (month/week/day/agenda) with status-coded interventions and access to the assigned intervention library.



Rather than building all content from scratch, the platform is designed to integrate and curate existing evidence-based educational and rehabilitation resources (eg, publicly available exercise instructions, videos, or rehabilitation applications) and to enable local tailoring to patient needs. This strategy aims to reduce duplication and support the extension to additional rehabilitation contexts.

Rather than building all content from scratch, the platform is designed to integrate and curate existing evidence-based educational and rehabilitation resources (eg, publicly available exercise instructions, videos, or rehabilitation applications) and to enable local tailoring to patient needs. This strategy aims to reduce duplication and support extension to additional rehabilitation contexts.

Figure 2. Patient interface for daily rehabilitation tasks and monitoring, shown across 4 screens of the patient-facing Progressive Web App. (A) Daily home screen ("Heute") showing the current date, the interventions recommended for that day, and today's synced activity data (eg, step count from a connected Fitbit device). (B) Intervention detail view for an individual educational ("Edukativ") intervention, including an embedded instructional video, which the patient can review and mark as completed. (C) Weekly plan view ("Wochenplan") listing the interventions scheduled across the days of the selected week. (D) Progress view ("Verlauf") summarizing longitudinal adherence and activity for a selected period (eg, the last 7 d), including the proportion of recommendations achieved, average daily step count, and active minutes, with status indicators showing which goals were met. Screen content is displayed in German, reflecting the platform's multilingual support.



XSL•FO
RenderX

Technical Implementation and Architecture

The platform is based on a modular server-client architecture designed to support scalability, maintainability, reproducible deployment, and local adaptation. The front end is built with React 18 (Meta Platforms, Inc) and TypeScript (Microsoft Corp), using Vite for optimized build performance, and is delivered as a Progressive Web App. This enables installation on patients' mobile and desktop devices directly from the browser and ensures the application functions reliably across varying network conditions. A service worker handles browser-based daily reminders that prompt patients to complete their assigned rehabilitation activities, using browser notification mechanisms compatible with the Web Notifications and Push API ecosystems [18]. Notification preferences are managed through a dedicated settings page and are compatible with browsers supporting the Web Notifications API. The back end is implemented with Django 5 (Django Software Foundation) and Django REST Framework (Encode OSS Ltd). Regarding data security, the back end employs token-based authentication, role-based access control, and encrypted data transport, with all data stored within the deploying institution's own infrastructure to support data sovereignty.

For data management, the platform uses MongoDB (version 8.0.3; MongoDB, Inc) for flexible, document-based storage of interventions and patient records, alongside Redis (version 8.4.0; Redis Ltd) for in-memory caching and as a message broker for Celery background tasks. All services are served behind an NGINX (version 1.28.2; F5, Inc) reverse proxy, which handles static file serving, Secure Sockets Layer (SSL) termination, and request routing. The full stack is orchestrated with Docker (version 27.3.1; Docker, Inc), enabling reproducible deployments across different institutional IT infrastructures. The separation between the presentation layer, back-end services, background processing, and data storage allows individual components to be maintained, extended, or replaced independently. To support code quality and deployment confidence, the repository includes comprehensive test suites for both the front end (Jest, version 29.7.0; OpenJS Foundation, and React Testing Library, version 16.3.2) and the back end (Pytest, version 9.0.3), with continuous integration via GitHub Actions and automated coverage reporting.

For data management, the platform uses MongoDB (version 8.0.3) for flexible, document-based storage of interventions and patient records, alongside Redis (version 8.4.0) for in-memory caching and as a message broker for Celery background tasks. All services are served behind an NGINX (version 1.28.2) reverse proxy, which handles static file serving, SSL termination, and request routing. The full stack is orchestrated with Docker (version 27.3.1), enabling reproducible deployments across different institutional IT infrastructures. The separation between the presentation layer, back-end services, background processing, and data storage allows individual components to be maintained, extended, or replaced independently. To support code quality and deployment confidence, the repository includes comprehensive test suites for both the front end (Jest, version 29.7.0 and React Testing Library, version 16.3.2) and the back end (Pytest, version 9.0.3), with continuous integration via GitHub Actions and automated coverage reporting.

Wearable Integration and Data Collection

To support longitudinal monitoring, the platform integrates with the Fitbit web service, which will be updated to the Google Health API due to the deprecation of the Fitbit web service in September 2026, via its publicly available REST API [19], which exposes activity, sleep, and physiological data from Fitbit wearable devices, using an OAuth 2.0 authorization flow [20]. Fitbit devices were selected because they operate independently of the user's smartphone brand, lowering the barrier to access for patients. Once a patient authorizes access, the system performs daily automatic background synchronization.

The platform specifically retrieves and stores the following health metrics in the database:

- Physical activity: daily step counts, distance traveled, and active minutes categorized by heart rate zones
- Physiological data: resting and exercise heart rates
- Sleep metrics: total sleep duration and sleep quality indicators

These passive data streams are complemented by active patient reporting, including the manual entry of vital signs (eg, blood pressure and weight) and responses to structured questionnaires.

Open-Source Strategy and Reproducibility

The software is released as open-source to support transparency, reuse, and reproducible research. The complete source code, documentation, and issue tracker are publicly available on GitHub and are released under the permissive Massachusetts Institute of Technology (MIT) License [16].

The proposed software's open-source strategy is based on several principles: permissive licensing, public access to the codebase, modular software design, reproducible deployment, transparent issue tracking, and documentation for installation and further development. The modular architecture separates the front end, back end, background processing, and data storage components, allowing individual modules to be maintained, extended, or replaced independently. This design supports the platform's adaptation to different rehabilitation contexts without requiring substantial changes to the overall system architecture.

Versioning of software artifacts is implemented using a Git-based repository structure. Source code, deployment files, documentation, and configuration artifacts, such as intervention templates or questionnaire definitions, when stored as repository files, can be tracked through commits, branches, release tags, and documented changes. This enables identifying the software state used in a given evaluation, reproducing deployments from a specific repository version, and tracing modifications over time. User-entered data are not versioned as source artifacts, but they can be exported in structured formats to support reproducibility, secondary analysis, and auditability. Stable software versions are documented using release tags, enabling specific platform versions to be referenced in publications, deployments, and evaluation studies.

XSL•FO
RenderX

To facilitate reuse across institutions, the repository includes developer documentation, deployment guidance, and contribution guidelines. These elements support external review, local adaptation, and collaborative development. Similar open-source strategies have been proposed to avoid reinventing existing components and to focus development efforts on extending capabilities and integrating reusable modules, as demonstrated by open-source telehealth frameworks such as OpenTera [14] and open-source telerehabilitation platforms such as OpenTeleRehab [15].

Usability Evaluation: Participants and Recruitment

Participants were recruited through the rehabilitation department of the University Hospital of Bern (Inselspital) and the Berner Reha Zentrum in Heiligenschwendi, both of which are part of the Insel Group. The recruited sample included health care professionals from the rehabilitation department, patients with rehabilitation experience, and additional healthy participants. Inclusion criteria were (1) aged 18 years or older and (2) the ability to provide written informed consent. Participants were not required to have prior experience with the platform.

The sample size was chosen pragmatically, in line with the evaluation's formative purpose. For early-stage usability testing, small samples are commonly used to identify recurring usability problems and inform iterative refinement. The Nielsen usability testing heuristic suggests that testing approximately 5 users can reveal many major usability issues within a user group, particularly when combined with iterative design cycles [21]. In this study, the inclusion of 24 participants, comprising 8 health care professionals, 9 patients, and 7 additional participants, was therefore considered appropriate for assessing perceived usability and identifying design improvements across the main intended user perspectives. For quantitative usability scoring with the System Usability Scale (SUS), this sample size also provides a pragmatic basis for estimating overall perceived usability, although group-specific estimates remain descriptive because of the limited number of participants per subgroup [22]. The study was not powered to test clinical effectiveness or to conduct inferential comparisons between participant groups.

The evaluation assessed 2 role-specific interfaces. Patients evaluated the patient-facing interface and health care professionals evaluated the clinician-facing interface, reflecting their respective intended uses. The additional participant group evaluated both interfaces, providing cross-interface usability coverage and a larger per-interface denominator (patient interface: n=16; clinician interface: n=15). This approach is appropriate for early-stage formative testing, where the priority is to surface recurring interface-level usability problems efficiently rather than to obtain population-representative estimates. Accordingly, we report SUS and Post-Study System Usability Questionnaire (PSSUQ) outcomes for health care professionals and patients as the primary user populations and present the additional participant group as supplementary, exploratory data.

Formative Usability Protocol

The usability evaluation was conducted on a prototype of the rehabilitation platform that included core workflows for rehabilitation planning, scheduling, patient task completion, and feedback. Features implemented after this evaluation are described separately as subsequent design refinements and were not included in the usability scores reported here. The formative evaluation used a moderated, task-based usability protocol to assess the software's core platform workflows for both user roles. Sessions were conducted in person at the rehabilitation department of the University Hospital of Bern (Inselspital) between August 2025 and December 2025. Before starting the tasks, participants received a short verbal introduction explaining the purpose of the evaluation, the voluntary nature of participation, and that the study aimed to evaluate the platform rather than the participants' performance. Participants were allowed to ask clarification questions before the task session started. During task completion, however, they were asked to work independently and without assistance, unless technical access problems required assistance.

Participants completed the tasks using study-provided hardware or their own devices. The platform was tested on computers and/or smartphones, reflecting the intended use of the clinician-facing and patient-facing interfaces. Clinician tasks were primarily performed on laptop devices, whereas patient-facing tasks were performed on smartphones, depending on participant availability and device preference. No strict maximum time limit was imposed, as the aim was to observe task completion and perceived usability under realistic formative conditions rather than to assess speed under time pressure. Task completion and notable usability issues were observed by the first author and/or the second author using a standardized task checklist.

Participants were asked to complete the following standardized task sets:

- Patient interface tasks: logging in, viewing the interventions scheduled for the current day, marking an intervention as completed, and submitting feedback after it is completed.
- Clinician interface tasks: creating a user account, logging in, assigning a recommendation to a patient with a given diagnosis, creating a new patient profile and a rehabilitation plan with 3 interventions, creating an intervention and assigning it to patient groups, and editing user profile information.

No formal think-aloud protocol was used. Instead, participants were encouraged to report difficulties or uncertainties if they occurred, and the observer documented task completion, usability issues, and spontaneous comments. After completing the tasks and exploring the platform, participants completed standardized usability questionnaires and a written semistructured posttask questionnaire.

The questionnaire combined closed-ended ratings with open-ended questions to capture perceived task difficulty, prior experience with digital technologies, perceived usefulness, barriers to use, and suggestions for improvement. Health care professionals received additional role-specific questions on expected usefulness in rehabilitation workflows, intention to use, perceived effects on time and workload, real-time monitoring, trust in platform-generated recommendations, and willingness to recommend the platform. The complete posttask

XSL•FO
RenderX

questionnaire, including the standardized usability instruments and the role-specific open-ended items, is provided as Multimedia Appendix 1.

The written questionnaire format was used to standardize data collection across participants and to minimize interviewer influence. Responses were entered into a structured electronic case report form. Because the open-ended responses were collected in written form, no audio recording or verbatim transcription was required. The questionnaire guide is provided in Multimedia Appendix 1.

Outcome Measures

Perceived usability was assessed using the SUS (10 items, 1 - 5 response scale, scored 0 - 100) [23] and the PSSUQ (16 items, 1 - 7 response scale) [24]. The SUS was scored according to the standard scoring procedure, with positively worded items scored as the item score minus 1, and negatively worded items scored as 5 minus the item score; the sum was then multiplied by 2.5 to obtain a score from 0 to 100.

For the PSSUQ, the original positively worded item content was retained. However, the response direction was aligned with the SUS and the other agreement-based items administered during the same usability assessment session. Specifically, the electronic questionnaire used a consistent agreement scale in which lower values indicated disagreement and higher values indicated agreement. This was done to reduce the risk of response errors caused by switching between opposite response directions across questionnaires within the same assessment. As a result, PSSUQ scores in this study are reported in a study-specific higher-is-better direction, rather than in the conventional PSSUQ lower-is-better direction. For comparability with conventional PSSUQ reporting, nonmissing item scores can be transformed using $x_{standard}=8 - x_{study}$. Responses coded as 99 were treated as missing. The full mapping is provided in Multimedia Appendix 2.

Additional questions assessed perceived task difficulty (1="very easy" to 5="very difficult"), and among health care professionals, perceived usefulness for follow-up care, intention to use, perceived effect on time and workload, perceived usefulness for real-time monitoring, trust requirements for platform recommendations, and willingness to recommend the platform. Open-ended questions captured positive aspects, barriers, missing functions, and suggested areas for improvement.

Statistical Analysis

We conducted descriptive analyses for SUS and PSSUQ scores overall and by participant group (including health care professionals, patients, and additional participants). Continuous scores are reported as mean and SD. Missing items were treated as missing. SUS scores were calculated only when all 10 items were completed. PSSUQ scale scores were computed as the mean of available items after excluding responses coded as 99 (not applicable). PSSUQ scores are reported in the study-specific direction, where higher values indicate better perceived usability. For transparency and comparability with conventional PSSUQ scoring, the conversion rule to the standard lower-is-better direction is reported in Multimedia Appendix 2. Task completion was summarized as the number and proportion of participants who completed each standardized task, computed separately for the patient-facing and clinician-facing interfaces. Perceived task difficulty was summarized as counts by response category, overall, and by group. The study was not powered for inferential comparisons between groups, and no hypothesis tests were performed.

Qualitative Analysis

Open-ended responses from the written posttask questionnaire were analyzed using rapid content analysis, an approach suited to formative and implementation-oriented evaluations that require timely identification of actionable themes [25]. Responses were exported from REDCap (Vanderbilt University) and organized in a structured matrix by question topic and participant group using Microsoft Excel. One researcher coded the responses, grouping them into recurring categories (perceived strengths, usability issues, missing functions, and improvement suggestions) and consolidating these into overarching themes; representative quotations were then selected to illustrate each theme across participant groups. The analysis focused on identifying actionable findings for the next development iteration rather than on generating a formal qualitative theory, and no formal interrater reliability statistic was computed. The limitation arising from single-researcher coding is addressed in the *Limitations* section.

Ethical Considerations

All procedures performed in the study involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee and with the Declaration of Helsinki and its later amendments [26]. Written informed consent was obtained from all participants before study participation. This study was approved by the Bern Cantonal Ethics Committee (KEK Bern: 2025 - 01071).

Results

Participants

In total, 24 participants completed the usability evaluation, including 8 health care professionals, 9 patients, and 7 additional participants. Participant characteristics are summarized in Table 3.

XSL•FO
RenderX

Table . Participant characteristics (N=24).

Characteristic	Overall (N=24)	Health care professionals (n=8)	Patients (n=9)	Other (n=7)
Female, n (%)	10 (41.7)	5 (62.5)	1 (11.1)	4 (57.1)
Male, n (%)	14 (58.3)	3 (37.5)	8 (88.9)	3 (42.9)
Age (y), mean (SD; range)	41.1 (12.9; 25-63)	33.9 (6.7; 27-44)	54.3 (8.5; 36-63)	32.4 (8.1; 25-48)

Usability Outcomes

Overall perceived usability was high. The mean SUS score was 82.8 (SD 14.1). The mean PSSUQ overall score, reported in the study-specific higher-is-better direction, was 6.02 (SD 0.77), corresponding to 1.98 in the conventional lower-is-better PSSUQ direction. Subscale results are shown in Table 4.

Table . Usability outcomes (N=24)[a].

Outcome (scale)	Overall (N=24), mean (SD)	Health care professionals (n=8), mean (SD)	Patients (n=9), mean (SD)	Regular (n=7), mean (SD)
SUS[b] (0 - 100)	82.8 (14.1)	79.4 (13.7)	83.1 (16.3)	86.4 (12.5)
PSSUQ[c] overall (1-7; higher=better, study direction)	6.02 (0.77)	5.74 (0.94)	6.23 (0.76)	6.06 (0.57)
System usefulness (items 1 - 6)	6.12 (0.80)	5.88 (0.92)	6.19 (0.87)	6.31 (0.56)
Information quality (items 7 - 12)	5.92 (0.88)	5.50 (1.17)	6.29 (0.61)	5.94 (0.65)
Interface quality (items 13 - 15)	5.95 (0.74)	5.81 (0.70)	6.15 (0.85)	5.86 (0.69)

[a]PSSUQ scores are reported in the study-specific higher-is-better direction. To convert to the conventional PSSUQ lower-is-better direction, use $x_{standard}=8-x_{study}$ for each nonmissing item.

[b]SUS: System Usability Scale.

[c]PSSUQ: Post-Study System Usability Questionnaire.

Most participants rated the task set as easy or very easy (20/24, 83%); the remaining participants reported neutral difficulty (4/24, 17%). No participant rated the tasks as difficult or very difficult. Perceived difficulty was consistent across groups: 6 out of 8 (75%) health care professionals, 7 out of 9 (78%) patients, and 7 out of 7 (100%) additional participants rated the platform as easy or very easy to use, with the remainder rating it as neutral.

Among 8 health care professionals, 7 (88%) reported that they would use the platform in practice. Six health care professionals reported that they would recommend it to colleagues, one was unsure, and one response was missing. This corresponds to 6 out of 8 (75%) of all health care professionals or 6 out of 7 (86%) of those with valid responses. The perceived impact on workload was mixed: 3 out of 8 reported that the platform would reduce time and workload, 1 out of 8 reported that it would not, 3 out of 8 were unsure, and 1 response was missing. Most health care professionals (6/8, 75%) reported that real-time monitoring features would be helpful. Among patients, 7 out of 9 (78%) found the rehabilitation platform easy or very easy to use, while 2 out of 9 (22%) reported neutral difficulty.

Task Completion

Task completion was high across both interfaces. All participants who attempted a task completed it, and no task failures were recorded against the standardized checklist (Table 5). The 2 patient-interface tasks were completed by all 16 participants (9 patients and 7 additional participants). The 6 clinician-interface tasks were completed by all participants (8 health care professionals and 7 additional participants); for the account-creation task, 1 record was missing, yielding 14 of 15 documented completions. Per-task details are provided in Multimedia Appendix 3. Because the evaluation used realistic, untimed conditions without a think-aloud protocol, time on task, error counts, and help requests were not systematically quantified; observed difficulties were documented descriptively and incorporated into the qualitative themes.

XSL•FO
RenderX

Table . Task completion by interface[a].

Task	Interface	Completed, n/N (%)
Log in and view interventions scheduled for today	Patient	16/16 (100)
Mark an intervention as completed and submit feedback	Patient	16/16 (100)
Create a user account	Clinician	14/14 documented (1 missing)[b]
Log in	Clinician	15/15 (100)
Assign a recommendation to a patient with a given diagnosis	Clinician	15/15 (100)
Create a patient profile and rehabilitation plan with 3 interventions	Clinician	15/15 (100)
Create an intervention and assign it to patient groups	Clinician	15/15 (100)
Edit user profile information	Clinician	15/15 (100)

[a]Patient-interface tasks were performed by patients and additional participants (n=16); clinician-interface tasks were performed by health care professionals and additional participants (n=15).

[b]One clinician-interface participant had no recorded result for this task (treated as missing). No task was recorded as failed.

Qualitative Feedback Analysis

Overview

Overall, the qualitative feedback was consistent with the high-quantitative usability ratings while highlighting specific points of friction. The reported frustrations are therefore interpreted as actionable usability issues identified during formative testing rather than as contradicting the positive usability scores.

Four recurring themes were identified, summarized below; representative quotations and their cross-group distribution are provided in Multimedia Appendix 4.

Value of Overview and Progress Tracking

Participants appreciated having an at-a-glance view of rehabilitation plans and daily activities, as well as the ability to review health and activity data over time, such as steps, sleep, and vital parameters. Several participants requested clearer progress indicators, such as "tiny wins," improved visualization of changes over time, and more actionable interpretation of collected parameters. This was the most frequently raised theme and was expressed across all 3 participant groups. One patient noted that "the overall picture is very useful" [Gesamtbild ist sehr nützlich], while others requested clearer progress feedback, for example, "a weekly overview; e.g. 80% achieved" (German: "Wochenübersicht; bspw. 80% erreicht") and, from an additional participant, "progress for the patients—tiny wins!"

Intervention Library and Structured Planning

Health care professionals valued the ability to create standard plans and templates, search and explore interventions, and schedule activities. Patients valued receiving structured exercise instructions and educational content. Several participants expressed interest in communication features, such as in-app messaging, and in tailoring interventions more explicitly to individual rehabilitation goals. One patient valued that "exercises are provided, and you do not have to assemble a program yourself" (German: "Gut, dass Übungen da sind …"), and an additional participant wished to "chat with my therapist through it."

Interface Simplification and Consistency

Suggested improvements included simplifying screens with many buttons, making views more compact, improving the calendar and navigation, and standardizing terminology and visual design. Some participants reported frustration with error messages and session handling (eg, unexpected logouts). This was the second-most-frequently raised theme. One health care professional observed that there were "currently still very many buttons; the system needs to be faster to operate" (German: "…sehr viele Buttons …"), and friction with error handling and session management was noted directly, for example, "sudden log out" and "bad error messages."

Trust and Clinical Validation

Participants emphasized that trust depends on clinician-validated content, transparent sourcing of educational materials, and clarity about who selected or approved interventions. Health care professionals noted the importance of integrating the platform into established clinical workflows and ensuring that recommendations and content are verified before patient release. This theme was raised predominantly by health care professionals; one emphasized the need for content "always confirmed by a therapist before being released" (German: "…von einem Therapeuten bestätigt …"), and an additional participant valued content that was "[personalised] from my therapist (approved)."

Design Refinements Identified Through UX Review

In parallel with the formative usability testing, an expert UX review workshop was conducted to identify priorities for subsequent design refinement. The usability scores reported in this study, therefore, reflect the prototype version available

XSL•FO
RenderX

during the evaluation, not all the features included in the current open-source release.

During usability testing, the evaluated prototype included the core clinician and patient workflows, including user log in, patient profile creation, intervention library access, rehabilitation plan creation, scheduled intervention assignment, patient-facing daily intervention views, completion marking, and postintervention feedback submission. The expert UX review focused on how these workflows could be improved in subsequent iterations.

Key recommendations included adding a concise landing page that clearly communicates the platform's purpose, streamlining the sitemap and navigation for both user roles, providing dashboards for at-a-glance status information, unifying patient data representations, and improving patient-profile views by combining plan adherence, questionnaire completion, and recent activity. The review also highlighted the potential value of notification support and motivational elements, such as visualization of the rehabilitation journey or gamification, to sustain patient engagement.

Some of these recommendations were implemented after the formative evaluation. In particular, browser-based daily reminders using a service worker were added after usability testing and are therefore part of the current platform version, but they were not reflected in the reported SUS and PSSUQ scores. Other recommendations, such as more advanced motivational elements and extended dashboard visualizations, remain priorities for future development.

Discussion

Principal Findings

This was a formative, single-session usability evaluation. It assessed perceived usability and task completion under short, moderated, task-based conditions and did not assess adherence, workflow burden, long-term engagement, clinical effectiveness, or real-world adoption; the following interpretations are bounded accordingly.

The rehabilitation platform achieved a mean SUS score of 82.8 (SD 14.1), placing it in the "excellent" usability range according to established SUS adjective-rating benchmarks [27]. The mean PSSUQ overall score was 6.02 (SD 0.77) on a 7-point scale, with all 3 subscales (system usefulness, information quality, and interface quality) scoring above 5.7. Scores were consistent across participant groups, with patients rating both instruments marginally higher than health care professionals did, suggesting that the patient-facing interface was particularly well received. Crucially, 20 out of 24 (83%) participants rated the task set as easy or very easy, and no participant rated any task as difficult.

The qualitative findings provide complementary clinical insight. The features participants valued most—at-a-glance rehabilitation plan overview, longitudinal progress tracking, and structured intervention scheduling—map onto broader telerehabilitation implementation priorities, including adherence, engagement, accessibility, satisfaction, and integration into routine workflows [8,10,28]. In our formative data, participants specifically emphasized the importance of visible progress and interpretable feedback. The repeated interest in "tiny wins" and actionable feedback may also be interpreted through self-determination theory, particularly the role of perceived competence in supporting sustained motivation and health behavior change [29]. Together, these themes suggest that the platform's design priorities were appropriate to the clinical context.

Among 8 health care professionals, 7 (88%) reported an intention to use the platform in clinical practice, and 6 out of 8 (75%) would recommend it to colleagues; these are encouraging early signals that require confirmation in deployment studies. The mixed workload impact ratings (3/8 expecting workload reduction and 3/8 uncertain) are expected for a formative prototype and point to specific areas requiring workflow optimization before deployment, particularly plan template creation and assignment workflows. The strong interest in real-time monitoring features (6/8, 75%), combined with the security and content validation concerns raised in qualitative feedback, underscores the importance of clear data governance communication and clinician-controlled content approval workflows in the next development iteration.

Comparison With Prior Work

The usability scores reported here add to an evidence base in which many telerehabilitation studies primarily emphasize clinical, physiological, adherence, or algorithmic outcomes. In contrast, detailed reporting of role-specific UX and formative usability findings is less consistently described [6,8,9,30]. The observed SUS score above 80 and the PSSUQ overall score above 6.0 suggest strong perceived usability at the formative stage. The consistency across patient and health care professional groups further suggests that the role-differentiated interface design did not introduce major usability asymmetries that would complicate deployment. These findings provide early empirical support for the feasibility of a workflow-oriented, open-source approach in this domain. For comparison, we focus on 2 commonly reported strands of digital rehabilitation research: telemonitoring systems and AI-focused rehabilitation tools. These areas were selected because they represent 2 major directions in current rehabilitation technology development, both directly relevant to the platform's intended scope. Telemonitoring systems typically focus on remote observation, physiological data capture, exercise supervision, or adherence monitoring [9,30]. AI-focused rehabilitation tools, in contrast, often emphasize prediction, activity recognition, movement assessment, or decision-support models [31-33]. Both strands are highly relevant to future rehabilitation services, but they often address specific functional components rather than the broader workflow infrastructure needed to plan, assign, monitor, and adapt rehabilitation interventions across patient and clinician roles. To make the platform's intended niche and its contribution relative to the state of the art more explicit, Table 6 contrasts the platform with the open-source systems introduced earlier (OpenTera and OpenTeleRehab) and with these 2 strands across capability, setting, deployment, and evaluation dimensions.

Compared with commonly reported telemonitoring systems and AI-focused rehabilitation tools, the proposed platform emphasizes workflow infrastructure. The platform does not aim

XSL•FO
RenderX

to replace remote monitoring systems or specialized AI models; rather, it provides an extensible environment in which such components can be embedded. Its current contribution lies in supporting intervention libraries, structured rehabilitation planning and scheduling, patient feedback capture, progress documentation, and data export. In this sense, the platform complements existing telemonitoring and AI-oriented approaches by addressing the organizational and interaction layers required to integrate digital rehabilitation tools into clinical workflows.

XSL•FO
RenderX

Table . Structured comparison of the proposed platform with related approaches[a].

Approach	Primary focus	Target setting/context	Rehabilitation workflow support	Data/wearable integration	Openness and deployment model	Evaluation status/evidence base	Relationship to proposed platform
Proposed platform	Open-source infrastructure for structured rehabilitation follow-up: plan management, patient self-management, and clinician-patient workflow support	High-resource hospital and outpatient rehabilitation settings; designed for institutional self-hosting	Intervention library, clinician-assigned plans, scheduling, patient task completion, postintervention feedback, progress documentation, and data export	Integrates Fitbit-derived activity and physiological data; supports manual patient-reported data and questionnaires	Open-source (MIT[b] License); containerized modular architecture for local institutional deployment under local data-governance frameworks	Formative, single-session usability evaluation (n=24; SUS[c], PSSUQ[d], task completion, qualitative feedback); adherence, long-term use, and clinical effectiveness not yet assessed	— (reference point for the comparison)
OpenTera [14]	Microservice architecture for rapid prototyping of telehealth and robotic applications	Care-facility and long-term-care contexts (originally a COVID-19 response)	General telehealth application framework oriented toward flexible prototyping rather than structured rehabilitation-plan management	Extensible architecture; wearable/Fitbit integration is not a described focus	Open-source telehealth framework	Published description of architecture and use cases [14]; not assessed in this study	A broader telehealth/robotics prototyping framework; complementary, as rehabilitation-planning workflows of the type provided here could be built on such an architecture
OpenTeleRehab [15]	Open-source telerehabilitation platform with treatment-plan templates, an exercise library, and therapist-patient communication	Primarily low- and middle-income countries and low-resource settings with limited digital infrastructure	Treatment-plan templates and communication features, with emphasis on low-infrastructure delivery (offline/SMS text messaging)	Optimized for low-infrastructure use (offline/SMS text messaging); wearable integration is not a described focus	Open-source; recognized as a Digital Public Good with documented real-world field deployment	Established field deployment as a Digital Public Good [15]; not assessed in this study	Complementary, with different design priorities: OpenTeleRehab optimizes for low-resource and offline contexts, whereas this platform targets self-hosted integration into high-resource hospital workflows with wearable-supported follow-up
Telemonitoring systems (research category)	Remote observation, physiological-data capture, exercise supervision, and/or adherence monitoring	Often disease- or program-specific remote rehabilitation contexts	Typically monitor activity, symptoms, or adherence; less commonly provide an integrated plan-assign-adapt-document workflow across patient and clinician roles	Frequently sensor-, device-, or physiological-data oriented	Varies across systems; often system-specific or proprietary	Individual systems report clinical, physiological, adherence, or monitoring outcomes (eg, [9,30])	Strong at generating follow-up data; this platform focuses on the organizational workflow layer needed to use such data within rehabilitation planning and could incorporate telemonitoring outputs

XSL•FO
RenderX

Approach	Primary focus	Target setting/context	Rehabilitation workflow support	Data/wearable integration	Openness and deployment model	Evaluation status/evidence base	Relationship to proposed platform
AI-focused rehabilitation tools (research category)	Prediction, activity recognition, movement assessment, or decision-support models	Often research prototypes or specialized digital rehabilitation modules	Usually address specific analytical or decision-support tasks rather than full patient-clinician workflow management	Often use sensor, video, wearable, or task-performance data depending on the model	Varies across tools; may be proprietary, experimental, or model-specific	Individual studies report algorithmic or model-performance outcomes (eg, [31-33])	This platform is not itself primarily an AI system; it provides extensible workflow infrastructure into which such models could be embedded as future modules

[a]OpenTera and OpenTeleRehab are specifically named systems, whereas "telemonitoring systems" and "AI-focused rehabilitation tools" are broad research categories included to situate the platform relative to 2 major directions in rehabilitation technology. For these 2 categories, the cells provide illustrative, nonexhaustive characterizations drawn from the cited examples rather than a systematic survey. Cells describe each approach's primary design orientation; the absence of a feature denotes a different design focus, not a judged deficiency.

[b]MIT: Massachusetts Institute of Technology.

[c]SUS: System Usability Scale.

[d]PSSUQ: Post-Study System Usability Questionnaire.

The Value of Open-Source Code in Rehabilitation Technology

A central contribution of this work is to release the software's source code under the permissive MIT License. This directly addresses the widespread issue of vendor lock-in in digital health, which systematic reviews have identified as a barrier to sustainable adoption and cross-institutional collaboration [12,13]. By making the entire codebase publicly available, we enable independent evaluation, adaptation, and extension by other research groups and clinical institutions without licensing constraints.

The containerized, modular architecture enables institutions to host the platform locally, supporting institutional data governance and compliance with applicable national frameworks, such as Swiss human research legislation [34], while contributing improvements to the shared code base. In addition to packaging the software components themselves, container-based deployment supports the explicit management of runtime state through persistent volumes. Volumes can store data generated and used by running services, such as database contents, uploaded media, configuration files, logs, or study-specific artifacts. When appropriately documented, anonymized, and governed, such volumes or volume exports can be archived and shared with a specific software version to reproduce complex study states, for example, a configured study instance, test dataset, or demonstration environment.

This distinction between container images and persistent volumes is relevant for reproducible rehabilitation research. Container images define the executable software environment, whereas volumes preserve selected runtime state across container restarts, updates, or redeployments. In research contexts, it is important to distinguish between sharing the platform implementation, sharing a configured deployment, and sharing study artifacts. Sensitive runtime data, such as patient records or study recordings, should not be included in public releases unless fully anonymized and ethically approved for sharing. Instead, public repositories can provide empty volume definitions, synthetic example data, documented import/export procedures, and where appropriate, restricted-access volume snapshots for reproducibility audits.

This positions the software as an extensible infrastructure rather than a finished product, a distinction that aligns with the iterative development philosophy described in the *Methods* section. Institutions can adopt the platform and adapt it to local clinical workflows, regulatory requirements, and patient populations without dependency on a single vendor's roadmap or pricing decisions.

An open-source release also supports reproducible research: other groups can inspect the implementation, replicate the evaluation, and build on the platform for their own clinical studies. This is particularly important in a field where proprietary systems often prevent independent verification of technical claims and limit cross-site comparability.

Implications of an Open-Source Rehabilitation Platform

Open-source development in digital health enhances transparency, reproducibility, and long-term sustainability by allowing stakeholders to inspect, modify, and extend software independently of vendor constraints. Reviews of open-source clinical software and electronic health record systems emphasize both opportunities and challenges, including variability in project maturity, documentation standards, and governance models [12,13]. Nevertheless, open-source ecosystems can reduce vendor lock-in, foster collaborative development, and support local adaptation, particularly in heterogeneous rehabilitation environments.

Rehabilitation workflows vary substantially across institutions, specialties, and patient populations. Proprietary systems may limit configurability and integration with external tools. A modular open-source platform can mitigate these constraints by enabling shared intervention libraries, standardized data structures, structured feedback capture, and extensible interfaces for integration with external services. In this context, the platform's architecture supports integration with wearable device

XSL•FO
RenderX

ecosystems, including Fitbit-based data exchange, enabling the incorporation of activity, sleep, and physiological metrics into structured rehabilitation follow-up. By separating core infrastructure (eg, user management, scheduling, and documentation export) from optional analytics and wearable integration modules, the proposed platform supports incremental extension without requiring monolithic deployment.

Beyond technical transparency, open-source architectures are intended to facilitate reproducible digital health research by enabling independent verification, replication studies, and cross-institutional reuse—areas that are frequently constrained by proprietary rehabilitation platforms. Cross-institutional reuse was not evaluated in this study.

Limitations and Future Work

This study reflects an early-stage formative usability evaluation and has several limitations.

First, the sample size was modest, and participants were recruited through the rehabilitation department of the University Hospital of Bern and affiliated rehabilitation settings, which may limit generalizability to broader rehabilitation populations and other institutional contexts. Second, participants reported relatively high familiarity with digital technologies, which may have contributed to favorable usability ratings. Third, the evaluation assessed perceived usability during short task-based sessions and did not measure long-term engagement, adherence, clinical outcomes, or integration into routine rehabilitation pathways. Fourth, the qualitative analysis of open-ended responses was conducted by one researcher. Although the analysis was structured around predefined categories and focused on identifying actionable usability issues for iterative development, the absence of independent dual coding or a formal interrater reliability assessment may have introduced interpretive bias. Future evaluations should include independent coding by multiple researchers or consensus-based qualitative analysis to strengthen the credibility of qualitative findings.

Future studies should include longitudinal field trials in real-world outpatient settings, assessment of sustained engagement and adherence, and evaluation of clinical effectiveness and workflow integration. Scalability under real-world load conditions, including concurrent multiuser access and high-volume data synchronization, has not yet been tested and should be evaluated in deployment studies.

Further development will focus on refinement, clinical integration, technical extension of the platform, and incorporation of feedback from the UX workshop. We will prioritize UX improvements, including simplifying navigation, optimizing the calendar, improving onboarding workflows, providing clearer progress summaries, and enhancing error handling, to further improve usability. Browser-based daily notification reminders have already been implemented in the current version.

Additional motivational features are planned to support adherence during longitudinal follow-up. These may include visualization of the rehabilitation journey, progress indicators, goal-setting functions, completion streaks, milestone badges, and feedback on small achievements or “tiny wins.” Such features should be designed carefully to support motivation without creating pressure, guilt, or inappropriate competition among patients. In rehabilitation contexts, gamification may therefore benefit from focusing on self-monitoring, encouragement, and personalized progress feedback rather than on competitive ranking. Future studies should examine whether these motivational components improve adherence, patient engagement, and perceived self-efficacy over time.

Subsequent work will evaluate the platform’s deployment in real-world outpatient rehabilitation workflows, including the assessment of sustained engagement, adherence patterns, and its impact on patient-reported and clinical outcomes. Implementation studies will examine barriers and facilitators to institutional adoption, including training requirements, governance processes, and integration with existing information systems.

From a technical perspective, future extensions will include deeper integration of wearable device ecosystems via standardized interfaces, such as Fitbit, enabling automated import of activity, sleep, and physiological data into structured rehabilitation plans. The platform architecture is designed to support the integration of AI-based motor assessment and decision-support modules while preserving modularity and reproducibility. Interoperability testing with electronic health record systems using standardized data exchange protocols, such as Fast Healthcare Interoperability Resources, will further support clinical integration and scalability [35].

Conclusions

This study describes the iterative development and formative usability evaluation of the platform, an open-source rehabilitation follow-up tool designed to support structured plan management, patient self-management, and clinician-patient workflow integration. In contrast to monitoring-centric telehealth systems and algorithm-focused AI rehabilitation tools, the platform emphasizes reproducible infrastructure, modular extensibility, and structured rehabilitation planning within real-world follow-up contexts.

The high usability scores observed in this formative evaluation suggest that an open-source, workflow-oriented rehabilitation platform can achieve strong perceived usability among both patients and health care professionals. By integrating structured intervention libraries, standardized feedback capture, and extensible interfaces for wearable device ecosystems, the platform is intended to provide a foundation for scalable, interoperable rehabilitation follow-up; this potential remains to be evaluated in deployment studies.

A central contribution of this work is to release the platform source code under the MIT License, directly addressing the widespread issue of vendor lock-in in digital health. By providing a containerized, modular architecture (Django or React or MongoDB), the architecture can help institutions host the platform locally, supporting data sovereignty and long-term research reproducibility. Future work will focus on longitudinal deployment within outpatient rehabilitation workflows and the evaluation of sustained engagement and clinical effectiveness.

XSL•FO
RenderX

Acknowledgments

The authors thank the patients, health care professionals, and other participants who participated in the usability evaluation. The authors used a generative AI assistant, Claude, for language editing to improve the readability of author-written text and to support them in the development of an open-source platform. It was not used to generate scientific content, analyze data, or draw conclusions. All AI-assisted text was reviewed and verified by the authors, who take full responsibility for the content of the manuscript.

Funding

This work was supported by the COPAIN (Co-creation and evaluation of a sensor-based closed-loop telecare program for multimorbid patients after inpatient rehabilitation) project, funded by the SF-Board of the University of Bern and the Insel Group (grant 750,000 CHF).

Data Availability

The software's source code, documentation, and issue tracker are publicly available on GitHub under the permissive

Massachusetts Institute of Technology (MIT) License [16]. Deidentified participant-level data from the usability evaluation are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Usability task protocol and posttask questionnaire.
[DOCX File, 22 KB - rehab_v13i1e100614_app1.docx]

Multimedia Appendix 2
Post-Study System Usability Questionnaire response direction and scoring.
[DOCX File, 15 KB - rehab_v13i1e100614_app2.docx]

Multimedia Appendix 3
Task completion detail.
[DOCX File, 14 KB - rehab_v13i1e100614_app3.docx]

Multimedia Appendix 4
Qualitative theme–quotation matrix.
[DOCX File, 15 KB - rehab_v13i1e100614_app4.docx]

Checklist 1
Complete REDCap data collection instrument for the Reha-Advisor usability and acceptance study (L2025_Reha-Advisor SUE). The bilingual (German/English) booklet comprises participant eligibility and demographic forms, a role-specific task-performance checklist for health care professionals and patients, two validated usability questionnaires (PSSUQ, System Usability Scale), a semistructured interview guide covering perceived usefulness, trust, and adoption intention, and end-of-study documentation.
[PDF File, 68 KB - rehab_v13i1e100614_app5.pdf]

References

1. Fares N, Sherratt RS, Elhajj IH. Directing and orienting ICT healthcare solutions to address the needs of the aging population. Healthcare (Basel) 2021 Feb 2;9(2):147. [doi: 10.3390/healthcare9020147] [Medline: 33540510]
2. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2021 Dec 19;396(10267):2006-2017. [doi: 10.1016/S0140-6736(20)32340-0] [Medline: 33275908]
3. Rehabilitation. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/rehabilitation [accessed 2026-04-27]
4. Dhakal A, Chowdhury MS. Tele-rehabilitation in rural Nepal: a systematic review of effectiveness, barriers, and strategic directions for digital health equity. Front Digit Health 2025;7:1681313. [doi: 10.3389/fdgth.2025.1681313] [Medline: 41782840]

XSL•FO
RenderX

5. Vellata C, Belli S, Balsamo F, Giordano A, Colombo R, Maggioni G. Effectiveness of telerehabilitation on motor impairments, non-motor symptoms and compliance in patients with Parkinson's disease: a systematic review. Front Neurol 2021;12:627999. [doi: 10.3389/fneur.2021.627999] [Medline: 34512495]
6. Lee AC, Deutsch JE, Holdsworth L, et al. Telerehabilitation in physical therapist practice: a clinical practice guideline from the American Physical Therapy Association. Phys Ther 2024 May 1;104(5):pzae045. [doi: 10.1093/ptj/pzae045] [Medline: 38513257]
7. Molina-Garcia P, Mora-Traverso M, Prieto-Moreno R, Díaz-Vásquez A, Antony B, Ariza-Vega P. Effectiveness and cost-effectiveness of telerehabilitation for musculoskeletal disorders: a systematic review and meta-analysis. Ann Phys Rehabil Med 2024 Feb;67(1):101791. [doi: 10.1016/j.rehab.2023.101791] [Medline: 38128150]
8. Seron P, Oliveros MJ, Gutierrez-Arias R, et al. Effectiveness of telerehabilitation in physical therapy: a rapid overview. Phys Ther 2021 Jun 1;101(6):pzab053. [doi: 10.1093/ptj/pzab053] [Medline: 33561280]
9. Shimbo M, Amiya E, Komuro I. Telemonitoring during exercise training in cardiac telerehabilitation: a review. Rev Cardiovasc Med 2023 Apr;24(4):104. [doi: 10.31083/j.rcm2404104] [Medline: 39076270]
10. Stampa S, Thienel C, Tokgöz P, Razum O, Dockweiler C. Factors facilitating and inhibiting the implementation of telerehabilitation—a scoping review. Healthcare (Basel) 2024 Mar 8;12(6):619. [doi: 10.3390/healthcare12060619] [Medline: 38540583]
11. Gropper A. Open-source health care software. Virtual Mentor 2011 Sep 1;13(9):632-636. [doi: 10.1001/virtualmentor.2011.13.9.stas1-1109] [Medline: 23137477]
12. Shen Z, Spruit M. A systematic review of open source clinical software on GitHub for improving software reuse in smart healthcare. Appl Sci 2019;9(1):150. [doi: 10.3390/app9010150]
13. Syzdykova A, Malta A, Zolfo M, Diro E, Oliveira JL. Open-source electronic health record systems for low-resource settings: systematic review. JMIR Med Inform 2017 Nov 13;5(4):e44. [doi: 10.2196/medinform.8131] [Medline: 29133283]
14. Panchea AM, Létourneau D, Brière S, et al. OpenTera: a microservice architecture solution for rapid prototyping of robotic solutions to COVID-19 challenges in care facilities. Health Technol (Berl) 2022;12(2):583-596. [doi: 10.1007/s12553-021-00636-5] [Medline: 35223361]
15. OpenTeleRehab recognized as digital public good. Humanity & Inclusion (HI). 2022. URL: https://www.hi.org/en/news/opentelerehab-recognized-as-digital-public-good [accessed 2026-04-27]
16. ARTORG-GERONTECHNOLOGY/Bearny-RehaAdvisor. GitHub. 2026 May. URL: https://github.com/ARTORG-GERONTECHNOLOGY/Bearny-RehaAdvisor [accessed 2026-07-10]
17. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (integrate, design, assess, and share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317. [doi: 10.2196/jmir.5927] [Medline: 27986647]
18. Push API. World Wide Web Consortium (W3C). 2025. URL: https://www.w3.org/TR/push-api [accessed 2026-04-27]
19. Web API reference. Fitbit. 2024. URL: https://dev.fitbit.com/build/reference/web-api [accessed 2024-05-21]
20. Hardt D. The OAuth 2.0 Authorization Framework. : Internet Engineering Task Force (IETF); 2012 URL: https://www.ietf.org/ietf-ftp/rfc/rfc6749.txt.pdf [accessed 2026-07-10]
21. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2026-07-10]
22. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research, 2nd edition: Morgan Kaufmann; 2016.
23. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry, 1st edition: CRC Press; 1996:189-194. [doi: 10.1201/9781498710411]
24. Lewis JR. Psychometric evaluation of the PSSUQ using data from five years of usability studies. Int J Hum Comput Interact 2002 Sep;14(3-4):463-488. [doi: 10.1080/10447318.2002.9669130]
25. Gale RC, Wu J, Erhardt T, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the Veterans Health Administration. Implement Sci 2019 Feb 1;14(1):11. [doi: 10.1186/s13012-019-0853-y] [Medline: 30709368]
26. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human participants. JAMA 2025 Jan 7;333(1):71-74. [doi: 10.1001/jama.2024.21972] [Medline: 39425955]
27. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
28. Simmich J, Ross MH, Russell TG. Real-time video telerehabilitation shows comparable satisfaction and similar or better attendance and adherence compared with in-person physiotherapy: a systematic review. J Physiother 2024 Jul;70(3):181-192. [doi: 10.1016/j.jphys.2024.06.001] [Medline: 38879432]
29. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
30. Palagin OV, Malakhov KS, Velychko VY, Semykopna TV. Hybrid e-rehabilitation services: SMART-system for remote support of rehabilitation activities and services. Int J Telerehabil 2022;14(Spec Issue):e6480. [doi: 10.5195/ijt.2022.6480] [Medline: 37063645]

XSL•FO
RenderX

31. Ciabattoni L, Ceravolo MG, Capecci M, Pepa L. Telerehabilitation outcome prediction via machine learning. In: Proceedings of the 2024 IEEE 3rd International Conference on Intelligent Reality (ICIR): IEEE; 2024. [doi: 10.1109/ICIR64558.2024.10976921]
32. Skaramagkas V, Kyprakis I, Karanasiou GS, Fotiadis DI, Tsiknakis M. A review on deep learning for quality of life assessment through the use of wearable data. IEEE Open J Eng Med Biol 2025;6:261-268. [doi: 10.1109/OJEMB.2025.3526457] [Medline: 39906266]
33. Tang K. Computer-aided integrated rehabilitation for post-stroke patients using deep learning techniques [Dissertation]. : McWilliams School of Biomedical Informatics at UTHealth Houston, University of Texas Health Science Center at Houston (DigitalCommons@TMC); 2025 URL: https://digitalcommons.library.tmc.edu/cgi/viewcontent.cgi?article=1073&context=uthshis_dissertations [accessed 2026-04-27]
34. Federal Act on Research involving Human Beings (Human Research Act, HRA), SR 810.30. : Swiss Confederation; 2011 URL: https://www.freedomofresearch.org/wp-content/uploads/2019/11/6-CC-810.30-Federal-Act-of-30-September-2011-on-Research-involving-Human-Beings-LRH.pdf [accessed 2026-04-27]
35. FHIR overview. HL7 FHIR Release 5. 2023. URL: https://fhir.hl7.org/fhir/overview.html [accessed 2026-04-27]

Abbreviations

IDEAS: integrate, design, assess, and share
KEK: Kantonale Ethikkommission (Cantonal Ethics Committee)
MIT: Massachusetts Institute of Technology
PSSUQ: Post-Study System Usability Questionnaire
SSL: Secure Sockets Layer
SUS: System Usability Scale
UX: user experience

Edited by J Sarvestan; submitted 07.May.2026; peer-reviewed by C Kapatais, H Kweon; revised version received 25.Jun.2026; accepted 01.Jul.2026; published 13.Aug.2026.

Please cite as:
Angelva NE, Gurschler F, Single M, Frischknecht V, Näf D, Kampus N, Cioffi E, Marcin T, Savic B, Wilhelm M, Scheel-Sailer A, Martin-Niedecken AL, Mosimann U, Nef T
Iterative Development and Usability Evaluation of an Open-Source Rehabilitation Follow-Up Platform: User-Centered Development Study
JMIR Rehabil Assist Technol 2026;13:e100614
URL: https://rehab.jmir.org/2026/1/e100614
doi:10.2196/100614

© Noora Emilia Angelva, Franziska Gurschler, Michael Single, Valerie Frischknecht, Dominik Näf, Noel Kampus, Eva Cioffi, Thimo Marcin, Branislav Savic, Matthias Wilhelm, Anke Scheel-Sailer, Anna Lisa Martin-Niedecken, Urs Mosimann, Tobias Nef. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 13.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Visual Telerehabilitation Program in Virtual Reality for Age-Related Macular Degeneration: Randomized Feasibility and Proof-of-Concept Trial

Yulia Pyatova[1], MD; Bill Zhang[2], MSci; Mariana Misawa[1], MD; Samuel N Markowitz[1], MD; Atri Sen[2], BSc; Lora Appel[3], PhD; Kyle Cheung[2], BSc; Pi Nasir[2], MSc; Danielle Tchao[4], MEng; Eduardo Garcia-Giler[2], MSci; Monica Daibert-Nido[1], MD; Michael Reber[2], PhD

[1]University Health Network, Toronto Western Hospital, Toronto, ON, Canada

[2]University Health Network, Donald K. Johnson Eye Institute, 60 Leonard Avenue, Toronto, ON, Canada

[3]School of Health Policy and Management, Faculty of Health, York University, Toronto, ON, Canada

[4]University Health Network, OpenLab, Toronto, ON, Canada

Corresponding Author:
Michael Reber, PhD
University Health Network, Donald K. Johnson Eye Institute, 60 Leonard Avenue, Toronto, ON, Canada

Abstract

Background: Age-related macular degeneration (AMD) causes progressive central vision loss in older adults. Low-vision rehabilitation can improve functional vision by training the use of a preferred retinal locus, commonly through clinic-based biofeedback training (BFT). However, repeated supervised rehabilitation is burdensome, and functional gains may be difficult to sustain without home practice. Stand-alone virtual reality (VR) may enable home-based, remotely monitored visual stimulation, but feasibility, safety, and usability in older adults with AMD remain insufficiently characterized.

Objective: This study aimed to evaluate the feasibility and safety of adding home-based VR 3D single-object tracking (3D-SOT-VR) to conventional BFT in older adults with dry AMD in a parallel, randomized, single-blind (to assessors), controlled, formative trial and to generate exploratory functional hypotheses for a future trial.

Methods: Adults with dry AMD were recruited at the Low Vision Clinic, Toronto Western Hospital, University Health Network, Toronto, Ontario, Canada, from September 2021 to October 2023. Participants were randomized to BFT once weekly for 4 weeks (BFT group) or BFT plus home-based 3D-SOT-VR (BFT-VR group) every other day for 4 weeks. Experimental intervention consisted of tracking a single object among distractors moving at different speeds in a 3D virtual space in a VR headset. Primary feasibility and safety outcomes included recruitment, adoption, adherence, compliance, intervention completion, remote data transfer, usability, and VR-induced symptoms and effects. Secondary outcomes included visual acuity, contrast sensitivity, fixation stability, retinal sensitivity, reading speed, and low-vision quality of life. Exploratory outcomes assessed performance at 3D-SOT-VR and usage. Analyses were descriptive and exploratory, with CIs and denominators reported to reflect limited precision and missingness.

Results: Fourteen individuals were randomized (BFT, n=6; BFT-VR, n=8), below the planned sample size of 32. Recruitment was not achieved because of COVID-19–related interruptions and reduced onsite access. Eleven individuals were analyzed for the primary outcome (BFT n=6, BFT-VR n=5). Intervention completion was 100% in the BFT arm and 75% in the BFT-VR arm, below the prespecified BFT-VR threshold. Among participants who used VR, adherence to scheduled home sessions was acceptable, completed VR-session files were transmitted without loss, and no participant met the predefined cybersickness stopping rule. One BFT-VR participant discontinued because headset weight caused neck fatigue. Group-level visual outcomes did not provide significant effectiveness. Reading speed showed a clinically meaningful individual-level improvement in the BFT-VR arm and correlated with VR-task performance. The findings were not clearly durable at follow-up.

Conclusions: This pilot study provides formative evidence that clinic-based BFT combined with home-based, remotely monitored VR visual stimulation can be implemented safely in older adults with dry AMD, while identifying major contextual feasibility barriers. The intervention is innovative because it extends low-vision rehabilitation into the home using a connected device and objective performance monitoring. Recruitment, retention, missing data handling, and sustainability of functional gains must be addressed before effectiveness testing.

Trial Registration: ClinicalTrials.gov NCT04685824; https://clinicaltrials.gov/study/NCT04685824

(JMIR Rehabil Assist Technol 2026;13:e87596) doi:10.2196/87596

XSL•FO
RenderX

KEYWORDS

macular degeneration; low vision; visual rehabilitation; biofeedback training; virtual reality

Introduction

Background

Age-related macular degeneration (AMD) is the main cause of irreversible vision loss in developed countries, responsible for the loss of central, or macular, vision, substantially affecting reading, face recognition, mobility, independence, and quality of life [1,2]. According to BrightFocus Foundation, the number of people living with macular degeneration is expected to reach 196 million worldwide in 2020 and to increase to 288 million by 2040 [3]. There are no treatments available to restore vision and cure the disease; therefore, low-vision rehabilitation remains a central component of supportive care. Contemporary rehabilitation aims not only to prescribe magnification or optical aids, but also to improve the functional use of residual peripheral retina and to reduce the activity limitations caused by central scotoma [4,5]. A core concept in AMD rehabilitation is the preferred retinal locus (PRL), an eccentric retinal area that patients use to substitute for foveal fixation after macular damage [4]. Following the degeneration of the macula and the loss of central vision, patients naturally relocate the fixation point for high visual acuity (VA) peripherally, in an eccentric and healthy part of the retina, to compensate for visual impairments [6]. Relocation to the PRL occurs over time [4,7] but, in more than 25% of cases, the naturally developed PRL lands on a suboptimal retinal area and becomes inefficient, contributing to impaired reading and daily visual tasks [4,6,7]. Modern standard low-vision care for patients with AMD includes biofeedback training (BFT) to relocate the PRL to a healthy retinal patch identified by the health care provider and to acquire better fixation skills [4,7-9]. BFT has been shown to improve distance and near vision, reading speed, contrast sensitivity, retinal sensitivity, and therefore quality of life [4,7,10-12]. However, in clinical practice, it is observed that the patients do not always use the PRL as trained for activities of daily living, losing the benefits of BFT over time, while also highlighting uncertainties regarding optimal training dose, treatment interval, durability, and patient selection [9,13]. The use of the new PRL requires an adjustment of the oculomotor control [4,6,7] and a brain adaptation [14]. In patients with AMD, not only is fixation stability impaired at the PRL [6], negatively affecting VA [15] and making daily tasks such as reading and grooming difficult, but the increased eccentricity of a visual target also extends reaction times and multiplies error rates, deteriorating performance in object detection and recognition tasks [1,2,16]. These observations provide a rationale for combining conventional BFT with repeated attentional and visuomotor training outside the clinic. However, standard BFT is typically delivered under professional supervision in weekly or biweekly sessions, creating access and treatment-burden constraints for older adults. Home-based digital rehabilitation could increase training opportunities, but feasibility must be established before clinical effectiveness can be inferred.

Objective

A new and promising vision rehabilitation technique using binocular 3D single object tracking (3D-SOT) deployed in virtual reality (VR) headsets can provide additional training of the ocular system [17-19]. This innovative technique is particularly well suited for training because the central features of the 3D-SOT closely match attentionally demanding real-life situations [20,21]. Three-dimensional object tracking stimulation programs displayed on monitors or television screens have been shown to increase brain capacity for complex processes such as anticipation, eye-tracking, field of view, and even decision-making in healthy and pathological participants [22-24]. Preliminary observations in low-vision patients with homonymous hemianopsia indicated a beneficial effect on visual perception of an audiovisual 3D object tracking VR stimulation program using stand-alone VR headsets at home [17,18,25]. We hypothesize that the intervention combining BFT and audiovisual 3D-SOT-VR stimulation may have an additive effect in restitution of visual functions and better visual skills resulting in an improved quality of life.

Methods

Study Design

This study was a 4-week, open-label, reference-controlled, randomized, feasibility, and proof-of-concept trial assessing feasibility (recruitment rates, adoption, adherence and compliance, data collection and retrieval, and safety) and potential effectiveness of BFT+3D-SOT-VR (BFT-VR group) versus BFT (BFT group) on visual perception in individuals with AMD and is reported following the CONSORT eHealth statement for pilot and feasibility studies.

Ethical Considerations

The study protocol has been approved by the Research Ethics Boards at the University Health Network, Toronto, Ontario, Canada (UHN REB# 20‐6143) and registered on ClinicalTrials.gov (NCT04685824; posted on December 28, 2020) following the principles outlined in the World Medical Association Declaration of Helsinki: Research involving human subjects. All participants provided written informed consent before study procedures. The original informed consent allows the secondary analysis without additional consent. Study data were deidentified and pseudonymized for analysis, stored on secure institutional systems, and transmitted through the predefined secure study data management pathway. Participants received US $36 per visit as compensation for commute and parking expenses. No identifiable participant images are included in the paper or multimedia appendices. The study was conducted at the Ophthalmology service, Toronto Western Hospital, University Health Network, Toronto, Ontario, Canada, between September 15, 2021, and October 6, 2023, with multiple interruptions and extensions due to the Ontario Emergency Management and Civil Protection Act related to the COVID-19 pandemic [26].

XSL•FO
RenderX

Participants

The study population included adults with dry AMD from the Toronto Western Hospital pool of patients who were referred by their ophthalmologists. Participants were enrolled by the co-principal investigator ophthalmologist. Inclusion criteria were diagnosed with dry AMD, male and female, VA (best-corrected visual acuity)<0.9 logMAR (logarithm of the minimum angle of resolution), ability to follow the visual and auditory stimuli and training instructions, online auditory test positive (−5 dBHL [decibels hearing level] to 30 dBHL range), home Wi-Fi access. Exclusion criteria were ocular disease not related to AMD physiopathology, wet AMD, both eyes with media opacity that impairs microperimetry testing, inability to perform during testing and training, consumption of psychoactive drugs, three consecutive Virtual Reality-Induced Symptoms and Effects (VRISE) questionnaire scores <25 at inclusion [27], history of vertigo or dizziness, and prior visual rehabilitation interventions.

Randomization and Intervention

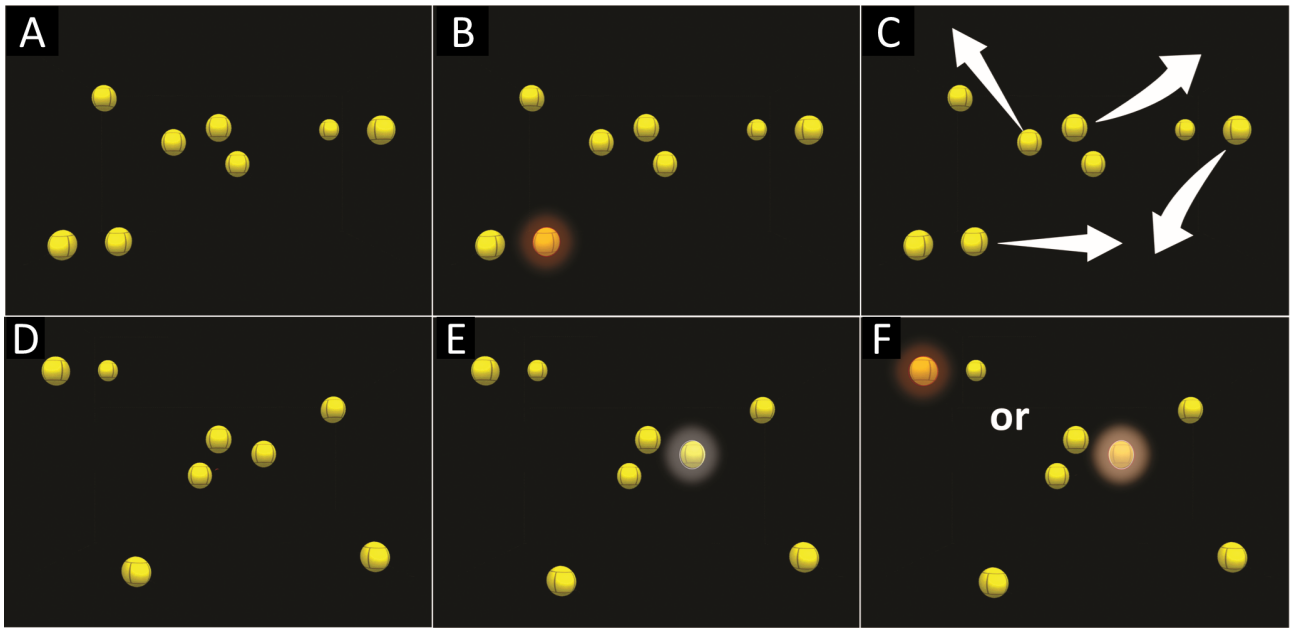
Following guidelines for feasibility and pilot studies [28,29], the sample size was determined at 16 participants per arm. The participants were randomized 1:1 (permuted block, block size=2, 4, sealedenveloppe.com generated by the principal investigator) to receive audiovisual monocular (either right -oculus dexter, or left -oculus sinister, eye) BFT for 20 minutes once a week for 4 weeks at Toronto Western Hospital (reference-controlled group, BFT group) as described earlier [7], or audiovisual monocular BFT once a week for 4 weeks and binocular 3D-SOT-VR every other day at home for 4 weeks (interventional group, BFT-VR group), using the stand-alone and remotely controlled head-mounted display Meta Quest 2 (allowing for additional wearing of glasses) provided by our laboratory (Figure 1). Each session of stimulation is composed of three blocks of 15 trials of 20 seconds, equivalent to 3 × 5 minutes of continuous audiovisual stimulation corresponding to 16 sessions for a total time of 3 hours and 15 minutes of stimulation. The audiovisual stimulation task corresponds to the object-tracking paradigm, developed in the 1980s to study visual attention in humans [20,21,30]. The 3D-SOT is composed of eight high-contrast yellow spheres on a black background (luminosity=80 cd/m^2, size=1°-3° of visual angle). After one of the spheres was cued (turning red for 5 s and switching back to yellow; Figure 2), the spheres moved for 20 seconds following random linear paths, bouncing off one another and off the walls of a virtual 3D cube when collisions occurred. The overall span of the movement of the spheres covers 80° and 70° horizontal and vertical visual angles, respectively, as measured using augmented reality and a geometric chart. The initial speed of the spheres was adjustable, from 3 to 240°/second, and determined individually for each participant at inclusion. A spatial sound (125 Hz, 0 - 30 dBHL) is emitted by the cued target and spatially and temporally correlated to its movement. The maximum delay between sound and target is equivalent to the refreshing rate of the headsets, corresponding to 16.6 ms (temporal tolerance), below the 100 ms temporal disparity above which multisensory response enhancement decreases [31]. After 20 seconds, the movement stops and the participants are asked to select, using a virtual laser pointer, the cued sphere (the target) among the other spheres (mark-all procedure) [21]. If the selection is correct (corresponding to the cued target), a positive feedback sound is provided, and the speed of the spheres in the next trial is increased. If the selection is incorrect (corresponding to a distractor), a negative feedback sound is provided, and the speed of the spheres in the next trial is decreased. Speed is modified following an adaptive simple up-down staircase:±0.05 log [32,33]. Participants were allowed to use head and eye movements to track the moving target (overt tracking). During the home-based procedures, head or target position was recorded during the task by the Meta Quest 2 built-in head tracking system (sampling rate 60 Hz) downsampled to 10 Hz for analysis.

XSL•FO
RenderX

Figure 1. Study flowchart according to CONSORT (Consolidated Standards of Reporting Trials) 2025 flow diagram. BFT: biofeedback training; VR: virtual reality; FU: follow-up.

XSL•FO
RenderX

Figure 2. Dynamic audiovisual 3D single-object tracking (3D-SOT-VR) stimulation. (A) Eight yellow still spheres are present in a virtual cube. (B) One of these spheres turns red for 15 seconds (cued target) and returns yellow. (C) All spheres randomly move following linear paths across the visual field encompassing the blind field and bouncing off one another and off the walls of the virtual 3D cube when collisions occur. (D) After 30 seconds, spheres stopped moving. (E) The patient had to select the cued target using a hand-guided 32 laser pointer. (F) Correct selection is considered a positive hit.



Outcome Measures

Primary outcomes correspond to feasibility (recruitment rate, dropouts, adoption, adherence, and compliance) and safety (cybersickness). Secondary outcomes correspond to visual function and functional vision measured at the Ophthalmology Low Vision Clinic at the Toronto Western Hospital (University Health Network) following standard procedures in low-vision rehabilitation [5,34]. These visual assessments, blind to the assessor, include VA (best-corrected visual acuity) measured using the Early Treatment Diabetic Retinopathy Study and Colenbrander charts at 4 m; contrast sensitivity measured using the Functional Acuity Contrast Test. Fixation stability (bivariate contour ellipse area 63%) and retinal sensitivity were measured using the Macular Integrity Assessment microperimeter (CentreVue). Reading speed was measured at critical print size using the Minnesota Low Vision Reading test [35]. Quality of vision was measured using the Veterans Affairs–Low-Vision Visual Functioning Questionnaire [36].

Exploratory outcomes in the BFT-VR group corresponded to home-based VR-task performance, including target speed for correct selection, reaction time, head movement, and session time of day. These exploratory variables were collected remotely during the VR task and were used to evaluate operational feasibility and generate mechanistic hypotheses.

Safety, Monitoring, Criteria for Removal, Early Stopping Rules, and Condition of Replacement

Monitoring of the at-home 3D-SOT-VR stimulation program occurred after each 5-minute block when the head-mounted display sent data relative to the participants' adherence and compliance (number of tasks and number of blocks initiated and completed, duration of the blocks or sessions), performance (correct or incorrect selection ratio, reaction time, speed of the target, head movement, and time of day), and safety (VRISE – questionnaire scores before and after intervention) to our secured laboratory computer through Wi-Fi. Criteria for removal and stopping rules were 3 consecutive home-based VRISE scores postintervention ≤25 [27] or no 3D-SOT-VR stimulation reports for 5 consecutive days. There was no replacement for withdrawn participants.

Data and Statistical Analysis

Data analysis of the primary outcomes was performed against predetermined feasibility and safety criteria. Data analysis on secondary outcomes was performed following both intention-to-treat (ITT) and per protocol (PP) methods, based on descriptive statistics including a measure of central tendency (mean) and dispersion (95% CI). Statistical comparisons between groups were performed using the Welch *t* test. Comparisons between pre- and postintervention data were performed using the Wilcoxon signed-rank test. Clinical meaningfulness was determined using the coefficient of repeatability (COR), specific to each test, to compare pre- and postintervention data [37]. The COR, also referred to as the smallest real difference, quantifies absolute reliability in the same unit as the assessment tool. The COR is directly related to the 95% limits of agreement proposed by Bland and Altman [38]. It corresponds to the value below which the absolute differences between two measurements would lie within 95% of probability. Therefore, measurement values strictly different from COR indicate an effect of the intervention [37,39]. Interaction effects were estimated using ANOVA.

Missing Data and Missing Completely At Random Assessment

Missingness was summarized by study arm, visit, and outcome category. Missing data occurred mainly at the participant or visit level because of withdrawal, missed postintervention

XSL•FO
RenderX

assessment, intervention discontinuation, and follow-up loss, rather than as isolated item-level missingness within completed visits. We assessed whether the observed missing data structure was compatible with a Missing Completely At Random (MCAR) mechanism, defined as missingness unrelated to both observed and unobserved data [40]. A formal Little MCAR test was considered [40], but the very small randomized sample, sparse missing data patterns, and number of repeated outcome variables made the test statistically underpowered and not reliably interpretable. Moreover, several missing observations had documented clinical or logistical explanations, making a strict MCAR assumption implausible. Multiple imputation was therefore considered but not implemented because the available sample was too small to support a stable imputation model across multiple outcomes and visits [41]. Analyses were consequently based on available observations, with denominators reported and secondary outcomes interpreted as exploratory.

Results

Feasibility and Safety

Feasibility criteria were defined based on previous studies using VR as an interventional tool [42,43] and CONSORT (Consolidated Standards of Reporting Trials) guidelines for feasibility trials [28]. Of 32 participants determined by power calculation, 18 were screened and 14 randomized in a 1:1 ratio to BFT (n=6) or BFT-VR (n=8; Figure 1 and Table 1).

Table . Participants' demographics.

Group	Sample size, n	Age (years), mean (SD)	Sex ratio	Treatment history
BFT[a]	6	81.5 (5.5)	4F[b], 2M[c]	Last anti-VEGF[d] injection >12 months
BFT-VR[e]	8	83.2 (5.3)	6F, 2M	Last anti-VEGF injection >12 months

[a]BFT: biofeedback training.

[b]F: female.

[c]M: male.

[d]VEGF: vascular endothelial growth factor.

[e]BFT-VR: biofeedback training in immersive virtual reality.

Recruitment rate was below feasibility criteria (≥1.5 participants/week) for both arms, at 0.13 and 0.11 participants/week for BFT-VR and BFT, respectively (Tables 2 and 3). Overall, 6/6 participants (100%) completed the reference-controlled intervention in the BFT arm, and 6/8 participants (75%) completed the investigational intervention in the BFT-VR arm, not reaching feasibility criteria for intervention completion (≥80% of the participants; Table 2). In the BFT-VR arm, 6/8 participants performed 90.6% (87/96) of the home-based scheduled stimulation sessions (45 trials of 20 s), reaching feasibility for adherence and compliance criteria (≥80%). Remote safety and exploratory data collection during the sessions reached the feasibility threshold (≥90%) with 100% of the files collected (96/96) through Wi-Fi, indicating no loss of data. Safety related to the stimulation was assessed using the VRISE questionnaire score, which evaluates cybersickness [27]. VRISE scores post stimulation showed a total average of 33.8 (SD 0.4) out of 35 points, with no scores ≤25 indicating no adverse event (cybersickness symptoms scores criteria ≤3 consecutive VRISE scores ≤25) [27,44] (Table 2). Detailed analysis showed that, for a total of 96 sessions, among the five types of VRISE symptoms (nausea, disorientation, dizziness, fatigue, and instability), fatigue ($\Delta_{post\text{-}pre}$= −0.37, 95% CI −0.46 to −0.28) and dizziness ($\Delta_{post\text{-}pre}$= −0.25, 95% CI −0.16 to −0.34) were the most pronounced symptoms after stimulation, with scores of 6.31 (95% CI 6.11-6.51) and 6.60 (95% CI 6.45-6.75) respectively, although above the cybersickness threshold (Tables 2 and 3).

One drop-out was due to the intervention: the weight of the headset (Meta Quest 2) was reported by the participant as too heavy for the neck muscles, which became sore after the 15-minute session. Usability, measured using a modified version of the virtual reality neuroscience questionnaire [27], showed an overall score >40 above the minimum threshold indicating adequate quality of the software (Table 2). However, three items related to the use of the VR device itself but not to the stimulation, VR device adjustment with straps and pointer use and selection using the remote controller, did not reach the acceptable threshold (Tables 2 and 3), indicating potential interference with the completion of the home-based program.

XSL•FO
RenderX

Table . Feasibility scores.

	Intervention (BFT-VR)[a,b] (n=8)	Reference-controlled (BFT; n=6)
Feasibility scores		
Recruitment rate, participants/wk[c]	0.13	0.11
Dropouts, n[d]	1	0
Adoption, (%, ratio)[e]	100% (8/8)	100% (6/6)
Adherence/compliance BFT (%, ratio)[e]	100% (28/28)	100% (24/24)
Adherence/compliance IVR (%, ratio)[e]	75% (96/128)	—[f]
Remote data collection through Wi-Fi (%, ratio)[g]	100% (96/96)	—

[a]BFT: biofeedback training.
[b]VR: virtual reality.
[c]criteria: ≥1.5.
[d]criteria: ≤3.
[e]criteria: ≥80%.
[f]Not applicable.
[g]criteria: ≥90%.

XSL•FO
RenderX

Table . Safety and usability scores.

Scores	Intervention (BFT-VR)[a,b], average score (95% CI)
Feasibility scores[c,d]	
Nausea	• −0.03 (−0.07 to 0.01) • 6.93 (6.86 to 7.00)
Disorientation	• −0.09 (0.00 to 0.18) • 6.84 (6.75 to 6.93)
Dizziness	• −0.25 (-0.34 to −0.16) • 6.60 (6.45 to 6.75)
Fatigue	• −0.37 (-0.46 to −0.28) • 6.31 (6.11 to 6.51)
Instability	• −0.02 (−0.07 to 0.03) • 6.92 (6.86 to 6.98)
Total scores post-IVR[e]	• −0.76 (−1.11 to −0.41) • 33.8 (33.4 to 34.2)
Usability scores[f,g]	
Putting device on	• 5.0 (4.1 to 5.9)
Adjusting device	• 4.8 (4.0 to 5.6)
Quality of graphics	• 5.6 (4.9 to 6.3)
Quality of sound	• 5.4 (4.7 to 6.1)
Quality of VR technology	• 5.8 (5.2 to 6.4)
Use of the pointer	• 4.4 (3.1 to 5.70)
Selection using the pointer	• 4.6 (3.6 to 5.6)
Device recharging process	• 6.1 (5.6 to 6.6)
Average total scores post-IVR	• 42.9 (39.6 to 46.2)

[a]BFT: biofeedback training.

[b]VR: virtual reality.

[c]Average score post-IVR (95% CI).

[d]Criteria average score: >5, criteria total score: >25.

[e]IVR: immersive virtual reality.

[f]Average score (virtual reality neuroscience questionnaire).

[g]Criteria average score: >5, criteria total score: >40.

Preliminary Effect of the Intervention

Effectiveness analyses followed the ITT principle, comprising all 16 randomized participants analyzed as allocated. Because 2 participants in the intervention arm discontinued the program after less than 80% of the sessions and one missed all posttreatment assessments, we prospectively defined a PP set that included participants who attended ≥80% of the sessions and contributed a valid primary-outcome measurement at 4 weeks. Performing both ITT and PP analyses is in line with ICH (International Conference on Harmonisation) E9 guidance (§5.2.3) on assessing the robustness of trial inferences [45], and with CONSORT 2010 recommendations [28] to report numbers analyzed for each set. The PP analysis is reported as a sensitivity check; the ITT analysis remains the primary basis for inference. Clinical relevance was determined using each test's COR; changes exceeding COR are clinically significant [37].

At baseline, groups did not differ in near or distance VA (ITT [Welch t test, 2-tailed]: near VA: $t_{12.1}$=0.75; P=.47; far VA: $t_{15.5}$=0.64; P=.53; PP [Welch t test, 2-tailed]: near VA: $t_{12.9}$=0.80; P=.44; far VA: t_{16}=0.32; P=.75). Post training, no statistically significant group-level gains emerged in either metric, and between-arm (ITT and PP analyses; Tables 4 and

XSL•FO
RenderX

5). Nevertheless, individual COR-positive improvement was observed: 3 of 6 BFT participants improved by ≥0.2 logMAR in near vision and 1 improved for distance; in BFT-VR, 3 of 8 participants showed near-vision gains and 2 showed distance-vision gains. One- and six-month follow-up revealed stable mean acuity but individual fluctuation (Table S1 in Multimedia Appendix 1).

Fixation stability (bivariate contour ellipse area 63%) showed no difference between groups at baseline (ITT: Welsch *t* test: $t_{9.85}$=1.63, *P*=.14; PP: Welsch *t* test: $t_{11.6}$=1.69, *P*=.12). After four weeks, BFT recipients showed no mean change (ITT: Welsch *t* test: $t_{18.1}$=1.37, *P*=.19; PP: Welsch *t* test: $t_{19.5}$=0.86, *P*=.40; Tables 4 and 5) whereas the PP subset of BFT-VR displayed a COR-level worsening (Tables 4 and 5). Individually, two BFT participants worsened (COR>0.61 °^{2}), and one was stable. In the BFT-VR group, three participants had normal fixation stability at baseline and one worsened after intervention (COR>0.61 °^{2}); the others did not show significant changes. Measures at 1- and 6-month post intervention showed no changes (Table S1 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Measures of visual outcomes and statistical analysis comparing experimental (biofeedback training in immersive virtual reality) and reference-controlled (biofeedback training) interventions (intention-to-treat).

Function	Intervention (BFT-VR)[a]				Reference controlled (BFT)[b]			
	Baseline, mean (95% CI)	Postintervention, mean (95% CI)	Mean difference (95% CI)	Wilcoxon ranked test	Baseline, mean (95% CI)	Postintervention, mean (95% CI)	Mean difference (95% CI)	Wilcoxon ranked test
Visual function								
Visual acuity (near-LogMAR)[c,d]	0.43 (0.29 to 0.56)	0.33 (0.20 to 0.46)	−0.06 (0.17 to 0.05)	—[e]	0.59 (0.19 to 0.99)	0.50 (0.21 to 0.79)	−0.14 (0.30 to 0.03)	8.5[f]
Visual acuity (far-LogMAR)[g]	0.63 (0.51 - 0.76)	0.65 (0.51 to 0.79)	−0.02 (−0.17 to 0.13)	7[h]	0.71 (0.4 to 1.01)	0.68 (0.40 to 0.95)	−0.04 (−0.21 to 0.14)	6[f]
Fixation stability (BCEA[i] 63 % to °2)[j]	1.92 (1.22 to 2.62)	2.16 (0.89 to 3.44)	0.35 (−0.93 to 1.63)	11[h]	4.25 (1.34 to 7.17)	4.06 (1.73 to 6.39)	−0.57 (−3.34 to 2.20)	9[f]
Retinal sensitivity, dB[k]	12.3 (9.74 to 14.9)	13.1 (10.4 to 15.7)	0.03 (−0.73 to 0.80)	12[h]	10.1 (5.79 to 14.4)	9.48 (4.99 to 14.0)	−0.62 (−1.39 to 0.15)	7[f]
Functional vision								
Reading speed (WPM)[l,m]	99.6 (44.8 to 154.3)	105.0 (43.5 to 166.8)	13.2 (−10.8 to 37.3)	10[h]	82.5 (20.6 to 144.4)	76.1 (31.3 to 120.9)	−6.4 (−53.3 to 40.6)	7[f]
Quality of life								
LV-VFQ-48[n,o]								
Visual ability	1.22 (0.39 to 2.04)	1.56 (0.75 to 2.36)	0.29 (-0.12 to 0.69)	8[h]	1.51 (0.50 to 2.51)	1.92 (0.92 to 2.93)	0.42 (0.05 to 0.78)	2[p]
Reading	0.82 (-0.27 to 1.92)	1.70 (0.3 to 3.00)	0.92 (0.17 to 1.66)	2[q]	1.84 (0.51 to 3.17)	3.01 (0.76 to 5.26)	1.17 (0.06 to 2.28)	1[r]
Mobility	1.47 (0.77 to 2.18)	1.59 (0.97 to 2.21)	0.08 (-0.23 to 0.39)	13[h]	1.24 (0.28 to 2.21)	1.59 (0.57 to 2.62)	0.35 (−0.18 to 0.88)	4[f]
Visual information	1.46 (0.61 to 2.31)	1.54 (0.68 to 2.40)	0.01 (-0.61 to 0.64)	13[h]	1.67 (0.44 to 2.90)	2.02 (1.06 to 2.97)	0.35 (−0.12 to 0.83)	3[f]
Visual motor	1.09 (0.06 to 2.12)	1.24 (0.44 to 2.05)	0.13 (-0.41 to 0.68)	9[h]	1.40 (0.50 to 2.30)	1.64 (0.86 to 2.43)	0.24 (−0.32 to 0.80)	3[f]
COR 2.2[s]								
Total	6.02 (1.84 to 10.2)	7.70 (3.7 to 11.6)	1.68 (-0.46 to 3.81)	8[h]	7.66 (2.55 to 12.8)	10.2 (4.72 to 15.7)	2.54 (0.48 to 4.59)	2[p]

[a]BFT-VR: biofeedback training in immersive virtual reality.

[b]BFT: biofeedback training.

[c]LogMAR: logarithm of the minimum angle of resolution.

[d]COR=0.24.

[e]Not applicable.

[f]n=6; *P*>.2.

[g]COR=0.11.

[h]n=8; *P*>.2.

[i]BCEA: bivariate contour ellipse area.

[j]COR=0.61.

[k]COR=1.6.

[l]WPM: words per minute.

[m]COR=8.6.

[n]LV-VFQ-48: Low vision visual function questionnaire.

XSL•FO
RenderX

[o]COR=0.44.

[p]n=6.05<*P*<.1.

[q]n=8.02<*P*<.05.

[r]n=6.1<*P*<.2.

[s]COR: coefficient of repeatability.

XSL•FO
RenderX

Table . Measures of visual outcomes and statistical analysis comparing experimental (biofeedback training in immersive virtual reality) and reference-controlled (biofeedback training) interventions (per protocol).

Function	Intervention (BFT-VR)[a]				Reference controlled (BFT)[b]			
	Baseline, mean (95% CI)	Postintervention, mean (95% CI)	Mean difference (95% CI)	Wilcoxon ranked test	Baseline, mean (95% CI)	Postintervention, mean (95% CI)	Mean difference (95% CI)	Wilcoxon ranked test
Visual function								
Visual acuity (near-LogMAR)[c,d]	0.41 (0.25 to 0.57)	0.33 (0.16 to 0.50)	−0.07 (−0.23 to 0.08)	5[e]	0.59 (0.19 to 0.99)	0.50 (0.21 to 0.79)	−0.15 (−0.33 to 0.03)	8.5[f]
Visual acuity (far-LogMAR)[g]	0.66 (0.51 to 0.81)	0.59 (0.45 to 0.73)	−0.07 (−0.13 to −0.01)	1[h]	0.71 (0.42 to 1.01)	0.68 (0.40 to 0.95)	−0.04 (−0.21 to 0.14)	6[f]
Fixation stability (BCEA[i] 63 % to °2)[j]	1.53 (0.65 to 2.41)	2.77 (0.98 to 4.56)	1.24 (-0.30 to 2.78)	2[h]	4.25 (1.34 to 7.17)	4.06 (1.73 to 6.39)	−0.57 (−3.34 to 2.20)	9[f]
Retinal sensitivity (dB)[k]	14.1 (11.3 to 16.9)	13.7 (10.3 to 17.1)	−0.05 (−1.13 to 1.03)	6[e]	10.1 (5.79 to 14.4)	9.48 (4.99 to 14.0)	−0.62 (−1.39 to 0.15)	7[f]
Functional vision								
Reading speed (WPM)[l,m]	96.1 (14.9 to 177.3)	117.4 (31.7 to 203)	21.3 (−10.1 to 52.6)	28[n]	82.5 (20.6 to 144.4)	76.1 (9531.3 to 120.9)	−6.4 (−53.3 to 40.6)	7[f]
Quality of life								
LV-VFQ-48[o]								
Visual ability	1.70 (0.73 to 2.68)	1.87 (0.91 to 2.83)	0.16 (−0.42 to 0.75)	7[e]	1.51 (0.50 to 2.51)	1.92 (0.92 to 2.93)	0.42 (−0.05 to 0.78)	2[p]
Reading	1.27 (−0.31 to 2.86)	2.21 (0.44 to 3.97)	0.93 (−0.20 to 2.07)	2[h]	1.84 (0.51 to 3.17)	3.01 (0.76 to 5.26)	1.17 (0.06 to 2.28)	1[q]
Mobility	1.72 (0.92 to 2.52)	1.84 (1.13 to 2.55)	0.13 (−0.33 to 0.58)	6[e]	1.24 (0.28 to 2.21)	1.59 (0.57 to 2.62)	0.35 (−0.18 to 0.88)	4[f]
Visual information	2.00 (0.97 to 3.03)	1.75 (0.63 to 2.86)	−0.25 (−1.07 to 0.57)	5[e]	1.67 (0.44 to 2.90)	2.02 (1.06 to 2.97)	0.35 (−0.12 to 0.83)	3[f]
Visual motor	1.75 (0.90 to 2.60)	1.57 (0.69 to 2.44)	−0.19 (−0.74 to 0.37)	7[e]	1.40 (0.50 to 2.30)	1.64 (0.86 to 2.43)	0.24 (−0.32 to 0.80)	3[f]
COR[r]=2.2								
Total	8.45 (3.43 to 13.5)	9.23 (3.96 to 14.5)	0.79 (−2.48 to 4.06)	7[e]	7.66 (2.55 to 12.8)	10.2 (4.72 to 15.7)	2.54 (0.48 to 4.59)	W=2[p]

[a]BFT: biofeedback training.

[b]BFT: biofeedback training.

[c]LogMAR: logarithm of the minimum angle of resolution.

[d]COR =0.24.

[e]n=5; *P*>.2.

[f]n=6; *P*>.2.

[g]COR =0.11.

[h]n=5.1<*P*<.2.

[i]BCEA: bivariate contour ellipse area.

[j]COR =0.61.

[k]COR =1.6.

[l]WPM: words per minute.

[m]COR =8.6.

[n]n=8; *P*>.2.

XSL•FO
RenderX

[o]COR =0.44.

[p]n=6.05<*P*<.1.

[q]n=6.1<*P*<.2.

[r]COR: coeficient of repeatability.

Neither baseline group difference (ITT: Welsch *t* test: *t*=1.32, df=19.1, *P*=.19; PP: Welsch *t* test: $t_{19.5}$=1.48, *P*=.17) nor postintervention changes (ITT: Welsch *t* test: $t_{18.5}$=0.87, *P*=.37; PP: Welsch *t* test: $t_{16.1}$=1.50, *P*=.16) reached significance in retinal sensitivity (Tables 4 and 5). Measures at 1- and 6-month post-intervention showed no changes between groups (Table S1 in Multimedia Appendix 1).

Contrast sensitivity was impaired in all participants, across all spatial frequencies at baseline, with no between-group differences (ITT: Welch *t* test, 2-tailed: $t_{12.0}$=0.73, *P*=.48; PP: Welch *t* test: t_9=1.52, *P*=.16). No differences were observed after intervention, nor at 1- and 6-month follow-up (Figure 1 in Multimedia Appendix 2).

Baseline reading speed heterogeneity was large: 2 participants were ≥200 wpm, three ranged 100 - 200 wpm, and the remainder <100 wpm. Group means did not differ initially (ITT: Welch *t* test: $t_{10.6}$=0.42, *P*=.68; PP: Welch *t* test: $t_{7.87}$=0.26, *P*=.79). Data analysis showed no significant difference between BFT and BFT-VR groups in reading speed after intervention (ITT: Welch *t* test: $t_{10.5}$=0.37, *P*=.48; PP: Welch *t* test: $t_{6.12}$=0.84, *P*=.40). Pre- and postintervention analysis showed no statistical difference in the BFT group (ITT and PP analyses; Tables 4 and 5). However, the BFT-VR group showed a clinically significant increase in reading speed after the intervention, above COR (ITT:+13.2, 95% CI −10.8 to 37.3; PP:+21.3, 95% CI −10.1 to 52.6; COR=8.6 wpm). Measures at 1- and 6-month postintervention showed no changes between groups (Table S1 in Multimedia Appendix 1); however, a clinically significant decrease in reading speed was observed between postintervention (117.4, 95% CI 31.7 to 203) and 1-month follow-up (99.2, 95% CI 13.8 to 185; Δ=−18.2, 95% CI -50.7 to 14.3; COR=8.6).

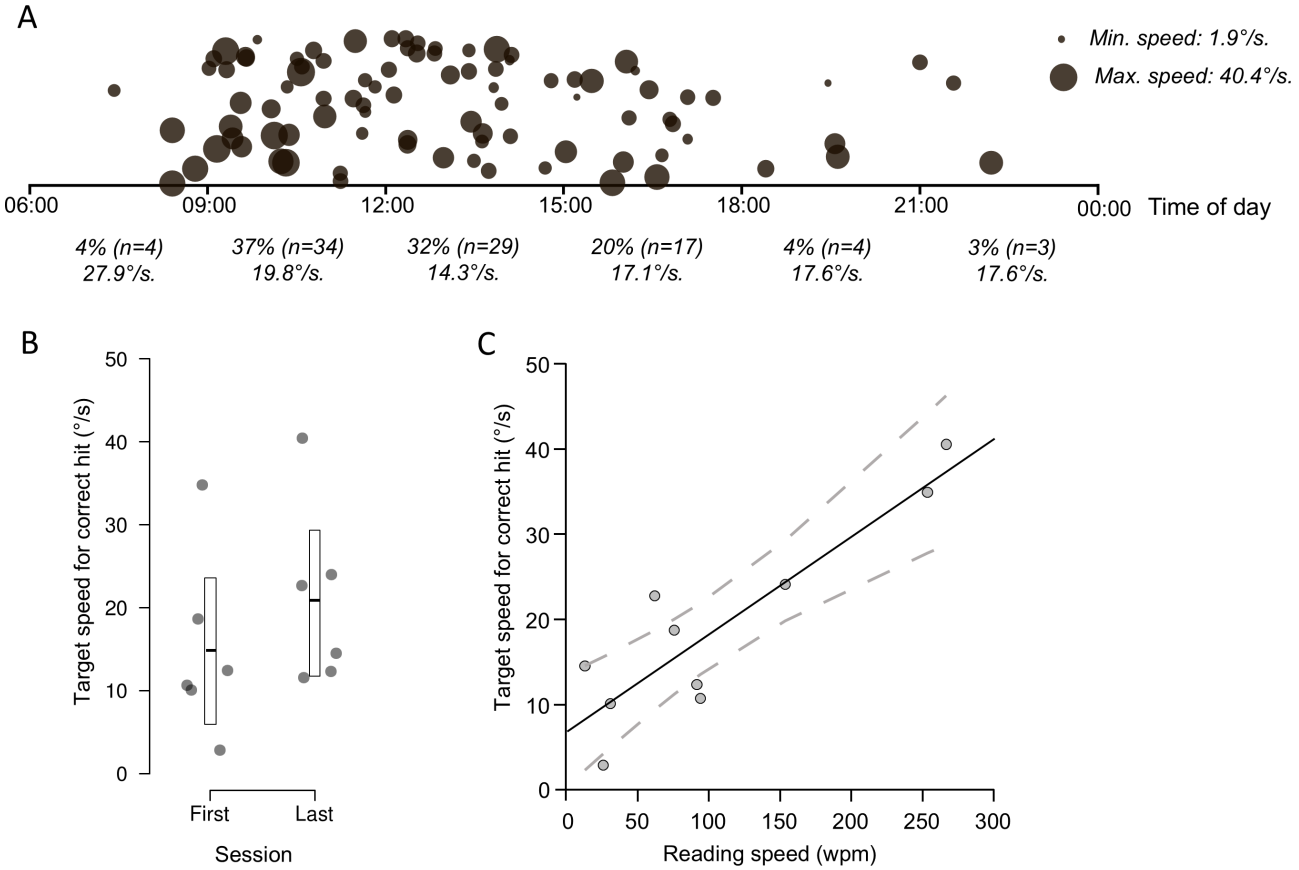
Five domain scores for daily living activities, visual ability, reading, mobility, visual information, and visual motor, and a composite were analyzed. At baseline, no domain differed between arms (Table S1 in Multimedia Appendix 2). After training, both groups exhibited clinically meaningful gains in the reading subscale (BFT:+1.17, 95% CI 0.06 to 2.28; COR>0.44; Table 4) statistically significant for the BFT-VR group (ITT:+0.92, 95% CI 0.17 to 1.66; Wilcoxon ranked test: W=2, n=8, .02<*P*<.05; PP:+0.93, 95% CI −0.20 to 2.07; COR>0.44; Tables 4 and 5). No changes between groups were observed at 1- and 6-month follow-up (Table S1 in Multimedia Appendix 1).

BFT-VR Group: Exploratory Outcomes

The potential use of VR as a home-based interventional tool is very innovative; therefore, data about its use in real-world conditions is lacking. We recorded the time of day when participants who completed the intervention in the BFT-VR group performed the 3D-SOT-VR sessions. Session timestamps revealed that 74% (63/91) were performed between 9 AM and 3 PM (Figure 3A). Average total session duration—pre-VRISE, 3 × 5-min blocks, post-VRISE—was 28 (SD 7) minutes. ANOVA showed no interaction between time of day and tracking performance ($F_{,2}$ =0.38; *P*=.83; Figure 3A).

XSL•FO
RenderX

Figure 3. Exploratory outcomes. (A) Distribution of the session during the day. The size of the circle is correlated to the average speed of the target for a correct answer "ts" during the corresponding session. Percentage and number of sessions during the 3-hour bin are represented with the average speed of the target for a correct answer (=performance). (B) Performance (average speed of the target for a correct answer) at the 3D-SOT-VR (3D single object tracking in immersive virtual reality) stimulation for participants who completed the intervention at the first and last session (n=6). (C) Graph of the correlation between performance (the speed of the target for a correct answer in °/s) as a function of the reading speed (wpm). Pearson R=0.89; P<.001. Max: maximum; Min: minimum; wpm: words per minute.



Participants who completed the intervention (6/8) followed the same 3D-SOT-VR stimulation program for 4 weeks. Aggregate tracking performance increased: mean target speed on correct trials rose from 14.9°/second (95% CI 6.1 to 23.7) at the first session to 20.9°/second (95% CI 12.1 to 29.6) at the final session (Wilcoxon W=1, n=6, .05<P<.10; Figure 3B).

Reading speed was the only functional vision outcome that showed a clinically significant increase in the BFT-VR group, and participants showed variable tracking performance at the 3D-SOT-VR program. We therefore explored whether there is a correlation between the overall tracking performance at the 3D-SOT-VR stimulation and the reading speed. Reading speed correlated strongly with mean 3D-SOT performance (Pearson r=0.89; P<.001) across five participants with complete data (Figure 3C).

Missing Data

Baseline data were complete for the five core clinical outcomes (far VA, near VA, retinal sensitivity, fixation stability, and reading speed). Missingness increased over follow-up, with 5/70 missing data at postintervention, 14/70 at 1-month follow-up, and 25/70 at 6-month follow-up. Little MCAR test on the baseline or postintervention core outcome matrix did not provide statistical evidence against the MCAR assumption (χ^2_5=5.8; P=.33). However, because of the very small sample size and documented reasons for missingness, including missed visits, withdrawal, and follow-up loss, this result was interpreted cautiously and not used as justification for multiple imputation.

Discussion

Principal Findings

This pilot feasibility study is the first to examine the combined use of clinical-based BFT and a tailored, home-delivered 3D-SOT-VR program for older adults with dry AMD. The main finding is one of partial feasibility rather than global feasibility. Adoption, safety monitoring, remote VR data transfer, and adherence among participants who continued the VR program were encouraging. In contrast, recruitment feasibility was not achieved, intervention completion in the BFT-VR arm did not reach the prespecified threshold, and follow-up attrition limited interpretation of longer-term outcomes. Recurrent suspensions under Ontario's Emergency Management and Civil Protection Act between February 2021 and September 2022 hampered the recruitment rate, limited onsite testing capacity, and increased participants' reluctance to travel and come to the hospital, fearing an increased risk of SARS-CoV-2 contamination. Altogether, this resulted in 10-fold slower enrollment than forecast. Similar setbacks have been reported in clinical trials embedded in health care systems during COVID-19 surges [46,47]. Post restriction, participants prioritized family gatherings and reconnecting with close relatives over the study's

XSL•FO
RenderX

requirements, leading to a loss of follow-up and increased no-shows. Nonetheless, the pandemic catalyzed broader public receptivity to digital and remote care solutions, even among older adults traditionally considered technophobic; our 75% completion rate underscores this shift. Previous studies, involving BFT and VR separately in older adults, have shown acceptable recruitment rates [7,10,48], although we cannot fully exclude that the combination of BFT and the audiovisual stimulation in VR might have increased the burden of the intervention.

Importantly, for participants enrolled during reopening periods and randomized in the BFT-VR group, a majority were able to complete the home-based 3D-SOT-VR intervention, in line with other studies indicating adoption of such technology in the older adult population [48,49]. A key benefit of home-based 3D-SOT-VR stimulation is its seamless integration into daily routines, as shown in other groups of participants [19], offering flexible and user-friendly use with reliable data transfer. Such benefits are even more important in the recent context of the pandemic, with the difficulty of travel, staff shortages, and related restrictions, as patients were unable to receive their treatment without delays. If VR potentially represents an adequate platform to deliver a visual rehabilitation program at home [18,19], it requires some adjustments tailored to the target population, especially older adults. Headset adjustments using straps and remote controllers seem to be troublesome for this population, hampering a full adoption or completion of the VR interventions.

Group-level VA, fixation stability, retinal sensitivity, contrast sensitivity, and quality-of-life estimates were imprecise and showed no consistent between-group benefit. Reading speed showed a clinically and statistically significant increase in the BFT-VR group, whereas no improvement was observed in the BFT, in contrast to previous studies [7,9,50,51]. Correspondingly, meaningful improvements were observed in patient-reported quality of vision questionnaires related to reading abilities in both groups. Individual reading-speed improvement in the BFT-VR group and its association with VR-task performance are potentially informative but should be interpreted as hypothesis-generating. A plausible mechanism is that dynamic target tracking may train visual attention, gaze allocation, head-eye coordination, and use of residual peripheral vision during eccentric viewing. However, the study did not directly measure PRL relocation stability, reading eye movements, cortical plasticity, or neural adaptation; therefore, the mechanism remains speculative. An important variable is the number of sessions of BFT delivered and total time of BFT. There is still no consensus on ideal exposure time for an optimal result [4], and previously we obtained significant improvement in visual functions and quality-of-life questionnaire using this scheme of four weekly sessions of 20 minutes (80 min in total) [7]. However, in a recent study that used two different intervals, 150 minutes in total delivered daily or on alternate days over 15 sessions, a statistically significant improvement was observed in VA, reading speed, fixation stability, and self-referred quality-of-life questionnaire in both groups [13]. The decrease in reading speed between postintervention and 1-month follow-up suggests that the short 4-week VR program may not be sufficient to induce sustained functional gains. This finding is compatible with clinical observations that patients may revert to less efficient viewing strategies after supervised rehabilitation if the trained strategy is not reinforced. It does not prove that indefinite treatment is required. Instead, future studies should test longer initial training, lower-burden maintenance sessions, or scheduled booster periods, while measuring whether PRL use and oculomotor behavior are stabilized over time.

The every-other-day schedule was chosen as a pragmatic compromise between repetition and tolerability. Weekly VR exposure was considered unlikely to add sufficient training beyond conventional weekly BFT, whereas daily exposure could increase fatigue, cybersickness risk, and nonadherence in older participants [48,49]. Future trials should include dose-finding components to compare training frequency, total exposure, and maintenance schedules. Such trials should also predefine whether the intended effect is short-term compensatory improvement, durable functional adaptation, or ongoing supportive therapy.

The ergonomics findings are clinically important. One participant discontinued the intervention because of headset weight and neck fatigue, and usability items related to headset adjustment and pointer selection were below threshold. This indicates that future implementation should prioritize lighter devices, improved weight distribution, simplified straps, larger and more accessible interfaces, gaze-based or controller-free selection, caregiver-assisted onboarding when needed, and remote troubleshooting. These design changes are particularly relevant for older adults with low vision, who may also have reduced manual dexterity, neck discomfort, or lower familiarity with immersive technologies [48,49].

Limitations

The major limitation is the small sample size. Recruitment and study conduct were heavily disrupted by the COVID-19 pandemic and associated restrictions in Ontario, producing a marked reduction in recruitment compared with the prespecified criterion. As a consequence, the planned target of 32 participants was not reached, the precision of all secondary estimates was low, and between-group comparisons should not be interpreted as evidence for or against clinical effectiveness.

Additional limitations include open-label design, possible selection bias toward motivated participants who were willing to use VR at home, incomplete follow-up, and missing data that were not demonstrably MCAR. Because missingness was partly attributable to withdrawal, missed assessments, device burden, and follow-up loss, available-case analyses may be biased. Multiple imputation was not used because the sample size and sparse missing data structure would make imputation unstable and potentially misleading. The trial also did not include mechanistic measures of PRL use during reading, fixation strategy outside microperimetry, or neural adaptation, limiting interpretation of the reading-speed signal.

Taken together, these limitations mean that this study should be viewed as a formative feasibility trial. It identifies safety, remote monitoring, adherence among users, and preliminary functional signals that justify further study, but it also identifies

XSL•FO
RenderX

recruitment, retention, missing data, long-term durability, and device ergonomics as barriers that must be solved before a definitive randomized trial.

Conclusions

In conclusion, home-based VR visual stimulation combined with clinic-based BFT can be deployed safely in selected older adults with dry AMD and can generate objective remote adherence, safety, and performance data. The approach differs from conventional low-vision rehabilitation by moving part of the intervention into the home while retaining clinical supervision and remote monitoring. These results should not be interpreted as evidence of effectiveness; rather, they define the operational requirements for the next phase of research. Future trials should use optimized lightweight hardware, simplified interaction, larger multicenter recruitment, predefined missing-data handling, dose-finding of VR exposure, and mechanistic endpoints such as PRL stability, eye and head tracking, and reading-related oculomotor behavior.

Acknowledgments

We thank the participants for their motivation and commitment. We thank NeuropowerTrain for their help with the initiation of this study. During revision, the authors used a generative AI tool to assist with language editing and clarity of phrasing. All scientific content, interpretation, statistical decisions, and final text were reviewed and approved by the authors. No generative AI tool was used to generate data, perform statistical analyses, or replace scientific judgment.

Funding

This work was supported by the Krembil Collaborative Seed grant at University Health Network (MR and SNM) and by the University Health Network Foundation (MR). The funders had no role in study design, data collection, analysis, interpretation, manuscript preparation, or the decision to submit the manuscript for publication.

Data Availability

Deidentified data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

YP, BZ, MM, MDN, AS, DT, and KC collected the clinical data. BZ, KC, and MR collected and analyzed the exploratory data. MDN, SNM, and MR analyzed the clinical data. MDN, LA, SNM, and MR designed the study. EGG and PN developed remote data collection in VR. MDN, SNM, LA, and MR wrote and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
1- and 6-month follow-up secondary outcome measures analyses (A. Intention-to-treat, B. Per Protocol).
[PDF File, 141 KB - rehab_v13i1e87596_app1.pdf]

Multimedia Appendix 2
Contrast sensitivity measures at 1- and 6-month follow-up in BFT and BFT-IVR groups.
[PNG File, 120 KB - rehab_v13i1e87596_app2.png]

Checklist 1
CONSORT e-HEALTH checklist.
[PDF File, 60407 KB - rehab_v13i1e87596_app3.pdf]

References

1. Taylor DJ, Hobby AE, Binns AM, Crabb DP. How does age-related macular degeneration affect real-world visual ability and quality of life? a systematic review. BMJ Open 2016 Dec 2;6(12):e011504. [doi: 10.1136/bmjopen-2016-011504] [Medline: 27913556]
2. Lane J, Rohan EMF, Sabeti F, et al. Impacts of impaired face perception on social interactions and quality of life in age-related macular degeneration: a qualitative study and new community resources. PLoS One 2018;13(12):e0209218. [doi: 10.1371/journal.pone.0209218] [Medline: 30596660]
3. World report on vision. World Health Organization. URL: https://www.who.int/publications-detail-redirect/world-report-on-vision [accessed 2020-09-20]

XSL•FO
RenderX

4. Deng Y, Jie CH, Wang JW, Li YY, Liu ZQ. Visual function and biofeedback training of patients with central vision loss: a review. Int J Ophthalmol 2023;16(5):824-831. [doi: 10.18240/ijo.2023.05.21] [Medline: 37206180]
5. Markowitz SN. State-of-the-art: low vision rehabilitation. Can J Ophthalmol 2016 Apr;51(2):59-66. [doi: 10.1016/j.jcjo.2015.11.002] [Medline: 27085259]
6. Tarita-Nistor L, González EG, Markowitz SN, Steinbach MJ. Plasticity of fixation in patients with central vision loss. Vis Neurosci 2009 Nov;26(5-6):487-494. [doi: 10.1017/S0952523809990265] [Medline: 20003597]
7. Daibert-Nido M, Patino B, Markowitz M, Markowitz SN. Rehabilitation with biofeedback training in age-related macular degeneration for improving distance vision. Can J Ophthalmol 2019 Jun;54(3):328-334. [doi: 10.1016/j.jcjo.2018.10.016] [Medline: 31109472]
8. Qian T, Xu X, Liu X, et al. Efficacy of MP-3 microperimeter biofeedback fixation training for low vision rehabilitation in patients with maculopathy. BMC Ophthalmol 2022 Apr 28;22(1):197. [doi: 10.1186/s12886-022-02419-6] [Medline: 35484529]
9. Widihastha SH, Iskandar E, Satari K, Irfani I, Virgana R, Amiruddin PO. Vision rehabilitation using microperimetric biofeedback training in age-related macular degeneration. Int J Ophthalmol 2023;16(6):933-938. [doi: 10.18240/ijo.2023.06.16] [Medline: 37332539]
10. Nido MD, Markowitz SN. Vision rehabilitation with biofeedback training. Can J Ophthalmol 2018 Jun;53(3):e83-e84. [doi: 10.1016/j.jcjo.2017.11.003] [Medline: 29784180]
11. Amore FM, Paliotta S, Silvestri V, Piscopo P, Turco S, Reibaldi A. Biofeedback stimulation in patients with age-related macular degeneration: comparison between 2 different methods. Can J Ophthalmol 2013 Oct;48(5):431-437. [doi: 10.1016/j.jcjo.2013.07.013] [Medline: 24093192]
12. Morales MU, Saker S, Wilde C, Rubinstein M, Limoli P, Amoaku WM. Biofeedback fixation training method for improving eccentric vision in patients with loss of foveal function secondary to different maculopathies. Int Ophthalmol 2020 Feb;40(2):305-312. [doi: 10.1007/s10792-019-01180-y] [Medline: 31583549]
13. Zhou J, Hou J, Li S, Zhang J. The effect of duration between sessions on microperimetric biofeedback training in patients with maculopathies. Sci Rep 2024 May 31;14(1):12524. [doi: 10.1038/s41598-024-63327-x]
14. Cheung SH, Legge GE. Functional and cortical adaptations to central vision loss. Vis Neurosci 2005;22(2):187-201. [doi: 10.1017/S0952523805222071] [Medline: 15935111]
15. Mandelcorn MS, Podbielski DW, Mandelcorn ED. Fixation stability as a goal in the treatment of macular disease. Can J Ophthalmol 2013 Oct;48(5):364-367. [doi: 10.1016/j.jcjo.2013.05.006] [Medline: 24093181]
16. Thibaut M, Tran THC, Szaffarczyk S, Boucart M. The contribution of central and peripheral vision in scene categorization: a study on people with central vision loss. Vision Res 2014 May;98:46-53. [doi: 10.1016/j.visres.2014.03.004] [Medline: 24657253]
17. Daibert-Nido M, Pyatova Y, Cheung KG, et al. An audiovisual 3D-immersive stimulation program in hemianopia using a connected device. Am J Case Rep 2021 Jun 9;22:e931079. [doi: 10.12659/AJCR.931079] [Medline: 34106907]
18. Daibert-Nido M, Pyatova Y, Cheung K, et al. Case report: visual rehabilitation in hemianopia patients. Home-based visual rehabilitation in patients with hemianopia consecutive to brain tumor treatment: feasibility and potential effectiveness. Front Neurol 2021;12(680211):680211. [doi: 10.3389/fneur.2021.680211] [Medline: 34354660]
19. Misawa M, Bajin IY, Zhang B, et al. A telerehabilitation program to improve visual perception in children and adolescents with hemianopia consecutive to a brain tumor: a single-arm feasibility and proof-of-concept trial. EClinicalMedicine 2024 Dec;78:102955. [doi: 10.1016/j.eclinm.2024.102955] [Medline: 39687429]
20. Pylyshyn ZW, Annan V. Dynamics of target selection in multiple object tracking (MOT). Spat Vis 2006;19(6):485-504. [doi: 10.1163/156856806779194017] [Medline: 17278524]
21. Meyerhoff HS, Papenmeier F, Huff M. Studying visual attention using the multiple object tracking paradigm: a tutorial review. Atten Percept Psychophys 2017 Jul;79(5):1255-1274. [doi: 10.3758/s13414-017-1338-1] [Medline: 28584953]
22. Romeas T, Guldner A, Faubert J. 3D-Multiple object tracking training task improves passing decision-making accuracy in soccer players. Psychol Sport Exerc 2016 Jan;22:1-9. [doi: 10.1016/j.psychsport.2015.06.002]
23. Legault I, Faubert J. Perceptual-cognitive training improves biological motion perception: evidence for transferability of training in healthy aging. Neuroreport 2012 May 30;23(8):469-473. [doi: 10.1097/WNR.0b013e328353e48a] [Medline: 22495038]
24. Corbin-Berrigan LA, Kowalski K, Faubert J, Christie B, Gagnon I. Three-dimensional multiple object tracking in the pediatric population: the NeuroTracker and its promising role in the management of mild traumatic brain injury. Neuroreport 2018 May 2;29(7):559-563. [doi: 10.1097/WNR.0000000000000988] [Medline: 29481522]
25. Misawa M, Yaman Bajin I, Zhang B, et al. A telerehabilitation program to improve visual perception in children and adolescents with hemianopia consecutive to a brain tumour: a single-arm feasibility and proof-of-concept trial. medRxiv. Preprint posted online on 2024. [doi: 10.2139/ssrn.4838991]
26. Report on amendments, extensions, and revocations of orders under the reopening ontario (a flexible response to COVID-19) act, 2020 from december 2, 2021, to march 28, 2022. : Government of Ontario; 2022 URL: http://www.ontario.ca/document/report-amendments-extensions-and-revocations-orders-under-reopening-ontario-flexible-3

XSL•FO
RenderX

27. Kourtesis P, Collina S, Doumas LAA, MacPherson SE. Validation of the Virtual Reality Neuroscience Questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology. Front Hum Neurosci 2019;13:417. [doi: 10.3389/fnhum.2019.00417] [Medline: 31849627]
28. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016;2:64. [doi: 10.1186/s40814-016-0105-8] [Medline: 27965879]
29. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073. [doi: 10.1177/0962280215588241] [Medline: 26092476]
30. Pylyshyn ZW, Storm RW. Tracking multiple independent targets: evidence for a parallel tracking mechanism. Spat Vis 1988;3(3):179-197. [doi: 10.1163/156856888x00122] [Medline: 3153671]
31. Meredith MA, Nemitz JW, Stein BE. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J Neurosci 1987 Oct;7(10):3215-3229. [doi: 10.1523/JNEUROSCI.07-10-03215.1987] [Medline: 3668625]
32. Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971 Feb;49(2):MID. [Medline: 5541744]
33. Leek MR. Adaptive procedures in psychophysical research. Percept Psychophys 2001 Nov;63(8):1279-1292. [doi: 10.3758/bf03194543] [Medline: 11800457]
34. Leat S. 2020 CAO Clinical practice guideline: optometric low vision rehabilitation FULL GUIDELINES. Can J Optom 2020;82(1):19-62. [doi: 10.15353/cjo.v82i1.1636] [Medline: 31078445]
35. Legge GE, Ross JA, Luebker A, Lamay JM. Psychophysics of reading. VIII. The Minnesota Low - Vision Reading Test. Optom Vis Sci 1989 Dec;66(12):843-853 [FREE Full text] [doi: 10.1002/j.1538-9235.1989.tb02415.x] [Medline: 2626251]
36. Stelmack JA, Szlyk JP, Stelmack TR, et al. Measuring outcomes of vision rehabilitation with the Veterans Affairs Low Vision Visual Functioning Questionnaire. Invest Ophthalmol Vis Sci 2006 Aug;47(8):3253-3261. [doi: 10.1167/iovs.05-1319] [Medline: 16877389]
37. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test-retest reliability. PLoS ONE 2013;8(9):e73990. [doi: 10.1371/journal.pone.0073990] [Medline: 24040139]
38. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol 2003 Jul;22(1):85-93. [doi: 10.1002/uog.122] [Medline: 12858311]
39. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek ALM. Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res 2001;10(7):571-578. [doi: 10.1023/a:1013138911638] [Medline: 11822790]
40. Little RJA. A test of missing completely at random for multivariate data with missing values. J Am Stat Assoc 1988 Dec;83(404):1198-1202. [doi: 10.1080/01621459.1988.10478722]
41. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009 Jun 29;338(jun29 1):b2393. [doi: 10.1136/bmj.b2393] [Medline: 19564179]
42. Huygelier H, Schraepen B, Lafosse C, et al. An immersive virtual reality game to train spatial attention orientation after stroke: a feasibility study. Appl Neuropsychol Adult 2022;29(5):915-935. [doi: 10.1080/23279095.2020.1821030] [Medline: 32945702]
43. Shim L, Lee J, Han JH, Jeon H, Hong SK, Lee HJ. Feasibility of virtual reality-based auditory localization training with binaurally recorded auditory stimuli for patients with single-sided deafness. Clin Exp Otorhinolaryngol 2023 Aug;16(3):217-224. [doi: 10.21053/ceo.2023.00206] [Medline: 37080730]
44. Saredakis D, Szpak A, Birckhead B, Keage HAD, Rizzo A, Loetscher T. Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Front Hum Neurosci 2020;14(96):96. [doi: 10.3389/fnhum.2020.00096] [Medline: 32300295]
45. ICH E9 statistical principles for clinical trials - scientific guideline. European Medicines Agency. URL: https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials-scientific-guideline [accessed 2025-05-07]
46. O'Brien EC, Sugarman J, Weinfurt KP, et al. The impact of COVID-19 on pragmatic clinical trials: lessons learned from the NIH health care systems research collaboratory. Trials 2022 May 21;23(1):424. [doi: 10.1186/s13063-022-06385-8] [Medline: 35597988]
47. Margas W, Wojciechowski P, Toumi M. Impact of the COVID-19 pandemic on the conduct of clinical trials: a quantitative analysis. J Mark Access Health Policy 2022;10(1):2106627. [doi: 10.1080/20016689.2022.2106627] [Medline: 35968522]
48. Huygelier H, Schraepen B, van Ee R, Vanden Abeele V, Gillebert CR. Acceptance of immersive head-mounted virtual reality in older adults. Sci Rep 2019 Mar 14;9(1):4519. [doi: 10.1038/s41598-019-41200-6] [Medline: 30872760]
49. Abeele VV, Schraepen B, Huygelier H, Gillebert CR, Gerling K. Immersive virtual reality for older adults: empirically grounded design guidelines. ACM Trans Access Comput 2021. [doi: 10.1145/3470743]
50. Raman R, Damkondwar D, Neriyanuri S, Sharma T. Microperimetry biofeedback training in a patient with bilateral myopic macular degeneration with central scotoma. Indian J Ophthalmol 2015 Jun;63(6):534-536. [doi: 10.4103/0301-4738.162609] [Medline: 26265646]
51. Vingolo EM, Cavarretta S, Domanico D, Parisi F, Malagola R. Microperimetric biofeedback in AMD patients. Appl Psychophysiol Biofeedback 2007 Dec;32(3-4):185-189. [doi: 10.1007/s10484-007-9038-6] [Medline: 17574525]

XSL•FO
RenderX

Abbreviations

3D-SOT-VR: 3D single object tracking in immersive virtual reality
AMD: age-related macular degeneration
BFT: biofeedback training
BFT-VR: biofeedback training in immersive virtual reality
CONSORT: Consolidated Standards of Reporting Trials
COR: coefficient of repeatability
dBHL: decibels hearing level
ICH: International Conference on Harmonisation
ITT: intention-to-treat
LogMAR: logarithm of the minimum angle of resolution
MCAR: Missing Completely At Random
PP: per protocol
PRL: preferred retinal locus
VA: visual acuity
VR: virtual reality
VRISE: Virtual Reality-Induced Symptoms and Effects

Edited by S Munce; submitted 11.Nov.2025; peer-reviewed by PDW Yoshinaga, R Lin; revised version received 02.Jun.2026; accepted 19.Jun.2026; published 03.Aug.2026.

Please cite as:
Pyatova Y, Zhang B, Misawa M, Markowitz SN, Sen A, Appel L, Cheung K, Nasir P, Tchao D, Garcia-Giler E, Daibert-Nido M, Reber M
A Visual Telerehabilitation Program in Virtual Reality for Age-Related Macular Degeneration: Randomized Feasibility and Proof-of-Concept Trial
JMIR Rehabil Assist Technol 2026;13:e87596
URL: https://rehab.jmir.org/2026/1/e87596
doi:10.2196/87596

© Yulia Pyatova, Bill Zhang, Mariana Misawa, Samuel N Markowitz, Atri Sen, Lora Appel, Kyle Cheung, Pi Nasir, Danielle Tchao, Eduardo Garcia-Giler, Monica Daibert-Nido, Michael Reber. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 3.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telerehabilitation Trends in Australian Physiotherapy and an Exploration of Factors That Influence Use After COVID-19 Restrictions: Qualitative Content Analysis

Megan H Ross[1], PhD; Joshua Simmich[1], PhD; Belinda J Lawford[2], PhD; Kim L Bennell[2], PhD; Rana S Hinman[2], PhD; Trevor Russell[1], PhD

[1]RECOVER Injury Research Centre, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Surgical, Treatment and Rehabilitation Service (STARS), 296 Herston Rd, Herston, Queensland, Australia

[2]Centre for Health, Exercise and Sports Medicine, Department of Physiotherapy, School of Health Sciences, Faculty of Medicine, Dentistry & Health Sciences, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Megan H Ross, PhD
RECOVER Injury Research Centre, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Surgical, Treatment and Rehabilitation Service (STARS), 296 Herston Rd, Herston, Queensland, Australia

Abstract

Background: Telerehabilitation is a safe and effective means of delivering physiotherapy services, but implementation in clinical practice has not been widespread.

Objective: This study aimed to explore the shifts in telerehabilitation use throughout the COVID-19 pandemic and the key factors that influenced telerehabilitation caseload after restrictions were eased.

Methods: Between September and November 2023, physiotherapists practicing in Australian private practice, hospital outpatient, or community settings completed an online survey. Data were collected regarding participants' use of telerehabilitation before, during, and after the COVID-19 pandemic restrictions to in-person physiotherapy. Qualitative content analysis of open-text questions was performed to garner more nuanced information about the use of telerehabilitation in clinical practice, and quantitative data were analyzed descriptively.

Results: The proportion of participants using telerehabilitation rose from 30% (44/148) before the pandemic to 94% (138/147) when restrictions to in-person physiotherapy were in place. Although 82% (118/144) of the sample continued to deliver telerehabilitation after COVID-19 restrictions were eased, telerehabilitation accounted for only 14% of the total caseload. Exploratory analyses suggest that despite increased confidence, satisfaction, and perceptions about the effectiveness of telerehabilitation, reduced patient demand, physiotherapists' perceptions about patient preference for in-person consultations, and the perception that in-person physiotherapy is easier continue to influence the use of telerehabilitation in the post-COVID era.

Conclusions: Despite increased uptake during the pandemic, telerehabilitation caseload after restrictions were eased was low. Physiotherapists' perceptions about telerehabilitation in clinical practice remain a substantial barrier to sustained adoption.

(JMIR Rehabil Assist Technol 2026;13:e81008) doi:10.2196/81008

KEYWORDS

telehealth; COVID-19; videoconferencing; survey; physiotherapy; qualitative

Introduction

Background

Evidence suggests that telerehabilitation is a safe, feasible, and effective means of delivering physiotherapy care that is at least as good as in-person physiotherapy in terms of patient outcomes [1,2]. Despite two decades of evidence supporting the effectiveness of telerehabilitation in the management of musculoskeletal [3], neurological [4], cardiorespiratory [5], and postsurgical rehabilitation [6], clinician acceptance and adoption have been low [7]. While telerehabilitation is a viable alternative to traditional in-person physiotherapy with the potential to overcome geographical barriers, improve access, and facilitate continuity of treatment, integration into routine physiotherapy practice before the COVID-19 pandemic remained limited [8].

Before the COVID-19 pandemic, physiotherapists were slow and reluctant to adopt telerehabilitation as a standard model of care [9,10]. Several barriers contributed to this limited uptake of telerehabilitation in physiotherapy practice, including limited acceptance of and low confidence in using telehealth technology,

XSL•FO
RenderX

perceived limitations in conducting physical assessments remotely, and reduced capacity to deliver hands-on interventions that are central to traditional physiotherapy practice [8,11,12]. Additional challenges included perceptions that telerehabilitation was less effective for certain clinical presentations and concerns about developing rapport and patient engagement [11,12].

The COVID-19 pandemic brought unprecedented challenges to health care systems worldwide, including the delivery of physiotherapy services. In response to the pandemic-related restrictions on in-person consultations, many physiotherapy practices turned to telerehabilitation [13] as an alternative method to continue providing necessary care [14]. Although physiotherapy was recognized as an essential health care service in Australia during the COVID-19 pandemic, practice was still notably restricted. Stringent infection control measures, such as mandatory use of personal protective equipment, rigorous patient screening, physical distancing requirements, density limits within clinical spaces, and group size limitations on group therapy sessions [15], impacted the delivery of care across multiple clinical settings. During initial lockdowns, some states, such as Victoria, further restricted in-person physiotherapy services, permitting face-to-face consultations only for urgent cases [16]. Community and aged care physiotherapy faced further barriers, including restrictions on therapists attending multiple sites and outright bans on external providers entering residential facilities [17]. Consequently, physiotherapy practice during the pandemic was markedly disrupted, forcing providers to rapidly transition to providing telerehabilitation services to adhere to public health guidelines and ensure continuity of care.

With the rapid transition to telerehabilitation in response to the pandemic came changes in regulatory frameworks to fund telerehabilitation [18], position statements advocating for the use of telerehabilitation [19], and increased infrastructure and clinical training to support the integration of telerehabilitation into clinical care [20]. During this period, uptake of telerehabilitation increased substantially, reflecting the necessity to maintain continuity of care. Research conducted at the time suggested that physiotherapists intended to continue offering services via telerehabilitation after the easing of restrictions to in-person physiotherapy [20,21]. However, international evidence suggests that uptake and usage have generally decreased from the pandemic peak [22].

Objectives

With the rapid transition to telerehabilitation in response to the pandemic came changes in regulatory frameworks to fund telerehabilitation [18], position statements advocating for the use of telerehabilitation [19], and increased infrastructure and clinical training to support the integration of telerehabilitation into clinical care [20]. During this period, uptake of telerehabilitation increased substantially, reflecting the necessity to maintain continuity of care. Research conducted at the time suggested that physiotherapists intended to continue offering services via telerehabilitation after the easing of restrictions to in-person physiotherapy [20,21]. However, international evidence suggests that uptake and usage have generally decreased from the pandemic peak [22].

The aim of this study was to investigate the use of telerehabilitation in Australian physiotherapy clinical practice throughout the COVID-19 pandemic, with a focus on telerehabilitation use after restrictions were eased.

The specific research questions for this study were as follows: (1) How did the use of telerehabilitation vary in physiotherapy clinical practice in Australia before, during, and after COVID-19 restrictions to in-person consultations? and (2) What are the key factors that influence physiotherapists' telerehabilitation caseload in the postrestrictions period?

Methods

Design

A descriptive, cross-sectional survey was conducted online with physiotherapists currently practicing in Australia. The study was primarily quantitative, with a small qualitative component to supplement descriptive analyses.

Ethical Considerations

The study was approved by The University of Queensland Human Research Ethics Committee (approval number: 2023/HE001802) and reported following the consensus-based CROSS (Checklist for Reporting of Survey Studies) [23]. Participants provided electronic informed consent after reviewing an information sheet and before completing the survey. Participants were entered into a draw for a AUD $1000 (US $667) gift voucher upon completion of the survey. Participants' privacy and confidentiality were maintained by storing nonidentifiable survey data separately from contact details on the University of Queensland Research Data Management System.

Participants

Participants were physiotherapists recruited from the community via online advertisements on social media (eg, Facebook, X, and LinkedIn), via targeted emails, and through Australian Physiotherapy Association member communications (eg, eComms). Physiotherapists were eligible to participate if they were registered with the Australian Health Practitioner Regulation Agency and currently practicing in an Australian private practice, hospital, or community setting. Participants who had not delivered telerehabilitation services were eligible to complete a short version of the questionnaire to explore reasons for not engaging with telerehabilitation and the circumstances that might influence uptake.

Procedure

An online survey was designed to capture information that was relevant to stakeholders and ensure readability and credibility (Multimedia Appendix 1). The survey was developed by the authors using Bennell [20] as a guide and adapted to capture information relevant to the different phases of the COVID-19 pandemic restrictions. The 3 phases were "Prior to the pandemic restrictions," "During the period of restrictions to in-person physiotherapy" (from the introduction of restrictions in 2020 to 2022), and "After restrictions were eased" (2022 onward). Questions were primarily multiple choice (checkbox questions), numerical rating scales (0 - 10), and 5-point Likert scales.

XSL•FO
RenderX

Respondents were asked to estimate their telerehabilitation caseload (individual video, group video, and telephone) for each phase using a sliding scale (0% - 100%). Free-text responses were sought for some questions to ascertain more nuanced and in-depth information about physiotherapists' perceptions of using telerehabilitation in clinical practice.

The survey was administered via an online secure platform (Qualtrics, LLC) and hosted by The University of Queensland. Participants were first invited to complete the online consent form and screening and, if eligible, proceeded to the survey. Participants were asked to provide demographic information, details of clinical practice, and experience with telerehabilitation. The second primary section of the survey comprised questions pertaining to the use of telerehabilitation during each phase of the COVID-19 pandemic restrictions (before, during, and after). All data were collected between September 15 and November 8, 2023.

Data Analysis

Data were exported from the online platform for analysis in R (version 4.3.3; R Core Team). Descriptive statistics, including frequencies (percentages) and means and standard deviations, were used to summarize the data. All responses (including partial responses) meeting eligibility criteria were included in analyses. When an "other" field was provided for additional response options, 2 researchers reviewed free-text responses and either aligned them with existing response options or designated them as unique responses that were added to the final list of response options. Any discrepancies in coding were resolved via discussion.

Responses to free-text questions were analyzed qualitatively using inductive content analysis in Microsoft Excel [24]. First, 2 researchers (MHR and JS) independently read the entire dataset, conducted open coding, and identified topics and initial patterns. The unit of analysis was meaning units, identified within individual responses. Codes were subsequently categorized and combined to form main categories or themes (abstraction), with both authors returning to the dataset to check that codes made sense in relation to the raw data. The 2 authors then met to compare and discuss their coding frameworks, and discrepancies were resolved through discussion. An audit trail was maintained to document coding decisions and category development. Themes with the highest number of individual data points were identified, reported, and described. To enhance trustworthiness, reflexivity was considered throughout the process, and attention was paid to credibility and transparency in coding and interpretation.

To explore which factors influenced physiotherapists' use of telerehabilitation in the postpandemic restrictions period, the total proportion of videoconferencing telerehabilitation caseload (individual and group consultations) was examined. Specifically, this proportion was plotted against the following five key postpandemic variables: confidence, satisfaction, and perceived effectiveness of telerehabilitation; physiotherapists' perception about how much patients like telerehabilitation; and how often patients are requesting it. Locally estimated scatterplot smoothing curves were fitted using the full span of the data (span=1). These smoothed trends, along with their corresponding 95% CIs, were used to visually explore apparent associations. No statistical correlation or regression analyses were performed on these trends.

Results

Sample Characteristics

A total of 222 physiotherapists responded to the survey, with 152 (68%) meeting eligibility criteria and providing sufficient data to be included in analyses (58/222, 26%, excluded for not being an Australian Health Practitioner Regulation Agency–registered physiotherapist currently practicing in an eligible setting [eg, private practice, hospital outpatient, or community] and 12/222, 5% not providing sufficient data to determine eligibility). Most participants (107/152, 70%) completed the survey in less than 20 minutes.

Respondents were primarily women (87/152, 57%); working in musculoskeletal (105/152, 69%) private practice (84/152, 55%) in Queensland (42/152, 28%), Victoria (42/152, 28%), or New South Wales (38/152, 25%); and held either a Bachelor's (70/152, 46%) or Master's (58/152, 38%) degree in physiotherapy. Physiotherapists primarily used Zoom (66/142, 47%), a telephone (59/142, 42%) or Microsoft Teams (47/142, 33%) to conduct telerehabilitation consultations. Only 40% (n=56) of respondents indicated that they had participated in telerehabilitation training. Additional participant characteristics are provided in Table 1.

XSL•FO
RenderX

Table . Participant characteristics (total N=152 unless otherwise specified).

Characteristic	Values, n (%)[a]
Gender	
Woman	87 (57)
Man	63 (41)
Prefer not to say	2 (1)
State or territory	
Queensland	42 (28)
Victoria	42 (28)
New South Wales	38 (25)
Western Australia	13 (9)
South Australia	10 (7)
Australian Capital Territory	6 (4)
Tasmania	1 (1)
Area of practice	
Private practice (primary care)	83 (55)
Public health outpatient center	45 (30)
Community health center	29 (19)
Private hospital	10 (7)
Other	20 (13)
Clinical focuses	
Musculoskeletal or orthopedic	105 (69)
Sports and exercise	43 (28)
Neurology	30 (20)
Gerontology	24 (16)
Pediatric	12 (8)
Other	35 (23)
Highest education	
Bachelor’s degree	70 (46)
Master’s by coursework	58 (38)
Masters by research	3 (2)
Postgraduate diploma	11 (7)
PhD	6 (4)
Other	4 (3)
Prior training in telehealth	
No	96 (63)
Yes, <6 mo ago	5 (3)
Yes, between 6 and 12 mo ago	6 (4)
Yes, between 12 mo and 2 y ago	17 (11)
Yes, between 2 and 3 y ago	17 (11)
Yes, longer than 3 y ago	11 (7)
Telerehabilitation software (recently used; n=142)	
Zoom	66 (47)
Telephone	59 (42)

XSL•FO
RenderX

Characteristic	Values, n (%)[a]
Microsoft Teams	47 (33)
Physitrack	30 (21)
Other	89 (63)

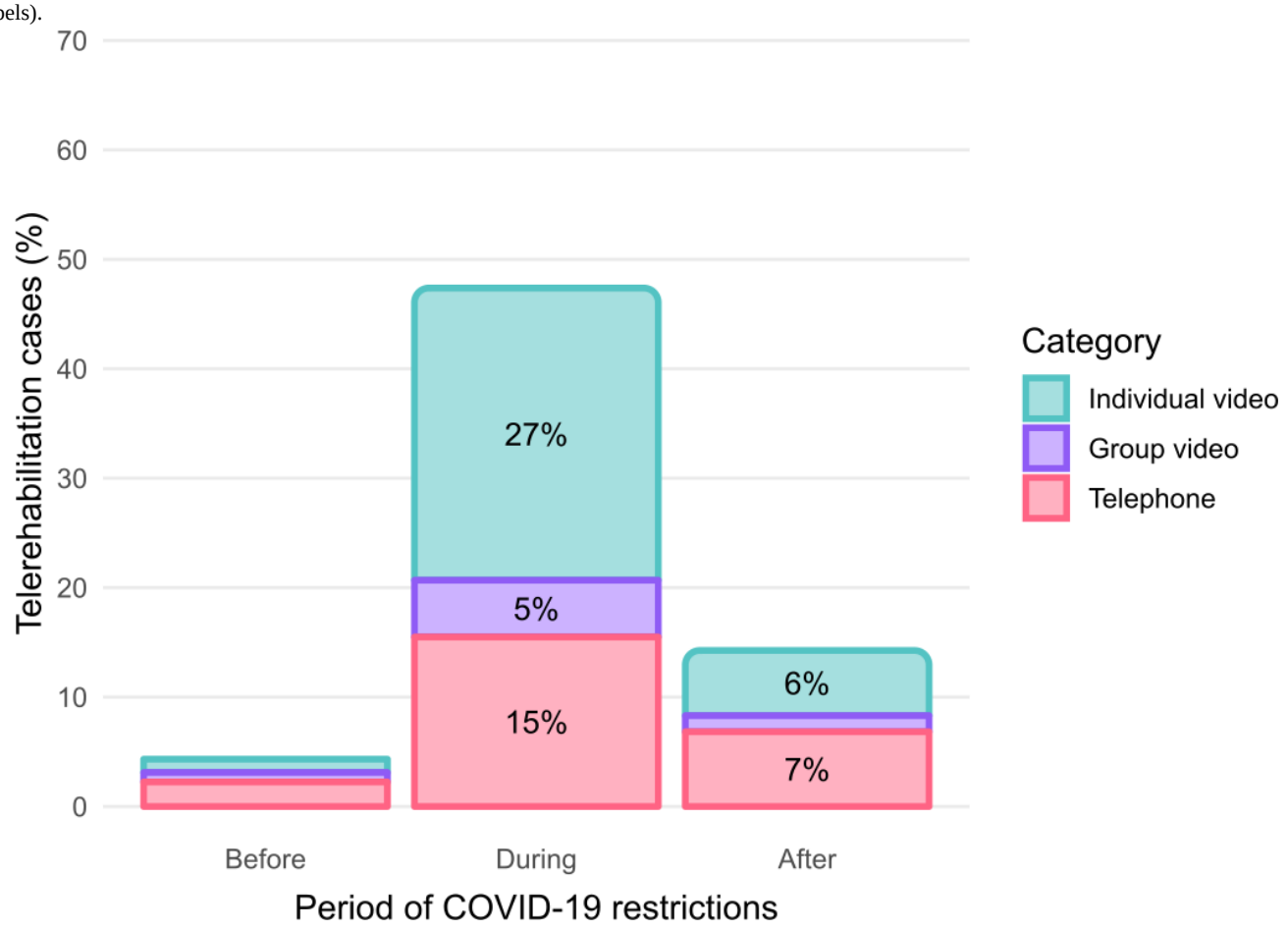
[a]Percentages may not sum to 100% because respondents could select multiple options.

Shifts in Telerehabilitation Use Through the Phases of the COVID-19 Pandemic

Thirty percent (44/148) of respondents indicated that they were using telerehabilitation in clinical practice before the pandemic. This rose to 94% (138/147) during the period of COVID-19 restrictions and reduced to 82% (118/144) after restrictions were lifted. Only 3% (4/152) of the sample indicated that they had never provided telerehabilitation consultations (individual or group videoconferencing, or telephone consultations). Total telerehabilitation caseload rose to account for almost 47% of the total caseload during the period of restrictions but dropped substantially to 14% once restrictions were lifted, but still remained above the prepandemic level of 4% (Figure 1). This pattern was fairly consistent across areas of practice over the COVID-19 pandemic (Figure S1 in Multimedia Appendix 2).

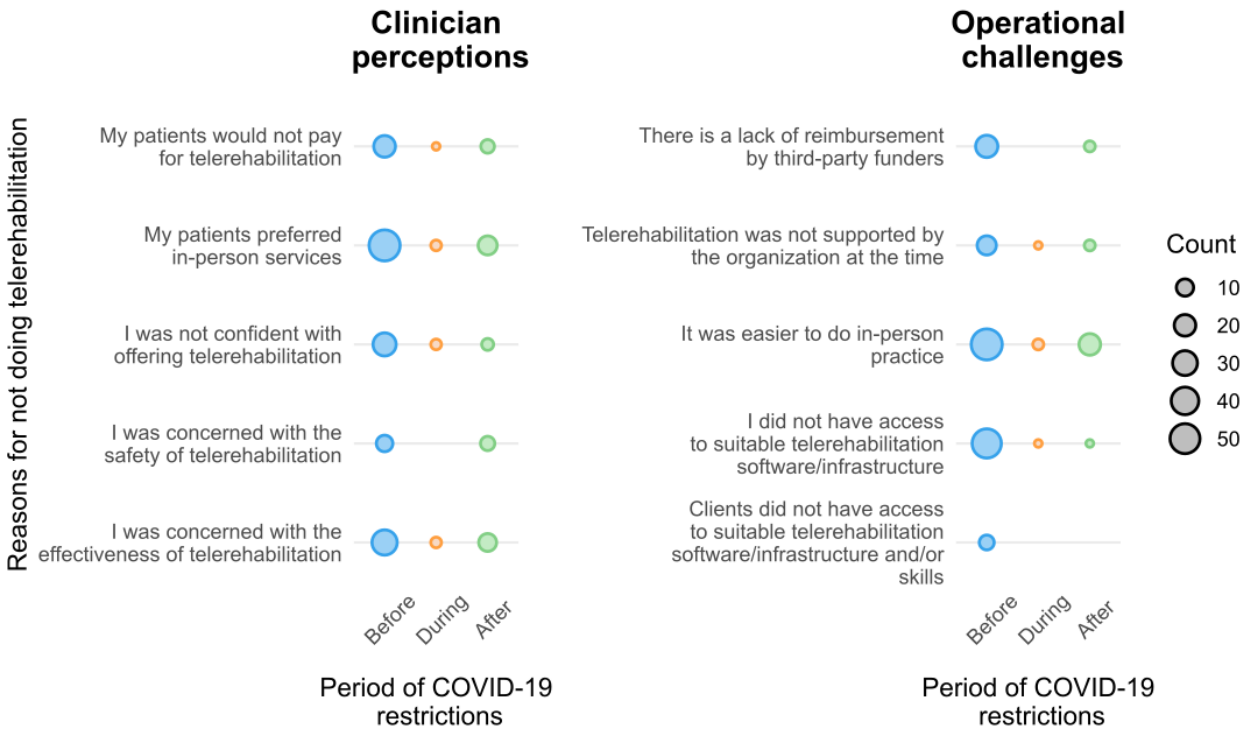
Reasons for not providing telerehabilitation consultations during each phase of the pandemic are provided in Figure 2. Across all 3 phases, the primary reasons were the perception that patients prefer in-person consultations (83/139, 60%) and that it was easier to do in-person consultations (55/139, 55%; Figure 2).

Figure 1. Shift in estimated telerehabilitation caseload before, during, and after the COVID-19 pandemic restrictions (values <2% are plotted without labels).



XSL•FO
RenderX

Figure 2. Reasons for not providing telerehabilitation throughout the phases of the COVID-19 pandemic, grouped by clinician perceptions and operational challenges. The area of circles represents the total count of respondents listing that reason for that point in time.



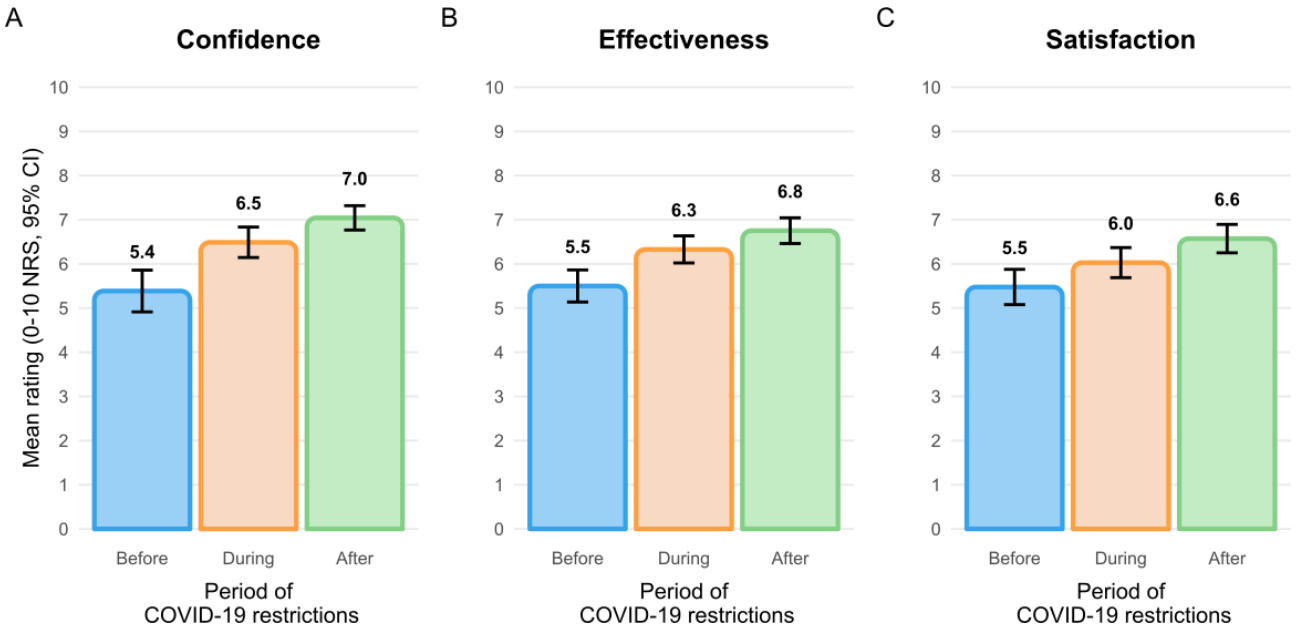
Before the pandemic, additional reasons for not offering telerehabilitation were primarily the perception that there was no need for telerehabilitation (57/104, 55%) or that physiotherapists did not have access to suitable telerehabilitation software or infrastructure (47/104, 45%; Table S1 in Multimedia Appendix 3). After restrictions were eased, the primary reasons for not providing telerehabilitation services were that respondents were concerned about the effectiveness of telerehabilitation (13/26, 50%) and did not like providing care via telerehabilitation (11/26, 42%; Table S1 in Multimedia Appendix 3).

Shifts in Confidence, Effectiveness, and Satisfaction With Telerehabilitation

Physiotherapist ratings of confidence in providing care via telerehabilitation, perceived effectiveness of telerehabilitation, and satisfaction with telerehabilitation progressively increased from before, during, to after restrictions associated with the COVID-19 pandemic (Figure 3). Almost 85% (120/142) of respondents indicated that providing telerehabilitation had become easier over time.

XSL•FO
RenderX

Figure 3. Participant ratings of (A) confidence, (B) perceived effectiveness, and (C) satisfaction with telerehabilitation across the pandemic. NRS: numerical rating scale.

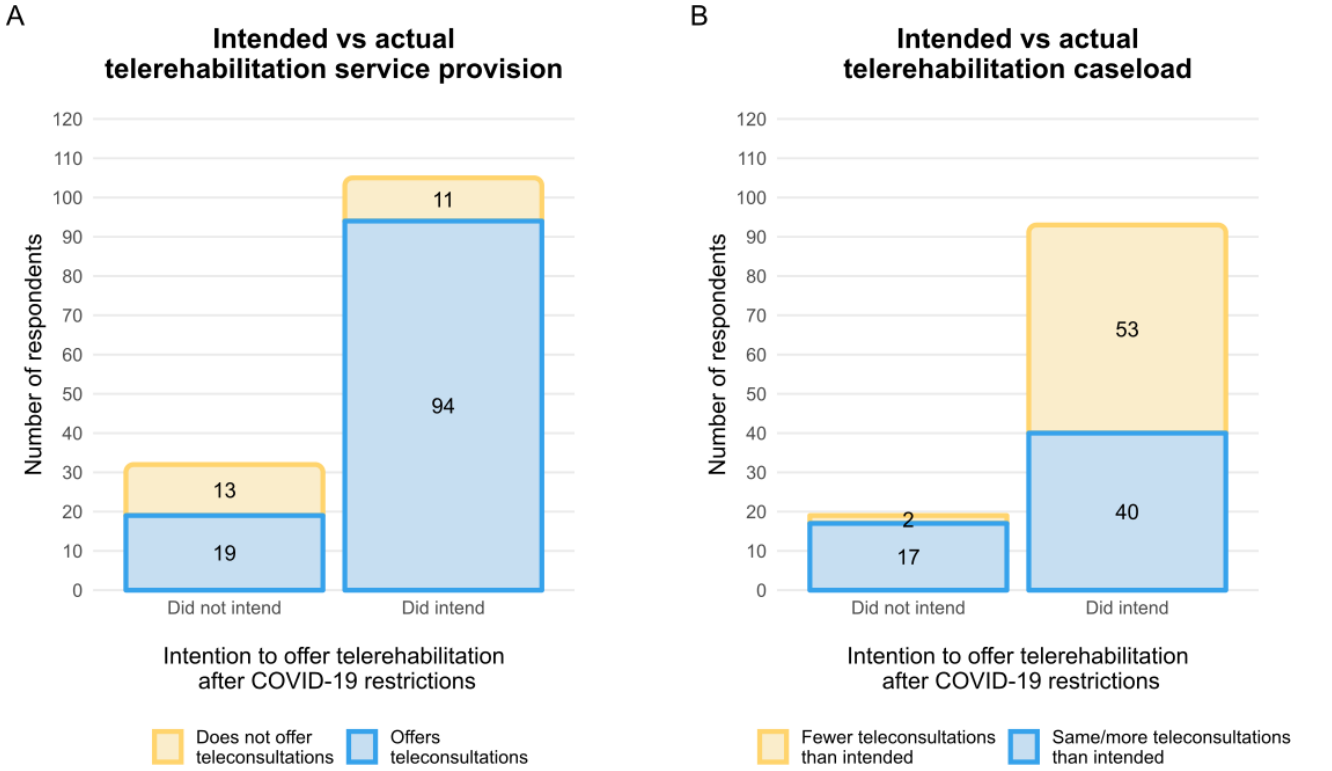


Intended Versus Actual Telerehabilitation Use

Most respondents (105/137, 77%) intended to offer telerehabilitation after the easing of COVID-19 restrictions (Figure 4). Of these, only 10% (11/105) did not offer telerehabilitation despite intending to do so (Figure 4A). Primary reasons for not intending to offer telerehabilitation were because it was "easier to do in-person" (23/32, 72%) and because "patients prefer in-person" (22/32, 69%; Table S2 in Multimedia Appendix 3). Of those who did not intend to, more than half (19/32, 59%) did continue to offer telerehabilitation after restrictions were eased (Figure 4A). Approximately 50% of respondents who intended to continue offering telerehabilitation consultations (53/93) or who actually continued offering them (58/118) after restrictions were eased were providing fewer consultations than initially intended (Figure 4B). Respondents indicated that this was because patients "prefer in-person services" (44/58, 76%), "patient demand reduced more than expected" (32/38, 55%) and because it was "easier to do in-person consultations" (27/38, 47%; Table S3 in Multimedia Appendix 3). Primary reasons for continuing to offer telerehabilitation services included that telerehabilitation allowed physiotherapists to offer services to patients who would not usually be able to attend their clinic (84/118, 71%), that patients like the option of receiving care via telerehabilitation (76/118, 64%), and that patients find telerehabilitation convenient (74/118, 63%; Table S4 in Multimedia Appendix 3).

XSL•FO
RenderX

Figure 4. Intended versus actual telerehabilitation service provision and caseload. (A) Number of respondents who intended to offer telerehabilitation after easing of pandemic restrictions compared to whether they do offer telerehabilitation now. (B) Number of respondents who intended to offer telerehabilitation after easing of pandemic restrictions compared to whether the frequency of telerehabilitation met intentions.



Factors That Influence Telerehabilitation Use Postpandemic Restrictions

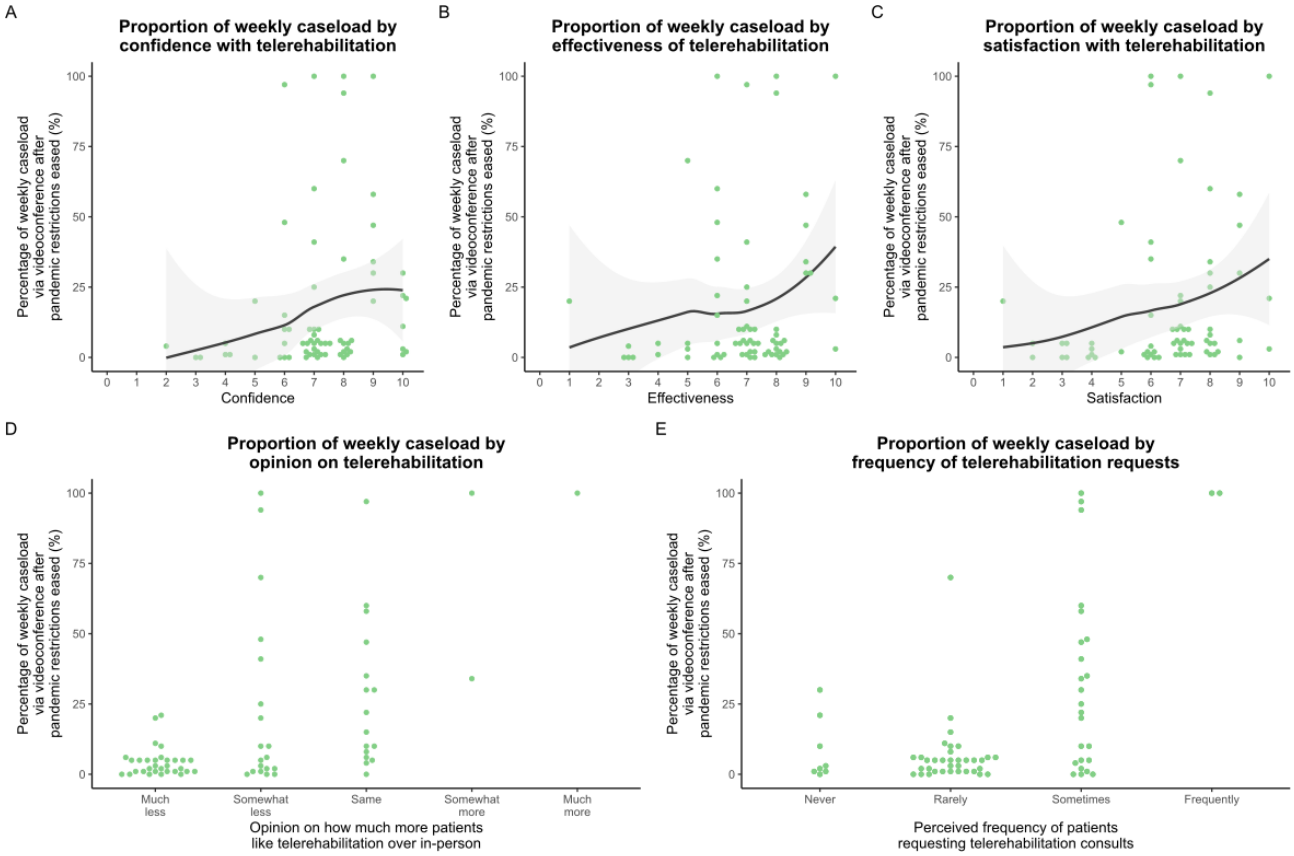
Positive correlations were noted between a higher proportion of weekly caseload conducted via telerehabilitation and higher ratings of confidence in using telerehabilitation (Figure 5A), perceived effectiveness of telerehabilitation (Figure 5B), and satisfaction with telerehabilitation (Figure 5C).

Almost half of the respondents (69/142, 49%) indicated that patients were “rarely” requesting telerehabilitation since the easing of restrictions, and in the opinion of approximately half of the respondents (70/142, 49%), patients like telerehabilitation “much less than in-person consultations.” Physiotherapists who believed that patients liked telerehabilitation much less than in-person consultations appeared to have a lower proportion of their weekly caseload conducted via telerehabilitation (Figure 5D). Similarly, physiotherapists who reported that their patients requested telerehabilitation at least sometimes seemed more likely to have a higher proportion of weekly cases conducted via telehealth (Figure 5E).

The median (IQR) percentage of patients considered unsuitable for telerehabilitation by the respondents was 50.5% (50). Patient complexity and conditions requiring hands-on treatment were the primary reasons that patients were “often” considered unsuitable (Figure S2 in Multimedia Appendix 4 and Table S5 in Multimedia Appendix 3). Additional reasons respondents provided for deeming patients unsuitable are provided in Table S6 in Multimedia Appendix 3, with the largest proportion being patient preference for in-person consultations (6/30, 20%), physical examination being indicated (4/30, 13%), and complex patient presentations (6/30, 14%).

XSL•FO
RenderX

Figure 5. noTotal weekly telerehabilitation caseload versus (A) confidence, (B) effectiveness, (C) satisfaction, (D) patient liking for telerehabilitation, and (E) patient requests for telerehabilitation. Dark gray lines represent locally estimated scatterplot smoothing fits (using the whole span of the data) with 95% CIs (light gray shading).



Telerehabilitation Clinical Practice Considerations

Although only 5% (8/142) of respondents reported never experiencing technical issues themselves, most (120/142, 84%) indicated encountering these issues rarely or sometimes, and just 9% (14/142) experienced them often. Likewise, only one respondent (1/142, 1%) reported that their patients had never encountered technical issues, whereas the majority (112/142, 79%) reported that patients experienced technical issues rarely or sometimes, and 19% (29/142) reported that patients often experienced technical issues. When technical issues were encountered, 74% (105/142) reported only moderate or less disruption to the consultation. Only 5% (7/142) reported that technical issues were extremely disruptive, and just 4% (6/142) reported often having to cancel or reschedule appointments due to technical issues (Figure S3 in Multimedia Appendix 5). To support the delivery of telerehabilitation consultations, physiotherapists used text message reminders (109/142, 77%); written or digital educational material about the condition (67/142, 47%); and written instructions, diagrams, or booklets (63/142, 44%; Table S7 in Multimedia Appendix 3).

Almost three-quarters (104/142, 73%) of respondents indicated that they used similar parameters of care for telerehabilitation as for in-person consultations (eg, similar consultation frequency, duration, and similar content). For the respondents indicating that parameters of care were different (38/142, 27%), the primary reasons were that physical assessment or treatment was limited via telerehabilitation (15/38, 39%), consultations were shorter (10/38, 26%), and consultations were more focused on exercise or education (8/38, 21%). Additional reasons are provided in Table S8 in Multimedia Appendix 3. Similarly, most respondents (128/142, 90%) indicated that telerehabilitation consultations were about the same duration or shorter than in-person consultations (Table S9 in Multimedia Appendix 3) and that consultation frequency was "about the same" as in person (72/142, 51%) or less often than in person (53/142, 37%; Table S9 in Multimedia Appendix 3).

Considering the cost of telerehabilitation consultations, almost three-quarters of physiotherapists indicated that they charged "about the same" as an in-person consultation (104/142, 73%), with very few respondents (5/142, 4%) charging more than in-person consultations. Responses about the cost to the business of providing telerehabilitation were similar, with 50% (71/142) of respondents considering telerehabilitation to cost about the same and 36% (51/142) indicating that telerehabilitation consultations cost the business less than in-person consultations (Table S9 in Multimedia Appendix 3).

The median proportion of patients offered hybrid care in a current weekly caseload was 5% (minimum=0, Q1=1, Q3=20, maximum=100). In hybrid models of care, 42% (48/113) of physiotherapists indicated that patients typically receive many more in-person than telerehabilitation consultations, 27% (30/113) receive the same, and 18% (21/113) receive fewer in-person visits compared to telerehabilitation consultations (Table S9 in Multimedia Appendix 3). Other ways in which telerehabilitation models of care differ from in-person models were coded qualitatively and provided in Table S10 in

XSL•FO
RenderX

Multimedia Appendix 3. When respondents used a hybrid model, 27% (8/27) offered telerehabilitation only after an initial in-person consult, 15% (4/27) described limiting the physical assessment or treatment component of consultations, and 11% (3/27) said telerehabilitation consultations in hybrid models had a greater case management focus (3/27, 11%).

Inductive content analysis of free-text responses identified four key themes that reflected respondents' perspectives on using telerehabilitation in clinical practice postpandemic (Table S11 in Multimedia Appendix 3): (1) concerns about telerehabilitation, (2) perceived benefits of telerehabilitation, (3) how telerehabilitation is used in practice, and (4) physiotherapists' willingness to provide telerehabilitation services.

Theme 1: Concerns About Telerehabilitation (n=28)

Physiotherapists expressed a range of concerns about the suitability and practicality of telerehabilitation in postpandemic physiotherapy care. The most commonly reported issue was that clients prefer or actively seek in-person consultations (n=12, 43%). For example, one participant said that despite telerehabilitation remaining available for their patients, they "*often prefer face-to-face*" and that "*people wanted to revert back to the 'usual' ways and leave the changes of COVID behind them moving forward once restrictions eased*" (musculoskeletal physiotherapist). Some respondents (n=4, 14%) emphasized that telerehabilitation is not suitable for all clients, particularly those with complex conditions, communication difficulties, or low digital literacy, and that they were selecting suitable clients for telerehabilitation, and "*not offering [it] for those not 'tech-savvy'*" (neurological physiotherapist). Concerns were also raised about the limitations of assessment via videoconference (n=2, 7%) and challenges related to internet connectivity and software reliability (n=3, 11%). Additional issues included payment and reimbursement barriers (n=2, 7%), difficulties building rapport remotely (n=2, 7%), and reduced referrals and attendance for telerehabilitation compared to in-person care.

Theme 2: Perceived Benefits of Telerehabilitation (n=20)

Despite concerns, respondents acknowledged several advantages of using telerehabilitation in their clinical practice postpandemic restrictions. The most frequently cited benefit was that telerehabilitation improved patient access to care, particularly for those in rural or remote areas or those with difficulties with travel or limited time (n=8, 40%). One musculoskeletal physiotherapist said that telerehabilitation "*has made physiotherapy much more accessible to a wider population and allows people greater flexibility with appointments*." Participants also noted an increased acceptance of telerehabilitation (among patients and providers; n=6, 30%), with some suggesting that it "*has become common practice now*" (musculoskeletal physiotherapist) and that it can be effective for certain presentations (eg, chronic musculoskeletal conditions; n=3, 15%), for supporting patient self-management (n=2, 10%) and providing greater flexibility in service delivery (n=1, 5%).

Theme 3: How Telerehabilitation Is Used in Practice (n=8)

Participants described integrating telerehabilitation into their clinical practice for subsequent consultations following initial in-person visits (n=2, 25%), for triaging (n=1, 12.5%) and case management (n=1, 12.5%), and as a tool for exercise prescription (n=1, 12.5%). Some participants (n=2, 25%) indicated that videoconferencing was preferred over telephone, and 1 (12.5%) participant noted that at times additional support is required at the patient end to effectively deliver telerehabilitation services. One cardiorespiratory, hospital-based physiotherapist described that telerehabilitation "*consultations have been effective in triaging patients and determining the appropriate level of care required*," whereas another described that telerehabilitation "*has been a great option for follow-up appointments, especially when you have already build rapport with patients...[and it]...has been a great way to check in with people who have busy schedules or live far away and find it difficult coming in*" (pelvic health and musculoskeletal physiotherapist).

Theme 4: Physiotherapists' Willingness to Provide Telerehabilitation Services (n=23)

Many participants (n=16, 70%) were willing to continue providing telerehabilitation services, driven by the perceived benefits and uses of telerehabilitation. For example, one private practice, musculoskeletal and mental health physiotherapist said that "*for the provision of exercise and movement based interventions, telerehabilitation has worked better than in-person as it provides easier access to more people given I live in a regional area*." Despite this willingness, some participants expressed low satisfaction with telerehabilitation (n=2, 9%) or a preference for in-person consultations (n=2, 9%). For example, a sports, exercise, and musculoskeletal physiotherapist working in private practice said that despite telerehabilitation "*opening up my practice to lots of different people around Australia and internationally... I still prefer to consult in-person...*" The need for ongoing education about the utility of telerehabilitation in physiotherapy was noted (n=1, 4%), despite the perception that education about how to deliver telerehabilitation had improved during the pandemic (n=1, 4%).

Discussion

Principal Findings

Despite an initial increase due to the COVID-19 pandemic restrictions and physiotherapists' intentions to continue offering telerehabilitation services, many physiotherapists were offering fewer telerehabilitation consultations than anticipated once restrictions were lifted. This was primarily due to a preference for in-person consultations, concerns about the effectiveness of telerehabilitation, and the perception that physiotherapy consultations are easier to conduct in person.

International postpandemic data across both physiotherapy and other health services show a similar "peak-to-plateau" pattern, where telehealth usage increased substantially during restrictions before falling and stabilizing at a lower level rather than returning to prepandemic levels. In a Polish national dataset,

XSL•FO
RenderX

telehealth in both outpatient health and rehabilitation services (excluding mental health) rose from prepandemic levels near zero to peak in 2020 before subsequently stabilizing at approximately one-fifth and one-third of their respective peak volumes [25]. Similarly, musculoskeletal physical therapists in the United States reported reduced telerehabilitation usage postpandemic, albeit at levels higher than prepandemic [22]. Likewise, across the US health system, overall telehealth usage peaked in 2020 and then declined but stabilized by 2023 [26,27]. Notably, however, telehealth usage was better sustained for services less dependent on hands-on care (eg, behavioral health and psychiatry) and in states where policies were put in place to ensure payment parity with comparable in-person services [26].

Although the perceived effectiveness of telerehabilitation had increased over the course of the pandemic, more than half of participants still identified concerns about the effectiveness of telerehabilitation as a primary reason to stop offering telerehabilitation consultations once able to resume in-person services. This is consistent with other studies conducted during the pandemic, where physiotherapists indicated concerns about the effectiveness of telerehabilitation for physiotherapy assessment and/or management [13,28].

Research indicates that outcomes for telerehabilitation are the same, if not better, than in-person physiotherapy for a range of conditions. For example, systematic reviews and randomized controlled trials in musculoskeletal, cardiac, and pulmonary populations demonstrate the noninferiority of telerehabilitation [29] and good validity for assessment conducted via telerehabilitation [30]. Physiotherapists' perceptions may be centered around occupational self-efficacy [31] or their own personal clinical experience of "effectiveness" rather than evidence of effectiveness in the published literature. However, our results suggest that physiotherapists were not concerned about their own ability to deliver services via telerehabilitation. Both perceived satisfaction with telerehabilitation and confidence in delivering telerehabilitation trended upward from the prepandemic to the postpandemic period, and "I was not confident with telerehabilitation" was not a key factor in our findings after restrictions were lifted (3/26, 12%) . However, large proportions of respondents who did not offer telerehabilitation at this stage said it was easier to do in-person consultations instead (20/26, 70%); they were concerned with the effectiveness of telerehabilitation (13/26, 50%), and they did not like providing care via telerehabilitation (11/26, 42%). These findings suggest that there are additional factors influencing physiotherapists' perceptions about the superiority of "hands-on" or "in-person" physiotherapy [32-34] that have not been comprehensively explored, such as the professional identity of a physiotherapist [31,35].

A qualitative study describing a successful, rapid transition to telerehabilitation during the pandemic challenges the perception that physiotherapy requires "hands-on" approaches and needs to be in person [36]. This study identified that physiotherapists' readiness and willingness to modify their approach influenced the success of telerehabilitation. In our study, physiotherapists preferred in-person consultations themselves and perceived that their patients also preferred in-person consultations, which is likely to influence whether they offer telerehabilitation to patients. While systematic reviews suggest that patient satisfaction with telerehabilitation is comparable to and often higher than in-person care [37,38], many patients report a preference for in-person physiotherapy if given a choice [37,39]. Although physiotherapists might have thought during the pandemic that patient demand for telerehabilitation would remain (eg, explaining their intention to offer it), if patient demand for it decreased (as 55% of our sample indicated), physiotherapists would likely perceive that patients prefer in-person care (and indeed 76% of our sample did).

Clinician preferences for providing in-person physiotherapy have also been explored and reported on in the literature. Despite high levels of clinician satisfaction when providing telerehabilitation in clinical trials [40,41], this does not appear to be the case for *in-practice* preference for, or satisfaction with, telerehabilitation [21,22,28]. Although satisfaction and confidence with telerehabilitation increased over time, participants in this study still perceived in-person physiotherapy to be easier. The rigorous planning or structured training required for telerehabilitation delivery in a randomized clinical trial, rather than day-to-day clinical practice, may explain this difference in perceptions, highlighting a need for training specific to the clinical implementation of telerehabilitation. Studies examining barriers to implementing telerehabilitation in routine physiotherapy practice consistently identify insufficient training for conducting telerehabilitation consultations as a primary concern [42]. To address these challenges, international clinical practice guidelines provide evidence-based recommendations and strategies for overcoming barriers, guiding the training of clinicians and facilitating effective implementation of telerehabilitation into physiotherapy practice [43].

Clinicians have long identified the technological illiteracy of clients as a barrier to the adoption of telerehabilitation in physiotherapy [42]. Despite advances in technology infrastructure, when transitioning to telerehabilitation during the COVID-19 pandemic period, clinicians still identified "technology concerns" (including clinician concerns about client ability to use technology) as a barrier to telerehabilitation use in clinical practice [28,34,36,42,44]. In our study, concerns about technical issues or patients being unable to use or access technology were not identified as primary reasons physiotherapists determined patients were unsuitable for telerehabilitation. Additionally, technical issues were only slightly or moderately disruptive to consultations. This is consistent with findings from an evaluation of consultations delivered in a randomized controlled trial, which found that technical issues occurred but were infrequent and minimally disruptive [45]. This could potentially be because, at the time data were collected for this study (2023), physiotherapists and clients had greater experience with and exposure to the technology required for telerehabilitation and had become more comfortable over time [46]. In other studies where data were collected earlier in the pandemic, it is possible that fewer people were familiar with telerehabilitation technology, hence it being a bigger barrier to delivering telerehabilitation services at the time [20,21,36].

XSL•FO
RenderX

Strengths and Limitations

The findings of this study should be interpreted with the following limitations in mind. First, this was a small convenience sample, and findings may have been skewed by self-selection bias (with those with strong opinions, either positive or negative, electing to complete the survey). Second, we asked participants to recall what they were doing before and during the pandemic several years after the fact. Therefore, it should be acknowledged that participants' responses may have been influenced by recall bias. However, our data pertaining to before and during the pandemic were consistent with other studies conducted during the pandemic and their intentions to continue (eg, 69% in our study said that during the pandemic, they intended to offer telerehabilitation after restrictions were eased). In a study by Bennell et al [20], 81% intended to continue offering telerehabilitation consultations after the pandemic, and in a study by Peng et al [28], 55% and 68% intended to continue offering phone and videoconferencing, respectively. If the opportunity arises (ie, another period of restrictions to in-person consultations), researchers should consider using prospective study designs. Moreover, because our survey encompassed both phone calls and videoconferencing, our findings may not reflect modality-specific differences in perceptions reported elsewhere [28]. This survey also only sampled physiotherapists operating within Australia's health care system, so its findings may not fully translate to other countries with different telerehabilitation policies, funding models, or cultural attitudes toward remote care. Finally, due to an error, questions about confidence, satisfaction, and perceived effectiveness after the pandemic restrictions were eased were misworded and instead asked about experiences during the pandemic. It is likely that, given that all questions before this were about easing restrictions, most respondents still answered according to the intention of the question, but we cannot discount that some answered more literally, thereby skewing the data.

Conclusions

Although telerehabilitation use surged with pandemic restrictions, it has subsequently decreased significantly, with telerehabilitation accounting for only a small proportion of the total caseload. Despite increased confidence and satisfaction with telerehabilitation, clinician preference, and physiotherapists' perceptions of patient preference for in-person care, reduced demand and the ease of in-person practice influence the use of telerehabilitation postrestrictions and suggest persistent barriers to frequent use. Addressing these barriers is crucial to enhance the long-term viability and effectiveness of telerehabilitation physiotherapy in Australia.

Funding

This work was supported by the National Health and Medical Research Council (NHMRC Project Grant 1157977); RSH is supported by an NHMRC Investigator grant (2025733); KLB is supported by a NHMRC Investigator Grant (1174431); BJL is supported by a University of Melbourne CR Roper Fellowship and Dame Kate Campbell Fellowship. For the purposes of open access, the author has applied a CC BY public copyright licence to any author-accepted manuscript version arising from this submission. The funders have no role in conduct, analysis, or reporting of this study.

Data Availability

The datasets generated during and analyzed in this study are not publicly available due to lack of ethical clearance to disclose data to third parties.

Authors' Contributions

Conceptualization: MHR, JS, BL, KLB, RSH, TR Data curation: MHR, JS Formal analysis: MHR, JS Visualization: JS Writing – original draft: MHR, JS Writing – review & editing: MHR, JS, BL, KLB, RSH, TR Funding acquisition: KLB, RSH, TR

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey.
[DOCX File, 59 KB - rehab_v13i1e81008_app1.docx]

Multimedia Appendix 2
Telerehabilitation caseload over the COVID-19 pandemic for each physiotherapy area of practice.
[PNG File, 220 KB - rehab_v13i1e81008_app2.png]

Multimedia Appendix 3
Detailed quantitative (n, %) and qualitative survey findings on physiotherapists' telerehabilitation practice across COVID-19 restrictions: reasons for not offering, reducing, ceasing, or continuing telerehabilitation; reasons patients were considered unsuitable

XSL•FO
RenderX

and the resources used to support videoconferencing; differences in consultation parameters and hybrid models of care compared with in-person practice; and postrestriction perspectives and willingness to provide telerehabilitation.
[DOCX File, 43 KB - rehab_v13i1e81008_app3.docx]

Multimedia Appendix 4
Patient suitability for telerehabilitation.
[PNG File, 41 KB - rehab_v13i1e81008_app4.png]

Multimedia Appendix 5
Technical issues.
[PNG File, 42 KB - rehab_v13i1e81008_app5.png]

References

1. Brown RC, Coombes JS, Rodriguez KJ, Hickman IJ, Keating SE. Effectiveness of exercise via telehealth for chronic disease: a systematic review and meta-analysis of exercise interventions delivered via videoconferencing. Br J Sports Med 2022 Sep;56(18):1042-1052. [doi: 10.1136/bjsports-2021-105118]
2. Dias JF, Oliveira VC, Borges PRT, et al. Effectiveness of exercises by telerehabilitation on pain, physical function and quality of life in people with physical disabilities: a systematic review of randomised controlled trials with GRADE recommendations. Br J Sports Med 2021 Feb;55(3):155-162. [doi: 10.1136/bjsports-2019-101375] [Medline: 33060156]
3. Dario AB, Cabral AM, Almeida L, et al. Effectiveness of telehealth-based interventions in the management of non-specific low back pain: a systematic review with meta-analysis. Spine J 2017 Sep;17(9):1342-1351. [doi: 10.1016/j.spinee.2017.04.008] [Medline: 28412562]
4. Laver KE, Adey-Wakeling Z, Crotty M, Lannin NA, George S, Sherrington C. Telerehabilitation services for stroke. Cochrane Database Syst Rev 2020 Jan 31;1(1):CD010255. [doi: 10.1002/14651858.CD010255.pub3] [Medline: 32002991]
5. Hwang R, Bruning J, Morris NR, Mandrusiak A, Russell T. Home-based telerehabilitation is not inferior to a centre-based program in patients with chronic heart failure: a randomised trial. J Physiother 2017 Apr;63(2):101-107. [doi: 10.1016/j.jphys.2017.02.017] [Medline: 28336297]
6. Zhou Z, Zhou X, Cui N, et al. Effectiveness of tele-rehabilitation after total hip replacement: a systematic review and meta-analysis of randomized controlled trials. Disabil Rehabil 2024 Oct;46(20):4611-4616. [doi: 10.1080/09638288.2023.2280070] [Medline: 37990882]
7. Wade V, Soar J, Gray L. Uptake of telehealth services funded by Medicare in Australia. Aust Health Rev 2014 Nov;38(5):528-532. [doi: 10.1071/AH14090] [Medline: 25219655]
8. Gray LC, Edirippulige S, Smith AC, et al. Telehealth for nursing homes: the utilization of specialist services for residential care. J Telemed Telecare 2012 Apr;18(3):142-146. [doi: 10.1258/jtt.2012.SFT105] [Medline: 22362837]
9. Brewster L, Mountain G, Wessels B, Kelly C, Hawley M. Factors affecting front line staff acceptance of telehealth technologies: a mixed-method systematic review. J Adv Nurs 2014 Jan;70(1):21-33. [doi: 10.1111/jan.12196] [Medline: 23786584]
10. Gregory WJ, Burchett S, McCrum C. National survey of the current clinical practices of the UK rheumatology physiotherapists: a short report. Musculoskeletal Care 2021 Mar;19(1):136-141. [doi: 10.1002/msc.1516] [Medline: 32975369]
11. Cottrell MA, Hill AJ, O'Leary SP, Raymer ME, Russell TG. Service provider perceptions of telerehabilitation as an additional service delivery option within an Australian neurosurgical and orthopaedic physiotherapy screening clinic: a qualitative study. Musculoskelet Sci Pract 2017 Dec;32:7-16. [doi: 10.1016/j.msksp.2017.07.008] [Medline: 28787636]
12. Lawford BJ, Bennell KL, Kasza J, Hinman RS. Physical therapists' perceptions of telephone- and internet video-mediated service models for exercise management of people with osteoarthritis. Arthritis Care Res (Hoboken) 2018 Mar;70(3):398-408. [doi: 10.1002/acr.23260] [Medline: 28437566]
13. Reynolds A, Awan N, Gallagher P. Physiotherapists' perspective of telehealth during the COVID-19 pandemic. Int J Med Inform 2021 Dec;156:104613. [doi: 10.1016/j.ijmedinf.2021.104613] [Medline: 34688969]
14. Bettger JP, Thoumi A, Marquevich V, et al. COVID-19: maintaining essential rehabilitation services across the care continuum. BMJ Glob Health 2020 May;5(5):e002670. [doi: 10.1136/bmjgh-2020-002670] [Medline: 32376777]
15. COVID-19 submission 325. : Australian Physiotherapy Association; 2020 URL: https://www.aph.gov.au/DocumentStore.ashx?id=52e7027c-1187-4d62-8f2b-53fcbcceb939&subId=685789 [accessed 2025-12-31]
16. COVID-19 (coronavirus disease 2019). Victorian Government Department of Health. 2020 Apr 1. URL: https://www.health.vic.gov.au/infectious-diseases/covid-19-coronavirus-disease-2019 [accessed 2025-12-31]
17. Aged care royal commission impact of COVID-19 on aged care supplementary submission. : Australian Physiotherapy Association; 2020 URL: https://australian.physio/sites/default/files/submission/APA_Supplementary_Submission_Impact_of_COVID-19_on_Aged_Care_August_2020.pdf [accessed 2025-12-31]

XSL•FO
RenderX

18. Report from the telehealth working group. : Medicare Benefits Schedule Review Taskforce; 2020 URL: https://www.health.gov.au/sites/default/files/documents/2021/06/final-report-from-the-telehealth-working-group.pdf [accessed 2025-12-31]
19. Telehealth guidelines – Response to COVID-19. : Australian Physiotherapy Association; 2020 URL: https://australian.physio/sites/default/files/APATelehealthGuidelinesCOVID190420FA.pdf [accessed 2025-12-31]
20. Bennell KL, Lawford BJ, Metcalf B, et al. Physiotherapists and patients report positive experiences overall with telehealth during the COVID-19 pandemic: a mixed-methods study. J Physiother 2021 Jul;67(3):201-209. [doi: 10.1016/j.jphys.2021.06.009] [Medline: 34147399]
21. Malliaras P, Merolli M, Williams CM, Caneiro JP, Haines T, Barton C. "It's not hands-on therapy, so it's very limited": telehealth use and views among allied health clinicians during the coronavirus pandemic. Musculoskelet Sci Pract 2021 Apr;52:102340. [doi: 10.1016/j.msksp.2021.102340] [Medline: 33571900]
22. McLaughlin K, Minick KI, Fritz JM, et al. Physical therapists' perceptions of telerehabilitation for patients with musculoskeletal conditions in a post-pandemic world. medRxiv. Preprint posted online on Jan 18, 2025. [doi: 10.1101/2025.01.17.25320739] [Medline: 39867367]
23. Sharma A, Minh Duc NT, Luu Lam Thang T, et al. A consensus-based Checklist for Reporting of Survey Studies (CROSS). J Gen Intern Med 2021 Oct;36(10):3179-3187. [doi: 10.1007/s11606-021-06737-1] [Medline: 33886027]
24. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
25. Krawczyk A, Marszałek M. A comparative analysis of telerehabilitation and telemedicine utilization during the COVID-19 pandemic in Poland: trends, patterns, and implications. Int J Telerehabil 2024;16(1):e6627. [doi: 10.5195/ijt.2024.6627] [Medline: 39022440]
26. Gage AD, Knight MA, Bintz C, et al. Disparities in telemedicine use and payment policies in the United States between 2019 and 2023. Commun Med (Lond) 2025 Feb 26;5(1):52. [doi: 10.1038/s43856-025-00757-2] [Medline: 40011624]
27. Mandal S, Wiesenfeld BM, Mann DM, Nov O. The "new" new normal: changes in telemedicine utilization since COVID-19. Am J Manag Care 2025 Mar 1;31(3):e74-e78. [doi: 10.37765/ajmc.2025.89700] [Medline: 40053411]
28. Peng TH, Eng JJ, Harris A, et al. A survey of the experiences of delivering physiotherapy services through telerehabilitation during the COVID-19 pandemic. Front Rehabil Sci 2024;5:1486801. [doi: 10.3389/fresc.2024.1486801] [Medline: 39512760]
29. Suso-Martí L, La Touche R, Herranz-Gómez A, Angulo-Díaz-Parreño S, Paris-Alemany A, Cuenca-Martínez F. Effectiveness of telerehabilitation in physical therapist practice: an umbrella and mapping review with meta-meta-analysis. Phys Ther 2021 May 4;101(5):pzab075. [doi: 10.1093/ptj/pzab075] [Medline: 33611598]
30. Bernhardsson S, Larsson A, Bergenheim A, et al. Digital physiotherapy assessment vs conventional face-to-face physiotherapy assessment of patients with musculoskeletal disorders: a systematic review. PLOS ONE 2023;18(3):e0283013. [doi: 10.1371/journal.pone.0283013] [Medline: 36943857]
31. Roitenberg N, Pincus T, Ben Ami N. Physiotherapy services during the COVID-19 pandemic: a mediated model of physiotherapists' self-efficacy, tele-physiotherapy role stressors, and motivation to provide tele-physiotherapy. Physiother Theory Pract 2024 Jun;40(6):1140-1149. [doi: 10.1080/09593985.2022.2138662] [Medline: 36305357]
32. Hasani F, Malliaras P, Haines T, et al. Telehealth sounds a bit challenging, but it has potential: participant and physiotherapist experiences of gym-based exercise intervention for Achilles tendinopathy monitored via telehealth. BMC Musculoskelet Disord 2021 Feb 4;22(1):138. [doi: 10.1186/s12891-020-03907-w] [Medline: 33541314]
33. Fernandes LG, Oliveira RFF, Barros PM, Fagundes FRC, Soares RJ, Saragiotto BT. Physical therapists and public perceptions of telerehabilitation: an online open survey on acceptability, preferences, and needs. Braz J Phys Ther 2022;26(6):100464. [doi: 10.1016/j.bjpt.2022.100464] [Medline: 36410257]
34. Sia LL, Sharma S, Kumar S, Ajit Singh DK. Exploring physiotherapists' perceptions of telerehabilitation for musculoskeletal disorders: insights from focus groups. Digit Health 2024;10:20552076241248916. [doi: 10.1177/20552076241248916] [Medline: 38665882]
35. Hawley-Hague H, Gluchowski A, Lasrado R, et al. Exploring the delivery of remote physiotherapy during the COVID-19 pandemic: UK wide service evaluation. Physiother Theory Pract 2024 Oct;40(10):2241-2255. [doi: 10.1080/09593985.2023.2247069] [Medline: 37610255]
36. Ross MH, Nelson M, Parravicini V, et al. Staff perspectives on the key elements to successful rapid uptake of telerehabilitation in medium-sized public hospital physiotherapy departments. Physiother Res Int 2023 Jul;28(3):e1991. [doi: 10.1002/pri.1991] [Medline: 36540908]
37. Zischke C, Simas V, Hing W, Milne N, Spittle A, Pope R. The utility of physiotherapy assessments delivered by telehealth: a systematic review. J Glob Health 2021;11:04072. [doi: 10.7189/jogh.11.04072] [Medline: 34956637]
38. Simmich J, Ross MH, Russell T. Real-time video telerehabilitation shows comparable satisfaction and similar or better attendance and adherence compared with in-person physiotherapy: a systematic review. J Physiother 2024 Jul;70(3):181-192. [doi: 10.1016/j.jphys.2024.06.001] [Medline: 38879432]
39. Lawford BJ, Bennell KL, Kimp A, Campbell PK, Hinman RS. Understanding negative and positive feelings about telerehabilitation in people with chronic knee pain: a mixed-methods study. J Orthop Sports Phys Ther 2024 Sep;54(9):594-607. [doi: 10.2519/jospt.2024.12383] [Medline: 39207737]

XSL•FO
RenderX

40. Hinman RS, Nelligan RK, Bennell KL, Delany C. “Sounds a bit crazy, but it was almost more personal:” a qualitative study of patient and clinician experiences of physical therapist-prescribed exercise for knee osteoarthritis via Skype. Arthritis Care Res (Hoboken) 2017 Dec;69(12):1834-1844. [doi: 10.1002/acr.23218] [Medline: 28217864]
41. Amin J, Ahmad B, Amin S, Siddiqui AA, Alam MK. Rehabilitation professional and patient satisfaction with telerehabilitation of musculoskeletal disorders: a systematic review. Biomed Res Int 2022;2022:7366063. [doi: 10.1155/2022/7366063] [Medline: 35958819]
42. Sia LL, Sharma S, Ing JBM, Kumar S, Singh DKA. Physiotherapists’ perceptions, readiness, enablers, and barriers to use telerehabilitation: a scoping review. J Back Musculoskelet Rehabil 2024;37(6):1441-1454. [doi: 10.3233/BMR-240009] [Medline: 38905032]
43. Lee AC, Deutsch JE, Holdsworth L, et al. Telerehabilitation in physical therapist practice: a clinical practice guideline from the American Physical Therapy Association. Phys Ther 2024 May 1;104(5):pzae045. [doi: 10.1093/ptj/pzae045] [Medline: 38513257]
44. Albahrouh SI, Buabbas AJ. Physiotherapists’ perceptions of and willingness to use telerehabilitation in Kuwait during the COVID-19 pandemic. BMC Med Inform Decis Mak 2021 Apr 8;21(1):122. [doi: 10.1186/s12911-021-01478-x] [Medline: 33832473]
45. Ross MH, Russell T, Bennell KL, et al. Technical issues occur but are infrequent and have little impact on physiotherapist-delivered videoconferencing consultations for knee osteoarthritis: a descriptive study. Musculoskelet Sci Pract 2023 Aug;66:102782. [doi: 10.1016/j.msksp.2023.102782] [Medline: 37269590]
46. Tousignant M, Boissy P, Moffet H, et al. Patients’ satisfaction of healthcare services and perception with in-home telerehabilitation and physiotherapists’ satisfaction toward technology for post-knee arthroplasty: an embedded study in a randomized trial. Telemed J E Health 2011 Jun;17(5):376-382. [doi: 10.1089/tmj.2010.0198] [Medline: 21492030]

Abbreviations

CROSS: Checklist for Reporting of Survey Studies

Edited by S Munce; submitted 21.Jul.2025; peer-reviewed by D Ceprnja, I Wilson; accepted 23.Dec.2025; published 27.Jan.2026.

Please cite as:
Ross MH, Simmich J, Lawford BJ, Bennell KL, Hinman RS, Russell T
Telerehabilitation Trends in Australian Physiotherapy and an Exploration of Factors That Influence Use After COVID-19 Restrictions: Qualitative Content Analysis
JMIR Rehabil Assist Technol 2026;13:e81008
URL: https://rehab.jmir.org/2026/1/e81008
doi:10.2196/81008

© Megan H Ross, Joshua Simmich, Belinda J Lawford, Kim L Bennell, Rana S Hinman, Trevor Russell. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Linguistic Validation and Cross-Cultural Adaptation of the Shoulder Telehealth Assessment Tool for Filipino Patients with Musculoskeletal Shoulder Condition: Cross-Sectional Study

Jeffrey Arboleda[1], PTRP, MD, MBA; Sharon Ignacio[1,2], MD; Jose Alvin Mojica[1], MD, MHPEd; Carl Froilan Leochico[2,3], PTRP, MD, MHSc

[1]Department of Rehabilitation Medicine, Philippine General Hospital, University of the Philippines Manila, Taft Avenue, Ermita, Metro Manila, Philippines

[2]Department of Physical Medicine and Rehabilitation, St. Luke's Medical Center, Metro Manila, Philippines

[3]Department of Medicine, Division of Physical Medicine and Rehabilitation, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

Corresponding Author:
Jeffrey Arboleda, PTRP, MD, MBA
Department of Rehabilitation Medicine, Philippine General Hospital, University of the Philippines Manila, Taft Avenue, Ermita, Metro Manila, Philippines

Abstract

Background: Telerehabilitation has been widely adopted to meet the growing rehabilitation demand, but it is often limited by unstable internet connection, poor audiovisual resolution, and difficult virtual assessment. The Shoulder Telehealth Assessment Tool (STAT), a comprehensive, patient-led, preconsultation shoulder physical examination pictorial guide, was developed to address these limitations by easing the communication of instruction during the consultation and potentially removing the need for video calls.

Objective: This study aimed to develop a linguistically valid and culturally appropriate Filipino version of STAT and to evaluate its content validity, internal consistency, understandability, and ease of use.

Methods: A cross-sectional study on the Filipino STAT was conducted in three phases: (1) linguistic validation by experts, (2) cross-cultural adaptation through pretesting of 12 participants diagnosed with a musculoskeletal shoulder condition at the Philippine General Hospital, and (3) pilot study on 47 participants of the same population.

Results: The Filipino STAT had an excellent content validity (scale validity index=0.80 - 0.97), excellent interrater reliability (κ coefficient=0.82 - 1.00), and good internal consistency (Cronbach α=0.87). Understandability was found to be excellent for pain and activity (98%), good for range of motion and special tests (85%), and poor for strength (37%). However, 24% (11/46) of participants perceived the tool difficult to understand with the use of some Tagalog words as the primary barrier, followed by non-familiarity with the tool and difficulty reading the text.

Conclusions: Development of the Filipino STAT through a rigorous linguistic validation and cultural adaptation has produced a culturally appropriate, valid, and reliable tool. Pain and activity, range of motions, and special test subdomains are suitable for clinical assessment, while strength subdomain needs further improvement in understandability.

(JMIR Rehabil Assist Technol 2026;13:e67974) doi:10.2196/67974

KEYWORDS

cross-cultural adaptation; linguistic validation; musculoskeletal rehabilitation; shoulder examination; Shoulder Telehealth Assessment Tool; telerehabilitation; STAT

Introduction

Telemedicine is the use of telecommunication technologies to deliver health care, public health, and health education services remotely [1]. Telerehabilitation is a branch of telemedicine specifically aimed at delivering rehabilitation services. Being an archipelago with limited health care professionals and facilities, the Philippines uses these innovative health care delivery services to improve health access for Filipinos. Telerehabilitation in the Philippines was first adopted in the context of community-based rehabilitation in 2017 [2] and was widely used in 2020 during the COVID-19 pandemic [3]. With the improving COVID-19 situation, telerehabilitation remains a viable solution to delivering rehabilitation services in far-flung areas of the country [4].

XSL•FO
RenderX

However, telerehabilitation is not without limitations, such as unstable internet connection, lack of confidence in establishing clinical diagnosis virtually, limited time allotment per patient, and poor audiovisual resolution [5]. Specifically, in the administration of virtual physical examination, rehabilitation professionals face difficulties in the conduct of special tests, range of motion (ROM) assessment, and strength testing, among others [3]. These challenges have opened opportunities for innovation in telerehabilitation.

One of these innovations was the development of the Shoulder Telehealth Assessment Tool (STAT), which is a comprehensive patient-led shoulder physical examination pictorial guide done prior to the actual teleconsultation, aimed to improve clinical efficiency [6]. It is the first published patient-reported outcome measure to simulate the performance of in-person physical examination, which includes special tests for screening of different shoulder pathologies [7]. A validated visual analog scale (VAS) [8], single assessment numeric evaluation (SANE) [9], and motion analysis and range of motion studies on activities of daily living [10] have also been integrated into the tool. The first subcategory of the STAT is composed of 3 questions for pain and activity. Both VAS and SANE are patient-reported. The VAS ranges from 0=no pain to 10=maximum pain, while the SANE is scored from 0% to 100% of normal. Meanwhile, the current level of daily activity is a nominal score (ie, unaffected sleep, full work, and full recreation or sport). The ROM testing has 9 questions that are answered with either yes (movement completed) or no (movement not completed). There are 5 items for the shoulder special tests answerable by yes or no, and finally 3 items for strength that are answered with either painful, weak, both, or none. Overall, there are a total of 20 questions in the STAT that can be completed in 30 to 45 minutes. In the Philippines and numerous other countries where Filipinos are found, the STAT can potentially improve the accuracy of virtual physical examination techniques for the shoulder, provide better rehabilitation care for patients who do not have stable internet connections or video call capacity, and optimize the delivery of telerehabilitation services. This may also be used by providers from various health care disciplines, such as physiatry, physical therapy, occupational therapy, orthopedics, rheumatology, and pain management, among others, which routinely conduct musculoskeletal examination of the shoulder.

In the Philippines, shoulder conditions have been found to be prevalent among Filipino office workers [11], as well as migrant workers [12]. Currently, the STAT is available in English [6] and has not been translated into any other language. Language and cultural differences call for a careful adaptation of health outcome measures to accurately reflect the cultural nuances and context of the target language version [13]. Hence, this study aimed to develop a linguistically valid and culturally appropriate Filipino version of the STAT and determine its content validity, internal consistency, understandability, and ease of use.

Methods

Study Design

This cross-sectional study used a mixed methods research design on the linguistic validation and cross-cultural adaptation of the Filipino STAT.

Participants

The target population was Filipino adults (aged at least 18 years) with unilateral shoulder pain for at least 6 weeks who consulted in-person or through telemedicine at the Philippine General Hospital (PGH)—Outpatient Department of Rehabilitation Medicine from January 2023 to June 2023. Based on the Department census during that period, there were 304 patients diagnosed with shoulder pathology, such as adhesive capsulitis, rotator cuff injury, and myofascial pain syndrome. The participants should have access to a stable internet and a device with video call capacity and be able to understand written instructions in Filipino or Tagalog, which is the Philippines' most commonly used language. Participants with a history of trauma, suspicion or diagnosis of upper extremity fracture, severe cognitive impairment, known psychiatric comorbidity, cerebrovascular disease, cervical radiculopathy, brachial plexus injury, upper extremity peripheral nerve injury, complex regional pain syndrome, and prior shoulder, neck, and breast surgery were excluded. The participants were allowed to withdraw from the study for any reason.

Sampling

The sample size for the pretest was based on a generally recommended sample size of 12 from the target population based on feasibility and precision of estimates for succeeding studies [14]. The sample size for a pilot study to determine internal consistency through Cronbach α was 46 [15]. Systematic randomized sampling was done where every seventh patient was selected.

Study Procedure

Overview

Adapting the recommendations on linguistic validation and cross-cultural adaptation of self-reported measures, the study procedure was divided into two phases [16,17]: (1) translation of the STAT from the original English to the Filipino version; and (2) cross-cultural adaptation through pretesting. An additional phase was added for internal consistency testing of the final version of the Filipino STAT (Figure 1). Meetings, where necessary, were all done virtually, and permission to record each online meeting was sought from all participants.

XSL•FO
RenderX

Figure 1. This is a flowchart adapted from Lorca et al [17] for linguistic validation and cross-cultural adaptation [6]. STAT: Shoulder Telehealth Assessment Tool.

XSL•FO
RenderX

Phase 1: Linguistic Validation

Step 1: Permission From the Instrument's Authors

Permission for linguistic validation and cross-cultural adaptation into the Filipino language was obtained from the STAT developers prior to initiation of translation.

Step 2: Forward Translation

The forward translation was independently done by 2 bilingual translators. The first forward translator (T1) was a licensed physical therapist with a Master's degree in physical therapy who provided a "reliable equivalence from a clinical perspective," while the second forward translator (T2) was a university instructor and a representative from the *Sentro ng Wikang Filipino* (National Center of the Filipino Language) with no medical background who served to "reflect the language used by the population."

Step 3: Review and Synthesis

The translators, together with a recording observer, convened to synthesize their versions and arrive at a consensus version. Translation and synthesis forms were used throughout the linguistic validation process by the translators and recording observer to document the translated version, rationale for changes, and any disputes or challenges in doing the translation or synthesis.

Step 4: Backward Translation

Backward translation was independently done by another pair of bilingual translators (B1 and B2), who happened to be high school Filipino language teachers and were not familiar with the original version. This step was done to prevent information bias and to detect any errors in the forward translation.

Step 5: Content Validation

The first forward- and backward-translated version of the Filipino STAT was then assessed for content validity independently by clinical experts using a content validation form. The experts were composed of 6 board-certified local physiatrists with expertise in the fields of sports medicine, musculoskeletal ultrasound, interventional physiatry, telerehabilitation, clinical anatomy, and kinesiology. The form was composed of a yes or no scale for clarity, 4-point Likert scale for relevance, and a comment section for each item.

Step 6: Synthesis and Conciliation

All versions of the Filipino STAT available so far and the completed content validation forms (Multimedia Appendix 1) were then reviewed by a consensus committee of experts to arrive at a prefinal version. The consensus committee consisted of all 4 translators (T1, T2, B1, and B2) from the previous steps and the 6 clinical experts who participated in the content validation step. They all met virtually and collectively decided and agreed on each item through the Expert Consensus Committee Guide on Equivalence Form (Multimedia Appendix 2), wherein semantic, idiomatic, experiential, and conceptual equivalence were assessed between the forward- and backward-translated English and Filipino versions. Semantic equivalence pertains to the use of words that have similar, unique meaning in both cultures. Idiomatic equivalence pertains to the use of equivalent colloquialisms and idioms in both cultures. Conceptual equivalence pertains to the use of words or phrases with equivalent conceptual meanings in both cultures, and experiential equivalence pertains to experiences elicited that are consistent or equivalent in both cultures. The form contained a yes or no scale on each type of equivalence, a column for rewording suggestions, and another column for other comments and suggestions to meet cultural equivalence.

Phase 2: Cross-Cultural Adaptation

Step 1: Pretesting

A pretest on 12 adult Filipino patients with shoulder pain was done using the prefinal version of the translated tool through video consultation to simulate the actual STAT procedure and to avoid unnecessary risk of contagion during the pandemic (time of study).

A think-aloud protocol during each video call with the principal investigator or research assistant using an encrypted platform, such as Zoom (Zoom Communications, Inc) or Viber (whichever the participant preferred), was used. The think-aloud protocol entailed each participant to read, perform, and answer the Filipino STAT while verbalizing their thoughts. This was done twice for each question: first without the pictorial guide; and second with the pictorial guide. The trial without the pictorial guide was done to assess the feasibility of the Filipino STAT Text Version for patients with no access to smartphones. An observational checklist (Multimedia Appendix 3) was filled out by the principal investigator to document the participant response, understandability of the tool, and ease and accuracy in performing the Filipino STAT, noting participants' quality of movements and compensatory movements. The principal investigator also asked each participant open-ended interview questions (Multimedia Appendix 4) on their experience (including perceived barriers and facilitators) in using the tool and their suggestions in improving it.

Step 2: Review of Results

General and question-specific errors and suggestions from the pre-test participants were reviewed by the committee in developing the final version of the Filipino STAT.

Step 3: Proofreading

The resulting version was then proofread by a Filipino language teacher to correct any spelling and grammatical errors. All translation forms and STAT versions were sent to the original authors for their review and feedback.

Phase 3: Internal Consistency Testing

Finally, all 46 participants underwent the same process as pretesting using the final version of the Filipino STAT for the pilot study to determine the internal consistency of the tool.

Data Analysis

Content validation was determined through computation of item-level and scale-level content validity indices based on average and universal agreement methods. Sociodemographic data were reported using descriptive statistics. The internal consistency of participant responses was determined through Cronbach α. This was evaluated separately for items in VAS,

XSL•FO
RenderX

SANE, ROM, strength tests, and special tests using STATA 16.1/IC (StataCorp LLC). Considering ≥90% as excellent, 89% to 70% as good, 69% to 50% as fair, and ≤49% as poor, understandability was assessed using the worst-performing item in the observational checklist.

Interview responses were uploaded to NVivo 12 (Lumivero) for organization and thematic analysis of unstructured data pertaining to understandability. An inductive approach was employed where the data content directed the development of themes. The members of the research team were all physiatrists and had broad experience in musculoskeletal evaluation and telerehabilitation in clinical practice and research. JA was a rehabilitation medicine resident trained during the height of telerehabilitation during the COVID-19 pandemic. CFL spearheaded several research endeavors on telerehabilitation in the Philippines. SI and JM were the chairs of the Department of Rehabilitation Medicine with extensive teaching, research, and clinical experiences.

Ethical Considerations

The study was approved by the University of the Philippines Manila Research Ethics Board (UPMREB Code 2022-0516-01). Informed consent was secured through Google Forms, written and explained in Filipino either by the principal investigator or research assistant. Study introduction, objectives, procedure, duration, risks, benefits, incentives, and contact information were discussed with the participants. Participants were also assured of the confidentiality of their information, the voluntary nature of participation, and their right to refuse at any point of the study. The translators, recording observers, and Filipino teacher were appropriately compensated with Philippine 1000 each (US $20). All phase 2 participants were compensated with Philippine 200 (US $4) worth of prepaid load as reimbursement for internet fee for the video call interview and answering of online forms. The clinical experts did not receive remuneration and were acknowledged in this paper.

Results

Forward Translation

Both forward translators had challenges in translating the strength and special tests questions due to absence of direct translation of some English words to Filipino. Hence, both translators resorted to rephrasing some English sentences in Filipino. During synthesis, T1 shared the technical context of the subtests, while T2 suggested the use of understandable albeit literal translations of some words. For item Q9 (ROM) (Table 1), the use of the phrase "*itupi ang siko*" was selected over "*itupi ang braso*" based on T1's input to pertain to the correct anatomical joint that does flexion. On the other hand, for item Q12 (special test), "*idiin ang kamay sa tiyan*" was selected over "*pindutin ang tiyan*" based on T2's input on the context of pressing the belly. Items that resulted in differences between T1 and T2, as well as the final terminology agreed upon during synthesis, are presented in Table 1.

XSL•FO
RenderX

Table . Results of the forward translation and synthesis steps.

Item	T1[a]	T2[b]	Prefinal STAT[c]
Pain and activity			
Q1	*Lebel* *Pinakamasakit na naramdaman sa iyong buhay*	*Antas* *Pinakamasakit*	*Lebel* *Pinakamasakit*
Q3	*Dalas ng pisikal na aktibidad* • *Hindi apektado ang pagtulog* • *Nagtatrabaho nang buong araw* • *Nagagawa ang panlibangang aktibidad o isports*	*Antas ng pang-araw-araw na aktibidad* • *Di naaabalang pagtulog* • *Panay trabaho* • *Panay lingan/ isports*	*Antas ng pang-araw-araw na aktibidad* • *'Di naaabalang pagtulog* • *Nakakapagtrabaho nang buong araw* • *Nagagawa ang panlibangang aktibidad o isports*
Range of motion	*Saklaw ng galaw*	*Saklaw ng paggalaw*	*Saklaw ng paggalaw*
Instructions	*Kilos*	*Paggalaw*	*Paggalaw*
Q4	*Taas ng ulo*	*Tuktok ng ulo*	*Taas ng ulo*
Q7	*Bulsa ng pantalon sa parehong panig*	*Bulsa sa likod*	*Bulsa sa likod ng apektadong balikat*
Q8	*Ibabang likod*	*Ibabang bahagi ng iyong likuran*	*Ibabang bahagi ng iyong likuran*
Q9	*Braso ay nakaharap sa pader* *Itupi ang siko* *Ilapat*	*Nakatagilid sa pader* *Itupi ang braso* *Idikit*	*Nakatagilid sa pader* *Itupi ang braso* *Ilapat*
Strength			
Instructions	*Bigyan ng pwersa* *Mahina ba ang pakiramdam?* *Wala sa mga nabanggit*	*Habang nilalabanan ang bigat* *Nanghihina ba?* *Wala*	*Bigyan ng pwersa* *Mahina ba ang pakiramdam?* *Wala*
Q10	*Gamit ang kamao* *Idiin ang palad ng kamay sa masakit na braso*	*Nakakuyom ang kamao* *Idiin ang mga ito sa isa't isa*	*Gamit ang kamao* *Idiin ang palad ng kamay sa masakit na braso*
Q11	*Hilahin* *Direksyon ng kabilang braso*	*Kapitan* *Gumalaw pakanan at pakaliwa*	*Hilahin* *Direksyon ng normal na braso*
Q12	*Labanan* *Ilayo*	*Kapitan* *Palayo*	*Kapitan* *Ilayo*
Special tests			
Instructions	*Nakasaad na kilos at sasabihin mo*	*Maniobra na makakapagsabi sa amin*	*Mga kilos na makakapagsabi sa amin*
Q13	*Iangat ang buong braso*	*Itaas ang kamay*	*Itaas ang kamay*
Q14	*Umabot lagpas sa kabilang balikat*	*Ilagay ang inyong apektadong braso sa harap ng inyong dibdib*	*Umabot lagpas sa kabilang balikat*
Q15	*Pindutin*	*Idiin ang kamay*	*Idiin ang kamay*
Q16	*Ilapat*	*Idikit*	*Ilapat*
Q17	*Itaas ang braso na parang may hawak na plato* *Dahan-dahang itulak pababa*	*Iangat ang braso sa harap ng inyong dibdib at ilahad ang palad* *Bahagyang diinan*	*Itaas ang braso na parang may hawak na plato* *Dahan-dahang itulak pababa*

[a]T1: forward translator 1.

[b] T2: forward translator 2.

[c]STAT: Shoulder Telehealth Assessment Tool.

XSL•FO
RenderX

Backward Translation

Similar to the forward translators, the back translators did not apply word-for-word translation in strength and special test questions to simplify the items and make them easier to understand. The back translation versions were generally consistent with the original English version, with no significant differences, as agreed upon by the consensus committee.

Content Validation of Prefinal STAT

The instructions in all subdomains, except for strength, were clear to the experts. The items on pain and activity (Q1) and range of motion (Q4, Q5, and Q6) were clear to all experts. Content validity through the scale validity index was measured using the item-level content validity index (0.97) and using universal agreement (0.80). κ values across all items ranged from 0.82 to 1.00, indicating excellent inter-rater reliability.

Cultural Equivalence

Pertinent findings from the discussion of the consensus experts on equivalence are summarized in Table 2.

Table . Pertinent results from the expert consensus discussion on equivalence[a].

	SE[b]	IE[c]	CE[d]	EE[e]	Comments
Pain and activity					
Q1	✓	✓	✓	✓	• Inconsistent: not a question
Q2	✓	✓	✓	✓	• Inconsistent: not a question
Q3	✓	✓	✓	X	• Inconsistent: not a question • Vague: choices not in one spectrum • EE: did not include informal work
Range of motion					
Q7	✓	✓	X	✓	• CE: not relatable to everyone
Special test					
Instruction	✓	X	✓	✓	• IE: Test demands active patient response
Q13	✓	✓	✓	✓	• Safety: for severely painful shoulder • Unclear: confusing literal translation
Q17	✓	✓	X	✓	• CE: incomplete instruction for the test

[a]Checks (✓) indicate equivalence, while cross marks (X) indicate non-equivalence.

[b]SE: semantic equivalence.

[c]IE: idiomatic equivalence.

[d]CE: conceptual equivalence.

[e]EE: experiential equivalence.

To facilitate consistency of the translations throughout the tool, items Q1 to Q3 (pain and activity) were converted from declarative statements into questions. Choices for item Q3 (sleep, work, and recreation or sports) were deemed to pertain to different aspects of activities and were thus converted to stand-alone questions for each activity (Q3A, Q3B, and Q3C). To facilitate clarity, the literal translation of item Q13 was rephrased for easier understanding. To ensure patient safety when using the tool, precautions for severely painful shoulders were added to the instructions for the special test subdomain.

XSL•FO
RenderX

To facilitate experiential equivalence, item Q3B's use of the term "*nakakapagtrabaho*" (able to work) was supplemented by the modifier "*gawaing bahay*" (house chores) to be more encompassing of the different types of work in the Filipino culture. To facilitate conceptual equivalence, item Q7's (ROM) use of "*bulsa sa likod ng apektadong balikat*" (back pocket on the affected side) was changed to "*pigi*" (buttock) as it pertains to the same area and is more relatable and easily understandable. Likewise, item Q17 was clarified with the phrase "*paharap sa lebel ng braso*" (forward raising to arm level) to demonstrate shoulder flexion to 90 degrees as in the Speed test. To facilitate idiomatic equivalence, instruction for the special test subdomain was improved from "*na makakapagsabi sa amin*" (that may tell us) to "*upang malaman namin*" (so that we will know) as the test demands active patient response. Finally, the expert committee all suggested taking new pictures for the tool with a Filipino-looking model for the pictorial guide for it to be more relatable and also to facilitate minor improvements in the angle of the shots and designation of the task of each arm.

Pretest

There were 12 participants in the pretest, with the majority being females (10/12, 83%), aged 50 - 59 years (5/12, 42%) with a mean age of 53 (SD 12) years, married or cohabiting (5/12, 42%), and having finished tertiary education (8/12, 67%; Table 3). Most were unemployed (6/12, 50%), while those employed were engaged in nonhealth-related work (8/12, 67%). The monthly family income of participants was positively skewed, with the majority (7/12, 58%) earning Philippine 5000 to 10,000 (US $100 - $200).

XSL•FO
RenderX

Table . Sociodemographic profile of the participants.

Characteristics	Pretest (n=12)	Pilot (n=47)
Age group (y), n (%)		
19 - 29	0 (0)	2 (4)
30 - 39	2 (2)	5 (11)
40 - 49	2 (2)	3 (6)
50 - 59	5 (42)	21 (45)
≥60	3 (25)	16 (34)
Age (y), mean (SD)	53 (12)	55 (14)
Sex, n (%)		
Female	10 (83)	35 (75)
Male	2 (17)	12 (26)
Civil status, n (%)		
Single	4 (33)	10 (21)
Married or cohabiting	5 (42)	29 (62)
Separated or divorced	1 (8)	2 (4)
Widowed	2 (17)	6 (13)
Educational status, n (%)		
Primary	0 (0)	4 (9)
Secondary	4 (33)	15 (32)
Tertiary	8 (67)	26 (55)
Postgraduate	0 (0)	2 (4)
Employment, n (%)		
Student	0 (0)	2 (4)
Employed	4 (33)	15 (32)
Unemployed	6 (50)	23 (49)
Retired	2 (17)	7 (15)
Type of work, n (%)		
Health-related work	4 (33)	7 (15)
Nonhealth-related work	8 (67)	40 (85)
Family monthly income (Philippine [US $]), n (%)		
<5000 (<100)	2 (17)	21 (45)
5000-10,000 (100 - 200)	7 (58)	9 (19)
10,001-20,000 (201 - 400)	2 (18)	8 (17)
20,001-50,000 (401 - 1000)	1 (8)	6 (13)
>50,000 (>1000)	0 (0)	3 (6)

Items on pain and activity (Q1 and Q3A-Q3C), ROM (Q2-Q6), and special tests (Q15) were performed correctly by at least 75% (9/12) of the participants independently. Up to 25% (3/12) of the participants needed cueing in the whole pain and activity subdomain and special tests (Q14). Less than 25% of the participants were able to perform ROM items Q8 and Q9, all strength items, and half of the items in the special test subdomain.

Six out of 12 (50%) participants reported an increase in pain upon performance of certain maneuvers (eg, items Q4, Q5, Q7, Q9, and Q14), although all were still able to complete the tasks. They were advised to seek outpatient consultation at the Rehabilitation Medicine Outpatient Clinic if the pain persisted.

Six (50%) participants found the tool easy to follow, while 5 (42%) found it difficult. Visual aids in the form of pictures and arrows were most helpful in making the tool easily

XSL•FO
RenderX

understandable. The presence of a caregiver was helpful for two of the participants.

The use of some Tagalog words (such as *pigi* or buttocks in English) was the primary barrier for understanding the tool (Table 4). Five (42%) participants were content with the prefinal version of the tool and had no suggested changes. One (8%) participant suggested that the presence of a physician or a caregiver (at the least) was still necessary to guide patients and ensure accuracy and safety in following the tool. Finally, 6 (50%) participants reiterated the need for pictures, and 1 (8%) participant suggested improvement in the portrayal of movement in the pictures.

Table . Results from the thematic analysis related to understandability of the Filipino Shoulder Telehealth Assessment Tool (STAT) and factors that make it easy or difficult to understand.

	Pretest (n=12), n (%)	Pilot (n=46), n (%)
Ease of understanding		
Tool is easy	6 (50)	13 (28)
Use of pictures and arrows	10 (83)	11 (24)
Presence of caregiver	2 (17)	3 (6)
Tool is difficult	5 (42)	11 (24)
Use of some Tagalog words	3 (25)	7 (15)
Nonfamiliarity with the tool	1 (8)	3 (6)
Difficulty reading the text	2 (17)	3 (6)
Poor internet connection	2 (17)	2 (4)
Poor audiovisual setup	1 (8)	1 (2)
Not adept with technology	1 (8)	1 (2)
Absence of caregiver	0 (0)	1 (2)
Nonideal venue	1 (8)	1 (2)
Use of arrows	1 (8)	0 (0)

Lessons From the Pretest and Development of the Final Filipino STAT

Difficulty reading the text of the tool and tendency to skip a specific question (Q3A) in a predominantly older adult population suggested inappropriate user interface design to the study team. While readers with advancing age have varying levels of possible age-related cognitive and visual decline, inclusivity was ensured in redesigning the tool [18]. These lessons were applied to the final version of the Filipino STAT (Multimedia Appendix 5) as follows: breaking down chunks of texts into bullet points, using Sans Serif typefaces with at least a 16-point font size, and providing white spaces and appropriate contrasts [18,19].

The conduct of special tests revealed an average of 73% (SD 14%) positive test findings (painful) that might have contributed to the increase in shoulder pain among pretest participants. This was addressed in the final version by placing the special test subdomain last as in clinical examination so any pain would not get in the way of the conduct of the other parts of the assessment.

Specific pretest errors could be classified as (1) misinterpretations observed in the subdomains of pain and activity (items Q2 and Q3C) and ROM (Q8); (2) differences in the semantic understanding of specific body parts, evident in ROM items Q1, Q4, Q5, Q7, and Q8; (3) difficulty understanding long instructions, which were evident in ROM item Q9 and all strength and special test items; and (4) inconsistent experiential equivalence in special test item Q17. Text qualifiers were added to misinterpreted items such as Q2, labeling 0=“normal” and 100=“hindi normal” (not normal), to avoid reversal of the scale. Items with different potential semantic meanings were significantly improved when pictorial guides were provided. Long, unclear instructions were broken down into short, step-by-step instructions with pictorial guides in each step. Finally, in item Q17 (Speed test), instead of holding a plate forward, most participants held a plate on the side, as a waiter would. The phrase was therefore changed from “*parang may hawak na plato*” (like holding a plate) to “*parang nanghihingi habang nakaunat ang siko*” (like reaching out or asking for something with an outstretched arm) to improve participants’ accuracy in performing the task and the relatability of the task in Filipino culture.

Final Filipino STAT Pilot Study

There were 47 participants in the pilot study; one of them did not complete the video consultation and withdrew due to technical difficulties (ie, unstable internet) on the patient’s end. The majority of the participants were females (35/47, 75%), aged 50 - 59 years (21/47, 45%) with a mean age of 55 (SD 14) years, married or cohabiting (29/47, 62%), and finished tertiary education (26/47, 55%; Table 3). Most were unemployed (23/47, 49%), while those employed were mostly engaged in nonhealth-related work (40/47, 85%). The monthly family income of the participants was positively skewed, with the majority (21/47, 45%) earning less than Philippine 5000 (US $<100).

XSL•FO
RenderX

For each of the items under the pain and activity subdomains, 35 (75%) to 43 (91%) participants were able to perform the tasks correctly. Meanwhile, some needed cueing from their caregiver to perform the tasks correctly, particularly for item Q2, where 11 (23%) participants needed help. With items Q3A and Q3C, 1 out of the 47 (2%) participants was not able to perform the task completely.

The ROM, strength, and special test subdomains were tested both without and with pictorial guides. For ROM-related items, a lot of participants were not able to perform the tasks correctly without a picture (Q2: chin, Q7: back pocket, Q8: lower back, and Q9: wall touching). In contrast, for those items with picture guides, there was a greater number of patients who were able to perform the tasks correctly, with little to no need of cueing from their caregivers. Four (8%) participants showed compensatory movements with some items, yet they answered yes when asked if they could do the tasks. Hence, they were considered incorrect task performances, with item Q8 being the most common item that resulted in compensatory movements.

Without pictorial guides, more than half of the participants were not able to perform all the tasks under the strength subdomain. Slightly more participants correctly performed the tasks for items Q11 and Q12 when shown pictorial guides. Moreover, the same scenario was observed for all items in the special test subdomain—having pictorial guides improved the number of participants able to perform the tasks correctly from 34 (74%) to 39 (87%) participants, while 2 (4%) to 6 (13%) participants needed cueing from their caregivers to be able to perform the tasks correctly.

Among the 46 participants who shared their experience with the final Filipino STAT, 13 (28%) found the tool easy to understand, while 11 (24%) found it difficult. The pictures and arrows were found to be most helpful for participants. Participant 47 remarked, "*maayos po binigay ang panuto at dahil sa larawan ay naintindihan po nang mabuti*" (the instructions were given properly, and because of the pictures, they were better understood). The use of some Tagalog words, however, was considered the primary barrier by the majority of the participants, followed by their nonfamiliarity with the tool and difficulty reading the text (Table 4). The same participant also remarked, "*May kahinaan po ako sa Pilipino, tulad po ng pigi po, hindi ko siya naintindihan*" (I have some difficulty in understanding Tagalog, for example, I did not understand the Tagalog word for buttock").

Content Validity and Internal Consistency Testing of the Final Filipino STAT

Content validity through scale validity indices was deemed excellent at 0.97 using the item-level content validity index and 0.80 using the universal agreement. The κ coefficients were excellent across all items, supporting that the degree of agreement among the experts was beyond chance [20]. A Cronbach α score of 0.87 indicated that all test items were unidimensional, or that performance on the items could be explained in terms of a single underlying factor (Multimedia Appendix 6).

Discussion

Principal Findings

The Filipino STAT has an excellent content validity (scale validity index=0.80 - 0.97), excellent interrater reliability (κ coefficient=0.82 - 1.00), and good internal consistency (Cronbach α=0.87). Its understandability is excellent for pain and activity (98%), good for ROM and special tests (85%), and poor for strength (37%).

Comparison With Prior Work

The findings of this study are consistent with the existing literature on shoulder teleconsultation psychometric properties. Internet assessment of rehabilitation outcomes such as pain, ROM, muscle strength, and functional assessment had good concurrent validity. There is a strong agreement between virtual and in-person examination for ROM (87.4% agreement; x^2=30.782; P<.001), and diagnosis (85.1% agreement; κ=0.82, 95% CI) [21-23]. Further studies on the Filipino STAT compared against in-person examination or imaging modalities may be done to ascertain concurrent validity of the tool.

Strengths and Limitations

Since a direct translation of a tool from an original English version may not reflect the cultural nuances and context of another culture [13], this study observed a rigorous and standard linguistic validation and cross-cultural adaptation to ensure a culturally appropriate, valid, and reliable translation. Although the original STAT has not been validated as a whole tool, the Filipino-translated tool has been subjected to content validity and internal consistency testing in this study. The collective inputs from the consensus committee of experts, the understandability questionnaire, and qualitative data from the participants ensured that the final tool can be used in the clinical setting.

There are some limitations to this study. First, the assessment of ease of understanding of the tool using the open-ended questionnaire may have been confounded by the consecutive performance of the tool without and with a pictorial guide, as well as the videoconsultation study setup. The test-taking that was twice as long, with no visual aid provided at the start, could have made the consultation more difficult for the participants. The barriers to the videoconsultation study setup were consistent with previous literature, such as unstable internet connection and poor audiovisual resolution [5]. In the envisioned clinical performance of the Filipino STAT, only the pictorial guide will be performed before the actual consultation, thereby eliminating these concerns.

Second, understandability per subdomain was excellent for pain and activity, good for ROM and special tests, and poor for strength. The incidence of compensatory or trick movement was accounted for in this study and was most frequently seen in the ROM subdomain, particularly item Q8 (50 degrees of internal rotation).

Finally, such as in an in-person physical examination, incorrect performance of a virtual test, from poor understandability or compensatory movement, makes its results invalid and

XSL•FO
RenderX

questionable. This shall serve as a reminder that the Filipino STAT is not meant to replace actual in-person consultations and should just aid the clinician in assessing the patient and make clinical work efficient. The clinician must be prudent in confirming the initial Filipino STAT findings during the actual consultation, as necessary. Further studies may consider the use of instructional videos to improve understandability of the tool, especially for the strength subdomain.

Future Directions

Both pretest and pilot study findings reveal significant improvements in the correct performance of the Filipino STAT tasks with pictorial guides, compared to none. The majority of the participants from both phases of the study also shared a positive perception of the use of pictures and arrows and the presence of a caregiver. Thus, it is the study's recommendation to use the Filipino STAT with a pictorial guide, as intended in the original tool [6]. The association of the presence of a caregiver in the successful performance of the Filipino STAT was beyond the scope of this present study. Nonetheless, having a caregiver around may help with the conduct of the Filipino STAT but should not discourage those who do not have an available caregiver.

Conclusions

The development of the Filipino STAT through a rigorous linguistic validation and cultural adaptation ensured a culturally appropriate, valid, and reliable translation. Understandability and ease of understanding by end-users are also critical to assess in patient-reported outcome measures. The pain and activity, ROM, and special test subdomains of the Filipino STAT may be used for clinical assessment, while the strength subdomain needs further improvement on understandability.

Acknowledgments

The authors would like to acknowledge the clinical experts Dorothy Dy Ching Bing-Agsaoay, Monalisa Lim-Dungca, Christopher Constantino, Nathaniel Napa, Mitchelle Gabuya, and Frances Carlos, as well as our Filipino STAT model Nancy Cheng, for their generous contributions to this study. They would also like to acknowledge the STAT original authors, Gregory Sprowls, Jaycen Brown, and Brett Robin, for allowing translation of their tool.

Funding

This research study was funded by the Philippine General Hospital—Expanded Hospital Research Office.

Data Availability

The data that support the findings of this study are uploaded online as supplementary files and may be requested from the corresponding author, JA, upon reasonable request.

Authors' Contributions

Conceptualization: JA, SI, JAM, CFL
Data curation: JA
Funding acquisition: SI
Implementation: JA
Methodology: JA, SI, JAM, CFL
Supervision: SI, JAM, CFL
Writing – original draft: JA
Writing – reviewing & editing: SI, JAM, CFL

Conflicts of Interest

None declared.

Multimedia Appendix 1
The content validation form.
[PDF File, 2546 KB - rehab_v13i1e67974_app1.pdf]

Multimedia Appendix 2
The expert consensus committee guide on equivalence.
[PDF File, 114 KB - rehab_v13i1e67974_app2.pdf]

Multimedia Appendix 3
The observational checklist.
[PDF File, 50 KB - rehab_v13i1e67974_app3.pdf]

XSL•FO
RenderX

Multimedia Appendix 4
The open-ended interview questions.
[PDF File, 48 KB - rehab_v13i1e67974_app4.pdf]

Multimedia Appendix 5
The final Filipino Shoulder Telehealth Assessment Tool (STAT).
[PDF File, 9454 KB - rehab_v13i1e67974_app5.pdf]

Multimedia Appendix 6
Supplemental dataset to a full manuscript published in the *J Med Internet Res*.
[DOCX File, 29 KB - rehab_v13i1e67974_app6.docx]

References

1. Bos WH, van Tubergen A, Vonkeman HE. Telemedicine for patients with rheumatic and musculoskeletal diseases during the COVID-19 pandemic; a positive experience in the Netherlands. Rheumatol Int 2021 Mar;41(3):565-573. [doi: 10.1007/s00296-020-04771-6] [Medline: 33449162]
2. D Leochico CF. Adoption of telerehabilitation in a developing country before and during the COVID-19 pandemic. Ann Phys Rehabil Med 2020 Nov;63(6):563-564. [doi: 10.1016/j.rehab.2020.06.001] [Medline: 32544528]
3. Leochico CFD, Mojica JAP, Rey-Matias RR, Supnet IE, Ignacio SD. Role of telerehabilitation in the rehabilitation medicine training program of a COVID-19 referral center in a developing country. Am J Phys Med Rehabil 2021 Jun;100(6):526-532. [doi: 10.1097/PHM.0000000000001755] [Medline: 33998606]
4. Leochico CFD, Montemayor GAM, Obeles AJT, Ong BAG. Telerehabilitation in a developing country toward the tail end of the COVID-19 pandemic: is it here to stay? Acta Med Philipp 2022 Nov. [doi: 10.47895/amp.vi0.5905]
5. Leochico CFD, Espiritu AI, Ignacio SD, Mojica JAP. Challenges to the emergence of telerehabilitation in a developing country: a systematic review. Front Neurol 2020;11:1007. [doi: 10.3389/fneur.2020.01007] [Medline: 33013666]
6. Sprowls GR, Brown JC, Robin BN. The Shoulder Telehealth Assessment Tool in transition to distance orthopedics. Arthrosc Tech 2020 Nov;9(11):e1673-e1681. [doi: 10.1016/j.eats.2020.07.008] [Medline: 33294325]
7. McLean JM, Awwad D, Lisle R, Besanko J, Shivakkumar D, Leith J. An international, multicenter cohort study comparing 6 shoulder clinical scores in an asymptomatic population. J Shoulder Elbow Surg 2018 Feb;27(2):306-314. [doi: 10.1016/j.jse.2017.08.016] [Medline: 29221757]
8. Reed MD, Van Nostran W. Assessing pain intensity with the visual analog scale: a plea for uniformity. J Clin Pharmacol 2014 Mar;54(3):241-244. [doi: 10.1002/jcph.250] [Medline: 24374753]
9. Retzky JS, Baker M, Hannan CV, Srikumaran U. Single Assessment Numeric Evaluation scores correlate positively with American Shoulder and Elbow Surgeons scores postoperatively in patients undergoing rotator cuff repair. J Shoulder Elbow Surg 2020 Jan;29(1):146-149. [doi: 10.1016/j.jse.2019.05.039] [Medline: 31401127]
10. Oosterwijk AM, Nieuwenhuis MK, van der Schans CP, Mouton LJ. Shoulder and elbow range of motion for the performance of activities of daily living: a systematic review. Physiother Theory Pract 2018 Jul;34(7):505-528. [doi: 10.1080/09593985.2017.1422206]
11. Capio CM, Maranion DZ, Pascua HE. Musculoskeletal pain and dysfunction among University of the Philippines Manila personnel. Univ Philippines Manila J 2001;6(4):25-23 [FREE Full text]
12. Labao HC, Faller EM, Bacayo MFD. “Aches and pains” of Filipino migrant workers in Malaysia: a profile of work-related musculoskeletal disorders. Ann Glob Health 2018 Aug;84(3):474-480. [doi: 10.29024/aogh.2331] [Medline: 30835403]
13. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993 Dec;46(12):1417-1432. [doi: 10.1016/0895-4356(93)90142-n] [Medline: 8263569]
14. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
15. Bonett DG. Sample size requirements for testing and estimating coefficient alpha. J Educ Behav Stat 2002 Dec;27(4):335-340. [doi: 10.3102/10769986027004335]
16. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. SPINE (Phila Pa 1976) 2000 Dec;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
17. Lorca LA, Torres-Castro R, Ribeiro IL, et al. Linguistic validation and cross-cultural adaptation of the post–COVID-19 functional status scale for the Chilean population. Am J Phys Med Rehabil 2021 Apr 1;100(4):313-320. [doi: 10.1097/PHM.0000000000001706] [Medline: 33496442]
18. Office of Disease Prevention and Health Promotion. Section 5.3 use a readable font that’s at least 16 pixels. Health Literacy Online, 3rd Edition. 2025. URL: https://odphp.health.gov/healthliteracyonline/design-easy-scanning/use-readable-font-thats-least-16-pixels [accessed 2025-12-01]

XSL•FO
RenderX

19. Hou G, Anicetus U, He J. How to design font size for older adults: a systematic literature review with a mobile device. Front Psychol 2022;13:931646. [doi: 10.3389/fpsyg.2022.931646] [Medline: 35978796]
20. Rodrigues IB, Adachi JD, Beattie KA, MacDermid JC. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. BMC Musculoskelet Disord 2017 Dec;18(1):540. [doi: 10.1186/s12891-017-1914-5] [Medline: 29258503]
21. Steele L, Lade H, McKenzie S, Russell TG. Assessment and diagnosis of musculoskeletal shoulder disorders over the internet. Int J Telemed Appl 2012;2012:945745. [doi: 10.1155/2012/945745] [Medline: 23193395]
22. Rabin A, Dolkart O, Kazum E, et al. Shoulder assessment by smartphone: a valid alternative for times of social distancing. Arch Orthop Trauma Surg 2022 Jun;142(6):979-985. [doi: 10.1007/s00402-021-03762-x] [Medline: 33439302]
23. Mani S, Sharma S, Omar B, Paungmali A, Joseph L. Validity and reliability of internet-based physiotherapy assessment for musculoskeletal disorders: a systematic review. J Telemed Telecare 2017 Apr;23(3):379-391. [doi: 10.1177/1357633X16642369] [Medline: 27036879]

Abbreviations

ROM: range of motion
SANE: single assessment numeric evaluation
STAT: Shoulder Telehealth Assessment Tool
VAS: visual analog scale

Edited by N Cahill; submitted 25.Oct.2024; peer-reviewed by T Lienou, V Prieto; revised version received 04.Nov.2025; accepted 11.Nov.2025; published 20.Jan.2026.

Please cite as:
Arboleda J, Ignacio S, Mojica JA, Leochico CF
Linguistic Validation and Cross-Cultural Adaptation of the Shoulder Telehealth Assessment Tool for Filipino Patients with Musculoskeletal Shoulder Condition: Cross-Sectional Study
JMIR Rehabil Assist Technol 2026;13:e67974
URL: https://rehab.jmir.org/2026/1/e67974
doi:10.2196/67974

© Jeffrey Arboleda, Sharon Ignacio, Jose Alvin Mojica, Carl Froilan Leochico. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Effects of Telerehabilitation on Quality of Life, Physical Capacity, and Symptom Burden Within the Post–COVID-19 Condition: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Xiaojing Huang[1], MS; Yidan Tang[2], MS; Dongsen Liu[3], PhD; Ryan Donald Burns[1], PhD; Wonwoo Byun[1], PhD

[1]Department of Health & Kinesiology, College of Health, University of Utah, Salt Lake City, UT, United States

[2]Yunnan Provincial Institute of Sports Science, Kunming, Yunnan, China

[3]Department of Sports Rehabilitation, School of Sports Medicine and Rehabilitation, Beijing Sport University, Beijing, China

Corresponding Author:
Wonwoo Byun, PhD
Department of Health & Kinesiology
College of Health
University of Utah
250 S 1850 E
Salt Lake City, UT, 84112
United States
Phone: 1 803 543 8132
Email: won.byun@utah.edu

Abstract

Background: Post–COVID-19 condition (PCC) is characterized by persistent symptoms and functional limitations that impair quality of life, and rehabilitation is an important component of its management. Because access to conventional face-to-face rehabilitation may be limited, telerehabilitation has emerged as an accessible alternative. However, the effectiveness of telerehabilitation across clinical outcomes remains uncertain, and the influence of different delivery modes is not well understood.

Objective: This study aimed to evaluate the effects of telerehabilitation on health-related quality of life (HRQoL), physical capacity, and symptom burden in adults with PCC, and to determine whether delivery mode (synchronous, asynchronous, or hybrid) modified these effects.

Methods: Seven databases (PubMed, Cochrane Library, Web of Science, Scopus, Embase, EBSCO, and PEDro) were searched for randomized controlled trials published from January 2020 through December 2025. Eligible trials enrolled adults with PCC and compared telerehabilitation with usual care or another control condition. The primary outcome was HRQoL; secondary outcomes included functional capacity assessed by the 6-minute walk test and sit-to-stand test, dyspnea, fatigue, and symptoms of anxiety and depression. Risk of bias was assessed using the Cochrane Risk of Bias 2 tool. Mean differences or standardized mean differences (SMDs) with 95% CIs were pooled using random-effects models with the Hartung-Knapp-Sidik-Jonkman adjustment. Heterogeneity was quantified using I^2 and τ^2, with 95% prediction intervals reported where estimable. Certainty of evidence was evaluated using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach.

Results: A total of 22 randomized controlled trials involving 2294 participants were included. Telerehabilitation was associated with significant improvements in HRQoL (SMD 1.26, 95% CI 0.07-2.45; *P*=.04), dyspnea (SMD 1.95, 95% CI 0.60-3.31; *P*=.005), and functional aerobic capacity (6-minute walk test : mean difference 77.79 meters, 95% CI 30.44-125.14; *P*=.001). Fatigue was also significantly reduced (SMD 0.89, 95% CI 0.16-1.62; *P*=.02), but no significant effects were observed for lower-limb strength (sit-to-stand test: SMD 0.55, 95% CI –0.15 to 1.25; *P*=.12) or symptoms of anxiety and depression (SMD 0.12, 95% CI –0.13 to 0.38; *P*=.36). No significant differences between delivery modes were observed for HRQoL (*P*=.93), functional aerobic capacity (*P*=.29), or dyspnea (*P*=.16). A significant subgroup difference was observed for lower-limb strength (*P*=.02).

Conclusions: Telerehabilitation may improve quality of life, aerobic capacity, and symptom burden in PCC and offer an accessible option for post-COVID rehabilitation, although the certainty of evidence was generally limited. Asynchronous delivery appears feasible for general reconditioning, whereas its benefit for lower-limb strength and the value of synchronous supervision remain uncertain.

XSL•FO
RenderX

Trial Registration: PROSPERO CRD42023490863; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023490863

(JMIR Rehabil Assist Technol 2026;13:e95368) doi:10.2196/95368

KEYWORDS

telerehabilitation; post–COVID-19 condition; health-related quality of life; digital rehabilitation; meta-analysis

Introduction

Post–COVID-19 condition (PCC), also referred to as long COVID, is a common long-term consequence of SARS-CoV-2 infection. The World Health Organization defines it as symptoms occurring about 3 months after probable or confirmed COVID-19. These symptoms last for at least 2 months and cannot be explained by another diagnosis [1,2]. Common manifestations include fatigue, dyspnea, postexertional symptoms, reduced exercise capacity, cognitive difficulties, and psychological symptoms. These problems often coexist and can lead to sustained limitations in daily functioning, participation, and health-related quality of life (HRQoL) [3-6]. Although the acute phase of the pandemic has passed, PCC remains common, and its prolonged course continues to create substantial health care needs for patients and health systems [2,7,8].

Physical rehabilitation, particularly pulmonary and exercise training, is widely recognized as a central component of PCC management [9,10]. Studies have demonstrated that structured rehabilitation can significantly improve exercise capacity and alleviate dyspnea in convalescent patients [11-14]. However, access to conventional face-to-face rehabilitation remains limited due to health care resource constraints, geographic barriers, and workforce shortages [15]. These challenges underscore the need for alternative rehabilitation delivery models that are both effective and scalable within existing health care systems.

Telerehabilitation, defined as the delivery of rehabilitation services through information and communication technologies, has emerged as a promising approach to address limited access to conventional rehabilitation in individuals with PCC [16,17]. By reducing travel requirements and allowing treatment to be completed at home, telerehabilitation may support continuity of care, scheduling flexibility, and access for patients who are unable to attend regular facility-based programs [18,19]. However, remote delivery also introduces important challenges. Participation and treatment response may be influenced by access to technology, digital literacy, platform usability, patient engagement, and the availability of timely clinical feedback. Existing digital health studies in PCC have also reported considerable variation in intervention design and limited reporting of implementation and safety outcomes [20,21].

Recent studies generally support the beneficial effects of telerehabilitation on physical function and symptom management in populations with PCC [22-24]. However, interpretation of these findings remains limited by differences in patient characteristics, intervention components, outcome instruments, and follow-up duration. Earlier reviews were also based largely on small or early trials and frequently combined different forms of remote rehabilitation within a single treatment category [23,25-27]. A recent updated meta-analysis expanded the evidence base to 28 randomized controlled trials (RCTs) and reported improvements in physical function and several symptom outcomes, but its primary quantitative synthesis remained focused on physical and symptom-based outcomes [28]. Consequently, uncertainty remains regarding the overall effect of telerehabilitation on HRQoL and the extent to which treatment effects vary across different digital delivery modes.

A further gap concerns the assessment of patient-centered outcomes. Because patients with PCC experience multiple and fluctuating symptoms, the overall impact of the condition cannot be adequately represented by a single objective test. HRQoL captures how symptoms and functional limitations affect patients' daily lives and was therefore selected as the primary outcome of this review. This emphasis is consistent with rehabilitation frameworks that prioritize functioning and participation [29,30] and with PCC core outcome initiatives calling for standardized, patient-centered assessment [31,32]. Objective measures of aerobic capacity and lower-limb function, including the 6-minute walk test (6MWT) [33] and sit-to-stand test (STS) [34,35], were included as complementary secondary outcomes.

The mode of telerehabilitation delivery represents another important but unresolved issue. Synchronous programs provide real-time supervision, immediate feedback, and opportunities to adjust treatment during a session, but they require greater clinician time and scheduling coordination. Asynchronous programs offer greater flexibility and may reduce resource demands, but they provide less immediate feedback and rely more heavily on patient self-management. Hybrid programs combine elements of both approaches, although their added value has not been clearly established [35,36]. Existing reviews have often pooled these delivery modes or examined telerehabilitation as a single category, making it difficult to determine whether the amount and timing of supervision influence treatment effects. Evidence comparing delivery approaches remains limited, and previous analyses have not consistently examined synchronous, asynchronous, and hybrid interventions across HRQoL, functional, and symptom outcomes [37,38].

This systematic review and meta-analysis evaluated the effects of telerehabilitation on HRQoL as the primary outcome and on functional capacity, dyspnea, fatigue, and symptoms of anxiety and depression as secondary outcomes in adults with PCC, compared with usual care or another control. We further examined whether intervention effects differed across synchronous, asynchronous, and hybrid delivery modes. By synthesizing HRQoL across validated instruments and assessing delivery mode within a single analytical framework, this review provides an updated, patient-centered synthesis of the available randomized evidence.

XSL•FO
RenderX

Methods

Design and Search Strategy

The research adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting guideline and was registered in PROSPERO (CRD42023490863) (Multimedia Appendix 1) [39]. A comprehensive electronic literature search was conducted from January 2020 to December 2025 in several databases including PubMed, the Cochrane Library, Web of Science, PEDro, Scopus, Embase, and EBSCO. Search terms and synonyms included adults with PCC or long COVID, along with Telerehabilitation, or Remote Rehabilitation, or Digital Rehabilitation, tailored to each search engine (Multimedia Appendix 2). We manually reviewed the reference lists of included studies and pertinent review to find additional papers not included. Study eligibility criteria were defined according to the PICOS (Population, Intervention, Comparison, Outcome, and Study design) framework (Table 1) [40].

Table 1. PICOS (Population, Intervention, Comparison, Outcome, and Study design) model.

Domain	Description
Population	Adults with post–COVID-19 condition or long COVID.
Intervention	Telerehabilitation interventions delivered remotely through digital or telecommunication technologies, including programs conducted in home-based or nonclinical settings.
Comparison	Contrast with usual care, no intervention group or placebo.
Outcomes	Reported data on health-related quality of life, physical capacity, symptom burden (fatigue and dyspnea), or psychological outcomes assessed using validated scales.
Study design	Randomized controlled design.

Study Selection and Data Extraction

We examined only RCTs with full text that compared telerehabilitation therapies with usual care, placebo, or control in adults (older than 18 years) with PCC. Telerehabilitation was defined as rehabilitation delivered remotely through digital or telecommunication technologies, allowing patients to perform rehabilitation in home-based or nonclinical settings, with guidance, feedback, monitoring, or education provided by health professionals or structured digital systems and without in-person contact [16]. Home-based programs that lacked ongoing remote supervision or interactive support were not classified as telerehabilitation. Based on the mode of delivery, interventions were classified into 3 categories: synchronous, delivered in real time with live, therapist-supervised interaction (eg, videoconferencing or live online sessions); asynchronous, delivered without real-time interaction through self-paced tools (eg, mobile apps, web platforms, prerecorded modules, or remote monitoring); and hybrid, combining real-time and self-paced components within the same program [41]. The delivery mode of each included study is shown in Table 2. We did not impose any limitations regarding comorbidities or medication usage during the rehabilitation procedure. Excluded were studies without sufficient statistical data for extraction, reviews, case reports, observational studies, and any other type of publications (eg, guidelines, protocols, comments, conference abstracts, and editorials). Two reviewers (XH and YT) conducted independent searches of papers in each database. Disagreements were resolved by consensus or by involving a third reviewer. We collected information on authors, publication year, participants, intervention details, and outcome data.

XSL•FO
RenderX

Table 2. Characteristics of included telerehabilitation interventions.

Study	Telerehabilitation intervention	Telerehabilitation protocol	Delivery mode	Outcome	Follow-up
Alsharidah et al [42], 2023	Supervised pulmonary telerehabilitation including breathing exercises, aerobic training, and resistance training delivered via live videoconferencing	6 weeks; 3 sessions per week, 60-80 minutes per session, intensity guided by ACSM[a] recommendations	Synchronous	6-minute walk test, pulmonary function, and health-related quality of life	Week 6
Bileviciute-Ljungar et al [43,44], 2024-2025	Group-based multidisciplinary telerehabilitation including physiotherapy, exercise therapy, psychoeducation (pain, sleep, and cardiopulmonary symptoms), and psychological and occupational therapy delivered online	8 weeks; group sessions 2 hours per day, 3 days per week; weekly 20-minute individual sessions; additional self-exercise ≥3 hours per week using ExorLive app	Hybrid	Health-related quality of life (SF-36[b]), pain intensity, pain-related ICF[c] categories, anxiety and depression	6 months
Calvo-Paniagua et al [45], 2024	A multicomponent program delivered via videoconferencing, including sanitary and health education, respiratory therapy and breathing exercises, aerobic exercise and physical conditioning, active mobilization, motor control, and balance exercises, and occupational therapy focused on activities of daily living	7 weeks, 18 sessions on alternate days, 40 minutes per session	Synchronous	Modified Borg Scale, mMRC[d] scale, SGRQ[e], 6MWT[f]	1 month and 3 months
Del Corral et al [46], 2023	Home-based respiratory muscle training using a threshold pressure device and remote supervision	8 weeks, 40 minutes per day, 6 times per week	Synchronous	EQ-5D-5L, Ruffier test, 1-minute STS[g], Anxiety, Depression	NR[h]
Demir et al [47], 2026	A comprehensive exercise-based telerehabilitation program including breathing and respiratory control exercises, inspiratory muscle training using incentive spirometry, upper- and lower-limb strengthening with elastic resistance bands, balance training, and walking exercises	6 weeks, 3 sessions per week, 60 minutes per session, progressive exercise intensity over time	Synchronous and asynchronous	6MWT, mMRC, Chalder Fatigue Scale	Week 12
Elyazed et al [48], 2024	Home-based pulmonary rehabilitation including walking training, diaphragmatic breathing exercises, and resistance exercises for upper and lower limbs	12 weeks; walking 30-60 minutes, 5 days per week; breathing exercises 10-15 minutes twice daily; low-load resistance training	Asynchronous	SF-36, 6MWT, Fatigue	Week 12
de Araújo Furtado et al [49], 2023	Multicomponent functional telerehabilitation training including aerobic, strength, agility, balance, speed, and power exercises delivered remotely with real-time supervision	8 weeks; 3 sessions per week; 45-60 min per session; Borg RPE[i] 6-8; progressive multicomponent training + remote health education lectures	Synchronous	6MWT, 30-second STS, flex-arm, get up and go); handgrip strength; depression (Beck Depression Inventory); anxiety (BAI[j])	Week 8
King et al [50], 2023	Walking training, upper- and lower-limb strengthening and combination aerobic movements	10 weeks, twice weekly, at least 16 sessions	Synchronous	SGRQ, 6MWT, 5STS, fatigue	NR

XSL•FO
RenderX

Study	Telerehabilitation intervention	Telerehabilitation protocol	Delivery mode	Outcome	Follow-up
Kerling et al [51], 2024	Online-supported individualized physical activity program including aerobic endurance, strength, recovery, breathing, and flexibility exercises with wearable-based monitoring	3 months; target ≥150 minutes per week moderate-intensity activity (60-75% HR_{max}[k]); individualized plan; weekly remote consultations (weeks 1, 4, and 8) via phone or video; Garmin activity tracker support	Asynchronous	Exercise capacity ($V\dot{O}_{2peak}$, peak power output); fatigue (FAS[l]); health-related quality of life (SF-36); anxiety and depression (HADS[m]); work ability; physical activity	3 months
León-Herrera et al [52], 2024	Multimodal online rehabilitation including physical activity, respiratory exercises, cognitive rehabilitation, diet, sleep hygiene, emotional management, and community resources	8 weeks; 1 session per week; 90 minutes per session; group videoconference via Google Meet + Moodle platform (educational materials, exercises, and forums)	Hybrid	Quality of life (SF-36 physical and mental), physical function (sit-to-stand), cognitive function (MoCA[n]), anxiety or depression (HADS), insomnia (ISI[o]), self-efficacy, health literacy, patient activation	3 months
Longobardi et al [53], 2023	Individualized home-based exercise training including aerobic, resistance, and flexibility exercises with partial live online supervision	16 weeks; 3 sessions per week; 60-80 minutes per session; intensity progressed by RPE	Hybrid	HRQoL[p] (SF-36), functional capacity (STS), cardiopulmonary fitness, pulmonary function, persistent symptoms	NR
Li et al [54], 2022	Breathing control, thoracic expansion, aerobic exercise, and LMS[q] exercises remotely monitoring	6 weeks; 3-4 sessions per week, 3 levels: 30%-60% HR[r]	Asynchronous	Quality of life, 6MWT, dyspnea, FVC[s]	NR
McNarry et al [55], 2022	Tele-education, using a handheld inspiratory flow resistive device, remote adherence monitoring	8 weeks; 3 sessions per week, 20 minutes each session, >80% SMIP[t], 6 set × 6 rep	Asynchronous	Quality of life, dyspnea index, maximal inspiratory pressure, $V\dot{O}_{2max}$	NR
McGregor et al [56], 2024	Live online supervised group exercise plus psychological support sessions delivered remotely	8 weeks; weekly live group exercise sessions and 6 psychological support sessions; home-based, equipment-free	Synchronous	Health-related quality of life (PROMIS-PROPr[u]), fatigue, depression, pain interference, physical function, mental health outcomes	3, 6, and 12 months
Okan et al [57], 2022	Respiratory control, pursed lip breathing, diaphragmatic breathing exercises, and telemedicine and supervision	5 weeks; everyday, 3 times per day, 1 set × 10 repetitions	Hybrid	6MWT, FVC, dyspnea, SGRQ[v], mMRC	NR
Philip et al [58], 2022	Tele-education and mood telemonitoring, teleconsultation, and telecommunication	6 weeks, on-demand and adapted to each participant	Synchronous	Quality of life, dyspnea, anxiety and depression, VAS[w], GAD-7[x], CAT[y]	NR
Paneroni et al [59], 2024	Home-based, self-directed aerobic and strengthening exercise program with progressive intensity, supplemented by scheduled teleconsultations and physiotherapist video supervision	4 weeks; 6 sessions per week; approximately 60 minutes per session; 24 total sessions; biweekly physiotherapist video calls; weekly nurse teleconsultations; exercise intensity tailored by SPPB[ad]; and oxygen desaturation	Asynchronous	6MWT, 1-minute-sit-to-stand (1MSTS), FEV_1[z], FVC, FEV_1/FVC, MIP[aa], FSS[ab] Fatigue, SF-12[ac], Barthel Index Dyspnea	2 months
Rodriguez-Blanco et al [60], 2023	10 breathing and strength-based exercises, telematic control via videoconference	14 days, once a day, 12 repetitions per exercise each day for 30 minutes	Synchronous	Fatigue, 6MWT, 30-second STS, Dyspnea-12, The modified Borg Scale of perceived effort	NR

XSL•FO
RenderX

Study	Telerehabilitation intervention	Telerehabilitation protocol	Delivery mode	Outcome	Follow-up
Samper-Pardo et al [61], 2023	Tele-education: 3 face-to-face sessions Tele-monitoring: symptom and individual need Health recommendations: diet and sleep Recovery exercises: physical exercise, respiratory physiotherapy exercises, and cognitive stimulation exercises Community resources	12 weeks; human-centered design by individual needs	Asynchronous	Quality of life and 30-second STS, cognitive domains (MoCA), HADS[ae], ISI[af], IPAQ GSES-12[ag]	3 months
da Silva et al [62], 2025	Individualized cardiopulmonary telerehabilitation based on functional exercises (stepping, stationary walking, sit-to-stand, upper-limb strengthening, and breathing and mobility exercises) delivered via videoconference	8 weeks; ≥3 sessions per week; 30-50 minutes per session; intensity individualized using 6MST[ah], 2MSWT[ai], Borg dyspnea scale, HR and SpO_2; some sessions supervised live by physiotherapist	Hybrid	Functional capacity (6MST, 2MSWT, and 30CST[aj]), estimated $\dot{V}O_{2peak}$, heart rate demand, physical functioning domain of SF-36	NR
Lai et al [63], 2024	Individualized aerobic exercise training guided by cardiopulmonary exercise testing, supported by wearable heart rate monitoring and mobile app	12 weeks; 3 sessions per week; 40 minutes per session; exercise intensity prescribed using heart rate reserve and ACSM guidelines; app-based monitoring and weekly reminders	Asynchronous	Cardiorespiratory fitness (CPET[ak]); physical activity (IPAQ); exercise self-efficacy; sleep quality (PSQI[al]); health-related quality of life (WHO-QOL-BREF[am])	NR

XSL•FO
RenderX

Study	Telerehabilitation intervention	Telerehabilitation protocol	Delivery mode	Outcome	Follow-up
Yasacı et al [64], 2025	Supervised multicomponent telerehabilitation including diaphragmatic and pursed-lip breathing, relaxation, range of motion, lower-limb strengthening, and aerobic walking exercises delivered via videoconferencing	6 weeks; 2 sessions per week; 45 minutes per session; intensity guided by ACSM and RPE 11-13	Synchronous	Dyspnea (mMRC); pain intensity (NPRS[an]); functional capacity (5-times sit-to-stand test); sleep quality (PSQI); anxiety and depression (HADS)	NR

[a]ACSM: American College of Sports Medicine.
[b]SF-36: 36-Item Short Form Health Survey.
[c]ICF: International Classification of Functioning, Disability and Health.
[d]mMRC: modified Medical Research Council Dyspnea Scale.
[e]SGRQ: St George Respiratory Questionnaire.
[f]6MWT: 6-minute walk test.
[g]STS: sit-to-stand test.
[h]NR: not reported.
[i]RPE: rating of perceived exertion.
[j]BAI: Beck Anxiety Inventory.
[k]HR_{max}: maximum heart rate.
[l]FAS: Fatigue Assessment Scale.
[m]HADS: Hospital Anxiety and Depression Scale.
[n]MoCA: Montreal Cognitive Assessment.
[o]ISI: Insomnia Severity Index.
[p]HRQoL: health-related quality of life.
[q]LMS: lower-limb muscle strength.
[r]HR: heart rate.
[s]FVC: forced vital capacity.
[t]SMIP: sustained maximal inspiratory pressure.
[u]PROMIS-PROPr: Patient-Reported Outcomes Measurement Information System–Preference score.
[v]SGRQ: St George Respiratory Questionnaire.
[w]VAS: visual analog scale.
[x]GAD-7: 7-item Generalized Anxiety Disorder scale.
[y]CAT: COPD Assessment Test.
[z]FEV_1: forced expiratory volume in 1 second.
[aa]MIP: maximal inspiratory pressure.
[ab]FSS: Fatigue Severity Scale.
[ac]SF-12: 12-Item Short Form.
[ad]SPPB: Short Physical Performance Battery.
[ae]HADS: Hospital Anxiety and Depression Scale.
[af]ISI: Insomnia Severity Index.
[ag]IPAQ GSES-12: International Physical Activity Questionnaire and 12-item General Self-Efficacy Scale.
[ah]6MST: 6-minute step test.
[ai]2MSWT: 2-minute stationary walk test.
[aj]30CST: 30-second chair stand test.
[ak]CPET: cardiopulmonary exercise testing.
[al]PSQI: Pittsburgh Sleep Quality Index.
[am]WHOQOL-BREF: World Health Organization Quality of Life Brief.
[an]NPRS: numerical pain rating scale.

Outcome Measures

The primary outcome was HRQoL. Because HRQoL was assessed using multiple instruments, we included generic measures where available, including the 36-Item Short Form Health Survey (SF-36), the 12-Item Short Form, the EQ-5D, the World Health Organization Quality of Life Brief, and the PROMIS Preference score, as well as respiratory

XSL•FO
RenderX

disease–specific measures such as the St George Respiratory Questionnaire and the King Brief Interstitial Lung Disease questionnaire. Secondary outcomes included functional aerobic capacity (6MWT), lower-limb function (STS), dyspnea (modified Medical Research Council scale), Dyspnea-12, Transition Dyspnea Index, and Barthel Dyspnea Index, fatigue (Chalder Fatigue Scale, Fatigue Assessment Scale, and Fatigue Severity Scale), and anxiety and depression (Hospital Anxiety and Depression Scale [HADS]). When a study reported more than 1 eligible instrument for the same outcome, we used the study's primary or prespecified measure and entered only 1 instrument per study to avoid double-counting. If no primary measure was specified, a generic instrument was prioritized over a disease-specific instrument. The instruments used in each study, together with the direction of scoring, are summarized in Multimedia Appendix 3.

Risk-of-Bias Assessment

Risk of bias was assessed independently by 2 reviewers (XH and YT) using the Cochrane Risk of Bias tool version 2.0 (RoB 2) [65]. The assessment covered 5 specific domains: the randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. For each domain, the risk of bias was classified as "low risk," "some concerns," or "high risk." An overall risk-of-bias judgment was derived for each study based on these domain-level assessments. Any discrepancies between the 2 reviewers were resolved through discussion or, when necessary, by consulting a third reviewer.

Statistical Analysis

Meta-analyses were performed using random-effects models to account for expected clinical and methodological heterogeneity across studies. For continuous outcomes, pooled effect estimates with 95% CIs were calculated as mean differences (MDs) when studies used the same measurement scale and as standardized mean differences (SMDs) when different instruments assessed the same construct (eg, HRQoL, dyspnea, and fatigue). For the SF-36, we used the physical component score rather than combining it with the mental component score; where a study reported only the total score, that value was used instead. Before pooling, all effect directions were harmonized so that positive values indicated clinical benefit: instruments on which higher scores reflect worse health were reverse-coded. The scoring direction of each instrument is summarized in Multimedia Appendix 3. Effect sizes were interpreted according to Cohen guidelines, where SMDs of 0.2, 0.5, and 0.8 represent small, medium, and large effects, respectively [66]. In addition to outcome-specific analyses, an overall pooled analysis was conducted to provide a global summary of the effects of telerehabilitation across outcome domains. Between-study heterogeneity was assessed using the I^2 statistic, the between-study variance (τ^2), and 95% prediction intervals (PIs). I^2 values of approximately 25%, 50%, and 75% were considered to indicate low, moderate, and high heterogeneity, respectively [67]. To obtain more conservative and robust confidence intervals, particularly in analyses including a limited number of studies, the Hartung-Knapp-Sidik-Jonkman (HKSJ) method was applied [68]. To explore potential sources of heterogeneity, we prespecified subgroup analyses by delivery mode (synchronous, asynchronous, or hybrid). For the primary outcome (HRQoL), we additionally performed exploratory subgroup analyses by participant's age, time since diagnosis, and intervention duration, with differences assessed using the test for subgroup differences. Sensitivity analyses were performed using a leave-one-out approach by sequentially excluding individual studies to assess the stability of pooled estimates (Multimedia Appendix 4 [42,45-50,52-64]). Publication bias was evaluated using funnel plots and Egger regression test when at least 10 studies were available for a given outcome [69]. Statistical significance was defined as a 2-sided P value <.05. All statistical analyses were conducted in R statistical software using metafor and meta packages (R Foundation for Statistical Computing).

Certainty of Evidence

To assess the certainty of evidence, we applied the GRADE (Grading of Recommendations Assessment, Development and Evaluation) [70-75] approach to the main outcomes, including HRQoL, functional capacity (6MWT and STS), dyspnea, fatigue, and symptoms of anxiety and depression. Two reviewers (XH and YT) independently evaluated the certainty of evidence across the domains of risk of bias, inconsistency, indirectness, imprecision, and publication bias, with disagreements resolved through discussion and consensus. The GRADE assessments are summarized in the Multimedia Appendix 5.

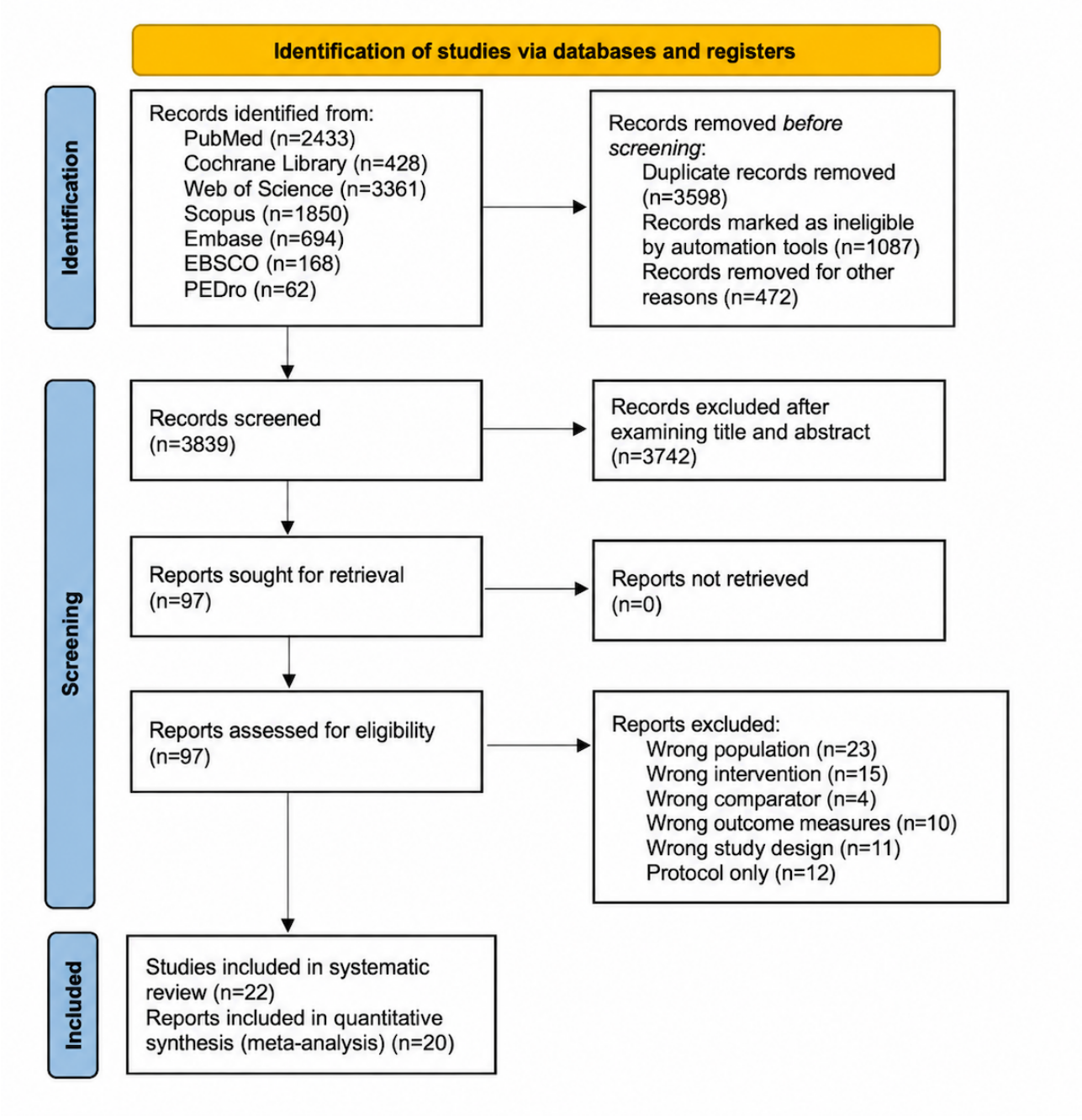
Results

Study Selection

A total of 8996 records were identified through database searching across PubMed, the Cochrane Library, Web of Science, Scopus, Embase, EBSCO, and PEDro. After removal of 5157 records before screening, 3839 records were screened by title and abstract, of which 3742 were excluded. In total, 97 full-text papers were assessed for eligibility, and 75 studies were excluded based on predefined criteria. Finally, 22 RCTs were included in the systematic review [42-64], with 20 studies contributing to the quantitative synthesis (meta-analysis). The study selection process is shown in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart detailing the selection process of randomized controlled trials comparing telerehabilitation with usual care. The search across 7 databases resulted in the inclusion of 22 studies.



Characteristics of the Included Studies

A total of 22 RCTs involving 2294 participants were included in this meta-analysis. Across studies, participants had mean ages ranging from the mid-40s to mid-50s, and female participants accounted for a higher proportion of the study populations. Sample sizes ranged from 21 to nearly 600 participants. The included trials enrolled clinically heterogeneous populations. Eight trials included patients who had been hospitalized or were recruited after hospital discharge; 3 enrolled community-dwelling, nonhospitalized individuals; and 11 did not clearly report hospitalization status. Acute illness severity ranged from mild to critical: 12 trials included patients with severe or critical illness, 9 enrolled only mild to moderate cases, and 1 did not clearly report severity. Time since infection or hospital discharge ranged from less than 1 month to approximately 24 months, with most studies assessing participants between 3 and 12 months after the acute phase. Sociodemographic and clinical characteristics of the participants across the included studies are summarized in Table 3.

XSL•FO
RenderX

Table 3. Sociodemographic characteristics.

Study	Intervention or comparator group	Values, n	Age (years), mean (SD)	Sex, n (%)	Time since diagnosis (months)	Inpatient stay, days	Severity of acute infection
Alsharidah et al [42], 2023	IG[a]: home-based pulmonary telerehabilitation; CG[b]: standardized educational instructions	IG: 24; CG: 24	IG: 23.33 (2.71); CG: 22.58 (2.51)	Female: 100%	4.5	Nonhospitalized	Mild to moderate
Bileviciute-Ljungar et al [43,44], 2024-2025	IG: multidisciplinary group telerehabilitation; CG: waiting list	IG: 67; CG: 42	IG: 43 (9); CG:47 (9)	IG: 15 (22%) males, 52 (78%) females; CG: 5 (12%) males; 37 (88%) females	IG: 11.7 (6.21); CG: 13.4 (7.13)	NR[c]	Mild
Calvo-Paniagua et al [45], 2024	IG: telerehabilitation program based on education, physical exercise, respiratory therapy, and occupational therapy; CG: conventional medical recommendations	IG: 32; CG: 32	IG: 50.8 (8.4); CG: 49.4 (10.0)	IG: 14 (44%) males, 18 (56%) females; CG: 10 (31%) males, 22 (69%) females	14.8 (1.7)	NR	Mild to moderate
Del Corral et al [46], 2023	IMT[d]: inspiratory muscle training IMTsham: sham training RMT[e]: respiratory muscle training RMTsham: sham training	IMT: 22; IMTsham: 22; RMT: 22; RMTsham: 22	IMT: 48.9 (8.3) IMTsham: 45.3 (12.8)RMT: 46.5 (9.6) RMTsham: 45 (10.2)	25 (28%) males; 63 (72%) females	IMT: 11.5 (2.8); IMTsham: 11 (3.5); RMT: 12 (2.2) days; RMTsham: 11.7 (3.1)	NR	Mild to critical
Demir et al [47], 2026	IG: supervised telerehabilitation program; HEG[f]: unsupervised home-based exercise program; CG: patient education only	TRG[g]:12; HEG: 12; CG:12	TRG: 49.50 (12.08); HEG: 38.33 (14.83); CG: 39.50 (17.82)	25 (69%) females; 11 (31%) males	2.3 (1.1)	NR	Mild to moderate
Elyazed et al [48], 2024	IG: home-based pulmonary rehabilitation; CG: usual medical care	IG: 30; CG: 30	56.2 (6.9)	27 (45%) females; 33 (55%) males	≥1 month after acute phase recovery	Nonhospitalized	Mild to moderate
de Araújo Furtado et al [49], 2023	IG: telerehabilitation functional training; CG: no intervention	IG: 16; CG: 16	48.35 (13.0)	18 (56%) females; 14 (44%) males	≥0.7	4.0 (5.3)	Mild to severe
King et al [50], 2023	IG: group telerehabilitation program; CG: usual care	IG:11; CG:10	53 (14)	10 (48%) males; 11 (52%) females	12 (2.2)	IG: 3 (27.3); CG: 1 (10.0)	Mild to critical
Kerling et al [51], 2024	IG: online-supported individualized physical activity intervention; CG: usual lifestyle	IG: 30; CG: 32	46.4 (11.2)	42 (68%) females; 20 (32%) males	≥3	NR	Mild to moderate

XSL•FO
RenderX

Study	Intervention or comparator group	Values, n	Age (years), mean (SD)	Sex, n (%)	Time since diagnosis (months)	Inpatient stay, days	Severity of acute infection
León-Herrera et al [52], 2024	IG: online multimodal telerehabilitation; CG: usual care	IG: 67; CG: 67	48.97 (7.64)	113 (84%) females; 21 (16%) males	23.77 (7.87)	NR	NR
Longobardi et al [53], 2023	IG: home-based exercise training; CG: usual care	IG: 25; CG: 25	IG: 60.8 (7.1); CG: 61.2 (7.7)	IG: 12 (48%) males, 13 (52%) females; CG: 13 (52%) males, 12 (48%) females	5.3 (1.2) months since hospital discharge	ICU LOS[h]: 9 (7); hospital LOS: 18 (13) days	Severe to critical
Li et al [54], 2022	IG: unsupervised home-based exercise program; CG: educational instructions	IG:59; CG:60	18-75; 50.61 (10.98)	53 (45%) males; 66 (55%) females	3.4	26.18 (15.25)	Severe to critical
McNarry et al [55], 2022	IG: inspiratory muscle training; CG: usual care	IG: 111; CG: 37	IG: 46.76 (12.03); CG: 46.13 (12.73)	IG: 14% males, 86% females; CG: 5% males, 95% females	9.0 (4.2)	NR	Mild to critical
McGregor et al [56], 2024	IG: online supervised group physical and mental health rehabilitation (REGAIN[i]); CG: best-practice usual care	IG: 298; CG: 287	IG: 56.1 (12.1);CG: 56.2 (12.3)	280 (48%) males; 305 (52%) females	10.6 (4.7) days since hospital discharge	ICU[j] or HDU[k] admission: 34% overall	Moderate to critical
Okan et al [57], 2022	IG: breathing exercises; CG: education brochure	IG: 26; CG: 26	IG: 48.85 (10.85); CG: 52.19 (14.84)	27 (52%) males; 25 (48%) females	IG: 2.8 (0.8); CG: 3 (1)	NR	Severe and critical
Philip et al [58], 2022	IG: breathing and well-being program; CG: usual care	IG:74; CG:76	18 years or older IG: 49 (12); CG: 50 (12)	121 (81%) females; 26 (17%) males; 3 (2%) others	IG: 10.9 (4.1); CG: 10.2 (4.3)	NR	Mild to severe
Paneroni et al [59], 2024	IG: self-directed exercise program with biweekly physiotherapist video calls; CG: remote teleconsultation from nursing staff	IG: 40; CG: 39	67.1 (10.3)	57 (72%) males; 22 (28%) females	Recruited at hospital discharge following inpatient rehabilitation	Acute stay: 26.7 (15.4)	Severe
Rodriguez-Blanco et al [60], 2023	IG: respiratory and strength exercises; CG: no intervention, relative rest at home	IG:24; CG:24	IG: 38.75 (15.40); CG:42.58 (11.40)	22 (46%) males; 26 (54%) females	NR	NR	Mild
Samper-Pardo et al [61], 2023	IG: APP ReCOVery+Usual carel; CG: usual care	IG:52; CG:48	18 years or older 48.28 (9.26)	20 (20%) males； 80 (80%) females	15.9	NR	Mild to critical

XSL•FO
RenderX

Study	Intervention or comparator group	Values, n	Age (years), mean (SD)	Sex, n (%)	Time since diagnosis (months)	Inpatient stay, days	Severity of acute infection
da Silva et al [62], 2025	IG: individualized exercise program (functional and accessible exercises) supervised by a physical therapist (videoconference); CG: received guidance on general care and self-monitoring of vital signs (videoconference).	IG: 33; CG: 34	55.0 (11.0)	32 (56%) males; 25 (44%) females	NR	12.0 (9.0)	Severe
Lai et al [63], 2024	IG: telerehabilitation exercise training; CG: physical activity counseling only	IG: 91; CG: 91	39.9 (12.6)	70 (38%) males; 112 (62%) females	3.1	Nonhospitalized	Predominantly mild;
Yasacı et al [64], 2025	IG: supervised telerehabilitation-based exercise program; CG: unsupervised home exercise program	IG: 32; CG: 28	56.0 (12)	31 (52%) males; 29 (48%) females	NR	IG:14.97 (9.22); CG:13.92 (8.25)	Mild to moderate

[a]IG: intervention group.

[b]CG: control group.

[c]NR: not reported.

[d]IMT: inspiratory muscle training.

[e]RMT: inspiratory/expiratory muscle training,

[f]HEG: home exercise group.

[g]TRG: telerehabilitation group.

[h]ICU LOS: intensive care unit length of stay.

[i]REGAIN: Rehabilitation Exercise and Psychological Support After COVID-19 Infection.

[j]ICU: intensive care unit.

[k]HDU: high dependency unit.

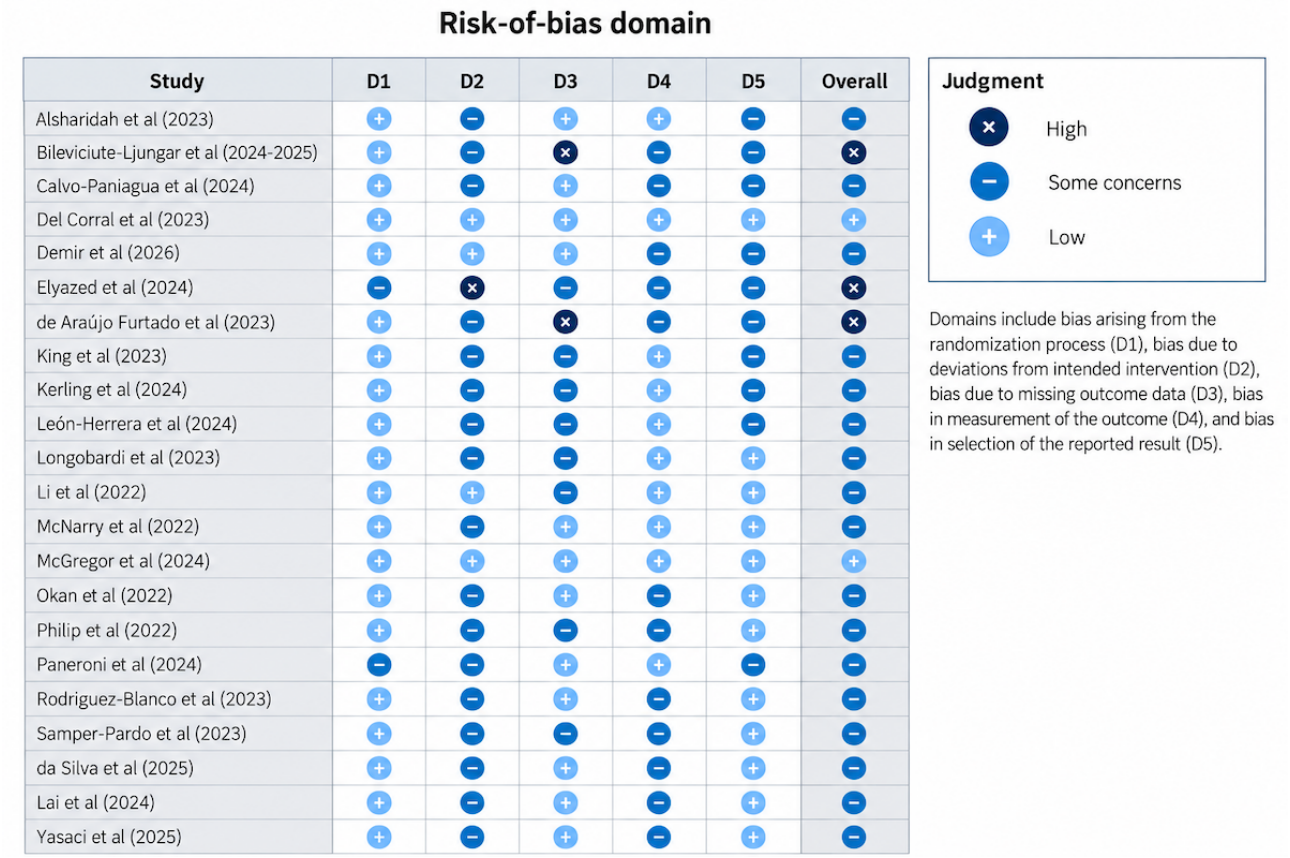
Telerehabilitation was delivered synchronously in 9 trials, asynchronously in 7, and through a hybrid approach in 5; 1 additional trial included both synchronous and asynchronous arms. Intervention duration ranged from 2 to 16 weeks, with 2-6 sessions per week. Most programs were multicomponent: aerobic or endurance training was included in 16 trials, breathing exercises or respiratory muscle training in 15, and resistance or strength training in 12, while 5 incorporated psychological or cognitive components. Comparators most commonly included usual care, educational materials, unsupervised exercise, waiting-list controls, or sham interventions. Outcomes were typically assessed at the end of the intervention, and a subset of trials included follow-up at 2, 3, 6, or up to 12 months. Full intervention details are provided in Table 2.

Risk of Bias and Certainty of the Evidence

All 22 studies were evaluated for bias using the RoB 2 tool (Figure 2), which indicated that 2 of the studies displayed a low risk of bias (9%) [46,56], 17 of them had some concerns (77%) [42,45,47,50-55,57-64], and 3 others had a high risk of bias (14%) [43,44,48,49]. The randomization process (domain 1) was the most methodologically robust domain, with most studies demonstrating a low risk of bias. This indicates that most trials used adequate random sequence generation (eg, computer-generated lists) and allocation concealment mechanisms. In contrast, deviations from the intended interventions (domain 2) and measurement of the outcome (domain 4) presented frequent sources of concern. Due to the nature of telerehabilitation (active exercise training), blinding of participants and personnel was generally not feasible, leading to a rating of “some concerns” or “high risk” in domain 2 for several studies. Similarly, for domain 4, while assessor blinding was implemented in some trials, the use of patient-reported outcome measures, such as quality of life or fatigue, in an unblinded population introduces potential detection bias. Missing outcome data (domain 3) and selection of the reported result (domain 5) were generally rated as low risk, suggesting that most studies performed intention-to-treat analyses and reported outcomes consistent with their protocols. Because these limitations primarily affected the subjective, patient-reported outcomes (HRQoL, dyspnea, and fatigue), the certainty of evidence for these outcomes was downgraded for risk of bias, and their pooled estimates should be interpreted with corresponding caution.

XSL•FO
RenderX

Figure 2. Risk of bias assessed using the Cochrane Risk of Bias tool (RoB 2).



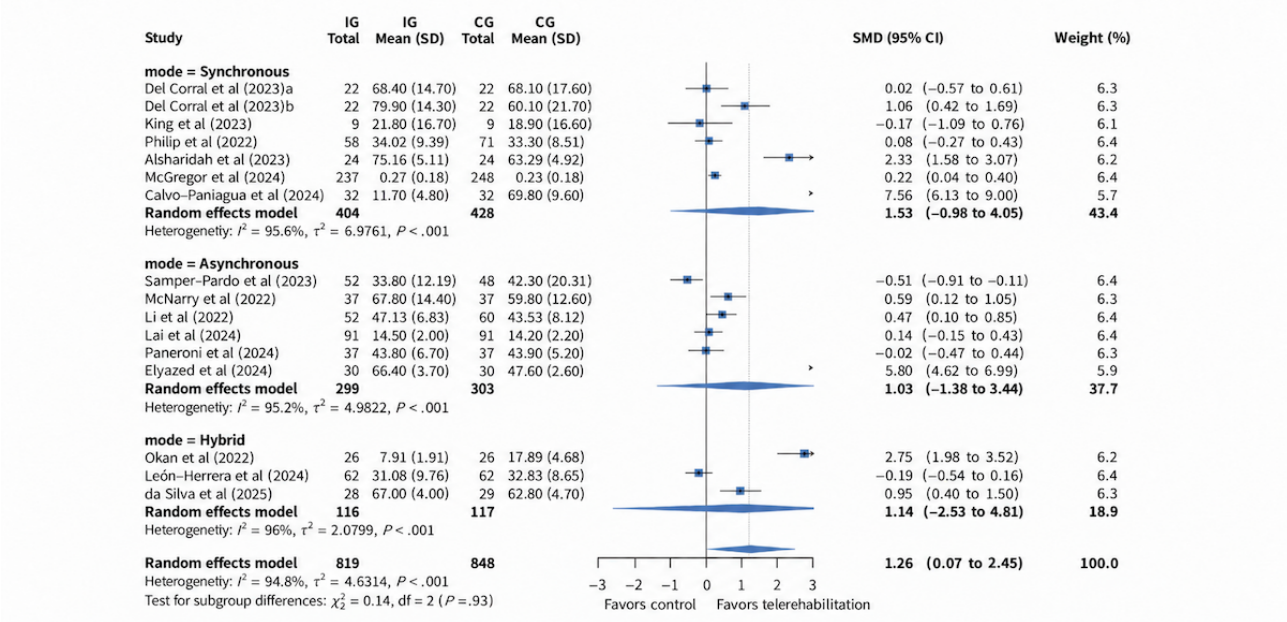
Health-Related Quality of Life

HRQoL was the primary outcome used to evaluate the overall efficacy of telerehabilitation. A total of 16 RCTs involving 1667 participants (819 in the intervention group and 848 in the control group) were included in the meta-analysis. In the pooled analysis using a random-effects model with HKSJ adjustment, telerehabilitation was associated with a statistically significant improvement in HRQoL compared with conventional care. The overall SMD was 1.26 (95% CI 0.07-2.45; P=.04). However, substantial heterogeneity was observed across the studies (I^2=94.8%; τ^2=4.63; P<.001; 95% PI –3.52 to 6.05), indicating considerable variability in the magnitude of the effect.

To investigate whether the delivery mode influenced efficacy, we performed a subgroup analysis stratified by delivery mode. The effect sizes for synchronous (k=7; SMD 1.53, 95% CI –0.98 to 4.05), asynchronous (k=6; SMD 1.03, 95% CI –1.38 to 3.44), and hybrid (k=3; SMD 1.14, 95% CI –2.53 to 4.81) telerehabilitation were comparable. The test for subgroup differences was not statistically significant (χ^2_2=0.14; P=.93), suggesting that the delivery mode was not a primary source of the observed heterogeneity. Notably, while the overall effect was significant, the confidence intervals for all individual subgroups crossed zero, likely attributable to the reduced sample size within subgroups and the conservative nature of the HKSJ adjustment (Figure 3) [42,45,46,48,50,52,54-59,61-63].

XSL•FO
RenderX

Figure 3. Forest plot of standardized mean differences for health-related quality of life comparing telerehabilitation with conventional care. Results are stratified by delivery. A random-effects model was applied. CG: control group; IG: intervention group; SMD: standardized mean difference.



We further explored potential sources of heterogeneity in HRQoL by participant age, time since diagnosis, and intervention duration (Table 4). None of these factors showed statistically significant subgroup differences (age, P=.11; time since diagnosis, P=.43; duration, P=.28), and substantial heterogeneity remained within all strata (I^2=88%-97%). Although point estimates tended to be larger among older participants, those with a shorter time since diagnosis, and in shorter duration programs, confidence intervals were wide and largely overlapping. These findings suggest that the examined characteristics did not meaningfully explain the observed heterogeneity.

Table 4. Subgroup analyses of health-related quality of life by patient and intervention characteristics[a].

Subgroup	Studies, n	SMD[b] (95% CI)	*P* value	I^2 (%)	*P* value for subgroup difference
Age (years)					.11
<50	8	0.40 (−0.33 to 1.13)	.23	88.3	
≥50	8	2.15 (−0.29 to 4.59)	.08	97.0	
Time since diagnosis (months)					.43
<6	6	1.87 (−0.45 to 4.19)	.09	96.6	
≥6	9	0.91 (−0.97 to 2.79)	.30	94.0	
Intervention duration (weeks)					.28
<8	6	2.14 (−0.84 to 5.12)	.12	96.9	
≥8	10	0.75 (−0.52 to 2.01)	.22	92.8	

[a]Values are pooled SMDs with 95% CIs from random-effects meta-analysis with Hartung-Knapp adjustment. Subgroup differences were tested using Cochran Q.

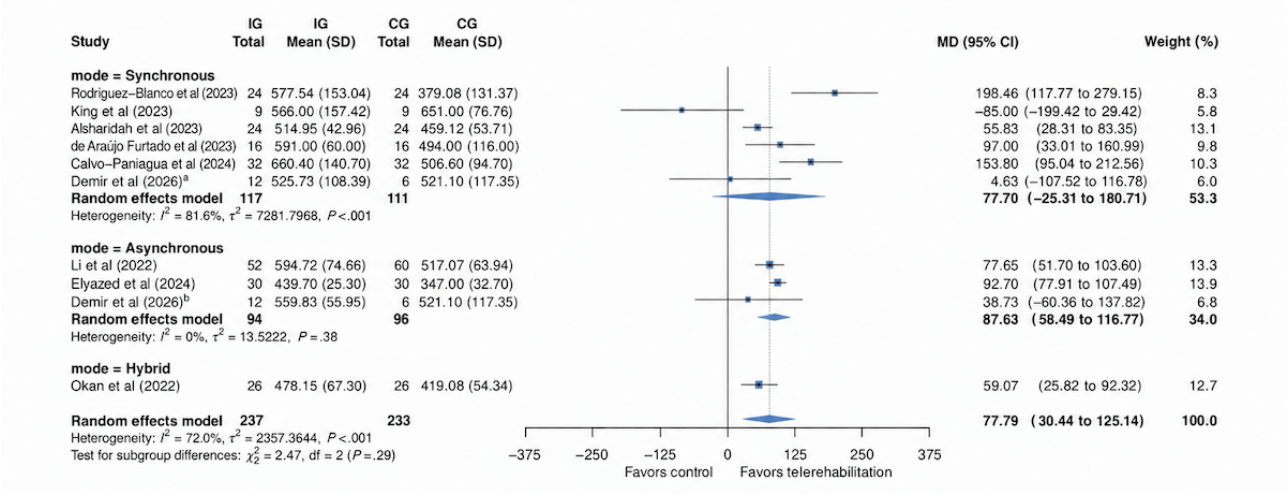
[b]SMD: standardized mean difference.

Functional Aerobic Capacity (6MWT)

Functional aerobic capacity was evaluated using the 6MWT across 10 RCTs involving 470 participants (Figure 4) [42,45,47-50,54,57,60]. In the overall pooled analysis using a random-effects model, telerehabilitation demonstrated a statistically significant improvement in functional aerobic capacity compared with conventional care (MD 77.79, 95% CI 30.44-125.14; P=.001). Although substantial heterogeneity was observed, a subgroup analysis stratified by delivery mode revealed comparable effects across synchronous (k=6; MD 77.70, 95% CI –25.31 to 180.71), asynchronous (k=3; MD 87.63, 95% CI 58.49-116.77), and hybrid (k=1; MD 59.07, 95% CI 25.82-92.32) interventions. The test for subgroup differences was not significant (χ^2_2=2.47; P=.29), suggesting that the efficacy of telerehabilitation in improving aerobic capacity did not differ significantly by delivery mode.

XSL•FO
RenderX

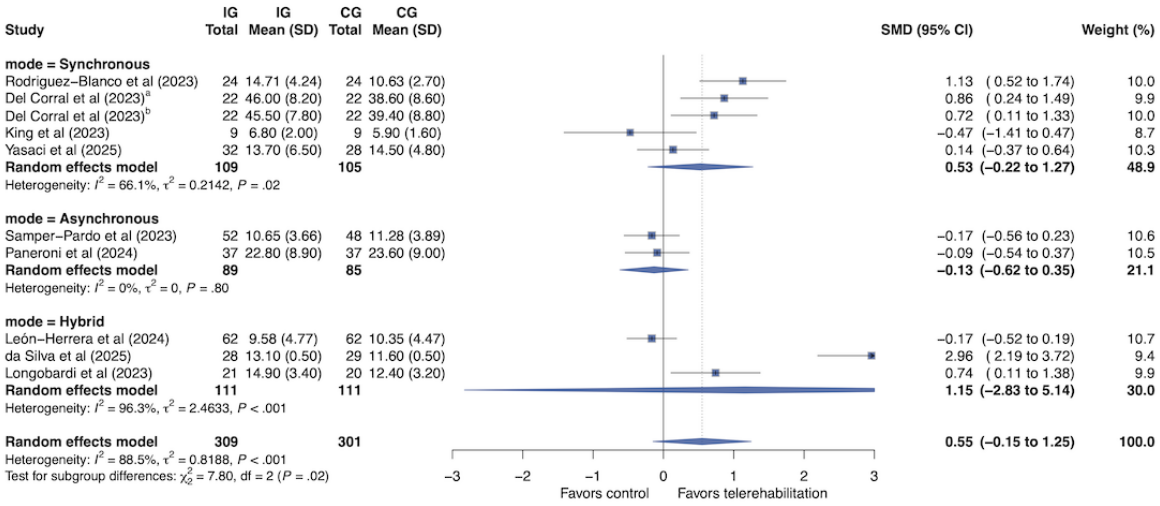
Figure 4. Forest plot of mean differences for functional aerobic capacity comparing telerehabilitation with conventional care. Results are stratified by delivery. A random-effects model was applied. CG: control group; IG: intervention group; MD: mean difference.



Lower-Limb Muscle Strength (STS)

Lower-limb function was assessed with the STS in 10 RCTs (610 participants). The overall pooled effect did not reach significance (SMD 0.55, 95% CI –0.15 to 1.25; *P*=.12), with substantial heterogeneity (I^2=88.5%). In the subgroup analysis by delivery mode, none of the 3 subgroups was individually significant (synchronous, k=5, SMD 0.53, 95% CI –0.22 to 1.27; asynchronous, k=2, SMD –0.13, 95% CI –0.62 to 0.35; hybrid, k=3, SMD 1.15, 95% CI –2.83 to 5.14). Although the test for subgroup differences was significant (χ^2_2=7.8; *P*=.02), the subgroup estimates were imprecise and within-subgroup heterogeneity was high, most notably in the hybrid group (I^2=96.3%), where the pooled estimate was strongly influenced by a single trial. This subgroup difference should therefore be interpreted with caution (Figure 5) [46,50,52,53,59-62,64].

Figure 5. Forest plot of standardized mean differences for lower-limb muscle comparing telerehabilitation with conventional care. Results are stratified by delivery. A random-effects model was applied. CG: control group; IG: intervention group; SMD: standardized mean difference.

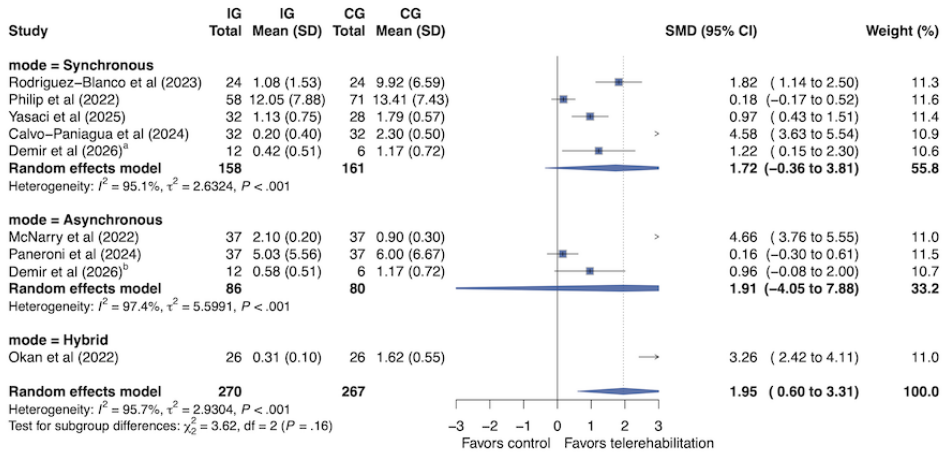


Dyspnea

Dyspnea was evaluated in 9 RCTs involving 537 participants. In the pooled analysis, telerehabilitation yielded a statistically significant reduction in dyspnea compared with conventional care (SMD 1.95, 95% CI 0.60-3.31; *P*=.005). Subgroup analysis stratified by delivery mode showed that the synchronous intervention (k=5) achieved a large effect size (SMD 1.72, 95% CI –0.36 to 3.81), and the asynchronous intervention (k=3) similarly demonstrated positive outcomes (SMD 1.91, 95% CI –4.05 to 7.88). Additionally, the hybrid intervention (k=1) showed a substantial effect (SMD 3.26, 95% CI 2.42-4.11). The test for subgroup differences was not statistically significant (χ^2_2=3.6; *P*=.16), indicating that the delivery mode did not significantly alter the therapeutic benefit (Figure 6) [45,47,55,57-60,64].

XSL•FO
RenderX

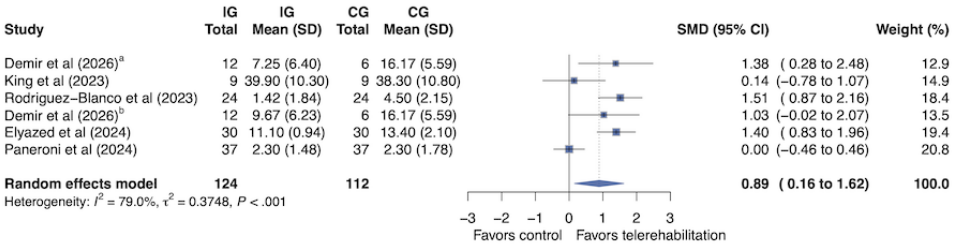
Figure 6. Forest plot of standardized mean differences for dyspnea comparing telerehabilitation with conventional care. Results are stratified by delivery mode. A random-effects model was applied. CG: control group; IG: intervention group; SMD: standardized mean difference.



Fatigue

Fatigue was evaluated in 6 RCTs (n=236). The pooled analysis indicated that telerehabilitation yielded a statistically significant reduction in fatigue compared with conventional care (SMD 0.89, 95% CI 0.16-1.62; *P*=.02) (Figure 7) [47,48,50,59,60].

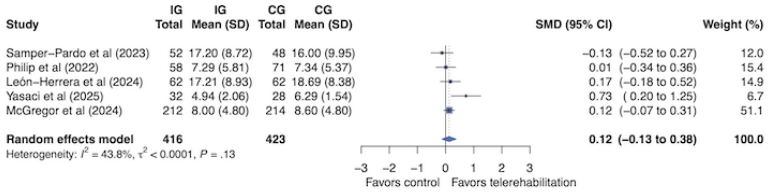
Figure 7. Forest plot of standardized mean differences for fatigue comparing telerehabilitation with conventional care. A random-effects model was applied. CG: control group; IG: intervention group; SMD: standardized mean difference.



Anxiety and Depression Symptoms

Five RCTs reported anxiety and depression outcomes. Because these were assessed with different instruments (the HADS total score in 2 trials, the HADS anxiety subscale in 2, and the GAD-7 in 1), effects were pooled as SMDs. The random-effects meta-analysis showed no significant difference between telerehabilitation and control interventions (SMD 0.12, 95% CI –0.13 to 0.38; P=.36). Between-study heterogeneity was low to moderate (I^2=43.8%, P=.13). Owing to the limited number of studies, no subgroup analyses were performed (Figure 8) [52,56,58,61,64].

Figure 8. Forest plot of standardized mean differences for anxiety and depression symptoms comparing telerehabilitation with conventional care. A random-effects model was applied. CG: control group; IG: intervention group; SMD: standardized mean difference.

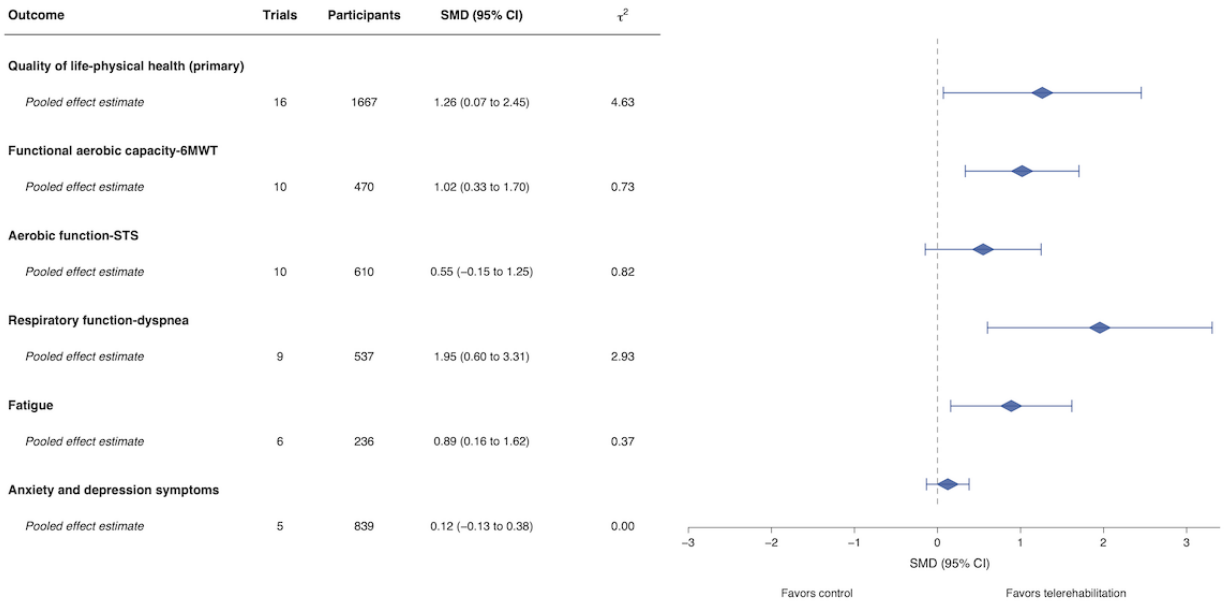


Overview of Effects Across Outcomes

An overview forest plot was used to summarize pooled effects across outcome domains. Overall, telerehabilitation showed a consistent favorable direction of effect for most physical and functional outcomes, with the largest effects observed for dyspnea and HRQoL, followed by improvements in aerobic capacity (6MWT) and fatigue. In contrast, mental health outcomes assessed by symptoms of anxiety and depression showed small and nonsignificant effects, with estimates centered around the null. Heterogeneity varied across outcomes and was generally more pronounced for physical and respiratory measures. Collectively, these findings suggest that telerehabilitation confers more robust benefits in physical and functional domains than in psychological outcomes (Figure 9).

XSL•FO
RenderX

Figure 9. Summary forest plot of pooled standardized mean differences for primary and secondary outcomes comparing telerehabilitation with conventional care. A random-effects model was applied. 6MWT: 6-minute walk test; SMD: standardized mean difference; STS: sit-to-stand test.



Sensitivity Analysis

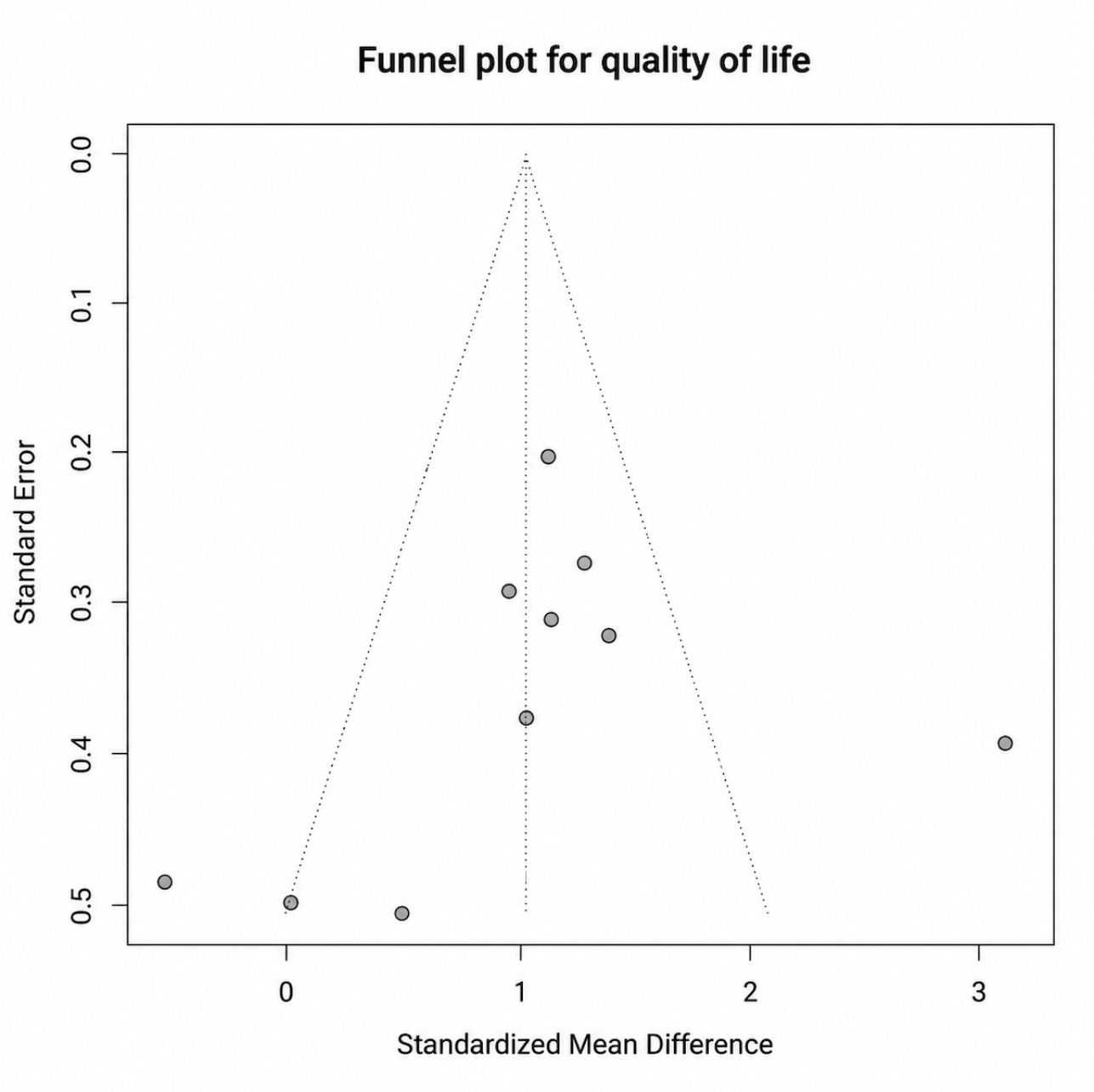
Given the high heterogeneity across outcomes, a leave-one-out sensitivity analysis was performed to test the stability of the pooled results. Detailed results are shown in Multimedia Appendix 4. The analysis confirmed the robustness of the primary findings, as omitting any single trial did not alter the statistical significance for HRQoL, 6MWT, or STS. For Fatigue and the anxiety and depression outcome, specific outliers were identified. Excluding 2 studies [59,64] significantly reduced heterogeneity (I^2 dropped to 39.0% for fatigue and 0% for anxiety and depression) while maintaining the direction of the therapeutic effect, thereby validating the reliability of our conclusions.

Publication Bias Test

Publication bias was assessed exclusively for the primary outcome (HRQoL) because other outcomes did not meet the recommended threshold of included studies (k<10) for reliable testing. The funnel plot for HRQoL (Figure 10) displays a symmetrical distribution of studies, and Egger linear regression test confirmed no evidence of significant small-study effects (t=–0.78; P=.46), indicating a low risk of publication bias.

XSL•FO
RenderX

Figure 10. Funnel plot for small-study effects in health-related quality of life meta-analysis.



Certainty of Evidence

According to the GRADE approach, the certainty of evidence was rated as low for all outcomes (Multimedia Appendix 5). For HRQoL, functional capacity (6MWT and STS), dyspnea, and fatigue, the certainty was downgraded primarily due to risk of bias and inconsistency, reflecting methodological limitations and substantial between-study heterogeneity. In contrast, the certainty of evidence for symptoms of anxiety and depression was downgraded due to risk of bias and imprecision, owing to the limited number of contributing trials and relatively wide confidence intervals. No serious concerns regarding indirectness or publication bias were identified across outcomes. Overall, these findings indicate that the pooled estimates should be interpreted cautiously.

Discussion

Principal Results

This systematic review and meta-analysis of 22 RCTs evaluated the effects of telerehabilitation in adults with PCC. Our synthesis indicates that telerehabilitation was associated with significant improvements in HRQoL, functional aerobic capacity (6MWT), dyspnea, and fatigue compared with control group. From a clinical perspective, the observed improvement in functional aerobic capacity also appears meaningful: the pooled increase in 6-minute walk distance (MD 77.79 meters) exceeded the minimal clinically important difference of approximately 30 meters reported for chronic respiratory disease [33], suggesting that the effect is likely to be clinically relevant. This estimate is broadly consistent with a recent updated meta-analysis, which reported a mean improvement of 85.83 meters in 6-minute walk

XSL•FO
RenderX

distance [28]. In contrast, HRQoL, dyspnea, and fatigue were synthesized using SMDs across multiple instruments and therefore could not be interpreted against a single minimal clinically important difference threshold. Beyond these overall benefits, the present findings suggest that the role of supervision in telerehabilitation may vary according to the targeted outcome. Stratified analyses suggested that the benefits for general aerobic capacity and HRQoL were broadly comparable across delivery modes, whereas a subgroup difference favoring synchronous delivery was observed for lower-limb muscle strength. These subgroup findings were exploratory and should be interpreted with caution.

Because the subgroup differences for HRQoL and 6MWT were not significant, our data do not suggest that continuous real-time supervision is required to improve general aerobic capacity, although the limited power of these comparisons means equivalence cannot be assumed. Asynchronous protocols may be sufficient for these outcomes, possibly because they allow patients to incorporate low-to-moderate intensity activity into daily routines, consistent with evidence that structured home-based rehabilitation can improve endurance regardless of direct supervision [76,77]. For lower-limb strength (STS), by contrast, a significant subgroup difference favored synchronous delivery; however, neither modality reached significance on its own, so synchronous supervision may play a more important role in strength training rather than being essential. A possible explanation is that real-time monitoring helps therapists gauge exertion and progress intensity safely [78], while supervised sessions may support adherence and correct exercise technique [79]. These factors may be particularly important for strength training, which requires adequate progression and training intensity to achieve meaningful gains.

Data on adherence and patient satisfaction were reported in limited studies, preventing a quantitative meta-analysis. Nevertheless, these factors are clinically relevant, as satisfaction is a key predictor of long-term adherence and outcomes [80]. Based on current evidence, the mode of delivery appears to influence patient engagement [76]. Asynchronous telerehabilitation provides temporal flexibility, allowing patients with fatigue to adjust exercise timing to their energy levels. This feasibility likely contributed to the QoL improvements observed in the self-managed group. Conversely, synchronous modes involve real-time interaction that may strengthen the therapeutic alliance and accountability. This supervision may help sustain participant engagement and adequate training intensity. A hybrid approach combining asynchronous delivery with periodic synchronous supervision may therefore provide a practical balance between accessibility and clinical support, although this hypothesis requires confirmation in future studies.

With respect to persistent symptoms, telerehabilitation was associated with significant improvements in both dyspnea and fatigue. The heterogeneity in dyspnea outcomes is consistent with prior evidence syntheses and may reflect differences in intervention focus [18]. Programs that included dyspnea-targeted components, such as inspiratory muscle training, appeared more likely to yield clinically meaningful benefit. The reduction in fatigue represented a large effect size (SMD 0.89), underscoring its clinical importance and aligning with accumulating evidence supporting structured exercise for PCC [26]. This benefit may reflect the counteracting of physical deconditioning through pacing and gradual, individualized reconditioning, without worsening symptoms. Conversely, consistent with prior evidence, we found no significant impact of physical telerehabilitation on symptoms of anxiety and depression [81]. Anxiety in PCC may involve complex neuroinflammatory and psychosocial mechanisms [82], which physical rehabilitation alone may not sufficiently address. Among the studies that reported adverse events, none were serious or intervention-related, although adverse event reporting was inconsistent across trials. The available evidence therefore suggests that telerehabilitation appears to be safe under the conditions evaluated, but trials with standardized safety reporting are warranted. This underscores the need for integrated, multicomponent interventions that combine physical training with structured psychological strategies, such as cognitive behavioral approaches.

Substantial heterogeneity was observed across several pooled outcomes, particularly HRQoL, and the findings should therefore be interpreted with caution. Subgroup analyses by delivery mode, participant age, time since diagnosis, and intervention duration did not identify statistically significant subgroup differences, and substantial residual heterogeneity remained within strata. Several other factors may contribute. First, HRQoL was assessed using different instruments, including generic measures (eg, SF-36 and EQ-5D) and disease-specific tools such as the St George Respiratory Questionnaire, which differ in construct, scaling, and responsiveness; such variation in measurement may contribute to between-study variability. Second, the populations ranged from mild symptoms to survivors of critical illness, and from community to posthospital settings, implying different baseline impairment and recovery trajectories. Third, the interventions differed in content, progression, and exercise dose, from single-component breathing or inspiratory muscle training to multicomponent aerobic and resistance programs (Table 2); their predominantly multicomponent nature limited the feasibility of a clean subgroup analysis by individual component, so these differences were considered narratively. Because none of these factors were formally tested as moderators, they should be regarded as plausible explanations rather than definitive sources of heterogeneity. Practically, this uncertainty is reflected in the wide PI, indicating that the true effect may vary substantially across future settings and patient populations, whose broader dynamics have also been characterized through population-level pandemic modeling [83]. These findings should therefore be interpreted as average effects across heterogeneous contexts rather than precise estimates applicable to every patient.

Although subgroup analysis by intervention component was not feasible because of the limited number of studies, the interventions can be broadly summarized using the frequency, intensity, time, and type framework (Table 2). Most programs prescribed exercise 2-6 times per week for 20-90 minutes over 2-16 weeks. Exercise intensity was generally individualized using Borg ratings of perceived exertion, percentages of maximal heart rate or heart rate reserve, American College of

XSL•FO
RenderX

Sports Medicine recommendations, or, for inspiratory muscle training, percentages of maximal inspiratory pressure. Interventions ranged from single-component breathing or inspiratory muscle training to multicomponent programs combining aerobic and resistance exercise. Overall, these findings suggest that individualized exercise programs delivered regularly over several weeks may be beneficial, although considerable variation remained in exercise intensity and program composition, and an optimal exercise prescription cannot be established from the current evidence.

Limitations

This systematic review and meta-analysis has several limitations that warrant consideration. First, substantial heterogeneity was present across the main pooled outcomes and, as discussed previously, likely reflected differences in measurement instruments, clinical populations, and intervention characteristics; although we used random-effects models and subgroup analyses, much of it remained unexplained. Second, the evidence base is constrained by the nature of rehabilitation interventions, where blinding of participants and therapists was largely infeasible, introducing potential performance and detection biases, particularly for subjective patient-reported outcomes such as HRQoL and dyspnea. Moreover, several key outcomes, including HRQoL, fatigue, and anxiety and depression, were assessed using patient-reported measures and are therefore inherently subjective. These outcomes may be influenced by participants' expectations, mood, social context, and the absence of blinding. Accordingly, they should be interpreted together with the objective functional outcomes and in light of the overall certainty of evidence. Third, the limited number of studies in certain subgroups constrained statistical power. Nevertheless, these exploratory analyses were included to discern potential trends regarding delivery modes; consequently, the findings should be interpreted as hypothesis-generating rather than definitive. Finally, most included trials had relatively short follow-up periods (typically 4-12 weeks) and often lacked rigorous assessments of patients' digital literacy, limiting our ability to determine the long-term sustainability of benefits and the generalizability of findings to older or less tech-savvy populations in real-world settings.

Future Perspectives

Collectively, these findings are consistent with a tentative stepped-care approach, although they should be regarded as hypothesis-generating given the low certainty of evidence and the wide, overlapping confidence intervals of the subgroup estimates. For most patients aiming to recover general fitness and quality of life, asynchronous telerehabilitation appears to be a reasonable and scalable option, given its broadly comparable effects and lower resource demands. For patients with marked muscle weakness or those requiring closer safety monitoring, synchronous supervision may be preferable. These suggestions require confirmation in adequately powered trials. Considering the limitations of the current evidence base, future research should prioritize 3 directions. First, to address the null effects observed for mental health outcomes, telerehabilitation protocols should move beyond a solely physical rehabilitation paradigm. The integration of structured physical exercise with remotely delivered psychological components, such as online cognitive behavioral therapy or mindfulness-based interventions, may offer a more comprehensive approach to the neuropsychiatric burden of PCC, consistent with emerging multidisciplinary models proposed in recent literature [81]. Second, given that most included trials were characterized by relatively short follow-up periods, there is a clear need to examine the long-term sustainability of observed benefits, particularly for outcomes such as muscle strength and functional capacity. Future studies should assess whether intervention effects are maintained at 6- and 12-month follow-up and explore whether incorporating periodic synchronous booster sessions into predominantly asynchronous programs can enhance adherence and prevent functional regression over time. Finally, although asynchronous telerehabilitation is often assumed to be more cost-efficient than conventional face-to-face rehabilitation [84], robust economic evaluations remain scarce. Rigorous cost-effectiveness and cost-utility analyses are therefore required to quantify potential economic advantages and to inform health policy decisions, particularly with respect to the inclusion of telerehabilitation within insurance reimbursement frameworks.

Conclusions

This review synthesized randomized evidence on telerehabilitation for adults with PCC, with a focus on HRQoL and delivery mode. Telerehabilitation may improve HRQoL, aerobic capacity, dyspnea, and fatigue, whereas no clear benefit was observed for symptoms of anxiety or depression. Evidence for lower-limb strength remained inconclusive, and any added value of synchronous supervision should be interpreted cautiously. Given the limited certainty of evidence for several outcomes and substantial heterogeneity across studies, telerehabilitation should be viewed as a promising and accessible rehabilitation option rather than an established standard of care. Future trials should use standardized outcomes, longer follow-up, and prespecified analyses of delivery mode to confirm benefits and clarify which patients are most likely to benefit.

Acknowledgments

The authors received no assistance in the preparation of this manuscript beyond that of the listed authors. During the preparation of this manuscript, the authors used generative artificial intelligence tools to assist with language editing and to support literature searching. All content was subsequently reviewed, verified, and edited by the authors, who take full responsibility for the integrity and accuracy of the manuscript.

XSL•FO
RenderX

Data Availability

All data generated or analyzed during this study are included in this published article and its multimedia appendices. The data analyzed in this systematic review and meta-analysis were extracted from previously published studies, which are cited in the manuscript. No new primary data were generated.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' Contributions

Conceptualization: XH (lead), DL (supporting)
Data curation: XH (lead), YT (supporting)
Formal analysis: XH (lead), RDB (supporting)
Investigation: XH (equal), YT (equal)
Methodology: XH (lead), DL (supporting)
Project administration: WB
Supervision: WB
Validation: RDB
Visualization: XH
Writing – original draft: XH
Writing – review & editing: RDB (equal), WB (equal)
All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA 2020 checklist.
[PDF File (Adobe PDF File), 122 KB - rehab_v13i1e95368_app1.pdf]

Multimedia Appendix 2
Full electronic search strategy for each database.
[DOCX File , 16 KB - rehab_v13i1e95368_app2.docx]

Multimedia Appendix 3
Outcome measures, instruments, scoring direction, and orientation in the meta-analysis.
[DOCX File , 30 KB - rehab_v13i1e95368_app3.docx]

Multimedia Appendix 4
Leave-One-Out Meta-Analysis for all outcomes.
[DOCX File , 1461 KB - rehab_v13i1e95368_app4.docx]

Multimedia Appendix 5
The GRADE assessments.
[PDF File (Adobe PDF File), 119 KB - rehab_v13i1e95368_app5.pdf]

References

1. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV, WHO Clinical Case Definition Working Group on Post-COVID-19 Condition. A clinical case definition of post-COVID-19 condition by a Delphi consensus. Lancet Infect Dis 2022;22(4):e102-e107 [FREE Full text] [doi: 10.1016/S1473-3099(21)00703-9] [Medline: 34951953]
2. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol 2023;21(3):133-146 [FREE Full text] [doi: 10.1038/s41579-022-00846-2] [Medline: 36639608]
3. Greenhalgh T, Sivan M, Perlowski A, Nikolich J. Long COVID: a clinical update. Lancet 2024;404(10453):707-724. [doi: 10.1016/S0140-6736(24)01136-X] [Medline: 39096925]
4. Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. Nat Med 2022;28(8):1706-1714 [FREE Full text] [doi: 10.1038/s41591-022-01909-w] [Medline: 35879616]

XSL•FO
RenderX

5. Ballering AV, van Zon SKR, Olde Hartman TC, Rosmalen JGM, Lifelines Corona Research Initiative. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. Lancet 2022;400(10350):452-461 [FREE Full text] [doi: 10.1016/S0140-6736(22)01214-4] [Medline: 35934007]
6. Global Burden of Disease Long COVID Collaborators, Wulf Hanson S, Abbafati C, Aerts JG, Al-Aly Z, Ashbaugh C, et al. Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. JAMA 2022;328(16):1604-1615 [FREE Full text] [doi: 10.1001/jama.2022.18931] [Medline: 36215063]
7. Xie Y, Choi T, Al-Aly Z. Mortality in patients hospitalized for COVID-19 vs influenza in fall-winter 2023-2024. JAMA 2024;331(22):1963-1965. [doi: 10.1001/jama.2024.7395] [Medline: 38748411]
8. Coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [accessed 2026-09-12]
9. Sánchez-García JC, Reinoso-Cobo A, Piqueras-Sola B, Cortés-Martín J, Menor-Rodríguez MJ, Alabau-Dasi R, et al. Long COVID and physical therapy: a systematic review. Diseases 2023;11(4):163 [FREE Full text] [doi: 10.3390/diseases11040163] [Medline: 37987274]
10. Imamura M, Mirisola AR, Ribeiro FDQ, De Pretto LR, Alfieri FM, Delgado VR, et al. Rehabilitation of patients after COVID-19 recovery: an experience at the Physical and Rehabilitation Medicine Institute and Lucy Montoro Rehabilitation Institute. Clinics (Sao Paulo) 2021;76:e2804 [FREE Full text] [doi: 10.6061/clinics/2021/e2804] [Medline: 34133481]
11. Zeraatkar D, Ling M, Kirsh S, Jassal T, Shahab M, Movahed H, et al. Interventions for the management of long covid (post-Covid condition): living systematic review. BMJ 2024;387:e081318 [FREE Full text] [doi: 10.1136/bmj-2024-081318] [Medline: 39603702]
12. Ahmadi Hekmatikar AH, Ferreira Júnior JB, Shahrbanian S, Suzuki K. Functional and psychological changes after exercise training in post-COVID-19 patients discharged from the hospital: a PRISMA-compliant systematic review. Int J Environ Res Public Health 2022;19(4):2290 [FREE Full text] [doi: 10.3390/ijerph19042290] [Medline: 35206483]
13. Neto MG, Suzart Coutinho de Araujo W, Pinto ACPN, Saquetto MB, Martinez BP, Gomes VA, et al. Effects of physical rehabilitation interventions on exercise performance, dyspnea, and health-related quality of life in acute and post-acute COVID-19 patients: systematic review and meta-analysis. Chronic Illn 2025;21(2):183-204. [doi: 10.1177/17423953241306275] [Medline: 40017108]
14. Pouliopoulou DV, Macdermid JC, Saunders E, Peters S, Brunton L, Miller E, et al. Rehabilitation interventions for physical capacity and quality of life in adults with post-COVID-19 condition: a systematic review and meta-analysis. JAMA Netw Open 2023;6(9):e2333838 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.33838] [Medline: 37725376]
15. Thomas P, Baldwin C, Bissett B, Boden I, Gosselink R, Granger CL, et al. Physiotherapy management for COVID-19 in the acute hospital setting: clinical practice recommendations. J Physiother 2020;66(2):73-82 [FREE Full text] [doi: 10.1016/j.jphys.2020.03.011] [Medline: 32312646]
16. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS. Telerehabilitation: review of the state-of-the-art and areas of application. JMIR Rehabil Assist Technol 2017;4(2):e7 [FREE Full text] [doi: 10.2196/rehab.7511] [Medline: 28733271]
17. Seron P, Oliveros M, Gutierrez-Arias R, Fuentes-Aspe R, Torres-Castro RC, Merino-Osorio C, et al. Effectiveness of telerehabilitation in physical therapy: a rapid overview. Phys Ther 2021;101(6):pzab053 [FREE Full text] [doi: 10.1093/ptj/pzab053] [Medline: 33561280]
18. Neto GEF, Prudente GD, de Oliveira HL, de Lima KBA, da Silva Menezes Junior A. Telerehabilitation as an innovative strategy for the management of anxiety and dyspnea in post-COVID-19: a scoping review. PM R 2025;17(11):1366-1380. [doi: 10.1002/pmrj.13403] [Medline: 40422292]
19. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020;8(8):e20359 [FREE Full text] [doi: 10.2196/20359] [Medline: 32784177]
20. Daniels K, Mourad J, Bonnechère B. Exploring the use of mobile health for the rehabilitation of long COVID patients: a scoping review. Healthcare (Basel) 2024;12(4):451 [FREE Full text] [doi: 10.3390/healthcare12040451] [Medline: 38391826]
21. Mfouth Kemajou P, Mbanya A, Coppieters Y. Digital approaches in post-COVID healthcare: a systematic review of technological innovations in disease management. Biol Methods Protoc 2024;9(1):bpae070 [FREE Full text] [doi: 10.1093/biomethods/bpae070] [Medline: 39440031]
22. Martínez-Pozas O, Corbellini C, Cuenca-Zaldívar JN, Meléndez-Oliva É, Sinatti P, Sánchez Romero EA. Effectiveness of telerehabilitation versus face-to-face pulmonary rehabilitation on physical function and quality of life in people with post COVID-19 condition: a systematic review and network meta-analysis. Eur J Phys Rehabil Med 2024;60(5):868-877 [FREE Full text] [doi: 10.23736/S1973-9087.24.08540-X] [Medline: 39235257]
23. Yang J, Li H, Zhao H, Xie Y, Li J, Wang M. Effectiveness of telerehabilitation in patients with post-COVID-19: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2024;14(7):e074325 [FREE Full text] [doi: 10.1136/bmjopen-2023-074325] [Medline: 38964791]
24. Oliveira MR, Hoffman M, Jones AW, Holland AE, Borghi-Silva A. Effect of pulmonary rehabilitation on exercise capacity, dyspnea, fatigue, and peripheral muscle strength in patients with post-COVID-19 syndrome: a systematic review and meta-analysis. Arch Phys Med Rehabil 2024;105(8):1559-1570. [doi: 10.1016/j.apmr.2024.01.007] [Medline: 38311096]

XSL•FO
RenderX

25. Huang J, Fan Y, Zhao K, Yang C, Zhao Z, Chen Y, et al. Do patients with and survivors of COVID-19 benefit from telerehabilitation? A meta-analysis of randomized controlled trials. Front Public Health 2022;10:954754 [FREE Full text] [doi: 10.3389/fpubh.2022.954754] [Medline: 36249181]
26. Du S, Cui Z, Xu X, Liu T, Ye J. Clinical efficacy of exercise in the treatment of post-COVID-19 syndrome: a systematic review and network meta-analysis. Front Physiol 2025;16:1656713 [FREE Full text] [doi: 10.3389/fphys.2025.1656713] [Medline: 41488929]
27. Calvache-Mateo A, Heredia-Ciuró A, Martín-Núñez J, Hernández-Hernández S, Reychler G, López-López L, et al. Efficacy and safety of respiratory telerehabilitation in patients with long COVID-19: a systematic review and meta-analysis. Healthcare (Basel) 2023;11(18):2519 [FREE Full text] [doi: 10.3390/healthcare11182519] [Medline: 37761716]
28. Javed MN, Khan SM, Hussain JM, Alam H, Sheikh M, Saleem Z, et al. Effectiveness of telerehabilitation in patients with post-COVID-19: an updated systematic review and meta-analysis of randomized controlled trials. BMC Pulm Med 2026;26(1):355 [FREE Full text] [doi: 10.1186/s12890-026-04377-x] [Medline: 42271357]
29. Gorst SL, Seylanova N, Dodd SR, Harman NL, O'Hara M, Terwee CB, PC-COS study group. Core outcome measurement instruments for use in clinical and research settings for adults with post-COVID-19 condition: an international Delphi consensus study. Lancet Respir Med 2023;11(12):1101-1114. [doi: 10.1016/S2213-2600(23)00370-3] [Medline: 37926103]
30. World Health Organization. International Classification of Functioning, Disability and Health: ICF. Geneva: World Health Organization; 2001.
31. Stucki G. International Classification of Functioning, Disability, and Health (ICF): a promising framework and classification for rehabilitation medicine. Am J Phys Med Rehabil 2005;84(10):733-740. [doi: 10.1097/01.phm.0000179521.70639.83] [Medline: 16205428]
32. Munblit D, Nicholson T, Akrami A, Apfelbacher C, Chen J, De Groote W, PC-COS Project Steering Committee. A core outcome set for post-COVID-19 condition in adults for use in clinical practice and research: an international Delphi consensus study. Lancet Respir Med 2022;10(7):715-724 [FREE Full text] [doi: 10.1016/S2213-2600(22)00169-2] [Medline: 35714658]
33. Holland AE, Spruit MA, Troosters T, Puhan MA, Pepin V, Saey D, et al. An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. Eur Respir J 2014;44(6):1428-1446 [FREE Full text] [doi: 10.1183/09031936.00150314] [Medline: 25359355]
34. Amput P, Tapanya W, Wongphon S, Naravejsakul K, Sritiyot T. Test-retest reliability and minimal detectable change of the 6-minute step test and 1-minute sit-to-stand test in post-COVID-19 patients. Adv Respir Med 2025;93(5):33 [FREE Full text] [doi: 10.3390/arm93050033] [Medline: 40981074]
35. Machado Marques J, Brasiliano da Paz M, Gomes Costa RR, Ribeiro Neto F. Validity, reliability and minimal detectable change in the sit-to-stand test for synchronous and asynchronous tele-assessment in post-COVID-19 condition. PeerJ 2025;13:e20211 [FREE Full text] [doi: 10.7717/peerj.20211] [Medline: 41328073]
36. Cigarroa I, Reyes-Molina D, Vargas-Rios F, López-Alarcón G, Jara-Aceituno S, Riquelme-Hernández C, et al. Effectiveness of synchronous telerehabilitation versus face-to-face physical therapy in older adults who are frail: protocol for a randomized controlled trial. JMIR Res Protoc 2025;14:e72318 [FREE Full text] [doi: 10.2196/72318] [Medline: 40957013]
37. Tawfek RA, Çil ET. Effects of synchronous versus asynchronous telerehabilitation programs for chronic nonspecific low back pain: a three-arm randomized controlled trial. Arch Phys Med Rehabil 2025;106(9):1392-1401. [doi: 10.1016/j.apmr.2025.05.012] [Medline: 40419012]
38. Tanhan A, Ozer AY, Timurtas E, Batirel A, Polat MG. Is asynchronous telerehabilitation equal to synchronous telerehabilitation in COVID-19 survivors with classes 4-6? J Telemed Telecare 2025;31(3):347-358. [doi: 10.1177/1357633X231189761] [Medline: 37545432]
39. Page M, Moher D, Bossuyt P, Boutron I, Hoffmann T, Mulrow C, et al. PRISMA 2020 explanation and elaboration. BMJ 2021;n160:372. [doi: 10.1136/bmj.n160]
40. Brown D. A review of the pubMed PICO tool: using evidence-based practice in health education. Health Promot Pract 2020;21(4):496-498. [doi: 10.1177/1524839919893361] [Medline: 31874567]
41. Hill AJ, Stangenberg-Gliss K, Aytür YK, Enderby P, Auger C, Gamal A, et al. Reaching a consensus on the definition of telerehabilitation: world federation of neurorehabilitation telerehabilitation special interest group. Int J Telerehabil 2025;17(1):6708 [FREE Full text] [doi: 10.63144/ijt.2025.6708] [Medline: 40575143]
42. Alsharidah AS, Kamel FH, Alanazi AA, Alhawsah EA, Alharbi HK, Alrshedi ZO, et al. A pulmonary telerehabilitation program improves exercise capacity and quality of life in young females post-COVID-19 patients. Ann Rehabil Med 2023;47(6):502-510 [FREE Full text] [doi: 10.5535/arm.23060] [Medline: 37980910]
43. Bileviciute-Ljungar I, Apelman A, Braconier L, Östhols S, Norrefalk J, Borg K. A first randomized eight-week multidisciplinary telerehabilitation study for the post-COVID-19 condition: improvements in health- and pain-related parameters. J Clin Med 2025;14(2):486 [FREE Full text] [doi: 10.3390/jcm14020486] [Medline: 39860492]
44. Bileviciute-Ljungar I, Norrefalk J, Borg K. Improved functioning and activity according to the International Classification of Functioning and Disability after multidisciplinary telerehabilitation for post-COVID-19 condition—a randomized control study. J Clin Med 2024;13(4):970 [FREE Full text] [doi: 10.3390/jcm13040970] [Medline: 38398284]
45. Calvo-Paniagua J, Díaz-Arribas MJ, Valera-Calero JA, Ramos-Sánchez M, Fernández-de-Las-Peñas C, Navarro-Santana MJ, et al. Educational, exercise, and occupational therapy-based telerehabilitation program versus "Wait-and-See" for

XSL•FO
RenderX

improving self-perceived exertion in patients with post-COVID fatigue and dyspnea: a randomized clinical trial. Am J Phys Med Rehabil 2024;103(9):797-804. [doi: 10.1097/PHM.0000000000002441] [Medline: 38320238]
46. Del Corral T, Fabero-Garrido R, Plaza-Manzano G, Fernández-de-Las-Peñas C, Navarro-Santana M, López-de-Uralde-Villanueva I. Home-based respiratory muscle training on quality of life and exercise tolerance in long-term post-COVID-19: randomized controlled trial. Ann Phys Rehabil Med 2023;66(1):101709 [FREE Full text] [doi: 10.1016/j.rehab.2022.101709] [Medline: 36191860]
47. Demir C, Aksoy CC, Gokmen GY, Durmaz D. Effect of telerehabilitation on post-COVID-19 individuals with long-term dyspnea: a randomized controlled study. J Telemed Telecare 2026;32(3):220-235. [doi: 10.1177/1357633X251333903] [Medline: 40325938]
48. Elyazed TIA, Alsharawy LA, Salem SE, Helmy NA, El-Hakim AAEA. Effect of home-based pulmonary rehabilitation on exercise capacity in post COVID-19 patients: a randomized controlled trail. J Neuroeng Rehabil 2024;21(1):40 [FREE Full text] [doi: 10.1186/s12984-024-01340-x] [Medline: 38528512]
49. de Araújo Furtado PL, do Socorro Brasileiro-Santos M, de Mello BLC, Andrade Araújo A, da Silva MAS, Arielly Suassuna J, et al. The effect of telerehabilitation on physical fitness and depression/anxiety in post-COVID-19 patients: a randomized controlled trial. Int J Telerehabil 2023;15(1):e6560 [FREE Full text] [doi: 10.5195/ijt.2023.6560] [Medline: 38046546]
50. King MF, Byrne A, Denehy L, Graham PL, Douglas B, de Toni P, et al. Feasibility of a group-based telerehabilitation intervention for post-COVID condition management. Phys Med Rehabil 2023;2(1):3152. [doi: 10.31532/physmedrehabil.2.1.005]
51. Kerling A, Beyer S, Dirks M, Scharbau M, Hennemann A, Dopfer-Jablonka A, et al. Effects of a randomized-controlled and online-supported physical activity intervention on exercise capacity, fatigue and health related quality of life in patients with post-COVID-19 syndrome. BMC Sports Sci Med Rehabil 2024;16(1):33 [FREE Full text] [doi: 10.1186/s13102-024-00817-5] [Medline: 38308307]
52. León-Herrera S, Oliván-Blázquez B, Sánchez-Recio R, Méndez-López F, Magallón-Botaya R, Sánchez-Arizcuren R. Effectiveness of an online multimodal rehabilitation program in long COVID patients: a randomized clinical trial. Arch Public Health 2024;82(1):159 [FREE Full text] [doi: 10.1186/s13690-024-01354-w] [Medline: 39294767]
53. Longobardi I, Goessler K, de Oliveira Júnior GN, Prado DMLD, Santos JVP, Meletti MM, et al. Effects of a 16-week home-based exercise training programme on health-related quality of life, functional capacity, and persistent symptoms in survivors of severe/critical COVID-19: a randomised controlled trial. Br J Sports Med 2023;57(20):1295-1303. [doi: 10.1136/bjsports-2022-106681] [Medline: 37164620]
54. Li J, Xia W, Zhan C, Liu S, Yin Z, Wang J, et al. A telerehabilitation programme in post-discharge COVID-19 patients (TERECO): a randomised controlled trial. Thorax 2022;77(7):697-706 [FREE Full text] [doi: 10.1136/thoraxjnl-2021-217382] [Medline: 34312316]
55. McNarry MA, Berg RMG, Shelley J, Hudson J, Saynor ZL, Duckers J, et al. Inspiratory muscle training enhances recovery post-COVID-19: a randomised controlled trial. Eur Respir J 2022;60(4):2103101 [FREE Full text] [doi: 10.1183/13993003.03101-2021] [Medline: 35236727]
56. McGregor G, Sandhu H, Bruce J, Sheehan B, McWilliams D, Yeung J, et al. Clinical effectiveness of an online supervised group physical and mental health rehabilitation programme for adults with post-covid-19 condition (REGAIN study): multicentre randomised controlled trial. BMJ 2024;384:e076506 [FREE Full text] [doi: 10.1136/bmj-2023-076506] [Medline: 38325873]
57. Okan F, Okan S, Duran Yücesoy F. Evaluating the efficiency of breathing exercises via telemedicine in post-Covid-19 patients: randomized controlled study. Clin Nurs Res 2022;31(5):771-781. [doi: 10.1177/10547738221097241] [Medline: 35485738]
58. Philip KEJ, Owles H, McVey S, Pagnuco T, Bruce K, Brunjes H, et al. An online breathing and wellbeing programme (ENO Breathe) for people with persistent symptoms following COVID-19: a parallel-group, single-blind, randomised controlled trial. Lancet Respir Med 2022;10(9):851-862 [FREE Full text] [doi: 10.1016/S2213-2600(22)00125-4] [Medline: 35489367]
59. Paneroni M, Scalvini S, Perger E, Zampogna E, Govetto S, Oliva FM, et al. Home-based exercise program for people with residual disability following hospitalization for COVID-19: randomized control trial. Ann Phys Rehabil Med 2024;67(2):101815 [FREE Full text] [doi: 10.1016/j.rehab.2023.101815] [Medline: 38479344]
60. Rodriguez-Blanco C, Bernal-Utrera C, Anarte-Lazo E, Gonzalez-Gerez JJ, Saavedra-Hernandez M. A 14-day therapeutic exercise telerehabilitation protocol of physiotherapy is effective in non-hospitalized post-COVID-19 conditions: a randomized controlled trial. J Clin Med 2023;12(3):776 [FREE Full text] [doi: 10.3390/jcm12030776] [Medline: 36769425]
61. Samper-Pardo M, León-Herrera S, Oliván-Blázquez B, Méndez-López F, Domínguez-García M, Sánchez-Recio R. Effectiveness of a telerehabilitation intervention using ReCOVery APP of long COVID patients: a randomized, 3-month follow-up clinical trial. Sci Rep 2023;13(1):7943 [FREE Full text] [doi: 10.1038/s41598-023-35058-y] [Medline: 37193738]
62. da Silva MMC, Viana DR, Colucci MG, Gonzaga LA, Arcuri JF, Frade MCM, et al. Effects of a cardiopulmonary telerehabilitation using functional exercises in individuals after COVID-19 hospital discharge: a randomized controlled trial. J Telemed Telecare 2025;31(3):311-319. [doi: 10.1177/1357633X231188394] [Medline: 37559399]

XSL•FO
RenderX

63. Lai CY, Lin CH, Chao TC, Lin CH, Chang CC, Huang CY, et al. Effectiveness of a 12-week telerehabilitation training in people with long COVID: a randomized controlled trial. Ann Phys Rehabil Med 2024;67(5):101853 [FREE Full text] [doi: 10.1016/j.rehab.2024.101853] [Medline: 38824899]
64. Yasacı Z, Mustafaoglu R, Ozgur O, Kuveloglu B, Esen Y, Ozmen O, et al. Virtual recovery: efficacy of telerehabilitation on dyspnea, pain, and functional capacity in post-COVID-19 syndrome. Ir J Med Sci 2025;194(2):631-640. [doi: 10.1007/s11845-025-03899-3] [Medline: 39920519]
65. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
66. Cohen J. A power primer. Psychol Bull 1992;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
67. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
68. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol 2014;14:25 [FREE Full text] [doi: 10.1186/1471-2288-14-25] [Medline: 24548571]
69. Lau J, Ioannidis JPA, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ 2006;333(7568):597-600 [FREE Full text] [doi: 10.1136/bmj.333.7568.597] [Medline: 16974018]
70. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. rating the quality of evidence. J Clin Epidemiol 2011;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]
71. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64(4):383-394. [doi: 10.1016/j.jclinepi.2010.04.026] [Medline: 21195583]
72. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). J Clin Epidemiol 2011;64(4):407-415. [doi: 10.1016/j.jclinepi.2010.07.017] [Medline: 21247734]
73. Kulig M, Perleth M, Langer G, Meerpohl JJ, Gartlehner G, Kaminski-Hartenthaler A, et al. GRADE guidelines: 6. Rating the quality of evidence: imprecision. Z Evid Fortbild Qual Gesundhwes 2012;106(9):677-688. [doi: 10.1016/j.zefq.2012.10.016] [Medline: 23200212]
74. Nolting A, Perleth M, Langer G, Meerpohl JJ, Gartlehner G, Kaminski-Hartenthaler A, et al. GRADE leitlinien: 5. einschätzung der qualität der evidenz—publikationsbias [GRADE guidelines: 5. Rating the quality of evidence: publication bias]. Z Evid Fortbild Qual Gesundhwes 2012;106(9):670-676. [doi: 10.1016/j.zefq.2012.10.015] [Medline: 23200211]
75. Perleth M, Langer G, Meerpohl JJ, Gartlehner G, Kaminski-Hartenthaler A, Schünemann HJ. GRADE leitlinien: 7. einschätzung der qualität der evidenz—inkonsistenz [GRADE guidelines: 7. Rating the quality of evidence—inconsistency]. Z Evid Fortbild Qual Gesundhwes 2012;106(10):733-744. [doi: 10.1016/j.zefq.2012.10.018] [Medline: 23217727]
76. Wang Y, Liu X, Wu Q, Zhu Q, Zhang M. Impact of telerehabilitation on rehabilitation efficacy and patient satisfaction after knee surgery: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2025;27:e76844 [FREE Full text] [doi: 10.2196/76844] [Medline: 41423944]
77. Li S, Dai B, Hou Y, Zhang L, Liu J, Hou H, et al. Effect of pulmonary rehabilitation for patients with long COVID-19: a systematic review and meta-analysis of randomized controlled trials. Ther Adv Respir Dis 2025;19:17534666251323482 [FREE Full text] [doi: 10.1177/17534666251323482] [Medline: 40083165]
78. Timoteo EF, Silva DF, Oliveira TMD, José A, Malaguti C. Real-time telerehabilitation for chronic respiratory disease and post-COVID-19: a systematic review and meta-analysis. J Telemed Telecare 2025;31(7):991-1004. [doi: 10.1177/1357633X241241572] [Medline: 38594927]
79. Santacaterina F, Campagnola B, Bressi F, Zollo L, Sterzi S, Bravi M. Telerehabilitation in postoperative breast cancer care: systematic review and meta-analysis. JMIR Rehabil Assist Technol 2025;12:e77161 [FREE Full text] [doi: 10.2196/77161] [Medline: 41364915]
80. Simmich J, Ross MH, Russell T. Real-time video telerehabilitation shows comparable satisfaction and similar or better attendance and adherence compared with in-person physiotherapy: a systematic review. J Physiother 2024;70(3):181-192 [FREE Full text] [doi: 10.1016/j.jphys.2024.06.001] [Medline: 38879432]
81. Tan C, Meng J, Dai X, He B, Liu P, Wu Y, et al. Effects of therapeutic interventions on long COVID: a meta-analysis of randomized controlled trials. EClinicalMedicine 2025;87:103412 [FREE Full text] [doi: 10.1016/j.eclinm.2025.103412] [Medline: 40808967]
82. Kubota T, Kuroda N, Sone D. Neuropsychiatric aspects of long COVID: a comprehensive review. Psychiatry Clin Neurosci 2023;77(2):84-93 [FREE Full text] [doi: 10.1111/pcn.13508] [Medline: 36385449]
83. Tutsoy O, Polat A, Colak S, Balikci K. Development of a multi-dimensional parametric model with non-pharmacological policies for predicting the COVID-19 pandemic casualties. IEEE Access 2020;8:225272-225283 [FREE Full text] [doi: 10.1109/ACCESS.2020.3044929] [Medline: 34812374]

XSL•FO
RenderX

84. Shambushankar AK, Jose J, Gnanasekaran S, Kaur G. Cost-effectiveness of telerehabilitation compared to traditional in-person rehabilitation: a systematic review and meta-analysis. Cureus 2025;17(2):e79028. [doi: 10.7759/cureus.79028] [Medline: 40099085]

Abbreviations

GRADE: Grading of Recommendations Assessment, Development and Evaluation
HADS: Hospital Anxiety and Depression Scale
HKSJ: Hartung-Knapp-Sidik-Jonkman
HRQoL: health-related quality of life
MD: mean difference
MoCA: Montreal Cognitive Assessment
PCC: post–COVID-19 condition
PI: prediction interval
PICOS: Population, Intervention, Comparison, Outcome, and Study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
RoB 2: Cochrane Risk of Bias 2 tool
6MWT: 6-minute walk test
SF-36: 36-Item Short Form Health Survey
SMD: standardized mean difference
STS: sit-to-stand test

Edited by S Munce; submitted 15.Mar.2026; peer-reviewed by K Longlalerng, O Tutsoy; comments to author 19.Jun.2026; revised version received 23.Jul.2026; accepted 27.Jul.2026; published 27.Oct.2026.

Please cite as:
Huang X, Tang Y, Liu D, Burns RD, Byun W
Effects of Telerehabilitation on Quality of Life, Physical Capacity, and Symptom Burden Within the Post–COVID-19 Condition: Systematic Review and Meta-Analysis of Randomized Controlled Trials
JMIR Rehabil Assist Technol 2026;13:e95368
URL: https://rehab.jmir.org/2026/1/e95368
doi:10.2196/95368
PMID:

©Xiaojing Huang, Yidan Tang, Dongsen Liu, Ryan Donald Burns, Wonwoo Byun. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Role of Medical Caption Technology to Support Physician-Patient Communication for Patients With Hearing Loss: Mixed Methods Pilot Study

Sarah E Hughes[1*], BA; Liang-Yuan Wu[2*], MS; Lindsay J Ma[1], MD; Dhruv Jain[2], PhD; Michael M McKee[3,4], MD, MPH

[1]University of Michigan Medical School, Ann Arbor, MI, United States

[2]Department of Computer Science and Engineering, University of Michigan, Ann Arbor, MI, United States

[3]Department of Family Medicine, University of Michigan, 1018 Fuller St., Ann Arbor, MI, United States

[4]Department of Physical Medicine and Rehabilitation, University of Michigan Medical School, Ann Arbor, MI, United States

[*]these authors contributed equally

Corresponding Author:
Michael M McKee, MD, MPH
Department of Family Medicine, University of Michigan, 1018 Fuller St., Ann Arbor, MI, United States

Abstract

Background: Speech recognition technology is widely used by individuals who are Deaf/deaf and hard-of-hearing (DHH) in everyday communication, but its clinical applications remain underexplored. Communication barriers in health care can compromise safety, understanding, and autonomy for individuals who are DHH.

Objective: This study aimed to evaluate a real-time speech recognition system (SRS) tailored for clinical settings, examining its usability, perceived effectiveness, and transcription accuracy among users who are DHH.

Methods: We conducted a pilot study with 10 adults who are DHH participating in mock outpatient encounters using a custom SRS powered by Google's speech-to-text application programming interface. We used a convergent parallel mixed-methods design, collecting quantitative usability ratings and qualitative interview data during the same study session. These datasets were subsequently merged and jointly interpreted. Participants completed postscenario surveys and structured exit interviews assessing distraction, trust, ease of use, satisfaction, and emotional response. Caption accuracy was benchmarked against professional communication access real-time translation transcripts using word error rate (WER). Because WER assigns equal weight to all tokens, it does not differentiate between routine transcription errors and those involving safety-critical clinical terms (eg, medications or diagnoses). Therefore, WER may underestimate the potential impact of certain errors in medical contexts.

Results: Across 29 clinical scenario simulations, 86% (25/29) of participants found captions nondistracting, 90% (26/29) reported them easy to follow and trustworthy, and 76% (22/29) were satisfied with the experience. Participants described the SRS as intuitive, emotionally grounding, and preferable to lip reading in masked settings. WER ranged from 12.7% to 22.8%, consistent with benchmarks for automated SRSs. Interviews revealed themes of increased confidence in following clinical conversations and staying engaged despite masked communication. Participants reported less anxiety about missing critical medical information and expressed a strong interest in expanding the tool to real-world settings, especially for older adults or those with cognitive impairments.

Conclusions: Our findings support the potential of real-time captioning to enhance accessibility and reduce the cognitive and mental burden of communication for individuals who are DHH in clinical care. Participants described the SRS as both functionally effective and personally empowering. While accuracy for complex medical terminology remains a limitation, participants consistently expressed trust in the system and a desire for its integration into clinical care. Future research should explore real-world implementation, domain-specific optimization, and the development of user-centered evaluation metrics that extend beyond transcription fidelity to include trust, autonomy, and communication equity.

(JMIR Rehabil Assist Technol 2026;13:e79073) doi:10.2196/79073

KEYWORDS

health communication; hearing loss; deafness; speech recognition software; usability testing; health care accessibility

XSL•FO
RenderX

Introduction

Effective communication is foundational to safe, equitable, and high-quality health care [1]. However, individuals who are Deaf/deaf and hard-of-hearing (DHH) often face communication barriers that compromise understanding and autonomy [2]. These barriers contribute to poor health outcomes and reduced patient engagement in real-time clinical settings [2]. The scale of this issue highlights the need to understand which communication support tools are available and provided, and to whom. In the United States, an estimated 48 million people live with some degree of hearing loss (HL), and 1 in 3 adults older than 65 years experiences disabling age-related hearing loss [3,4]. Despite this growing population, access to communication supports remains inconsistent [5,6].

Deaf individuals who use American Sign Language often receive interpreter services [7]. In contrast, oral communicators with people with HL who normally rely on spoken English are less likely to receive accommodations such as captioning, assistive listening devices, or environmental modifications [7]. Especially in clinical workflows, interpreter services are systematically implemented, whereas accommodations for oral communicators are likely not [8-12]. This gap persists despite longstanding mandates under the Americans with Disabilities Act, which mandates effective communication in health care [13]. As a result, many patients who are DHH still receive incomplete or delayed health information [5]. These gaps undermine informed decision-making, autonomy, and overall care outcomes [14,15]. Far from logistical oversights, these structural inequities perpetuate persistent disparities in care for individuals who are DHH.

These long-standing disparities became even more visible during the COVID-19 pandemic [14]. Universal masking eliminated lip reading and facial cues, which were essential supports for many individuals who are DHH and rely on oral communication [16]. This shift underscored the need for scalable solutions to maintain accessible communication in high-stakes settings [14,17].

Real-time captioning is 1 solution for improving communication access for individuals who are DHH when traditional strategies (eg, lip reading or interpreters) are unavailable [18]. Captioning tools can be deployed quickly and readily support both in-person and virtual communication [19]. However, captioning accuracy of clinical conversations may be affected by terminology unique to the medical field or speaker attribution and is understudied [19,20]. This has left a critical gap in the development of effective and equitable access tools.

By allowing both conversation partners to see each other's faces while reading the same captions, transparent or dual-visibility captioning preserves the natural flow of spoken interaction and is a promising solution for clinical communication. Prior work, such as See-Through Captions [21], See-Through Captions in a Museum Guided Tour [22], and Wearable Subtitles [23], has primarily focused on general or educational settings. Our study extends this line of research into medical contexts, where communication accuracy can directly affect patient safety and outcomes. It also emphasizes the emotional and psychological impact of captioning during clinical interactions and addresses the unique technical challenges posed by medical vocabulary and workflow integration.

In summary, we developed and evaluated a real-time captioning tool using Google's speech-to-text engine to generate live captions during simulated clinical encounters. We tested this system in dynamic, medically relevant scenarios designed to simulate typical ambulatory care encounters. In this pilot study, we explored how individuals who are DHH experienced the captioning system in these simulated encounters, focusing on usability, accuracy, and communication access.

Methods

Background

The pilot took place in a patient room at one of the Department of Family Medicine clinics. The primary goal was to assess the feasibility and acceptability of a real-time captioning tool in a clinical setting. Secondary objectives included evaluating ease of use, distraction, trust, and satisfaction, factors critical to determining whether the tool supports communication access. Quantitative and qualitative data were collected concurrently within the same study session using a convergent parallel mixed-methods design. Participants completed postscenario surveys and a brief structured exit interview during the same visit, allowing us to analyze both datasets in parallel before merging findings during the interpretation phase.

Recruitment

We recruited participants who self-identified as DHH through internal email lists compiled from prior studies, social media, and snowball sampling. Inclusion criteria included people who were DHH, preferred to communicate in spoken English, and were at least 18 years old. Recruitment materials explained that the study evaluated a real-time captioning system in simulated medical scenarios.

Mock Clinical Scenarios

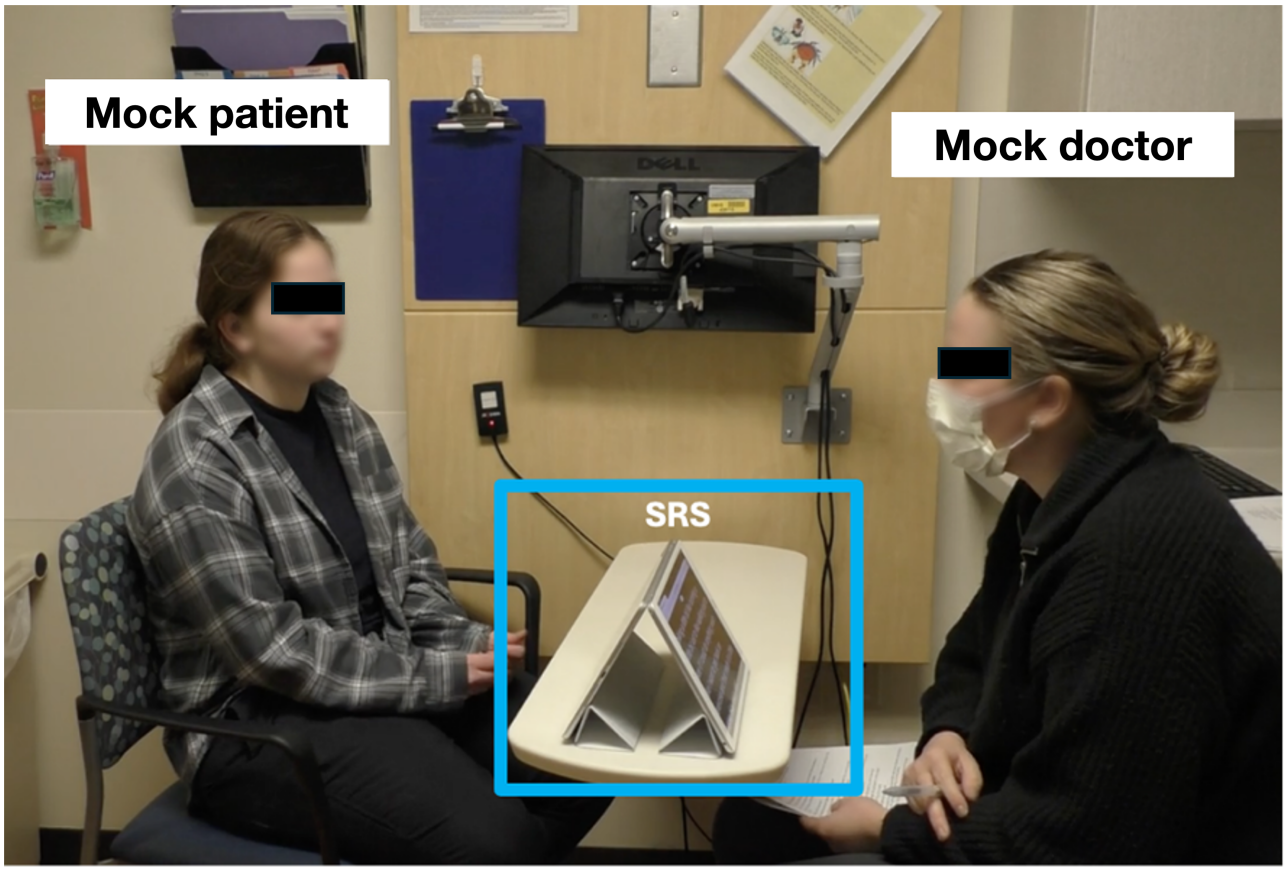
Participants completed 3 mock clinical scenarios using the automated speech recognition system (SRS) which was developed by us. The SRS used Google's speech-to-text application programming interface to transcribe speech to text with low latency and competitive accuracy [24]. The setup included 2 iPads arranged in a tented position so that each device faced either the participant or the mock doctor. Both iPads displayed the generated captions simultaneously (Figure 1).

Before each experiment, we used a random number generator to assign scenario order for each participant. Two team members (both medical students) alternated between serving as the mock doctor (administering scenarios) or facilitator (administering postscenario surveys and exit interviews).

The scenarios were based on commonly reported primary care concerns: (1) back pain, (2) headache, and (3) high blood pressure. Scenario scripts were designed by trained medical students and a clinical faculty member to closely replicate real clinical conversations. The mock doctors wore surgical masks to simulate real-life communication barriers, such as muffled sound and loss of visual cues.

XSL•FO
RenderX

Figure 1. An example of a mock clinical scenario with the real-time speech recognition system set up on a table between the participant (left) and the mock doctor (right). A microphone on the iPad facing the mock doctor detects audio during interviews. Transcripts are displayed on both iPads in real time. SRS: speech recognition system.



Postscenario SRS Assessments

Following each scenario, participants provided feedback on the captioning system, rating it across 4 domains: distraction, ease of use, trust, and overall satisfaction (Multimedia Appendix 1). Scenario-specific questions included: “In your discussion with the mock doctor, how distracting were the captions?” “How easy or difficult was it to watch the caption while talking with the mock doctor?” “How much did you trust the accuracy of the generated captions?” “In this scenario, how satisfied were you with the captioning technology?” To reduce response bias, we alternated the direction of the scales: ease of use and trust rated from 1 (strong agreement) to 5 (strong disagreement), and satisfaction rated from 1 (strong disagreement) to 5 (strong agreement). Distraction was scaled separately from 1 (strong disagreement) to 3 (strong agreement).

Participant Survey Questions

To evaluate user experience with the SRS, participants completed a structured exit interview consisting of 9 questions (5 scalar and 4 open-ended items; Multimedia Appendix 1). To ensure accessibility, a study team member read all questions aloud while they were displayed on an iPad (Apple Inc). We audio-recorded and transcribed responses verbatim using a third-party service, then deidentified the transcripts. We reviewed audio files to clarify unclear segments. Given the brief interviews, we organized and analyzed responses in Microsoft Excel (version 16.77).

Open-ended responses were reviewed using a structured framework aligned with predefined domains: ease of use, comfort, satisfaction, trust, emotional response, and the captioning system’s ability to support or replace lip reading. Overall, 3 team members (SEH, LJM, and LW) independently applied initial codes to a subset of transcripts. Coding discrepancies were resolved through discussion, and the codebook was refined iteratively. Consistency was maintained through regular team meetings, and reflexive discussions were used to address potential bias.

Themes were identified based on frequency, relevance to study aims, and salience across participants. Representative participant comments were selected to illustrate key insights. Thematic saturation was reached when no new concepts emerged from successive interviews.

Mixed Methods Integration

To integrate quantitative and qualitative data, we used a convergent parallel approach in which both datasets were collected during the same phase, analyzed separately, and then merged during the interpretation phase. Integration occurred through (1) narrative weaving of findings across domains and (2) construction of a joint display that juxtaposed quantitative ratings with representative qualitative insights to generate meta-inferences. This approach allowed identification of areas of convergence and divergence between usability ratings and participants’ lived communication experiences.

XSL•FO
RenderX

Closed Captioning Accuracy

In addition to participant feedback, we analyzed the accuracy of the system's transcriptions. We compiled all transcripts generated by the mock doctors and compared them to professional communication access real-time translation transcripts.

We used word error rate (WER), a standard metric in automatic speech recognition (ASR) that calculates errors as the ratio of insertions, deletions, and substitutions required to align the system output with the reference [25,26]. We implemented WER calculations using the Python-based *jiwer* library, which provides standardized scoring for automated SRSs. This approach allowed us to assess how closely the SRS-generated captions matched professional-level transcription, validating the system's effectiveness in realistic use cases.

Statistical Analysis

We performed univariate analyses on demographic data and postscenario survey responses. Because of the small sample size, the study was not powered to detect subgroup differences. For transcript analysis, we segmented transcripts from mock sessions into 3 distinct scenarios. To focus on the primary use case, captioning clinician speech, we excluded utterances from participants who are DHH and analyzed only the mock doctors' speech.

Ethical Considerations

This study was approved by the University of Michigan Institutional Review Board (IRB; HUM00240244). All participants provided informed consent prior to participation. Participants were informed of the study purpose, procedures, potential risks, and their right to withdraw at any time without penalty. All study data were deidentified prior to analysis, and transcripts were reviewed to remove personally identifiable information. Audio recordings and transcripts were stored on secure, password-protected institutional servers accessible only to the study team. Participants received a US $25 Amazon gift card for their participation. The individuals depicted in the figure provided explicit written consent for publication of their images. The individuals shown in Figure 1 provided explicit written consent for their images to be published.

Results

Participant Characteristics

Overall, 11 participants who are DHH enrolled and participated in the pilot study. Due to equipment failure resulting in complete data recording loss with Participant 5, this participant was excluded from the analysis. The 10 remaining participants had an even distribution of genders (Table 1).

Table . Study participant demographics.

ID	Age (years)	Sex	Identity	Hearing loss levels[a]	Wearable technology[b]	Lip reader
P01	61	Female	HoH[c]	Severe	Yes	All of the time
P02	66	Male	HoH	Severe	Yes	Sometimes
P03	66	Male	HoH	Severe	Yes	No
P04	66	Female	HoH	Moderately severe	Yes	Sometimes
P06	43	Female	Deaf	Profound	Yes	All of the time
P07	21	Female	deaf	Moderately severe	Yes	Sometimes
P08	39	Female	HoH	Severe	Yes	All of the time
P09	24	Male	Deaf	Profound	Yes	No
P10	56	Male	deaf	Profound	Yes	Sometimes
P11	20	Male	HoH	Mild	Yes	Sometimes

[a]Hearing loss levels were self-identified, and all participants reported equal hearing loss levels bilaterally.

[b]Wearable technology includes hearing aids and cochlear implants.

[c]HoH: hard of hearing.

The mean age of the participants was 46.2 (SD 19.3) years. Six participants identified as "hard of hearing," 22 as "Deaf," and 2 as "deaf." Seven participants self-reported severe to profound HL, and all participants had bilateral HL. Five participants self-reported congenital HL, 2 reported childhood onsets of HL (<12 y old), and 2 reported HL as adults (>18 y old). Hearing aids were used by 8 participants, and 2 participants used cochlear implants. Seven participants used captioning services in the past. Seven also incorporated smartphone-based hearing assistive technology. Three used "other" tools, including using cupped hands behind ears to assist in hearing. Eight participants reported varying degrees of dependence on lip reading, but 5 participants depended sometimes on lip reading and 5 depended fully on lip reading.

Postscenario SRS Assessments

There were 29 postscenario SRS assessment surveys, 3 survey responses each from 9 participants and 2 survey responses from 1 participant. One survey response from participant P11 was not collected due to a technician error. Overall, participants found the captioning technology not distracting in 86% (25/29) of scenarios (Table 2). In 90% (26/29) of scenarios, participants

XSL•FO
RenderX

trusted the accuracy of generated transcription and felt the captions were easy to watch while conversing with the mock doctor. In 76% (22/29) of scenarios, participants were satisfied with the captioning technology. The technology was least satisfying to participants in the back pain scenarios (70% satisfaction) compared to the high blood pressure (78% satisfaction) and headache (80% satisfaction) scenarios.

Table . Summary of participant assessments regarding live captioning technology compiled from all 3 scenarios and dichotomized.

Questions[a]	Assessments	Values, n (%)
In your discussion with the mock doctor, how distracting were the captions?	Not distracting[b]	25 (86)
How easy or difficult was it to watch the caption while talking with the mock doctor?	Easy[c]	26 (90)
How much did you trust the accuracy of the generated captions?	Trusted[d]	26 (90)
In this scenario, how satisfied were you with the captioning technology?	Satisfied[e]	22 (76)

[a]For all 4 questions, n=29 since 1 of the 10 participants did not participate in 1 of the 3 scenarios.

[b]Not distracting: not at all distracting.

[c]Easy: very easy + somewhat easy.

[d]Trusted: completely trusted + somewhat trusted.

[e]Satisfied: very satisfied + somewhat satisfied.

Participant Experience Surveys

All 10 participants completed structured exit interviews following the captioning scenarios, providing reflections on their overall experience with the SRS (Table 3). Interview responses were analyzed using a predefined framework aligned with domains explored in the postscenario ratings (eg, ease of use, comfort, satisfaction, trust, emotional impact, and support for lip reading). This section summarizes participant perspectives and provides representative quotes to contextualize the quantitative results described above.

Table . Representative participant reflections by theme.

Themes	Relevant quotes[a]	Interpretation
Ease of use	“At first I wasn’t sure what to expect, but after a few lines of text I stopped even thinking about it—it just worked. That made me feel more in control.” (P04)	Participants found the system intuitive and accessible.
Comfort	“I didn’t have to strain or overthink. It just flowed naturally and I didn’t even realize how relaxed I was until the end.” (P08)	Technology reduced cognitive effort and fostered emotional ease.
Satisfaction	“I was happy. I wish all the doctors would have something like this. It made me feel like my experience mattered.” (P03)	Participant expresses satisfaction and a sense of being valued.
Safety and trust	“Because it’s live, it feels very safe. You’re not left guessing, and I felt confident nothing important was missed.” (P01)	Real-time functionality enhanced user confidence and perception of safety.
Emotional response	“I didn’t realize how much stress I usually carry during appointments. This made me feel heard and like I could finally breathe.” (P09)	System reduced communication-related anxiety and supported emotional well-being.
Support or replace lip reading	“With the mask on, it would have been extremely difficult to follow—and with the captioning, it was just leaps better. I wasn’t exhausted from trying to read lips the whole time.” (P07)	The technology was viewed as a vital alternative to lip reading, especially in masked settings.

[a]Relevant quotes from individual participants illustrating each core theme, including insight into perceived usability, comfort, satisfaction, emotional impact, and the role of real-time captions in supporting communication.

Most participants (9/10) described the system as easy to use, frequently using phrases like “very easy” or “easier than usual.” One participant remarked,

XSL•FO
RenderX

> *After a few lines of text I stopped even thinking about it—it just worked. That made me feel more in control.*

Another noted,

> *It was easier than usual because we don't have captioning. It's always nice to have it just in case you miss something.*

Participants also reported high comfort with the system. Descriptions included "very comfortable," "easy to work with," and "high comfortability."

As 1 participant shared,

> *It just flowed naturally, and I didn't even realize how relaxed I was until the end.*

Satisfaction was also high across interviews. While 76% of scenario ratings reflected satisfaction, all participants described themselves as satisfied or very satisfied in exit interviews. One stated,

> *I was happy. I wish all the doctors would have something like this.*"

Another shared,

> *I was pretty satisfied, and the captioning was spot-on.*

When asked about trust in the system, participants frequently described the captions as reliable. One participant reflected,

> *Because it's live, it feels very safe. You're not left guessing.*

A few raised questions about data privacy, with one noting,

> *I would also want to know what happens to the transcript and who has access to it.*

Participants also described emotional benefits from the technology. In total, 9 of 10 participants used words like "reassured," "relaxed," and "comfortable" to describe how the SRS made them feel. One participant shared,

> *This made me feel heard and like I could finally breathe.*

Perceptions of the captions' ability to support or replace lip reading were more varied. Several participants described the system as a helpful supplement or improvement, particularly in masked settings. As one noted,

> *With the mask on, I definitely depended on it more.*

Another stated,

> *I think it's better than lip reading.*

Others expressed that lip reading remained important, with one participant saying,

> *Not going to replace lip reading... captions help, but I still rely on visual cues.*

Beyond these predefined domains, participants spontaneously shared reflections on broader applications of the SRS. Several expressed enthusiasm for expanding its use in real-world clinical settings, with one stating,

> *I would like to see that in many doctor's offices tomorrow.*

Others suggested the system may be particularly helpful for patients who are older, have cognitive impairments, or use interpreters. A few noted that having real-time captions reduced the pressure to maintain constant visual attention, allowing for more natural communication and less fatigue.

Closed Captioning Accuracy

We collected and preprocessed transcripts from 10 mock clinical sessions. Due to varying levels of verbosity among the participants, the total transcript lengths varied substantially, ranging from 1144 to 4704 words.

Overall, participants found the SRS to be sufficiently accurate (Table 4). For instance, P04 noted that the system was *"more accurate than the phone captions"* she typically uses in daily conversations. Similarly, P06 commented on the system's effectiveness compared to human captioners, stating,

> *A lot of the captions I had were court reporters—they caption fast, but sometimes they make mistakes. ... And this one [the SRS], it's more accurate and I see words better.*

Nonetheless, participants expressed concerns about the system's ability to handle more complex or specialized medical vocabulary. For example, P10 questioned *"how it would be with more complex medical terminologies,"* in real clinical settings where more technical jargon and medication names were frequently used.

XSL•FO
RenderX

Table . The word error rate for each scenario, along with the accumulated word error rate for each participant across all 3 scenarios. These word error rate scores specifically reflect the accuracy of the automated speech recognition system in transcribing the mock doctors' speech.

ID	Mock doctor	Scenario 1[a]	Scenario 2[a]	Scenario 3[a]	Accumulated (range: 0.127-0.167)
P01	M1	0.136	0.133	0.125	0.131
P02	M2	0.193	0.141	0.133	0.153
P03	M1	0.129	0.122	0.133	0.127
P04	M2	0.228	0.128	0.151	0.167
P06	M2	0.137	0.152	0.135	0.144
P07	M1	0.127	0.134	0.132	0.133
P08	M2	0.155	0.142	0.152	0.149
P09	M1	0.136	0.131	0.141	0.137
P10	M1	0.127	0.126	0.133	0.129
P11	M2	0.147	0.185	0.134	0.151

[a]The scenario-level word error rates ranged between 0.122 and 0.228.

Joint Display of Integrated Findings

To illustrate convergence between quantitative usability ratings and qualitative interview themes, we constructed a joint display summarizing merged findings and resulting meta-inferences across key domains (Table 5).

Table . Joint display of integrated quantitative and qualitative findings

Domain	Quantitative result	Representative quote	Integrated meta-inference
Ease of use	90% rated captions "easy"	"After a few lines of text I stopped even thinking about it—it just worked."	High usability with minimal cognitive load Captions supported natural conversational flow
Comfort	Not directly measured	"It just flowed naturally, and I didn't realize how relaxed I was."	Technology reduced strain and fostered emotional ease during communication
Satisfaction	76% satisfied	"I wish all the doctors would have something like this."	Satisfaction tied to both functional value and feeling understood and supported
Safety and trust	90% trusted accuracy	"Because it's live, it feels very safe. You're not left guessing."	Real-time display strengthened perceived safety and reliability despite minor errors
Emotional response	Not directly measured	"This made me feel heard and like I could finally breathe."	Captions enhanced psychological safety and reduced anxiety—benefits not captured numerically
Support or replace lip reading	Not directly measured	"With the mask on, I depended on it more… it was leaps better."	Captions supplemented or replaced lip reading, reducing fatigue in masked settings

Discussion

Principal Results

To successfully deploy SRS in clinical settings, it is essential that the system accurately captures and reflects clinicians' speech. Our findings show that although the SRS output was not flawless, its WERs fell between 0.10 and 0.20, a range generally considered acceptable for real-world ASR use [19,26]. Furthermore, participants understood the captions with relative ease, suggesting that transcription quality was sufficient to support comprehension in simulated outpatient scenarios. However, stricter accuracy standards may be required in high-stakes contexts, such as discussions of medications or treatment options, where small errors can have serious consequences.

Although WER is widely used to evaluate ASR performance, it weighs all error types equally, regardless of their impact on comprehension [27]. Prior work has proposed alternative evaluation approaches that aim to capture semantic accuracy or user-centric measures of intelligibility and usefulness [20]. In clinical communication, we support developing evaluation metrics that align more closely with safety-critical requirements. Such metrics would be instrumental in determining when ASR

XSL•FO
RenderX

systems are truly ready for deployment in health care environments. In clinical settings, misrecognition of medical terminology can have consequences far more serious than common transcription errors, especially when involving medication names, diagnoses, or treatment instructions. Because of this, future work should consider safety-critical evaluation frameworks that go beyond traditional WER. Approaches, such as semantic error analysis, comprehension-based scoring, or accuracy, weighting for medically significant terms could better capture the real-world implications of captioning errors in health care communication.

Our participants represented a variety of ages, genders, HL levels, and degrees of dependence on lip reading. However, most participants had previously used captioning technology as an accommodation, so our usability findings may be less generalizable to individuals who are DHH with no prior captioning experience. Also, only 2 participants preferred written communication with hearing people. Therefore, satisfaction with our captioning technology may be higher than our results suggest for people who are DHH and depend more on written communication. Regardless of the scenario, most participants were satisfied with the SRS, trusted its accuracy, found it easy to watch, and were not distracted.

Participants trusted the captioning system despite occasional transcription errors, which embodies the concept of trust-in-automation frameworks, where user reliance is shaped by perceived system reliability and predictability [28]. Exit interviews revealed that beyond meeting technical expectations, the captioning system also meaningfully supported emotional connection, trust, and autonomy during clinical interactions. Participants described the captions as easy to use and grounding. They also noted reduced stress, lower cognitive fatigue, better understanding, and a stronger sense of being heard. Encouragingly, the observed reduction in stress and fatigue is consistent with prior work where assistive technology helped manage cognitive effort during information processing [29,30]. These findings suggest that accessibility tools should be evaluated not only by their accuracy but by their ability to support psychological safety and communication equity [31].

Additionally, although participants generally trusted the captioning system, a few raised concerns about transcript privacy and data handling. These concerns highlight the ethical need for transparency when implementing automated captioning in health care. This pilot used secure, locally stored recordings without identifiable data, but clinical deployment will require Health Insurance Portability and Accountability Act (HIPAA)-compliant encryption and explicit consent protocols. Adding user controls, such as options to delete transcripts or disable storage, could further strengthen trust among users who are DHH and other vulnerable populations. Nevertheless, participants recommended broader adoption of SRS, particularly for older adults and others facing progressive hearing-related communication barriers, underscoring the system's potential to improve care for a heterogeneous population of DHH patients.

Limitations

While our findings are promising, this study has several limitations. Most participants were experienced caption users and had prior familiarity with assistive communication technologies, which may have positively influenced usability and satisfaction ratings. As a result, these findings may not fully represent the experiences of individuals who are DHH and are less familiar with captioning or other accessibility tools or primary American Sign Language users. Future research should include participants with varying levels of captioning experience and a broader demographic range to better assess generalizability and identify barriers for first-time users. This study was conducted in controlled, simulated settings, which may not fully reflect the complexity and spontaneity of real-world medical encounters. Because these mock scenarios involved medical students rather than practicing clinicians, the communication dynamics may differ from authentic physician–patient interactions. Future work should therefore include real-world clinical deployments to evaluate how captioning systems perform in active care settings and adapt to diverse communication styles and environmental conditions.

Second, although our participant pool included individuals with diverse hearing identities and varying degrees of familiarity with assistive technologies, it does not capture the full range of experiences within the broader community who are DHH. Future work should include longitudinal application in various clinical settings and recruitment of a more diverse participant population to better assess long-term usability and impact.

In addition, our SRS was not specifically optimized for medical vocabulary. This limitation was evident in the system's tendency to misrecognize medical terminology, words that are infrequent in everyday speech yet crucial for accurate clinical communication. Furthermore, while we used WER as a standard quantitative evaluation metric, it does not fully capture how users who are DHH interpret and understand captions, particularly in high-stakes contexts. Future research should explore the development of domain-specific SRS trained on medical speech and adopt evaluation metrics that better reflect comprehension and user experience among individuals who are DHH. Finally, since WER assigns equal weight to all tokens, it does not differentiate between routine transcription errors and those involving safety-critical clinical terms (eg, medications or diagnoses). Therefore, WER may underestimate the potential impact of certain errors in medical contexts.

Future Directions

Improving SRS accuracy for medical terminology remains a key technical priority for clinical use. Strategies may include (1) speech recognition models on deidentified clinical audio to capture the acoustic variability of real-world medical speech [32], (2) embedding domain-specific medical dictionaries and medication name libraries into the language model of the SRS systems to reduce substitution errors [33], (3) leveraging context-aware large language models that can infer meaning from partial or uncertain input [34], and (4) integrating clinician feedback loops for rapid correction of recurring misinterpretations [35]. These enhancements would not only improve accuracy for technical vocabulary but also strengthen user trust and perceived reliability in clinical environments.

Building on these preliminary findings, future work should also explore integration with medical-domain ASR models to

XSL•FO
RenderX

enhance accuracy for specialized terminology and complex clinical dialog. Longitudinal studies will be valuable for assessing maintained usability, user trust, and performance over time. Additionally, testing captioning systems in broader clinical contexts, such as emergency care, geriatrics, and among patients with cognitive impairment, will help determine their adaptability and impact across diverse care settings.

Implications for Clinical Workflow Integration

Our findings demonstrate that real-time captioning is usable and beneficial in clinical settings for patients who are DHH, aligning with prior evidence that captioning improved recall of anesthesia-related consent conversations [36]. Given this demonstrated value, practical integration of captioning tools into clinical workflows will require thoughtful design to minimize disruption while enhancing accessibility. Participants envisioned use cases in which SRS displays could be embedded within existing electronic health record systems or mirrored on clinician tablets to preserve natural eye contact and conversational flow. Integration will also depend on clear institutional protocols for activating captioning on demand, ensuring confidentiality, and providing clinician training on how to engage with patients who are DHH using this technology. Establishing these processes could enable captioning to function as a routine accessibility feature rather than an exception, supporting both efficiency and equitable communication in care delivery.

Conclusions

This pilot study demonstrates that artificial intelligence–enhanced captioning can meaningfully improve communication experiences for individuals who are DHH in clinical settings. Participants found the system intuitive, emotionally supportive, and effective in bridging common communication barriers, especially those worsened by face masks and unfamiliar environments. While traditional captioning tools often fall short in medical contexts, integrating large language models into the speech recognition process offers a promising path toward more coherent, accurate, and human-centered accessibility. By centering on user perspectives, this study highlights the importance of evaluating assistive technologies not only for transcription quality, but for their impact on trust, inclusion, and psychological safety. Future research should build on these early insights to further refine captioning systems, examine their use in real-world clinical care, and ensure that patients who are DHH are active partners in the design of accessible digital health solutions.

Acknowledgments

The authors thank the Deaf/deaf and hard-of-hearing community members who participated in this study for generously sharing their insights and experiences. We are grateful to the Dexter Family Medicine clinic at the University of Michigan for their support in facilitating pilot testing.

Funding

This work was supported by funding from the Blue Cross Blue Shield of Michigan Foundation (2024010111) and by the University of Michigan's e-HAIL (e-Health and Artificial Intelligence Laboratory) program.

Data Availability

All data generated or analyzed during this study are included in this published article. Further inquiries can be directed to the corresponding author.

Authors' Contributions

Methodology (equal), participant recruitment (equal), investigation (equal), data curation (lead), formal analysis (lead), writing – original draft (lead), and writing – review and editing (equal): SEH
Methodology (equal), participant recruitment (equal), technology development (lead), investigation (equal), data curation (equal), formal analysis (equal), writing – original draft (equal), and writing – review and editing (equal): LYW
Methodology (equal), project administration (lead), investigation (equal), data curation (equal), formal analysis (equal), writing – original draft (equal), and writing – review and editing (equal): LJM
Conceptualization (equal), technology development (supporting), methodology (supporting), writing – review and editing (equal), and supervision (equal): DJ
Conceptualization (equal), methodology (equal), writing – original draft (supporting), writing – review and editing (equal), and supervision (lead): MMM

Conflicts of Interest

None declared.

Multimedia Appendix 1
Structured exit interview questions.
[DOCX File, 17 KB - rehab_v13i1e79073_app1.docx]

XSL•FO
RenderX

References

1. Guttman OT, Lazzara EH, Keebler JR, Webster KLW, Gisick LM, Baker AL. Dissecting communication barriers in healthcare: a path to enhancing communication resiliency, reliability, and patient safety. J Patient Saf 2021 Dec 1;17(8):e1465-e1471. [doi: 10.1097/PTS.0000000000000541] [Medline: 30418425]
2. McKee M, James TG, Helm KVT, et al. Reframing our health care system for patients with hearing loss. J Speech Lang Hear Res 2022 Oct 17;65(10):3633-3645. [doi: 10.1044/2022_JSLHR-22-00052] [Medline: 35969852]
3. Quick statistics about hearing, balance, & dizziness. NIH. URL: https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing [accessed 2025-05-26]
4. Culturally affirmative and linguistically accessible services. National Association of the Deaf. URL: https://www.nad.org/resources/health-care-and-mental-health-services/mental-health-services/culturally-affirmative-and-linguistically-accessible-services/ [accessed 2025-05-10]
5. James TG, Coady KA, Stacciarini JMR, et al. "They're not willing to accommodate deaf patients": communication experiences of deaf american sign language users in the emergency department. Qual Health Res 2022 Jan;32(1):48-63. [doi: 10.1177/10497323211046238] [Medline: 34823402]
6. Marlow NM, Samuels SK, Jo A, Mainous AG. Patient-provider communication quality for persons with disabilities: a cross-sectional analysis of the Health Information National Trends Survey. Disabil Health J 2019 Oct;12(4):732-737. [doi: 10.1016/j.dhjo.2019.03.010] [Medline: 30995967]
7. Diamond L, Izquierdo K, Canfield D, Matsoukas K, Gany F. A systematic review of the impact of patient-physician non-english language concordance on quality of care and outcomes. J Gen Intern Med 2019 Aug;34(8):1591-1606. [doi: 10.1007/s11606-019-04847-5] [Medline: 31147980]
8. Barnett S, McKee M, Smith SR, Pearson TA. Deaf sign language users, health inequities, and public health: opportunity for social justice. Prev Chronic Dis 2011 Mar;8(2):A45. [Medline: 21324259]
9. Barnett SL, Matthews KA, Sutter EJ, et al. Collaboration with deaf communities to conduct accessible health surveillance. Am J Prev Med 2017 Mar;52(3 Suppl 3):S250-S254. [doi: 10.1016/j.amepre.2016.10.011] [Medline: 28215374]
10. Pollard RQ, Barnett S. Health-related vocabulary knowledge among deaf adults. Rehabil Psychol 2009 May;54(2):182-185. [doi: 10.1037/a0015771] [Medline: 19469608]
11. Mitra M, Akobirshoev I, McKee MM, Iezzoni LI. Birth outcomes among U.S. women with hearing loss. Am J Prev Med 2016 Dec;51(6):865-873. [doi: 10.1016/j.amepre.2016.08.001] [Medline: 27687529]
12. Alexander A, Ladd P, Powell S. Deafness might damage your health. Lancet 2012 Mar 17;379(9820):979-981. [doi: 10.1016/S0140-6736(11)61670-X] [Medline: 22423872]
13. ADA.gov. URL: https://www.ada.gov [accessed 2025-05-27]
14. McKee M, Moran C, Zazove P. Overcoming additional barriers to care for deaf and hard of hearing patients during COVID-19. JAMA Otolaryngol Head Neck Surg 2020 Sep 1;146(9):781-782. [doi: 10.1001/jamaoto.2020.1705] [Medline: 32692807]
15. Iezzoni LI, O'Day BL, Killeen M, Harker H. Communicating about health care: observations from persons who are deaf or hard of hearing. Ann Intern Med 2004 Mar 2;140(5):356-362. [doi: 10.7326/0003-4819-140-5-200403020-00011] [Medline: 14996677]
16. Moreland CJ, Ruffin CV, Morris MA, McKee M. Unmasked: how the COVID-19 pandemic exacerbates disparities for people with communication-based disabilities. J Hosp Med 2021 Mar;16(3):185-188. [doi: 10.12788/jhm.3562] [Medline: 33617440]
17. Bernard A, Weiss S, Rahman M, et al. The impact of COVID-19 and pandemic mitigation measures on persons with sensory impairment. Am J Ophthalmol 2022 Feb;234:49-58. [doi: 10.1016/j.ajo.2021.06.019] [Medline: 34197781]
18. Kawas S, Karalis G, Wen T, Ladner RE. Improving real-time captioning experiences for deaf and hard of hearing students. Presented at: ASSETS '16: Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 23-26, 2016. [doi: 10.1145/2982142.2982164]
19. Kafle S, Huenerfauth M. Evaluating the usability of automatically generated captions for people who are deaf or hard of hearing. Presented at: ASSETS '17: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 20 to Nov 1, 2017. [doi: 10.1145/3132525.3132542]
20. Favre B, Cheung K, Kazemian S, et al. Automatic human utility evaluation of ASR systems: does WER really predict performance? Presented at: Interspeech 2013; Aug 25-29, 2013 URL: https://www.isca-archive.org/interspeech_2013/favre13_interspeech.pdf [accessed 2025-12-15]
21. Yamamoto K, Suzuki I, Shitara A, Ochiai Y. See-through captions: real-time captioning on transparent display for deaf and hard-of-hearing people. Presented at: ASSETS '21: Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibilit; Oct 18-22, 2021. [doi: 10.1145/3441852.3476551]
22. Suzuki I, Yamamoto K, Shitara A, Hyakuta R, Iijima R, Ochiai Y. See-through captions in a museum guided tour: exploring museum guided tour for deaf and hard-of-hearing people with real-time captioning on transparent display. Presented at: Computers Helping People with Special Needs: 18th International Conference, ICCHP-AAATE 2022; Jul 11-15, 2022. [doi: 10.1007/978-3-031-08648-9_64]

XSL•FO
RenderX

23. Olwal A, Balke K, Votintcev D. Wearable subtitles: augmenting spoken communication with lightweight eyewear for all-day captioning. Presented at: UIST '20: Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology; Oct 20-23, 2020 URL: https://dl.acm.org/doi/proceedings/10.1145/3379337 [accessed 2025-12-15] [doi: 10.1145/3379337.3415817]
24. Speech-to-Text. Google Cloud. URL: https://cloud.google.com/speech-to-text [accessed 2025-05-27]
25. Morris AC, Maier V, Green P. From WER and RIL to MER and WIL: improved evaluation measures for connected speech recognition. Presented at: Interspeech 2004; Oct 4-8, 2004 URL: https://www.isca-archive.org/interspeech_2004/morris04_interspeech.pdf [accessed 2025-12-15] [doi: 10.21437/Interspeech.2004-668]
26. jitsi/jiwer. GitHub. URL: https://github.com/jitsi/jiwer [accessed 2025-05-27]
27. Ye-Yi W, Acero A, Chelba C. Is word error rate a good indicator for spoken language understanding accuracy. Presented at: 2003 IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat No03EX721); Nov 30 to Dec 4, 2003. [doi: 10.1109/ASRU.2003.1318504]
28. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors 2004;46(1):50-80. [doi: 10.1518/hfes.46.1.50_30392] [Medline: 15151155]
29. Sweller J. Cognitive load during problem solving: effects on learning. Cogn Sci 1988 Jun;12(2):257-285. [doi: 10.1016/0364-0213(88)90023-7]
30. Paas F, Renkl A, Sweller J. Cognitive load theory and instructional design: recent developments. Educ Psychol 2003 Jan 1;38(1):1-4. [doi: 10.1207/S15326985EP3801_1]
31. Bickmore TW, Paasche-Orlow MK. The role of information technology in health literacy research. J Health Commun 2012;17 Suppl 3(sup3):23-29. [doi: 10.1080/10810730.2012.712626] [Medline: 23030559]
32. Afonja T, Olatunji T, Ogun S, Etori NA, Owodunni A, Yekini M. Performant ASR models for medical entities in accented speech. Presented at: Interspeech 2024; Sep 1-5, 2024 URL: https://www.isca-archive.org/interspeech_2024/afonja24_interspeech.pdf [accessed 2025-12-15]
33. Michalopoulos G, Wang Y, Kaka H, Chen H, Wong A. UmlsBERT: clinical domain knowledge augmentation of contextual embeddings using the unified medical language system metathesaurus. Presented at: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics; Jun 6-11, 2021 URL: https://aclanthology.org/2021.naacl-main [accessed 2025-12-15] [doi: 10.18653/v1/2021.naacl-main.139]
34. Hsie C, Moreira C, Nobre IB, et al. DALL-m: context-aware clinical data augmentation with llms. arXiv. Preprint posted online on May 1, 2025. [doi: 10.48550/ARXIV.2407.08227]
35. Wu LY, Jain D. EvolveCaptions: empowering DHH users through real-time collaborative captioning. arXiv. Preprint posted online on Oct 2, 2025. [doi: 10.48550/ARXIV.2510.02181]
36. Spehar B, Tye-Murray N, Myerson J, Murray DJ. Real- time captioning for improving informed consent: patient and physician benefits. Reg Anesth Pain Med 2016;41(1):65-68. [doi: 10.1097/AAP.0000000000000347] [Medline: 26650432]

Abbreviations

ASR: automatic speech recognition
DHH: Deaf/deaf or hard of hearing
HIPAA: Health Insurance Portability and Accountability Act
HL: hearing loss
SRS: speech recognition system
WER: word error rate

Edited by S Munce; submitted 14.Jun.2025; peer-reviewed by A Roundtree, SD Ubur; revised version received 21.Nov.2025; accepted 03.Dec.2025; published 15.Jan.2026.

Please cite as:
Hughes SE, Wu LY, Ma LJ, Jain D, McKee MM
Assessing the Role of Medical Caption Technology to Support Physician-Patient Communication for Patients With Hearing Loss: Mixed Methods Pilot Study
JMIR Rehabil Assist Technol 2026;13:e79073
URL: https://rehab.jmir.org/2026/1/e79073
doi:10.2196/79073

© Sarah E Hughes, Liang-Yuan Wu, Lindsay J Ma, Dhruv Jain, Michael M McKee. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

XSL•FO
RenderX

and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Therapeutic Conversational Agent (Solace) for Management of Chronic Pain: Acceptability and Usability Study

P Maxwell Slepian[1,2,3,4], MS, PhD; Stephanie Buryk-Iggers[1], PhD; Anna M Lomanowska[1], PhD; Binh Nguyen[4], PhD; Tahir Janmohamed[4], PEng, MBA; Hance Clarke[1,2,3,4], MD, PhD; Joel Katz[1,2,3,4,5], PhD; Nils G Niederstrasser[6], PhD

[1]Department of Anesthesia and Pain Management, University Health Network, Toronto, ON, Canada

[2]Department of Anesthesiology and Pain Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[3]University of Toronto Centre for the Study of Pain, University of Toronto, Toronto, ON, Canada

[4]ManagingLife, Inc, Toronto, ON, Canada

[5]Department of Psychology, Faculty of Health, York University, Toronto, ON, Canada

[6]School of Psychology, Sport, and Health Sciences, University of Portsmouth, Portsmouth, England, United Kingdom

Corresponding Author:
P Maxwell Slepian, MS, PhD
Department of Anesthesia and Pain Management
University Health Network
3 EN 317A
200 Elizabeth St.
Toronto, ON, M5G 2C4
Canada
Phone: 1 7242890586
Email: maxwell.slepian@uhn.ca

Abstract

Background: Chronic pain is a critical cause of personal suffering and societal concern. However, treatment options remain inadequate, and access to efficacious treatment is limited by geography, economics, and scale. Digital health interventions for chronic pain are easily scaled solutions to this problem, and autonomous conversational agents represent a new frontier in this treatment domain. Despite their potential impact, conversational agents powered by generative artificial intelligence (AI) have yet to be developed or examined for the treatment of chronic pain.

Objective: We aimed to develop and test Solace, a first-of-its-kind, expert-trained generative AI conversational agent designed to deliver support grounded in principles of evidence-based pain psychology.

Methods: We conducted an acceptability and usability study of Solace in a group of individuals with chronic pain. Participants (n=175) were recruited from Prolific, an online crowdsourcing platform, and interacted with Solace for at least 25 minutes. Self-report measures of system usability, treatment acceptability, and therapeutic alliance were completed after the interaction, and clinically relevant pain-related measures were completed before and after the interaction.

Results: Participants rated the usability of Solace to be excellent (System Usability Scale mean score 85.04, SD 13.6) and found it to be an acceptable intervention for chronic pain. The therapeutic alliance between participants and Solace was rated highly. Participants also demonstrated improvements in anxiety, pain interference, kinesiophobia, and pain resilience. Safety guardrails designed to identify and manage instances of suicidal ideation, injury, or requests for medication recommendations performed appropriately during the study. Solace is a usable and acceptable treatment for chronic pain that facilitates forming a strong therapeutic alliance with users. A single 30-minute conversation with Solace was also associated with improvements in several clinically relevant domains.

Conclusions: Solace is a usable and acceptable expert-trained generative AI conversational agent for pain management. Randomized clinical trials are needed to evaluate the efficacy of Solace as a strategy for the treatment of chronic pain.

Trial Registration: OSF Registries osf.io/9c7tv; https://osf.io/9c7tv

(JMIR Rehabil Assist Technol 2026;13:e87689) doi:10.2196/87689

XSL•FO
RenderX

KEYWORDS

artificial intelligence; chatbot; chronic pain; digital health; pain psychology

Introduction

Chronic pain affects approximately 20% of adults worldwide, and is an enormous source of personal suffering as well as a major public health challenge [1,2]. Psychological interventions are efficacious for chronic pain. Across numerous large-scale randomized controlled trials, cognitive behavioral therapy, in particular, has been robustly associated with improvements in pain intensity and pain-related disability [3]. There is also accumulating evidence that other psychological treatment approaches, including mindfulness-based interventions and acceptance and commitment therapy (ACT), are similarly effective [4-7].

Despite the wealth of evidence for pain psychology interventions, widespread implementation of these interventions has been limited. This may be due, in part, to systemic barriers to adoption, including insufficient finances, time constraints, poor accessibility, and social stigma [8,9]. However, one of the key limitations is simply the lack of competent providers. For instance, difficulty accessing competent providers is the most common barrier reported by individuals with chronic pain [8]. This is further complicated by the fact that most licensed health care providers practicing therapy (eg, psychiatrists, psychologists, counselors, and social workers) report little to no education or training in pain treatment [8].

Digital health interventions (DHIs) have emerged as one of the primary means to overcome these barriers and bring psychological interventions for pain to scale [10,11]. DHIs exist in many forms, ranging from distance-based replications of services via telephone or videoconferencing to web-based treatment programs and mobile applications. Across these services, content can be delivered in either a fully self-guided format or supported by a provider or "coach," either synchronously or asynchronously [12]. These interventions have broadly demonstrated efficacy, but uptake remains limited, with studies reporting high dropout rates and low adherence [10,13].

More recently, researchers have begun to examine the provision of therapy by conversational agents, or chatbots, that enable patients to have interactive conversations in real time [13]. A recent scoping review identified beneficial effects of chatbots on pain intensity, but results were equivocal for pain interference and mental health outcomes [13]. These chatbots have generally been designed according to rule-based response algorithms rather than large language model (LLM)–based conversational support [13-16]. Such approaches have demonstrated that conversational agents can build a strong therapeutic alliance with users [15].

The use of LLM-based generative artificial intelligence (AI) conversational agents offers an alternative to algorithm-based chatbots. To date, not one study has examined fully generative AI conversational agents for the treatment of chronic pain. However, such generative AI chatbots have been studied for general mental health treatment. A clinical trial of a generative AI chatbot, Therabot, identified clinically significant improvement in symptoms of depression, anxiety, and feeding/eating disorder risk after 4 and 8 weeks, compared to waitlist controls [17]. Generative AI chatbots are also being applied to coping with medical concerns. In a recent study, health care providers identified a generative AI chatbot designed to provide coping support for adolescents with cancer as a feasible option for support provision, although the impact on clinical outcomes remains to be examined [18].

We sought to develop and test Solace, an expert-trained, generative AI chatbot built by ManagingLife, Inc. Solace was created by pairing a purpose-built knowledge base with a multiagent architecture leveraging a variety of commercially available foundational models. This study describes a prospective feasibility evaluation of the acceptability and usability of Solace in a group of participants reporting chronic pain (n=175). Participants engaged in an open-ended 30-minute interaction with Solace and completed measures of treatment acceptability, system usability, and therapeutic alliance after the conversation. Pain-related self-report measures were also completed before and after the participant's interaction with Solace. We hypothesized that individuals would find Solace to be a usable and acceptable treatment for chronic pain.

Methods

Study Design

This study was a prospective feasibility evaluation [19,20] of the AI-driven conversational agent, Solace. The design was guided by the Developmental and Exploratory Clinical Investigation of Decision-Support Systems Driven by Artificial Intelligence (DECIDE-AI), providing an early-stage clinical evaluation checklist of AI-specific reporting items (Multimedia Appendix 1) [21].

AI Chatbot

Overview

Solace is a stand-alone, web-based digital health application developed by ManagingLife [22]. It interacts with users to gather information related to the user's experience of pain and delivers appropriate, structured, interactive support grounded in evidence-based pain psychology techniques (eg, psychoeducation and ACT) to bolster emotional coping and self-management in individuals living with chronic pain. The development process focused on ensuring appropriate psychological support for adults with chronic pain through systematic programming and testing phases.

Technical Infrastructure

This study used the Solace v1.0 beta system version. The technical backbone of the chatbot system was built on top of a cloud-based digital health platform, Manage My Pain, developed by ManagingLife [22]. The main components of the Manage My Pain platform included user authentication, secure cloud hosting, and compliance with international data privacy

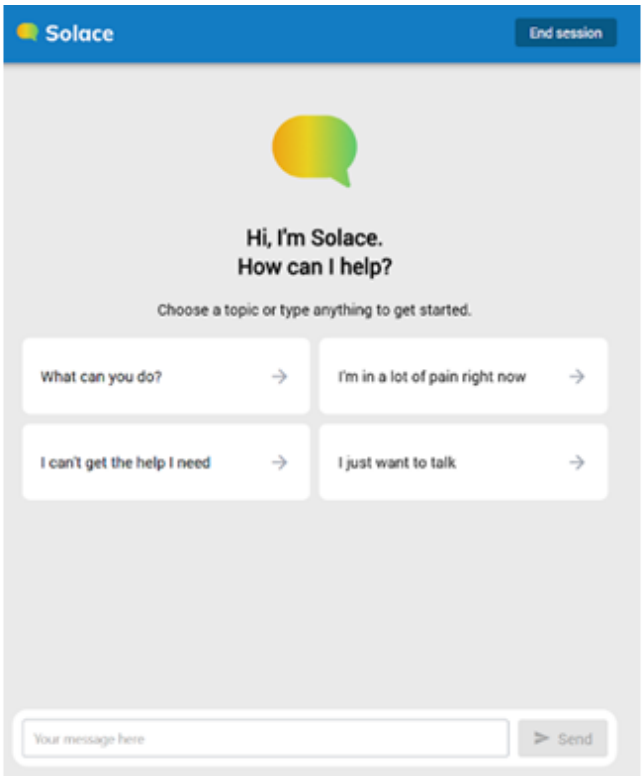
XSL•FO
RenderX

regulations, including end-to-end encryption and role-based access controls. Solace was designed to extend this platform by leveraging Amazon Web Services Bedrock to build a multiagent architecture using a library of LLMs. Each agent is responsible for a specialized task, and, together, these agents are coordinated by a centralized orchestration layer that manages conversational flow, memory retention, response timing, and system-level safety controls [23]. The foundational model versions, underlying platform, and all settings were fixed throughout the study period.

Conversational Design and Rules

Users began their experience with Solace by typing freely or selecting from starter prompts (Figure 1). User inputs were accepted without restrictions on length or content. Solace followed a set of core design principles developed through iterative collaboration with subject matter experts (SMEs) to ensure therapeutic integrity and user safety. These included the following:

Figure 1. The user options for starting a Solace conversation, including the use of the free-text field or starter prompts.

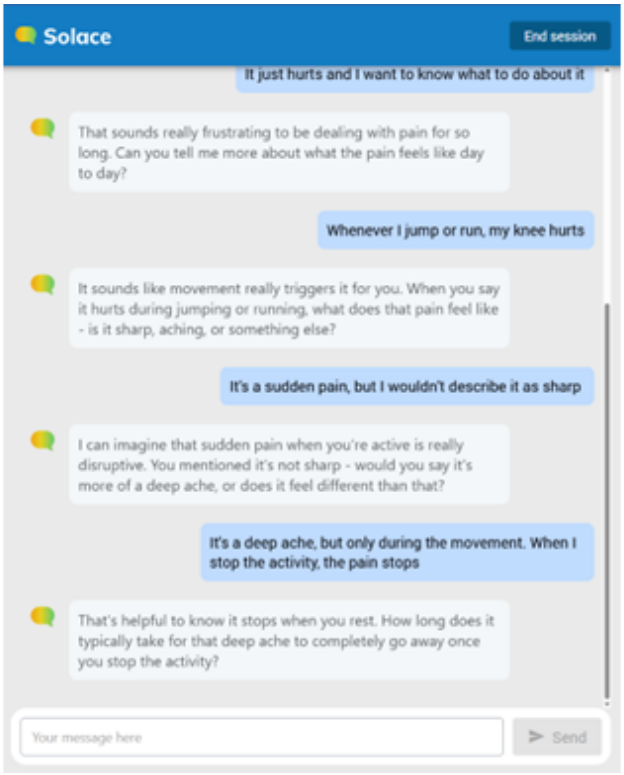


1. Open-ended engagement, encouraging users to freely express their thoughts and emotions
2. Plain language to ensure accessibility and reduce cognitive burden
3. Emotional sensitivity, with empathic responses embedded throughout the conversation
4. Therapeutic guidance toward evidence-based strategies and self-understanding
5. Context continuity, using memory variables to support coherent, personalized dialogue
6. Multilevel safety guardrails, with screening for sensitive content (eg, mention of self-harm, harm to others, and opioid use) and fallback responses designed in collaboration with clinical experts.

Solace's outputs drew from knowledge bases composed from psychological frameworks, including but not limited to ACT, CBT, and dialectical behavior therapy and were presented to the user as (typically) 1-4 sentences (Figure 2).

XSL•FO
RenderX

Figure 2. Examples of Solace outputs during a user conversation.



Testing and Optimization

Initial development and testing were led by a clinical pain psychologist (PMS) and included a cohort of 8 SMEs, including other clinical psychologists, pain physicians, and pain researchers. SMEs conducted simulated therapy sessions with Solace and provided both real-time "think-aloud" feedback and postsession feedback to the development team, focusing on the conversational tone of Solace, therapeutic alignment, and emotional sensitivity. The development team worked with SMEs to translate their feedback into Solace's conversational logic, which included refining pacing, language, and safety mechanisms.

Within Solace, user interactions are processed as string-based text inputs and routed through an orchestration layer that assigns each request to an appropriate specialized agent, which generates a response. Each specialized agent represents a specific element within the structure and therapeutic style of a clinical psychology session (eg, 1 agent for managing transitions between stages within a session), and has been fine-tuned and reviewed during development by an SME.

All generated responses are subsequently evaluated by a guardrail validation component that performs safety and compliance checks. If the validator fails, the response is discarded, and a specialized guardrail agent generates an alternative response (see "Safety and Errors" below); this agent and its outputs are likewise reviewed and refined by the SME. If validation succeeds, the approved response is returned to the user.

Stress-testing scenarios were used to explore how the system responded to ambiguous or sensitive inputs. System-level performance metrics, including response latency, session duration, and output length, were recorded during internal testing. This series of tests was part of an iterative process, informing and optimizing prompt development and the design of response templates.

Safety and Errors

System-level errors were defined as failure to generate a response, excessive latency, or loss of conversational flow. Errors were identified using a series of validator checkpoints embedded within the system architecture. Checkpoints initiated fallback mechanisms to trigger predefined safe response templates. System anomalies and platform reliability were further monitored using session-level logs and performance data. Detection of risks was supported by multilevel safety guardrails built into Solace's conversation logic and prompt structures and relied on a combination of rule-based screening and automated classification techniques. The safety guardrails were designed to identify and mitigate risks in the following safety-sensitive domains: (1) disclosure of content indicating potential harm to self, suicidal ideation, or harm to others; (2) active psychological or medical emergencies; and (3) recommendations related to prescription medication use.

When high-risk content was detected, Solace generated predefined fallback responses designed to redirect the user to seek professional help, reinforce system boundaries, and discourage further engagement with unsafe topics.

Participant Eligibility and Enrollment

Participants were recruited online through Prolific [24], an online research participation platform, and assigned an anonymous Prolific identification number. The target sample size for this study was 145 participants based on sufficiently approximating the population means across multiple continuous

XSL•FO
RenderX

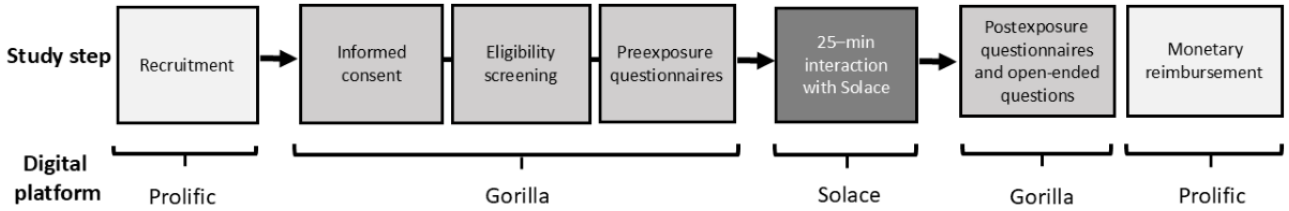
variables that are expected to have normal distributions. We aimed to recruit 195 participants to accommodate an anticipated 25% rate of attrition [25].

Those who consented were redirected to Gorilla [26], a cloud-based research platform hosting the questionnaires, to screen for eligibility and to access the survey. Participants were considered eligible based on the following criteria: (1) 18 years of age or older, (2) English speaking, and (3) living with chronic pain (ie, pain that has persisted for longer than 3 months). Chronic pain status and participant demographics were provided via the Prolific platform. Those eligible were asked to enter demographic information (age and sex) and to complete the baseline self-report measures.

Chatbot Engagement

Participants were informed that Solace is a nonclinical investigational tool under development and is not intended to provide medical advice or replace professional care. Training or an opportunity for familiarization with Solace was not provided to the participants prior to the study. Participants were prompted to engage with Solace, which was described as a "conversational agent, developed by ManagingLife. Solace provides pain management support and evidence-based strategies to help you understand and manage pain in your daily life." Participants were instructed to engage with Solace for at least 30 minutes from the start of the conversation, and then manually end the session. If users ended the session earlier than 30 minutes, they were informed how much time was left. Participants were then congratulated for completing their time with Solace and were provided with a link to return to Gorilla to complete the same self-report measures as at baseline, as well as a set of postexposure self-report measures. After completing the questionnaires, they were thanked for their participation, debriefed about the study hypotheses, and signposted to additional support if needed. The full participant flow through the study is depicted in Figure 3.

Figure 3. Participant flow through the study steps and associated digital platforms.



Measures

Participant Demographics

Participant demographics, including sex, ethnicity, employment status, and pain duration, were obtained from Prolific records.

System Usability

The System Usability Scale (SUS) is a 10-item questionnaire designed to measure the usability of a system, including digital products or digital systems [27]. Participants are asked to respond to items regarding usability (eg, "I thought the system was easy to use") on a scale from 1=strongly disagree to 5=strongly agree. The SUS is scored by subtracting the sum of the negatively worded items (eg, "I thought there was too much inconsistency in this system") from the sum of positively worded items (eg, "I think that I would use this system frequently"), adding 20, and multiplying by 2.5 to get a score out of 100. An SUS score of ≥85 has been identified as "excellent."

Treatment Acceptability

Two self-report inventories were included to examine treatment acceptability. The Treatment Acceptability/Adherence Scale (TAAS) is a 10-item self-report questionnaire designed to assess treatment acceptability and anticipated adherence to psychological interventions for anxiety and related disorders [28]. The TAAS has shown strong convergent and discriminant validity as well as excellent internal consistency [28]. In this study, the TAAS also demonstrated excellent internal consistency (α=.85). The Treatment Evaluation Inventory-Short Form (TEI-SF) is a 9-item measure of the perceived acceptability of behavioral treatments, and captures both positive (eg, "I would find this treatment to be an acceptable way of dealing with my chronic pain") and negative (eg, "I believe this treatment is likely to harm or injure me") perceptions of treatment [29]. The TEI-SF had excellent internal consistency in this study (α=.85). Both measures were adapted to the context of chronic pain treatment, which has been done in previous studies.

Therapeutic Alliance

The Working Alliance Inventory-Client Version (WAI-C) is a 36-item questionnaire used to measure therapeutic alliance in face-to-face therapy [30]. The WAI-C captures the client's perspective on 3 elements of therapeutic alliance, including the bond between therapist and client, agreement on the goals of therapy, and agreement on tasks to accomplish those goals. The WAI-C has excellent internal consistency (α=.87) and is strongly related to psychotherapy treatment satisfaction and outcomes. It has also been shown to predict treatment response to therapeutic interventions for pain. This instrument was adjusted for use within the context of an interaction with an AI-agent to align with the study's exploratory aims.

Pain-Related Measures

Patient Reported Outcomes Measurement Information System-29, version 2.0 (PROMIS-29) provides a profile of physical and mental health and assesses 7 health domains, including physical function, pain interference, pain intensity, anxiety, depression, fatigue, sleep disturbance, and social role integration. The PROMIS-29 has demonstrated strong validity in chronic pain populations. Internal consistency of subscales ranged from α=.88 to α=.96, indicating excellent consistency.

XSL•FO
RenderX

The questionnaire was administered using the static short forms of the domains [31].

The Tampa Scale of Kinesiophobia (TSK) was included as a measure of fear of movement (and reinjury) [32]. Participants rate items relevant to this fear (eg, "Pain always means I injured my body") on a scale from 0=strongly disagree to 4=strongly agree. The TSK has been extensively validated for individuals with chronic pain, and internal consistency in this study was adequate (α=.70)

The Pain Resilience Scale (PRS) is a 14-item measure of an individual's ability to maintain positive physical and emotional functioning despite pain. Participants are asked to rate agreement to items (eg, "I push through the pain") on a Likert scale from 0=not at all to 5=all the time [33,34]. The PRS has 2 subscales, behavioral perseverance and cognitive-affective positivity. The PRS has demonstrated psychometric validity among individuals with chronic pain [33], and internal consistency in this sample was excellent (α=.95).

The Chronic Pain Acceptance Questionnaire (CPAQ) is a 20-item measure with two subscales: (1) the degree to which one engages in life activities regardless of pain and (2) willingness to experience pain [35]. The CPAQ has been extensively validated for individuals with chronic pain [36,37], and the internal consistency in this study was adequate (α=.70).

System Metrics

User engagement and system performance were abstracted from transcripts of each participant session and associated data. User engagement characteristics included the number and length of prompts entered into Solace, the number and length of Solace's responses, and total session duration. Response readability was calculated using the Flesch Reading Ease Score and the Flesch-Kincaid Grade Level [38,39]. Safety was assessed by manual review of all conversations that contained content relevant to the guardrails.

Analysis

Analyses were primarily descriptive in nature. Participant responses on measures of usability, treatment acceptability, and therapeutic alliance were averaged (rather than summed) to give an indication of the overall level of agreement with measure items. Median (IQR) ratings for individual items were also reported. Averages with SD or counts, where applicable, were reported for system-level metrics. Only participants who completed at least 25 minutes of interaction with Solace were included for analyses of usability, treatment acceptability, and therapeutic alliance, whereas intent-to-treat analyses (eg, including all participants) were conducted using linear mixed effects models to analyze changes in clinically relevant self-report outcomes. Of note, due to a technical issue, 1 item was missed from the WAI-C (item 31) and 1 item from the CPAQ (item 8).

Ethical Considerations

The University of Portsmouth Ethics Board approved this study (REB SHFEC 2025-073), and the study has been registered with the Open Science Framework. Written informed consent was obtained from all participants prior to taking part in the study. Participants were presented with information about the study, made aware of their right to withdraw from the study at any time, and provided consent. Participants were compensated £18 (approximately US $24) for completing the study.

All identifying information was removed from the data before storage, and the data were further anonymized by assigning ID codes. Data pertaining to participants' responses to the structured and open-ended questionnaires and to their interactions with Solace were anonymized. Anonymized data were temporarily stored on the servers of Gorilla and ManagingLife, respectively. Data were extracted and transferred to the University of Portsmouth's secure server. The data will be retained for a minimum of 10 years in accordance with the University of Portsmouth Retention Schedule for Research Data.

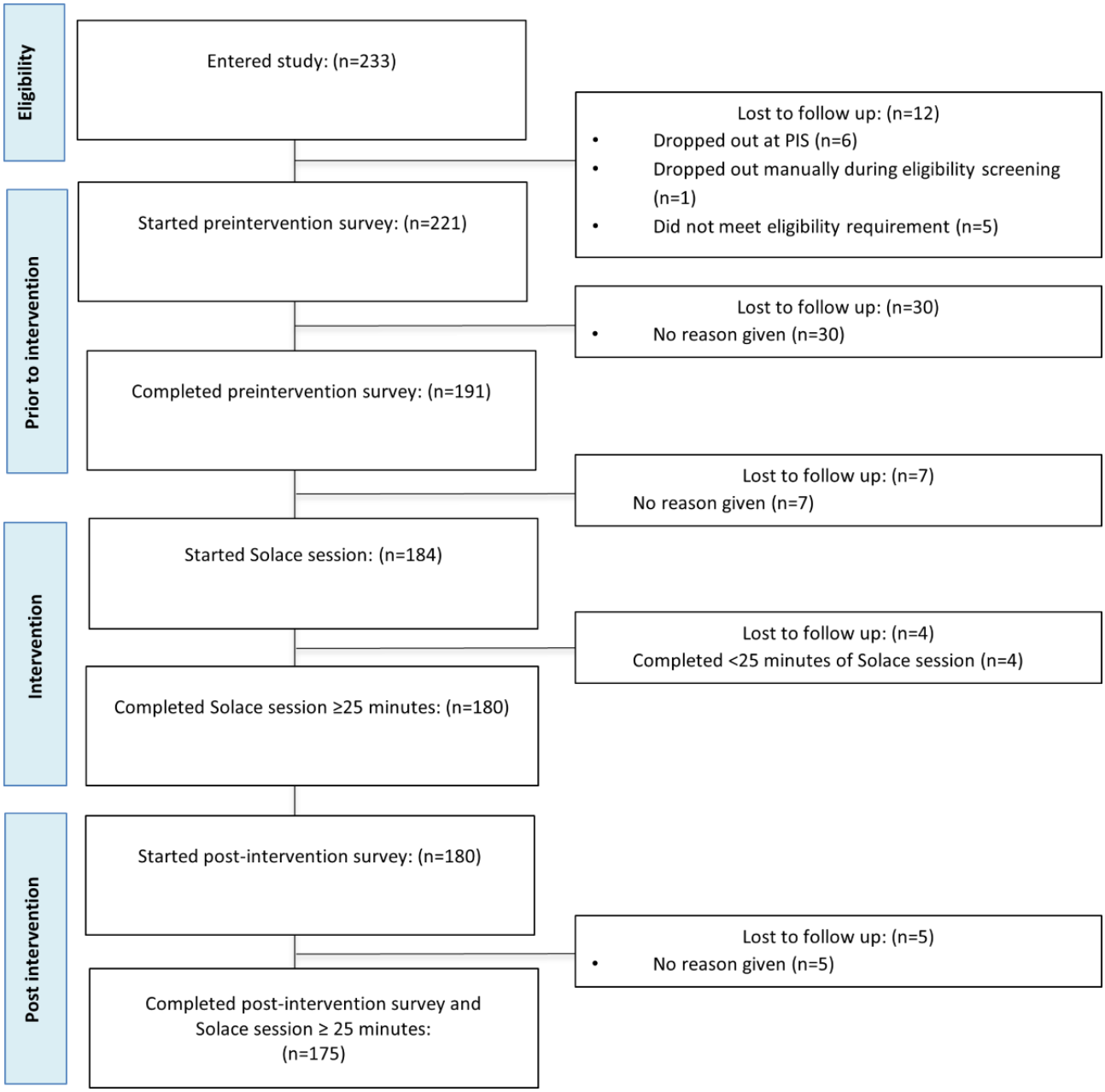
Results

Participant Characteristics and Flow

The flow of participants through the study is depicted in Figure 4. A total of 221 participants consented and completed the preinteraction survey. Of these, 184 began a session with Solace, and 175 engaged with Solace for at least 25 minutes. Participant demographics are listed in Table 1. All participants passed attention checks.

XSL•FO
RenderX

Figure 4. Flow diagram of participants who completed a ≥25-minute Solace session and self-report constructs. PIS: preintervention survey.



XSL•FO
RenderX

Table 1. Participant characteristics (n=175).

Characteristic	Value
Age (years), mean (SD)	42.54 (12.62)
Sex, n (%)	
Male	81 (46.3)
Female	93 (53.1)
Prefer not to say	1 (0.6)
Ethnicity, n (%)	
Asian	12 (6.9)
Black	13 (7.4)
Mixed	9 (5.1)
White	139 (80.8)
Other	2 (1.1)
Employment status, n (%)	
Full time	93 (53.1)
Part time	28 (16.0)
Unemployed and job seeking	9 (5.1)
Not in paid work (eg, homemaker, retired, or disabled)	23 (13.1)
Other	8 (4.6)
No data	14 (9.7)
Pain duration, n (%)	
3-6 months	9 (5.1)
6-12 months	10 (5.7)
1-2 years	16 (9.1)
2-5 years	52 (29.7)
5-10 years	32 (18.3)
10-20 years	43 (24.7)
>20 years	12 (6.9)
No data	1 (0.6)

Usage Characteristics and Safety Performance

The interaction sessions with Solace lasted an average of 38.74 (SD 11.7) minutes. On average, participants provided 29.2 (SD 10.1) input prompts to the system with an average length of 19.8 (SD 12.4) words, and Solace gave responses averaging 49.73 (SD 9.8) words on each occasion. On average, the readability of Solace's responses was at the early high school level (Flesch-Kincaid Grade Level mean 9.56, SD 2.8; Flesch Reading Ease mean 56.44, SD 15.6). The distribution of Flesch-Kincaid Grade Levels and Flesch Reading Ease scores is depicted in Figures 5 and 6, respectively.

XSL•FO
RenderX

Figure 5. Distribution of Flesch-Kincaid Grade Levels across all responses generated by Solace.

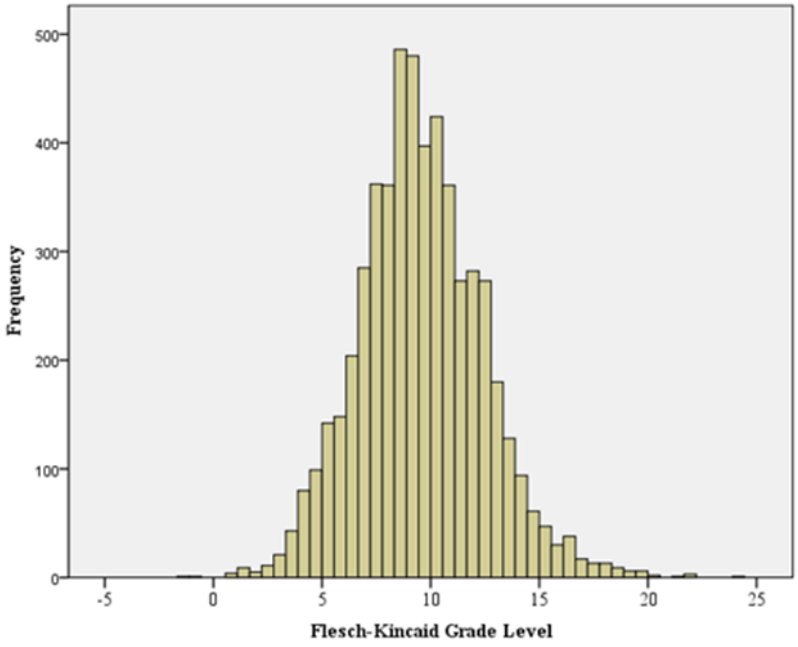
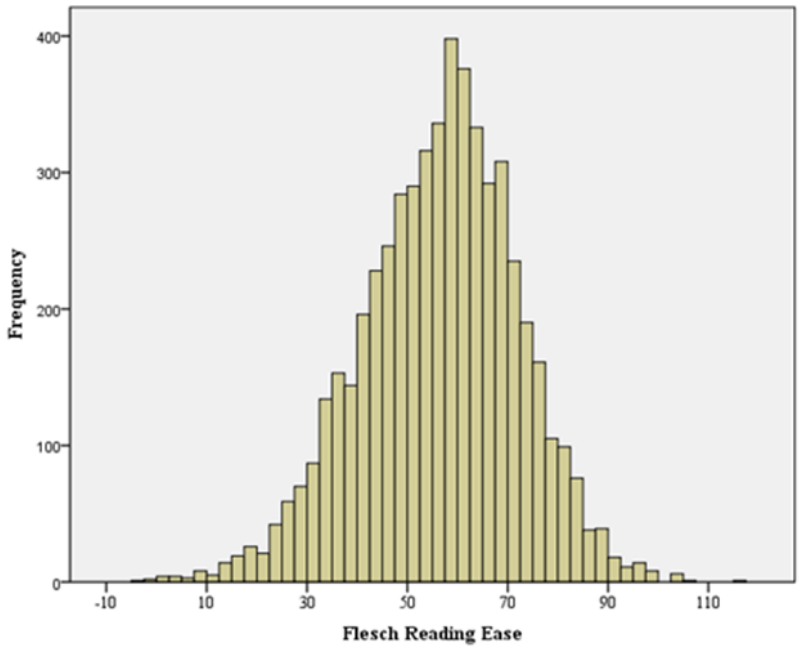


Figure 6. Distribution of Flesch Reading Ease scores across all responses generated by Solace.



Notably, 2 transcripts featured discussions of death, and 89 transcripts featured discussions of medication. Upon review by a clinical health psychologist (PMS), neither instance where death was discussed constituted suicidal ideation, self-harm, or harm to others, and Solace's responses were appropriate. Likewise, none of the conversations regarding medication included elicitation of recommendations, and, as such, guardrails were not activated.

System Usability

The SUS was completed as a measure of how easy participants found Solace to use. Participants rated the usability of Solace as excellent (mean 85.04, SD 13.6). On a scale from 1=strongly disagree to 5=strongly agree, participants rated all items as "strongly agree" except "I would like to use this system frequently," and "I found the various functions in this system were well-integrated," which were both rated as "agree." Item-level results are presented in Table 2.

XSL•FO
RenderX

Table 2. System Usability Scale item-level results (n=175). Scale ranges from 1=strongly disagree to 5=strongly agree.

Items	Median (IQR)
1. I think that I would like to use this system frequently	4.00 (2.00)
2. I found the system unnecessarily complex	1.00 (1.00)
3. I thought the system was easy to use	5.00 (1.00)
4. I think that I would need the support of a technical person to be able to use this system	1.00 (0.00)
5. I found the various functions in this system were well integrated	4.00 (2.00)
6. I thought there was too much inconsistency in this system	1.00 (1.00)
7. I would imagine that most people would learn to use this system very quickly	5.00 (1.00)
8. I found the system very cumbersome to use	1.00 (1.00)
9. I felt very confident using the system	5.00 (1.00)
10. I needed to learn a lot of things before I could get going with this system	1.00 (0.00)

Treatment Acceptability

Participants completed 2 measures of treatment acceptability, the TEI-SF and the TAAS. Average scores for the TEI-SF and TAAS, as well as individual item characteristics (median, IQR), are presented in Tables 3 and 4, respectively. On a scale from 1=strongly disagree to 5=strongly agree, participants reported an average rating of 3.77 (SD 0.7) on the TEI-SF. On the TAAS, participants, on average, rated treatment acceptability as 5.47 (SD 0.92) on a scale from 1=disagree strongly to 7=agree strongly.

Table 3. Treatment Evaluation Inventory-Short Form item-level results (n=175). Scale ranges from 1=strongly disagree to 5=strongly agree.

Items	Median (IQR)
1. I would find this treatment to be an acceptable way of dealing with my chronic pain	4.00 (1.00)
2. I would be willing to use this procedure if prescribed or recommended by a physician	4.00 (1.00)
3. I like the procedures that may be used in this treatment	4.00 (1.00)
4. I believe this treatment is likely to be effective	4.00 (2.00)
5. I believe that I will experience discomfort during the treatment	3.00 (1.00)
6. I believe this treatment is likely to result in permanent improvement	3.00 (1.00)
7. I believe this treatment is likely to harm or injure me	1.00 (1.00)
8. I believe it would be acceptable to use this treatment with individuals who cannot choose treatments for themselves	4.00 (1.00)
9. Overall, I have a positive reaction to this treatment	4.00 (1.00)

Table 4. Treatment Adherence/Acceptability Scale item-level results (n=175). Scale ranges from 1=disagree strongly to 7=agree strongly.

Items	Median (IQR)
1. If I began this treatment, I would be able to complete it	6.00 (2.00)
2. If I participated in this treatment, I would adhere to its requirements	6.00 (2.00)
3. I would find this treatment exhausting	3.00 (3.00)
4. It would be distressing for me to participate in this treatment	2.00 (2.00)
5. Overall, I would find this treatment intrusive	2.00 (2.00)
6. This treatment would provide effective ways to help me cope with my chronic pain	5.00 (1.00)
7. I would prefer to try another type of psychological treatment instead of this one	3.00 (3.00)
8. I would prefer to receive medication for my chronic pain instead of this treatment	3.00 (2.00)
9. I would recommend this treatment to a friend with chronic pain	5.00 (2.00)
10. If I began this treatment, I would likely drop out	2.00 (3.00)

XSL•FO
RenderX

Therapeutic Alliance

The WAI-C was completed to gauge the therapeutic alliance between participants and Solace. Overall, participants reported a strong working alliance with Solace (mean 188.03, SD 36.1), with a shared sense of goals rated as the strongest component of the therapeutic alliance (mean 67.86, SD 13.3), followed by agreement on tasks (mean 61.6, SD 12.7), and the therapeutic bond (mean 58.6, SD 12.6). WAI-C individual item scores are presented in Table 5.

Table 5. Working Alliance Inventory-Client Version item-level results (n=175). Scale ranges from 1=never to 7=always.

Items	Median (IQR)
1. I felt uncomfortable with Solace	1.00 (1.00)
2. Solace and I agreed about the things I will need to do in therapy to help improve my situation	6.00 (2.00)
3. I was worried about the outcome of the session	1.00 (1.00)
4. What I was doing in therapy gave me new ways of looking at my problem	5.00 (2.00)
5. Solace and I understood each other	6.00 (2.00)
6. Solace perceived accurately what my goals were	6.00 (2.00)
7. I find what I was doing in therapy confusing	1.00 (1.00)
8. I believe that Solace liked me	5.00 (2.00)
9. I wish Solace and I could have clarified the purpose of our session	2.00 (4.00)
10. I disagreed with Solace about what I ought to get out of therapy	1.00 (1.00)
11. I believe the time Solace and I were spending together was not spent efficiently	1.00 (2.00)
12. Solace did not understand what I was trying to accomplish in therapy	1.00 (1.00)
13. I was clear on what my responsibilities were in therapy	6.00 (2.00)
14. The goals of the session were important for me	6.00 (2.00)
15. I find what Solace and I were doing in therapy were unrelated to my concerns	1.00 (1.00)
16. I feel that the things I did in therapy helped me to accomplish the changes I wanted	5.00 (2.00)
17. I believe Solace was genuinely concerned for my welfare	5.00 (3.00)
18. I was clear as to what Solace wanted me to do in that session	6.00 (2.00)
19. Solace and I respected each other	7.00 (2.00)
20. I feel that Solace was not totally honest about its feelings toward me	1.00 (0.00)
21. I was confident in Solace's ability to help me	5.00 (3.00)
22. Solace and I were working toward mutually agreed upon goals	6.00 (2.00)
23. I feel that Solace appreciated me	5.00 (3.00)
24. We agreed on what was important for me to work on	6.00 (2.00)
25. As a result of therapy, I became clearer as to how I might be able to change	6.00 (3.00)
26. Solace and I trusted each other	6.00 (3.00)
27. Solace and I had different ideas about what my problems were	1.00 (1.00)
28. My relationship with Solace was very important to me	4.00 (4.00)
29. I had the feeling that if I said or did the wrong thing Solace would stop working with me	1.00 (0.00)
30. Solace and I collaborated on setting goals for my therapy	6.00 (2.00)
31. I was frustrated by the things I was doing in therapy	—[a]
32. We had a good understanding of the kind of changes that would be good for me	6.00 (2.00)
33. The things that Solace was asking me to do did not make sense to me	1.00 (1.00)
34. I did not know what to expect as the result of my therapy	2.00 (3.00)
35. I believe the way were working with my problem was correct	6.00 (2.00)
36. I feel Solace cared about me even when I did things it did not approve of	5.00 (3.00)

[a]Not available.

XSL•FO
RenderX

Clinical Outcomes

Clinically relevant outcomes were completed before and after the conversation with Solace and were analyzed using linear mixed effects models. Descriptive statistics and test results are presented in Table 6. Following Bonferroni correction to account for multiple comparisons, participants reported significant improvement in kinesiophobia, pain resilience, and pain willingness. On the PROMIS-29, participants reported improvement in anxiety and pain interference.

Table 6. Changes in clinically relevant outcomes.

Outcomes	Baseline, mean (SD)	Posttest, mean (SD)	*F* test (*df*)	*P* value
PRS[a]	28.07 (10.6)	30.57 (10.6)	41.62 (1, 193.6)	<.001
PRS-Behavioral Perseverance	11.64 (4.2)	12.1 (3.9)	6.79 (1, 196.8)	.01
PRS-Cognitive-Affective Positivity	16.43 (7.7)	18.48 (7.6)	51.95 (1, 194.1)	<.001
Kinesiophobia	43.4 (7.1)	41.79 (8.0)	24.73 (1, 191.5)	<.001
CPAQ[b]	61.66 (12.6)	62.5 (12.2)	1.71 (1, 201.1)	.19
CPAQ-Activity Engagement	31.71 (8.0)	33.02 (8.1)	16.46 (1,194.05)	<.001
CPAQ-Pain Willingness	29.95 (8.5)	29.48 (9.1)	0.37 (1, 194.05)	.54
PROMIS-29[c]				
Physical Function	15.15 (3.5)	15.0 (3.6)	2.38 (1, 191.7)	.13
Anxiety	10.46 (3.9)	10.0 (4.1)	17.04 (1, 191.2)	<.001
Depressive Symptoms	9.58 (4.6)	9.42 (4.6)	2.82 (1, 192.2)	.09
Sleep	12.87 (4.0)	12.86 (3.9)	0.01 (1, 193.6)	.93
Social Function	13.0 (3.9)	13.08 (3.8)	0.37 (1, 195.1)	.55
Pain Interference	11.60 (4.0)	11.15 (3.9)	10.89 (1, 195.1)	.001
Fatigue	12.95 (4.1)	12.7 (4.4)	3.78 (1, 191.2)	.05

[a]PRS: Pain Resilience Scale.

[b]CPAQ: Chronic Pain Acceptance Questionnaire.

[c]PROMIS-29: Patient Reported Outcome Measurement Information System-29.

Discussion

This study conducted a feasibility evaluation of Solace, an expert-trained generative AI conversational agent designed to use pain psychology–informed self-management strategies for chronic pain, following the DECIDE-AI framework [40]. Participants reporting chronic pain completed self-report measures before and after at least 25 minutes of conversation with Solace. Broadly, participants rated Solace to have excellent usability, good treatment acceptability, and reported a strong therapeutic alliance. Moreover, participants reported improvement in several clinically relevant domains from before to after the interaction with Solace.

The reported system usability (mean 85.04, SD 13.6) was rated well above the benchmark established for digital health applications (SUS score >65), indicating excellent usability [41,42]. Notably, SUS scores were also higher than those reported for other generative AI tools, such as ChatGPT, when rated by either general users (mean 67.44) or health care professionals (mean 64.52, SD 13.91) [43,44]. Although the SUS has been frequently used to evaluate AI applications [45], the SUS may not capture the nuance of using a generative AI conversational agent such as Solace [41]. Future studies should include more specific measures of usability designed to analyze generative AI conversational agents, such as the Bot Usability Scale [46].

Solace was deemed an acceptable health application for chronic pain and was well-received by participants. Participant responses on the TEI-SF and the TAAS were largely on par with other DHIs for chronic pain [47,48]. Participants responded with agreement to almost all items. This includes statements such as “I believe this treatment is likely to be effective,” “I would find this treatment to be an acceptable way of dealing with my chronic pain,” and “this treatment would provide effective ways to help me cope with my chronic pain.” Participants also endorsed items related to the belief that they could adhere to treatment with Solace, such as “If I began this treatment, I would be able to complete it,” and beliefs that Solace is a safe intervention (ie, strongly disagreeing with “I believe this

XSL•FO
RenderX

treatment is likely to harm or injure me"). It is worth noting that participants rarely expressed the strongest agreement with the TEI-SF or TAAS items. This may be due to the noninterventional nature of the evaluation and the lack of set treatment expectations. It will be important to further evaluate acceptability in the context of a treatment trial or clinical practice.

Participants reported developing a strong therapeutic alliance with Solace over the course of the 25-minute conversation. In particular, participants reported a strong shared sense of treatment goals (ie, strongly disagreeing with the statement "I find what Solace and I were doing in therapy was unrelated to my concerns"). Several previous studies have examined the experience of therapeutic alliance with a conversational agent, and have found similarly strong working alliance [15,17]. It is perhaps understandable that participants reported the therapeutic bond as the weakest aspect of the alliance. Items designed to capture this on the WAI-C reflect a personal relationship that might not be relevant for interactions with a conversational agent, such as items asking if participants felt Solace liked or trusted them. New tools might be needed in this area to thoroughly assess therapeutic alliance with nonhuman conversational agents [49]. It is also notable that this study was not presented to participants as a treatment designed to help them with their pain, despite the conversational focus on treatment-related topics. It will be important to replicate these promising indications of therapeutic alliance with Solace in a treatment setting, and to examine if alliance improves with longer treatment duration, an important predictor of treatment success [50,51].

Safety of Solace and other generative AI conversational agents is a critical consideration in their development as DHIs. However, this could not be fully assessed as guardrails designed to manage suicidality, self-harm, medical emergencies, or medication recommendations were developed, but user prompts did not include safety concerns, and therefore, guardrails were not activated in this study. It will be necessary to analyze a much larger volume of interactions to fully test guardrail performance.

After the 25-minute conversation with Solace, participants demonstrated improvement in several meaningful clinical domains, including kinesiophobia, resilience, activity engagement, as well as anxiety and pain interference on the PROMIS-29 measure. Across these factors, there was 4%-9% improvement, with the greatest improvement seen in pain resilience. There are several difficulties in interpreting these changes. Notably, although participants were told Solace was a treatment in development and self-report constructs inherently included reference to treatment, they were also told that this was not a treatment study. There was also a mismatch between the wording on several of the measures and the time frame of the study. In particular, instructions for the subscales of the PROMIS-29 ask respondents to reference how they felt "in the last 7 days," whereas the study might have lasted as little as 25 minutes. Despite these challenges in interpretation, improvements were all in the beneficial direction, held up to statistical correction, and likely reflect the participants' feelings at the time of responding. Although the change was modest in magnitude, this is a promising indication for further treatment-focused research on Solace.

Future research on Solace should take several directions. As described above, a program of clinical research should be undertaken to evaluate the efficacy and effectiveness of Solace, either as a stand-alone program or integrated into the Manage My Pain application, a digital pain self-management application developed by ManagingLife. This will include a fuller test of the effectiveness of the safety guardrails in identifying discussions involving recommendations for medication use, suicidal ideation, and harm to self or others. Further evaluation of the nature of users' interactions with Solace will also be critical. These should focus on both the therapeutic relationship as well as evaluation of safety features of the application. Moreover, algorithmic fairness was not assessed in this study. Subsequent studies reporting on Solace will aim to compare model performance (sensitivity and specificity) across demographic subgroups, including sex, age, and ethnicity, to ensure equitable accuracy and lack of bias. Patient engagement is also needed to more fully address the lived experience of individuals with chronic pain.

There is room to further improve Solace's functionality. In particular, Solace's responses required a reading comprehension level considered roughly equivalent to early-middle high school. Future iterations of Solace should focus on the accessibility of language to improve utility by the general population. More detailed analysis of conversations with Solace is also warranted to ensure that conversational content is adherent to prompting.

There are several notable limitations of this study. First, the study was conducted via an online crowdsourcing platform. Although the sample recruited in this study is similar in several respects (ie, age, sex, and duration of pain) to individuals with chronic pain recruited from community samples [52], there were several notable distinctions, and also limited demographic data collection. In particular, participants in this study were primarily White and working full-time. As such, it is possible that community or treatment-seeking individuals with chronic pain would have a different experience of Solace. Although extensive research on chronic pain has been conducted via the Prolific platform [53,54], future research needs to be conducted in clinical samples with verified diagnoses. Second, the study was not designed as a test of Solace as a treatment. As such, participants did not necessarily have a treatment-related interaction with Solace, which could impact perceptions of acceptability and therapeutic alliance. Third, the sensitivity and specificity of the guardrail system were not fully tested, given the few instances that required guardrail activation. Finally, the short time frame of the study and the lack of a control condition limit the interpretation of the clinical impact of Solace. It is possible that alternative explanations, such as regression to the mean or social desirability, could account for the change associated with Solace use. Subsequent studies should examine prolonged or repeated interactions and feature comparisons to a control condition to enhance internal validity.

In conclusion, participants found Solace to be a usable system and an acceptable treatment for chronic pain after a 25-minute conversation. Participants reported a strong therapeutic alliance

XSL•FO
RenderX

and demonstrated improvement in several clinically relevant domains. Furthermore, examination of safety guardrails for Solace indicated appropriate performance. Solace is a promising therapeutic tool, and clinical trials are needed to fully examine its clinical efficacy as a DHI for chronic pain.

Data Availability

The data are available from the University of Portsmouth Research Portal, which functions as a data repository.

Funding

This research was funded by the University of Portsmouth's Higher Education Innovation Fund.

Conflicts of Interest

TJ is the founder and CEO of ManagingLife. BN is an employee of ManagingLife. PMS is an independent consultant for ManagingLife. JK and HC are unpaid members of the ManagingLife Advisory Board, serving as research director and medical director, respectively, and providing guidance on the product and the company's research initiatives. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Developmental and Exploratory Clinical Investigation of Decision-Support Systems Driven by Artificial Intelligence (DECIDE-AI) checklist.
[DOCX File , 2338 KB - rehab_v13i1e87689_app1.docx]

References

1. Reitsma ML, Tranmer JE, Buchanan DM, Vandenkerkhof EG. The prevalence of chronic pain and pain-related interference in the Canadian population from 1994 to 2008. Chronic Dis Inj Can 2011;31(4):157-164. [doi: 10.24095/hpcdp.31.4.04]
2. Hogan ME, Taddio A, Katz J, Shah V, Krahn M. Incremental health care costs for chronic pain in Ontario, Canada: a population-based matched cohort study of adolescents and adults using administrative data. Pain 2016;157(8):1626-1633. [doi: 10.1097/j.pain.0000000000000561] [Medline: 26989805]
3. Williams ACDC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2020;8:CD007407. [doi: 10.1002/14651858.CD007407.pub4] [Medline: 32794606]
4. Zhu M, Wong SY, Zhong CC, Zeng Y, Xie L, Lee EK, et al. Which type and dosage of mindfulness-based interventions are most effective for chronic pain? A systematic review and network meta-analysis. J Psychosom Res 2025;191:112061 [FREE Full text] [doi: 10.1016/j.jpsychores.2025.112061] [Medline: 40010103]
5. Martinez-Calderon J, García-Muñoz C, Rufo-Barbero C, Matias-Soto J, Cano-García FJ. Acceptance and commitment therapy for chronic pain: an overview of systematic reviews with meta-analysis of randomized clinical trials. J Pain 2024;25(3):595-617. [doi: 10.1016/j.jpain.2023.09.013]
6. Ma TW, Yuen ASK, Yang Z. The efficacy of acceptance and commitment therapy for chronic pain: a systematic review and meta-analysis. Clin J Pain 2023;39(3):147. [doi: 10.1097/ajp.0000000000001096]
7. Hilton L, Hempel S, Ewing BA, Apaydin E, Xenakis L, Newberry S, et al. Mindfulness meditation for chronic pain: systematic review and meta-analysis. Ann Behav Med 2017;51(2):199-213 [FREE Full text] [doi: 10.1007/s12160-016-9844-2] [Medline: 27658913]
8. Darnall BD, Scheman J, Davin S, Burns JW, Murphy JL, Wilson AC, et al. Pain psychology: a global needs assessment and national call to action. Pain Med 2016;17(2):250-263 [FREE Full text] [doi: 10.1093/pm/pnv095] [Medline: 26803844]
9. Driscoll MA, Edwards RR, Becker WC, Kaptchuk TJ, Kerns RD. Psychological interventions for the treatment of chronic pain in adults. Psychol Sci Public Interest 2021;22(2):52-95. [doi: 10.1177/15291006211008157] [Medline: 34541967]
10. Valentijn PP, Tymchenko L, Jacobson T, Kromann J, Biermann CW, AlMoslemany MA, et al. Digital health interventions for musculoskeletal pain conditions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022;24(9):e37869 [FREE Full text] [doi: 10.2196/37869] [Medline: 36066943]
11. Kazdin AE. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav Res Ther 2017;88:7-18. [doi: 10.1016/j.brat.2016.06.004] [Medline: 28110678]
12. El-Tallawy SN, Pergolizzi JV, Vasiliu-Feltes I, Ahmed RS, LeQuang JK, Alzahrani T, et al. Innovative applications of telemedicine and other digital health solutions in pain management: a literature review. Pain Ther 2024;13(4):791-812. [doi: 10.1007/s40122-024-00620-7] [Medline: 38869690]
13. Souza FL, Bowman H, Yang F, Hesam-Shariati N, Linke J, Gilanyi YL, et al. Conversational agents to support pain management: a scoping review. Eur J Pain 2025;29(5):e70016. [doi: 10.1002/ejp.70016] [Medline: 40170347]

XSL•FO
RenderX

14. Sinha C, Cheng AL, Kadaba M. Adherence and engagement with a cognitive behavioral therapy-based conversational agent (Wysa for Chronic Pain) among adults with chronic pain: survival analysis. JMIR Form Res 2022;6(5):e37302 [FREE Full text] [doi: 10.2196/37302] [Medline: 35526201]
15. Beatty C, Malik T, Meheli S, Sinha C. Evaluating the therapeutic alliance with a free-text CBT conversational agent (Wysa): a mixed-methods study. Front Digit Health 2022;4:847991 [FREE Full text] [doi: 10.3389/fdgth.2022.847991] [Medline: 35480848]
16. Meheli S, Sinha C, Kadaba M. Understanding people with chronic pain who use a cognitive behavioral therapy-based artificial intelligence mental health app (Wysa): mixed methods retrospective observational study. JMIR Hum Factors 2022;9(2):e35671 [FREE Full text] [doi: 10.2196/35671] [Medline: 35314422]
17. Heinz MV, Mackin DM, Trudeau BM, Bhattacharya S, Wang Y, Banta HA, et al. Randomized trial of a generative AI Chatbot for mental health treatment. NEJM AI 2025;2(4):AIoa2400802. [doi: 10.1056/aioa2400802]
18. Hasei J, Hanzawa M, Nagano A, Maeda N, Yoshida S, Endo M, et al. Empowering pediatric, adolescent, and young adult patients with cancer utilizing generative AI chatbots to reduce psychological burden and enhance treatment engagement: a pilot study. Front Digit Health 2025;7:1543543 [FREE Full text] [doi: 10.3389/fdgth.2025.1543543] [Medline: 40070545]
19. McCradden MD, Stephenson EA, Anderson JA. Clinical research underlies ethical integration of healthcare artificial intelligence. Nat Med 2020;26(9):1325-1326. [doi: 10.1038/s41591-020-1035-9] [Medline: 32908273]
20. Van der Vegt AH, Scott I, Dermawan K, Schnetler R, Kalke V, Lane P. Implementation frameworks for end-to-end clinical AI: derivation of the SALIENT framework. J Am Med Inform Assoc 2023;30(9):1503-1515 [FREE Full text] [doi: 10.1093/jamia/ocad088] [Medline: 37208863]
21. The DECIDE-AI Steering Group. DECIDE-AI: new reporting guidelines to bridge the development-to-implementation gap in clinical artificial intelligence. Nat Med 2021;27(2):186-187. [doi: 10.1038/s41591-021-01229-5] [Medline: 33526932]
22. ManagingLife Inc. 2025. URL: https://www.managinglife.com/ [accessed 2026-02-21]
23. IEEE Computer Society, Standards Committee Working Group. Standard for large language model agents for AI-powered education. AI Standards Hub. 2023. URL: https://aistandardshub.org/forums/topic/standard-for-large-language-model-agents-for-ai-powered-education-3/ [accessed 2026-02-21]
24. Prolific. URL: https://www.prolific.com [accessed 2026-02-28]
25. Macea DD, Gajos K, Daglia Calil YA, Fregni F. The efficacy of web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010;11(10):917-929 [FREE Full text] [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
26. GORILLA. URL: https://www.gorilla.sc/ [accessed 2026-02-28]
27. Brooke J. SUS: a retrospective. J Usability Stud 2013;2:29-40 [FREE Full text]
28. Milosevic I, Levy HC, Alcolado GM, Radomsky AS. The treatment acceptability/adherence scale: moving beyond the assessment of treatment effectiveness. Cogn Behav Ther 2015;44(6):456-469. [doi: 10.1080/16506073.2015.1053407] [Medline: 26091250]
29. Kelley ML, Heffer RW, Gresham FM, Elliott SN. Treatment Evaluation Inventory—Short Form. APA PsycNet 2016. [doi: 10.1037/t49940-000]
30. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. Psychother Theory Res Pract 1979;16(3):252-260. [doi: 10.1037/h0085885]
31. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS-29 v2.0 profile physical and mental health summary scores. Qual Life Res 2018;27(7):1885-1891 [FREE Full text] [doi: 10.1007/s11136-018-1842-3] [Medline: 29569016]
32. Swinkels-Meewisse EJCM, Zuyd H, Vlaeyen JW, Leuven KU, Oostendorp RAB. Psychometric properties of the Tampa Scale for Kinesiophobia and the Fear-Avoidance Beliefs Questionnaire in acute low back pain. Man Ther 2003;8(1):29-36. [doi: 10.1054/math.2002.0484] [Medline: 12586559]
33. Ankawi B, Slepian PM, Himawan LK, France CR. Validation of the pain resilience scale in a chronic pain sample. J Pain 2017;18(8):984-993 [FREE Full text] [doi: 10.1016/j.jpain.2017.03.013] [Medline: 28428092]
34. Slepian PM, Ankawi B, Himawan LK, France CR. Development and initial validation of the pain resilience scale. J Pain 2016;17(4):462-472 [FREE Full text] [doi: 10.1016/j.jpain.2015.12.010] [Medline: 26748044]
35. McCracken LM, Vowles KE, Eccleston C. Acceptance of chronic pain: Component analysis and a revised assessment method. Pain 2004;107(1-2):159-166. [doi: 10.1016/j.pain.2003.10.012] [Medline: 14715402]
36. Fish RA, McGuire B, Hogan M, Morrison T, Stewart I. Validation of the chronic pain acceptance questionnaire (CPAQ) in an internet sample and development and preliminary validation of the CPAQ-8. Pain 2010;149(3):435-443. [doi: 10.1016/j.pain.2009.12.016] [Medline: 20188472]
37. Vowles KE, McCracken LM, McLeod C, Eccleston C. The Chronic Pain Acceptance Questionnaire: confirmatory factor analysis and identification of patient subgroups. Pain 2008;140(2):284-291. [doi: 10.1016/j.pain.2008.08.012] [Medline: 18824301]
38. Flesch R. A new readability yardstick. J Appl Psychol 1948;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
39. Eleyan D, Othman A, Eleyan A. Enhancing software comments readability using Flesch Reading Ease Score. Information 2020;11(9):430. [doi: 10.3390/info11090430]

XSL•FO
RenderX

40. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, DECIDE-AI expert group. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ 2022;377:e070904 [FREE Full text] [doi: 10.1136/bmj-2022-070904] [Medline: 35584845]
41. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
42. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
43. Aljamaan F, Malki KH, Alhasan K, Jamal A, Altamimi I, Khayat A, et al. ChatGPT-3.5 system usability scale early assessment among healthcare workers: horizons of adoption in medical practice. Heliyon 2024;10(7):e28962 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e28962] [Medline: 38623218]
44. Mulia AP, Piri PR, Tho C. Usability analysis of text generation by ChatGPT OpenAI using system usability scale method. Procedia Comput Sci 2023;227:381-388. [doi: 10.1016/j.procs.2023.10.537]
45. Nikou SA, Guliya A, Van VS, Chang M. A generative artificial intelligence empowered chatbot: system usability and student teachers' experience. In: Sifaleras A, Lin F, editors. Generative Intelligence and Intelligent Tutoring Systems. Gener Intell Intell Tutoring Syst Cham: Springer Nature Switzerland; 2024:330-340.
46. Borsci S, Malizia A, Schmettow M, van der Velde F, Tariverdiyeva G, Balaji D, et al. The chatbot usability scale: the design and pilot of a usability scale for interaction with AI-based conversational agents. Pers Ubiquit Comput 2021;26(1):95-119. [doi: 10.1007/s00779-021-01582-9]
47. Scheutzow J, Attoe C, Harwood J. Acceptability of web-based mental health interventions in the workplace: systematic review. JMIR Ment Health 2022;9(5):e34655 [FREE Full text] [doi: 10.2196/34655] [Medline: 35544305]
48. Hennessy RW, Rumble D, Christian M, Brown DA, Trost Z. A graded exposure, locomotion-enabled virtual reality app during walking and reaching for individuals with chronic low back pain: cohort gaming design. JMIR Serious Games 2020;8(3):e17799 [FREE Full text] [doi: 10.2196/17799] [Medline: 32773381]
49. Malouin-Lachance A, Capolupo J, Laplante C, Hudon A. Does the digital therapeutic alliance exist? Integrative review. JMIR Ment Health 2025;12:e69294-e69294. [doi: 10.2196/69294]
50. Holmes MB, Scott A, Camarinos J, Marinko L, George SZ. Working Alliance Inventory (WAI) and its relationship to patient-reported outcomes in painful musculoskeletal conditions. Disabil Rehabil 2023;45(8):1363-1369. [doi: 10.1080/09638288.2022.2060337] [Medline: 35416110]
51. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. J Couns Psychol 1989;36(2):223-233. [doi: 10.1037/0022-0167.36.2.223]
52. Mills SEE, Nicolson KP, Smith BH. Chronic pain: a review of its epidemiology and associated factors in population-based studies. Br J Anaesth 2019;123(2):e273-e283 [FREE Full text] [doi: 10.1016/j.bja.2019.03.023] [Medline: 31079836]
53. Niederstrasser NG, Attridge N. Associations between pain and physical activity among older adults. PLoS One 2022;17(1):e0263356 [FREE Full text] [doi: 10.1371/journal.pone.0263356] [Medline: 35089966]
54. Todd J, Pickup B, Sharpe L. Interpretation bias and the transition from acute to chronic pain. Pain 2024;165(2):357-364. [doi: 10.1097/j.pain.0000000000003016] [Medline: 37624880]

Abbreviations

ACT: acceptance and commitment therapy
AI: artificial intelligence
CPAQ: Chronic Pain Acceptance Questionnaire
DECIDE-AI: Developmental and Exploratory Clinical Investigation of Decision-Support Systems Driven by Artificial Intelligence
DHI: digital health intervention
LLM: large language model
PROMIS-29: Patient Reported Outcomes Measurement Information System-29, version 2.0
PRS: Pain Resilience Scale
SME: subject matter expert
SUS: System Usability Scale
TAAS: Treatment Acceptability/Adherence Scale
TEI-SF: Treatment Evaluation Inventory-Short Form
WAI-C: Working Alliance Inventory-Client Version
TSK: Tampa Scale of Kinesiophobia

XSL•FO
RenderX

Edited by S Munce; submitted 12.Nov.2025; peer-reviewed by J Greenberg, KH Lin; comments to author 19.Dec.2025; revised version received 27.Jan.2026; accepted 10.Feb.2026; published 20.Mar.2026.

Please cite as:
Slepian PM, Buryk-Iggers S, Lomanowska AM, Nguyen B, Janmohamed T, Clarke H, Katz J, Niederstrasser NG
A Therapeutic Conversational Agent (Solace) for Management of Chronic Pain: Acceptability and Usability Study
JMIR Rehabil Assist Technol 2026;13:e87689
URL: https://rehab.jmir.org/2026/1/e87689
doi:10.2196/87689
PMID:41711568

©P Maxwell Slepian, Stephanie Buryk-Iggers, Anna M Lomanowska, Binh Nguyen, Tahir Janmohamed, Hance Clarke, Joel Katz, Nils G Niederstrasser. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Smartphone App–Delivered Mindfulness-Based Intervention for Concussion in Adolescents (MBI-4-mTBI): Feasibility Randomized Controlled Trial

Veronik Sicard[1], PhD; Molly Cairncross[2,3,4], PhD; Roger Zemek[1,5], MD; Noah Silverberg[4,6,7], PhD; Gary S Goldfield[1,5,8], PhD; Nick Reed[9], PhD; Bechara J Saab[10], PhD; Andra Smith[8], PhD; Andrée-Anne Ledoux[1,11,12], PhD

[1]Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, ON, Canada
[10]Mobio Interactive, Singapore, Singapore
[11]Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada
[12]University of Ottawa Brain and Mind Research Institute, University of Ottawa, Ottawa, ON, Canada
[2]Department of Psychology, Simon Fraser University, Burnaby, BC, Canada
[3]BC Children's Hospital Research Institute, Vancouver, BC, Canada
[4]Rehabilitation Research Program, Centre for Aging SMART, Vancouver Coastal Health Research Institute, Vancouver, BC, Canada
[5]Department of Pediatrics, University of Ottawa, Ottawa, ON, Canada
[6]Department of Psychology, University of British Columbia, Vancouver, BC, Canada
[7]Djavad Mowafaghian Centre for Brain Health, Vancouver, BC, Canada
[8]School of Psychology, University of Ottawa, Ottawa, ON, Canada
[9]Department of Occupational Science & Occupational Therapy, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Andrée-Anne Ledoux, PhD
Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, ON, Canada

Related Article:

Correction of: https://rehab.jmir.org/2026/1/e84623

(JMIR Rehabil Assist Technol 2026;13:e110902) doi:10.2196/110902

In "Smartphone App–Delivered Mindfulness-Based Intervention for Concussion in Adolescents (MBI-4-mTBI): Feasibility Randomized Controlled Trial" [1], the authors noted 6 errors.

The outcome measure used to assess anxiety was the Severity Measure for Generalized Anxiety Disorder–Child Age 11 - 17 (DSM-5 GAD-C), a validated tool developed alongside the *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition; *DSM-5*) for use in pediatric populations, not the Generalized Anxiety Disorder 7-item scale (GAD-7), an adult measure. This affects 5 statements in the article and 1 reference. The corrections below reflect the accurate instrument used throughout data collection and analysis.

First, the following bullet point in the Secondary Outcomes section of the Methods was revised:

Anxiety: Generalized Anxiety Disorder 7-item scale

The corrected bullet point now reads:

Anxiety: Severity Measure for Generalized Anxiety Disorder–Child Age 11 - 17 (DSM-5 GAD-C)

Second, the following row heading in Table 3 was revised:

GAD-7[j] (anxiety), median (IQR)

The row heading now reads:

DSM-5 GAD-C[j] (anxiety), median (IQR)

Third, footnote j in Table 3 was revised from the following:

[j] GAD-7: Generalized Anxiety Disorder 7-item scale.

Footnote j now reads:

[j]DSM-5 GAD-C: Severity Measure for Generalized Anxiety Disorder–Child Age 11 - 17.

Fourth, the following row heading in Table 4 was revised:

GAD-7[b] (anxiety)

The row heading now reads:

DSM-5 GAD-C[b] (anxiety)

Fifth, footnote b of Table 4 was revised from the following:

[b]GAD-7: Generalized Anxiety Disorder 7-item scale.

Footnote b now reads:

[b]DSM-5 GAD-C: Severity Measure for Generalized Anxiety Disorder–Child Age 11 - 17.

Finally, reference 69 was revised from the following:

69. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. May 22, 2006;166(10):1092 - 1097. URL:

XSL•FO
RenderX

https://jamanetwork.com/ [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]

Reference 69 now reads:

69. Craske M, Wittchen U, Bogels S, Stein M, Andrews G, Lebeu R. Severity measure for generalized anxiety disorder—child age 11–17. American Psychiatric Association. 2013. URL: [illegible] [Accessed 2026-08-31]

The correction will appear in the online version of the paper on the JMIR Publications website together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Sicard V, Cairncross M, Zemek R, et al. Smartphone app–delivered mindfulness-based intervention for concussion in adolescents (MBI-4-mTBI): feasibility randomized controlled trial. JMIR Rehabil Assist Technol 2026 May 11;13:e84623. [doi: 10.2196/84623] [Medline: 42114021]

Abbreviations

***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition)
DSM-5 GAD-C: Severity Measure for Generalized Anxiety Disorder–Child Age 11-17
GAD-7: Generalized Anxiety Disorder 7-item scale

Submitted 31.Aug.2026; this is a non–peer-reviewed article;accepted 14.Sep.2026; published 09.Nov.2026.

Please cite as:
Sicard V, Cairncross M, Zemek R, Silverberg N, Goldfield GS, Reed N, Saab BJ, Smith A, Ledoux AA
Correction: Smartphone App–Delivered Mindfulness-Based Intervention for Concussion in Adolescents (MBI-4-mTBI): Feasibility Randomized Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e110902
URL: https://rehab.jmir.org/2026/1/e110902
doi:10.2196/110902

© Veronik Sicard, Molly Cairncross, Roger Zemek, Noah Silverberg, Gary Goldfield, Nick Reed, Bechara J Saab, Andra Smith, Andrée-Anne Ledoux. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 9.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Addendum: Managing Musculoskeletal Pain in Older Adults Through a Digital Care Solution: Secondary Analysis of a Prospective Clinical Study

Anabela C Areias[1], PhD; Dora Janela[1], PT; Maria Molinos[1], PhD; Robert G Moulder[1,2], PhD; Virgílio Bento[1], PhD; Vijay Yanamadala[1,3,4], MBA, MD; Steven P Cohen[5,6], MD, PhD; Fernando Dias Correia[1,7], MD, PhD; Fabíola Costa[1], PhD

[1]Sword Health, Inc, New York, NY, United States

[2]Institute for Cognitive Science, University of Colorado Boulder, Boulder, CO, United States

[3]Department of Surgery, Frank H Netter School of Medicine, Quinnipiac University, Hamden, CT, United States

[4]Department of Neurosurgery, Hartford Healthcare Medical Group, Westport, CT, United States

[5]Departments of Anesthesiology & Critical Care Medicine, Physical Medicine and Rehabilitation, Neurology, and Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD, United States

[6]Departments of Anesthesiology and Physical Medicine and Rehabilitation, Uniformed Services University of the Health Sciences, Bethesda, CT, United States

[7]Neurology Department, Centro Hospitalar e Universitário do Porto, Porto, Portugal

Corresponding Author:
Fabíola Costa, PhD
Sword Health, Inc
169 Madison Ave, Suite 15501
New York, NY, 10016
United States
Phone: 1 385 308 8034
Email: f.costa@swordhealth.com

Related Article:

Correction of: https://rehab.jmir.org/2023/1/e49673

(JMIR Rehabil Assist Technol 2026;13:e97495) doi:10.2196/97495

The authors of the publication “Managing Musculoskeletal Pain in Older Adults Through a Digital Care Solution: Secondary Analysis of a Prospective Clinical Study” [1] have introduced further details regarding the exercise component of the intervention to improve transparency and interpretability. To provide a generalized framework underlying the exercise prescription within the digital care program, Multimedia Appendix 2 was added (included here as Multimedia Appendix 1), including representative exercise categories, progression principles, and dosage parameter ranges. Additionally, a detailed description of a condition-specific implementation of digital care programs was reported in a previously published study [2], cited in the Introduction section of the original article. These clarifications do not impact the results or conclusions of the original study.

Multimedia Appendix 1
Overview of the exercise prescription framework, including exercise characteristics, dosage parameters, and progression principles.
[DOCX File , 16 KB - rehab_v13i1e97495_app1.docx]

References

1. Areias AC, Janela D, Molinos M, Moulder RG, Bento V, Yanamadala V, et al. Managing musculoskeletal pain in older adults through a digital care solution: secondary analysis of a prospective clinical study. JMIR Rehabil Assist Technol 2023 Aug 15;10:e49673 [FREE Full text] [doi: 10.2196/49673] [Medline: 37465960]

XSL•FO
RenderX

2. Correia FD, Nogueira A, Magalhães I, Magalhães J, Moreira M, Barradas I, et al. Home-based rehabilitation with a novel digital biofeedback system versus conventional in-person rehabilitation after total knee replacement: a feasibility study. Sci Rep 2018 Jul 26;8(1):11299 [FREE Full text] [doi: 10.1038/s41598-018-29668-0] [Medline: 30050087]

Edited by JMIR Publications Editorial Office; submitted 08.Apr.2026; this is a non–peer-reviewed article;accepted 18.Aug.2026; published 27.Aug.2026.

Please cite as:
Areias AC, Janela D, Molinos M, Moulder RG, Bento V, Yanamadala V, Cohen SP, Correia FD, Costa F
Addendum: Managing Musculoskeletal Pain in Older Adults Through a Digital Care Solution: Secondary Analysis of a Prospective Clinical Study
JMIR Rehabil Assist Technol 2026;13:e97495
URL: https://rehab.jmir.org/2026/1/e97495
doi:10.2196/97495
PMID:

©Anabela C Areias, Dora Janela, Maria Molinos, Robert G Moulder, Virgílio Bento, Vijay Yanamadala, Steven P Cohen, Fernando Dias Correia, Fabíola Costa. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 27.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Crossover Analysis Is Needed in Smartphone-Supported Vestibular Rehabilitation

Jin Yan[1*], MMed; Kaiyuan Yang[2*], MD, PhD; Xixin Wang[2], MD, PhD; Boyan Su[2], MD, PhD

[1]Clinical Cognitive Neuroscience, Advanced Occupational Therapy, Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Kyoto, Japan

[2]Department of Neurosurgery, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, 168 Litang Road, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Boyan Su, MD, PhD
Department of Neurosurgery, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, 168 Litang Road, Beijing, China

Related Articles:

https://rehab.jmir.org/2026/1/e84207

Comment in: https://rehab.jmir.org/2026/1/e109396

(JMIR Rehabil Assist Technol 2026;13:e107851) doi:10.2196/107851

KEYWORDS

vestibular rehabilitation; smartphone; crossover trial; digital health; statistical analysis

Kaplan et al [1] evaluated smartphone-supported vestibular rehabilitation in a randomized 2-period, 2-sequence crossover pilot trial and interpreted their findings as suggesting app support conferred an advantage. However, the reported analyses did not directly estimate the incremental effect of the app.

Participants were randomized to app-first or conventional-first sequences, crossed over after 3 weeks, and had no washout period; residual effects may therefore have influenced period 2 [1]. The CONSORT (Consolidated Standards of Reporting Trials) crossover extension emphasizes paired participant-level analyses that account for within-participant correlation [2]. In contrast, Mann-Whitney *U* tests were used to compare sequences, whereas Friedman and Wilcoxon tests were used to compare time points within sequences. A comparison between baseline and week 6 provides an estimate of the total 6-week change, and a between-sequence comparison of that change provides an estimate of the order difference after both sequences received both delivery modes. Neither isolates app support from conventional delivery.

The Instrumented Timed Up and Go Test result illustrates the problem. In the app-first sequence, the only significant post hoc change was from baseline to week 6, spanning both treatment periods, not from baseline to week 3, the app-supported period. The results therefore cannot establish an early benefit of smartphone guidance. Moreover, significance in one sequence but not the other does not establish a difference between treatments or sequences [3].

The Dizziness Handicap Inventory responder analyses have the same limitation. At week 6, 7 of 9 eligible app-first participants and 2 of 8 conventional-first participants exceeded the minimal clinically important difference, but both sequences had received both modes [1]. The period-specific counts (6/17 vs 3/17) are paired outcomes, so discordant-pair counts are needed. Two *P* values also require clarification. For the week 6 table, 7 of 9 versus 2 of 8 yields a likelihood-ratio χ^2_1=4.98 (*P*=.026), whereas the 2-sided Fisher exact test gives *P*=.057; the manuscript reports χ^2_1=4.97 (*P*=.06). For period-specific responders, treating 6 of 17 and 3 of 17 as independent observations yields *P*=.438 with the 2-sided Fisher exact test, not *P*=.057. However, because these are paired outcomes, the appropriate analysis requires the discordant-pair counts, which were not reported.

We recommend reporting period-specific outcomes using a participant-level model that estimates treatment and period effects and considers sequence and carryover. Given the absence of washout and plausible persistence of rehabilitation effects, an exploratory baseline-adjusted first-period comparison preserving randomization may be more interpretable than a full crossover contrast. Until then, the study supports feasibility and overall improvement across 6 weeks, but does not establish an incremental benefit of app support.

XSL•FO
RenderX

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Kaplan A, Kalderon L, Wolfovitz A, Gimmon Y, Levy-Tzedek S. Smartphone-supported vestibular rehabilitation in individuals with vestibular dysfunction: pilot randomized crossover trial assessing functional clinical outcomes and anxiety. JMIR Rehabil Assist Technol 2026 Mar 24;13:e84207. [doi: 10.2196/84207] [Medline: 41889357]
2. Dwan K, Li T, Altman DG, Elbourne D. CONSORT 2010 statement: extension to randomised crossover trials. BMJ 2019 Jul 31;366:l4378. [doi: 10.1136/bmj.l4378] [Medline: 31366597]
3. Gelman A, Stern H. The difference between "significant" and "not significant" is not itself statistically significant. Am Stat 2006 Nov;60(4):328-331. [doi: 10.1198/000313006X152649]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials

Edited by J Sarvestan; submitted 24.Jul.2026; this is a non–peer-reviewed article;accepted 15.Sep.2026; published 30.Sep.2026.

Please cite as:
Yan J, Yang K, Wang X, Su B
Crossover Analysis Is Needed in Smartphone-Supported Vestibular Rehabilitation
JMIR Rehabil Assist Technol 2026;13:e107851
URL: https://rehab.jmir.org/2026/1/e107851
doi:10.2196/107851

© Jin Yan, Kaiyuan Yang, Xixin Wang, Boyan Su. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clarifying Methodological Principles and Preliminary Trends in Digital Vestibular Rehabilitation: Authors' Reply

Azriel Kaplan[1], MSc, PT; Liran Kalderon[1], PT, PhD; Amit Wolfovitz[2], MD; Yoav Gimmon[2,3*], PT, PhD; Shelly Levy-Tzedek[1,4,5*], PhD

[1]Department of Physical Therapy, Recanati School for Community Health Professions, Ben-Gurion University of the Negev, David Ben-Gurion Blvd 1, Beer-Sheva, Southern District, Israel

[2]Department of Otolaryngology—Head and Neck Surgery, Sheba Medical Center, Ramat Gan, Tel Aviv, Israel

[3]Department of Physical Therapy, Faculty of Social Welfare & Health Sciences, University of Haifa, Haifa, Haifa, Israel

[4]Zelman Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Southern District, Israel

[5]Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Freiburg, Baden-Wurttemberg, Germany

[*]these authors contributed equally

Corresponding Author:
Azriel Kaplan, MSc, PT
Department of Physical Therapy, Recanati School for Community Health Professions, Ben-Gurion University of the Negev, David Ben-Gurion Blvd 1, Beer-Sheva, Southern District, Israel

Related Articles:

https://rehab.jmir.org/2026/1/e84207

https://rehab.jmir.org/2026/1/e107851

(JMIR Rehabil Assist Technol 2026;13:e109396) doi:10.2196/109396

KEYWORDS

vestibular rehabilitation; smartphone; crossover trial; digital health; statistical analysis

To the Editor

We welcome the commentary [1] regarding our publication [2] evaluating the Vestibulon app in vestibular rehabilitation. We appreciate the opportunity to clarify the clinical and statistical considerations guiding our pilot trial and to contextualize our findings.

The author raises a concern regarding the absence of a washout period. The absence of a washout period was a deliberate clinical decision. Vestibular rehabilitation involves central compensation, adaptation, habituation, and motor learning, which may not be reliably reversed by a short interruption. Because participants were symptomatic and receiving clinically indicated care, continuity of rehabilitation was prioritized. We nevertheless acknowledge that carryover from the first period may have limited the interpretation of period-2 comparisons, as noted in our Limitations section [2].

Given the pilot nature of the study and the small sample size, we used prespecified nonparametric methods for the exploratory comparisons. This approach was appropriate for the bounded and potentially skewed Dizziness Handicap Inventory (DHI) and State-Trait Anxiety Inventory scores, which derive from ordinal response categories, and was applied consistently to the continuous Instrumented Timed Up and Go (iTUG) measure. We acknowledge that these tests do not constitute a complete crossover analysis accounting for treatment, period, sequence, within-participant correlation, and potential carryover. This exploratory analysis was not intended to provide definitive evidence of the app's efficacy, but rather to provide initial clinical insights and establish feasibility for future studies.

Across all sections, we explicitly refrained from making definitive claims of superiority, including regarding early iTUG changes, repeatedly framing results as "preliminary trends." Importantly, our Limitations section explicitly outlined the constraints noted by the author: the pilot sample size, potential carryover effects due to the absence of a washout period, and the essential need for a future fully powered randomized controlled trial (RCT).

We thank the author for identifying the typographical error on page 9, where P=.057 was inadvertently reported twice. This P value (.057) corresponds to the 2-sided Fisher exact test for the week 6 comparison of 7 of 9 versus 2 of 8 minimal clinically important difference (MCID) responders. Furthermore, for period-specific responder outcomes (6/17 versus 3/17), a valid paired analysis relied on discordant pairs, with 5 participants achieving the MCID strictly during the app period and 2 strictly during the conventional period (exact McNemar test: P=.453). Therefore, they should not be analyzed as independent samples alone. As a supplementary analysis less affected by carryover, we analyzed period 1 using an analysis of covariance. Baseline-adjusted week 1 DHI scores were 38.82 in the app-first

XSL•FO
RenderX

group and 40.33 in the conventional-first group, a difference that was not statistically significant (F=0.04, P=.840, partial η^2=.002). The DHI MCID was achieved by 5 of 9 versus 1 of 8 eligible participants (55.6% vs 12.5%; P=.131). Although exploratory for definitive inference, these findings provide a useful trend and preliminary data for an RCT.

In conclusion, our pilot trial supports the feasibility of integrating smartphone support into vestibular rehabilitation and observed improvement in functional outcomes over the 6-week rehabilitation program. We agree that establishing definitive incremental efficacy requires a parallel-group RCT, and we thank the author for this constructive scientific dialogue.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Yan J, Yang K, Wang X, Su B. Crossover analysis is needed in smartphone-supported vestibular rehabilitation. JMIR Rehabil Assist Technol 2026;13:e107851. [doi: 10.2196/107851]
2. Kaplan A, Kalderon L, Wolfovitz A, Gimmon Y, Levy-Tzedek S. Smartphone-supported vestibular rehabilitation in individuals with vestibular dysfunction: pilot randomized crossover trial assessing functional clinical outcomes and anxiety. JMIR Rehabil Assist Technol 2026 Mar 24;13:e84207. [doi: 10.2196/84207] [Medline: 41889357]

Abbreviations

DHI: Dizziness Handicap Inventory
iTUG: Instrumented Timed Up and Go
MCID: minimal clinically important difference
RCT: randomized controlled trial

Edited by J Sarvestan; submitted 12.Aug.2026; this is a non–peer-reviewed article;accepted 15.Sep.2026; published 30.Sep.2026.

Please cite as:
Kaplan A, Kalderon L, Wolfovitz A, Gimmon Y, Levy-Tzedek S
Clarifying Methodological Principles and Preliminary Trends in Digital Vestibular Rehabilitation: Authors' Reply
JMIR Rehabil Assist Technol 2026;13:e109396
URL: https://rehab.jmir.org/2026/1/e109396
doi:10.2196/109396

© Azriel Kaplan, Liran Kalderon, Amit Wolfovitz, Yoav Gimmon, Shelly Levy-Tzedek. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Barriers to Adoption of Electronic Low Vision Aids Among Eye Care Professionals in Jordan: Descriptive Cross-Sectional Study

Areej Okasheh-Otoom, PhD

Faculty of Applied Medical Sciences, Jordan University of Science and Technology (JUST), PO Box 3030, Irbid, Jordan

Corresponding Author:
Areej Okasheh-Otoom, PhD
Faculty of Applied Medical Sciences, Jordan University of Science and Technology (JUST), PO Box 3030, Irbid, Jordan

Abstract

Background: Digital, smart, and electronic low vision aids (LVAs) have expanded options for visual rehabilitation and functional independence among people with visual impairment. However, adoption of these technologies remains limited, particularly in low- and middle-income countries such as Jordan, where access, affordability, and training resources may be constrained.

Objective: To examine barriers to the adoption of electronic LVAs and identify factors associated with their use among eye care professionals in Jordan.

Methods: A descriptive cross-sectional survey was conducted among 270 eye care professionals working in hospitals, rehabilitation centers, and private clinics across Jordan. The questionnaire assessed awareness, training exposure, institutional support, and perceived barriers related to electronic LVAs. Descriptive statistics and inferential analyses were used to examine adoption patterns and predictors, with statistical significance set at $P<.05$.

Results: Of the 270 participants, 156 (57.8%) were optometrists, 78 (28.9%) were ophthalmologists, and 36 (13.3%) were low vision specialists. The mean age was 36 (SD 8) years, and the mean professional experience was 12 (SD 6) years. Overall, 117 of 270 (42.2%) participants reported current use or recommendation of electronic LVAs. The most frequently reported barriers were high device cost (n=213, 79%), lack of training (n=184, 68.1%), limited institutional support (n=173, 64%), and low patient awareness (n=154, 57%). In multivariable analysis, greater training exposure (odds ratio [OR] 1.82, 95% CI 1.31 - 2.53; $P<.001$), stronger institutional support (OR 1.48, 95% CI 1.12 - 1.96; $P=.008$), and higher awareness scores (OR 1.35, 95% CI 1.05 - 1.72; $P=.02$) were positively associated with aid adoption, whereas high device cost was negatively associated with aid adoption (OR 0.41, 95% CI 0.27 - 0.62; $P<.001$).

Conclusions: Adoption of electronic LVAs among eye care professionals in Jordan remains limited. Cost, training exposure, and institutional support are key factors influencing uptake. These findings suggest that strengthening professional training and institutional support may facilitate broader integration of electronic LVAs into low vision rehabilitation services.

(JMIR Rehabil Assist Technol 2026;13:e87685) doi:10.2196/87685

KEYWORDS

low vision; electronic low vision aids; assistive technology; digital rehabilitation; adoption barriers; Jordan; optometry; ophthalmology

Introduction

Visual impairment affects more than 2.2 billion people worldwide, and at least 1 billion cases could have been prevented or remain unaddressed [1]. Low vision rehabilitation aims to enhance residual vision and functional independence using optical and electronic assistive technologies [2]. Recent advances in portable, digital, and smart low vision aids (LVAs), including magnifiers, wearable displays, and smartphone-based systems, have significantly improved navigation and access to information [2-5]. These devices, collectively termed electronic LVAs, complement traditional optical aids by offering adjustable magnification, contrast enhancement, and speech feedback [6,7].

Despite growing evidence supporting their functional and psychosocial benefits [8-11], adoption of electronic LVAs remains limited globally [12]. Reported barriers include high device cost, inadequate professional training, limited awareness among clinicians and patients, and insufficient institutional and policy support [13-15]. These challenges are particularly pronounced in low- and middle-income countries, where limited service infrastructure, fragmented referral pathways, and absence of sustainable funding mechanisms further restrict access to electronic low vision technologies [2].

In Jordan, visual impairment represents the most prevalent form of disability, affecting approximately 6% to 7% of the population aged 5 years or older, according to national statistics [16]. Although optometric education and eye care infrastructure

XSL•FO
RenderX

have expanded in recent years, integration of structured low vision rehabilitation services within the health care system remains inconsistent [7,15,17]. Available local evidence indicates that assistive and digital technologies are used predominantly in educational or professional contexts, with limited incorporation into formal low vision rehabilitation pathways. Regional studies further suggest that fewer than half of eye care professionals routinely prescribe or recommend electronic LVAs, underscoring persistent gaps in training, access, and service delivery [3,18,19]. Understanding context-specific barriers in Jordan is therefore essential for developing targeted professional training and service-level interventions.

Technology adoption frameworks, including the technology acceptance model and the unified theory of acceptance and use of technology, provide a useful lens for examining these challenges [20]. These models emphasize perceived usefulness, ease of use, and facilitating conditions as key determinants of adoption [21]. In low vision rehabilitation, these constructs translate into clinician awareness, perceived patient benefit, affordability, and availability of institutional support. While prior studies have examined assistive technology use in various contexts, limited evidence exists regarding how these adoption factors interact among eye care professionals involved in low vision care in Jordan. The aim of this study was to examine awareness, training exposure, institutional support, perceived barriers, and factors associated with the adoption and recommendation of electronic low vision aids among eye care professionals in Jordan.

Methods

Study Design

A descriptive cross-sectional study was conducted among eye care professionals working in public and private hospitals, university eye clinics, and vision rehabilitation centers across Jordan. This study was implemented as an online cross-sectional survey and is reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guidelines (Checklist 1). Invitations were distributed electronically through established professional networks, including national professional associations, institutional mailing lists, and closed professional communication groups (eg, email lists and secure messaging platforms) commonly used by ophthalmologists, optometrists, and low-vision specialists in Jordan. These networks included clinicians working in public, private, and academic settings across Jordan. Distribution was not stratified by geographic region or sector; therefore, participation depended on individual access to and engagement with these professional channels. Survey items were presented in a fixed order. Completion of all mandatory closed-ended items was required for submission, whereas open-ended questions were optional. Only fully completed questionnaires were included in the analysis.

Eligible participants included optometrists, ophthalmologists, and low-vision specialists currently practicing in Jordan. Inclusion criteria were (1) possession of a valid professional license from the Ministry of Health, (2) at least 1 year of clinical experience, and (3) direct involvement in patient care. Professionals in administrative roles or not directly engaged in patient management were excluded. A total of 300 invitations were distributed electronically, of which 270 completed responses were received (response rate 90%). The response rate was calculated at the overall survey level based on completed questionnaires relative to the total number of invitations distributed. The survey was distributed through professional online groups and networks of eye care clinicians. Because the questionnaire was not sent to a fixed list of individuals, accurate denominators for each professional group were not available; therefore, group-specific response rates could not be calculated. An a priori sample size estimation was conducted using G*Power (version 3.1) [22] for multiple linear regression analysis. Assuming a medium effect size (f^2=0.15), an α level of .05, statistical power of 0.80, and up to 10 predictors, the minimum required sample size was 118 participants. The final sample of 270 respondents exceeded this requirement, providing adequate statistical power for the planned analyses.

A self-administered electronic questionnaire was developed following a review of existing validated surveys on assistive technology adoption [10-12,14,18,23]. The instrument comprised five domains: (1) awareness, (2) training exposure, (3) institutional support, (4) perceived barriers, and (5) adoption of electronic low vision aids (Table 1). Items were rated on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). Optional open-ended questions were included at the end of the survey to allow participants to provide additional comments or perspectives not captured by the closed-ended items. Selected anonymized comments are presented in the Results section to illustrate key quantitative findings; no formal qualitative coding or thematic analysis was performed.

XSL•FO
RenderX

Table . Survey domains, representative items, and mean Likert scores (5-point scale; 1=strongly disagree, 5=strongly agree; N=270 respondents).

Domain and representative item		Score, mean (SD)
Awareness		
	I am familiar with the main types of electronic LVAs.[a]	3.6 (0.7)
	I understand the difference between optical and electronic devices.	3.5 (0.8)
	I know about digital magnifiers and smart glasses.	3.8 (0.8)
	I am aware of the benefits of electronic LVAs in improving reading and mobility.	3.7 (0.7)
	I have read or heard about electronic LVAs through professional sources.	3.4 (0.9)
	I know where to obtain information about new low vision technologies.	3.5 (0.8)
Training exposure		
	I attended workshops or training sessions on electronic LVAs.	2.7 (0.9)
	I received hands-on demonstration during my education.	2.8 (0.8)
	I participated in online webinars related to electronic LVAs.	2.9 (0.9)
	I practiced using electronic magnifiers with patients.	2.6 (0.9)
	I have self-trained using online tutorials or videos.	2.7 (0.8)
Institutional support		
	My workplace provides access to low vision devices for demonstration.	2.9 (0.8)
	My department encourages use of electronic aids in patient management.	3.0 (0.9)
	Institutional funding is available for electronic LVA training.	2.7 (0.9)
	My clinic has a designated low vision service or room.	2.8 (0.8)
	Technical support is available for device maintenance.	2.8 (0.9)
Perceived barriers		
	The high cost of devices limits their use in clinical practice.	4.2 (0.7)
	Lack of professional training prevents effective use of electronic LVAs.	3.9 (0.8)
	Limited institutional support restricts adoption.	4.1 (0.8)
	Low patient awareness	3.8 (0.8)
	There is a lack of Arabic-language software support.	3.7 (0.9)
	Devices are difficult to use for older adults or individuals with limited formal education.	3.9 (0.7)
	There is limited technical support for device maintenance.	3.6 (0.8)
	Import regulations make devices unavailable.	3.5 (0.9)

XSL•FO
RenderX

Domain and representative item		Score, mean (SD)
Adoption behavior		
	I currently prescribe or recommend electronic LVAs to my patients.	3.5 (0.8)
	I routinely demonstrate electronic LVAs in my clinic.	3.4 (0.8)
	I feel confident teaching patients how to use electronic LVAs.	3.6 (0.7)
	I intend to integrate electronic LVAs more frequently in future practice.	3.8 (0.8)
	I believe electronic LVAs improve my patients' quality of life.	4.1 (0.7)

[a]LVA: low vision aid.

The survey was administered in English, as English is the primary language of instruction and clinical documentation among eye care professionals in Jordan. Participants were able to request clarification for any items if needed. The questionnaire underwent expert review by 3 senior vision rehabilitation specialists to establish content validity, followed by pilot testing with 25 clinicians. Pilot data were excluded from the final analysis. Minor linguistic refinements were made to improve clarity and internal consistency. Internal consistency was assessed using Cronbach α for each domain (Table 2). Values ≥0.70 were considered acceptable.

Table . Internal consistency (Cronbach α) of questionnaire domains based on the final survey sample (N=270).

Domain	Cronbach α
Awareness	0.86
Training exposure	0.81
Institutional support	0.83
Perceived barriers	0.88
Adoption behavior	0.79

Data were analyzed using SPSS software (version 29; IBM Corp). Descriptive statistics summarized participant characteristics and mean domain scores. Inferential analyses included chi-square tests to compare aid adoption rates across professional groups, independent-sample *t*-tests to compare training hours between adopters and nonadopters, and multivariable binary logistic regression to identify predictors of electronic low vision aid adoption. Multicollinearity was assessed using variance inflation factors (<2). Model fit was evaluated using the Hosmer-Lemeshow test (P>.05). Statistical significance was set at P<.05 (2-tailed).

Professional experience was assessed using a single self-reported item asking participants to indicate their years of clinical practice in eye care. Awareness of electronic and digital low vision aids was measured using a composite awareness score derived from multiple Likert scale items assessing familiarity with device types, indications, and clinical applications. Training exposure was assessed by asking participants whether they had received formal or informal training related to electronic LVAs (yes or no), as well as the total number of training hours completed.

Internal consistency was calculated using Cronbach α for each domain (Table 2). Values ≥0.70 were considered acceptable. The α coefficients were calculated using responses from the final survey participants. Expert review (n=3) was used solely for content validity and was not included in the final internal consistency analysis.

Ethical Considerations

This study was approved by the Institutional Review Board of Jordan University of Science and Technology (IRB/642) and was conducted in accordance with the principles of the Declaration of Helsinki. Participation was voluntary, and electronic informed consent was obtained from all participants prior to survey initiation. The survey was administered anonymously through a secure online platform. No personally identifiable information (such as names, identification numbers, or contact details) was collected. Data were stored on password-protected devices accessible only to the principal investigator and were analyzed in aggregate form to ensure participant confidentiality. Measures were implemented to minimize duplicate submissions.

Results

A total of 270 eye care professionals completed the survey, representing a 90% response rate. Participants included 156 optometrists (57.8%), 78 ophthalmologists (28.9%), and 36 low vision specialists (13.3%) working across multiple practice settings in Jordan. The mean age was 36 (SD 8) years, and 149 of 270 (55.2%) participants were female. Respondents reported

XSL•FO
RenderX

a mean of 12 (SD 6) years of professional experience, indicating a clinically experienced cohort.

Across all participants, the mean awareness score regarding electronic LVAs was 3.2 (SD 0.7) on a 5-point Likert scale, suggesting moderate familiarity. The mean training exposure time was 18.4 (SD 12.5) hours. However, substantial variability was observed, with many clinicians reporting fewer than 10 hours of formal training. These data are summarized in Table 3.

Table . Participant characteristics (N=270).

Characteristic	Value
Profession, n (%)	
Optometrist	156 (57.8)
Ophthalmologist	78 (28.9)
Low vision specialist	36 (13.3)
Gender, n (%)	
Female	149 (55.2)
Male	121 (44.8)
Age (y), mean (SD)	36 (8)
Professional experience (y), mean (SD)	12 (6)
Awareness score (1-5), mean (SD)	3.2 (0.7)
Training hours, mean (SD)	18.4 (12.5)

Profession-specific analyses revealed that low vision specialists had the highest training exposure (mean 28.5, SD 9.2 hours) and awareness score (mean 3.8, SD 0.6), followed by optometrists (mean 17.6, SD 11.3 training hours; mean 3.3, SD 0.7 awareness score) and ophthalmologists (mean 12.1, SD 8.5 training hours; mean 3.0, SD 0.6 awareness score). Gender and years of experience were not significantly associated with awareness or training exposure (P>.05).

Among all respondents, 117 of 270 professionals (42.2%) reported that they currently prescribe or recommend electronic LVAs to patients. Adoption was significantly higher among low vision specialists (22/36, 61%) than optometrists (69/156, 44.2%) and ophthalmologists (26/78, 33%) (χ^2=9.64; P=.008). Eye care professionals who reported active aid adoption also demonstrated substantially greater training exposure (mean 23.6, SD 11.2 hours) than nonadopters (mean 14.9, SD 9.8 hours), a difference that was statistically significant (t=5.74; P<.001). Additionally, adopters also had higher awareness scores than nonadopters (mean 3.6, SD 0.6 vs mean 3.0, SD 0.7; P<.001) and higher institutional support scores (mean 3.4, SD 0.8 vs mean 2.9, SD 0.7; P=.002).

Overall, these results suggest that professional experience alone does not drive aid adoption; rather, the presence of structured training and institutional support are key determinants of electronic LVA use. The perceived barriers domain comprised 8 Likert scale items representing 8 distinct barriers to electronic LVA adoption (Table 1). Respondents most frequently cited high device cost (213/270, 79%), followed by lack of training opportunities (184/270, 68.1%), limited institutional support (173/270, 64%), and low patient awareness (154/270, 57%; Table 4) as barriers to electronic LVA adoption.

Table . Barriers to electronic low vision aid adoption among Jordanian eye care professionals (N=270).

Perceived barriers	Respondents, n (%)
Import regulations make devices unavailable	111 (41.1)
There is limited technical support for device maintenance	124 (45.9)
Devices are difficult to use for older adults or individuals with limited formal education	130 (48.1)
There is a lack of Arabic-language software support	140 (51.9)
Low patient awareness	154 (57)
Limited institutional support restricts adoption	173 (64)
Lack of professional training prevents effective use of electronic low vision aids	184 (68.1)
The high cost of devices limits their use in clinical practice	213 (79)

Subgroup analysis revealed that cost concerns were particularly pronounced among optometrists (131/156, 84.0%), whereas low vision specialists reported greater concern regarding limited patient awareness (22/36, 61%) and lack of Arabic-language

XSL•FO
RenderX

device interfaces (20/36, 56%). In contrast, ophthalmologists more frequently emphasized institutional limitations (56/78, 72%) and unclear referral mechanisms (50/78, 64%) as major obstacles.

One-way ANOVA demonstrated significant between-group differences for lack of professional training ($F_{2,267}$=6.21; P=.002) and limited maintenance or technical support ($F_{2,267}$=4.89; P=.008). Post hoc analyses indicated that optometrists perceived these barriers more strongly than ophthalmologists and low vision specialists. Although cost- and patient-related barriers were commonly reported across all professional groups, no statistically significant between-group differences were observed for these domains (Multimedia Appendix 1). Comments from participants reinforced these findings, as illustrated by the following representative quotes:

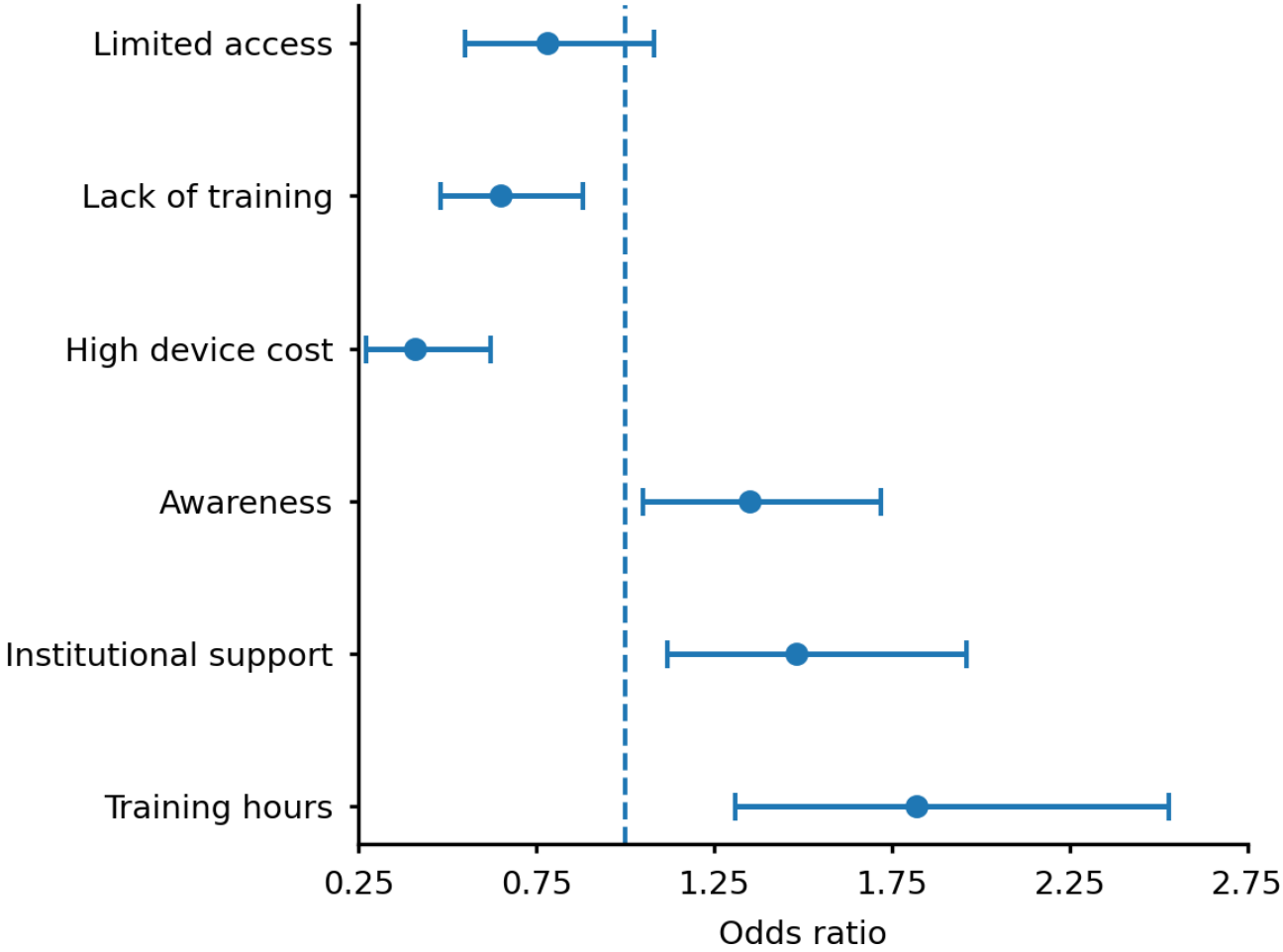
> *Even when devices are available, few clinicians are trained to demonstrate them effectively.* [Participant 47, optometrist]
>
> *Cost remains prohibitive for many families, especially without Ministry of Health support.* [Participant 112, ophthalmologist]

These results point to a multifactorial set of barriers (financial, educational, and systemic) that collectively hinder widespread electronic LVA integration into clinical practice.

A binary logistic regression model was fitted to identify predictors of electronic LVA adoption (Multimedia Appendix 2). The model explained 42% of the variance in adoption behavior (Nagelkerke R^2=0.42) and correctly classified 79% of cases. Training hours emerged as the strongest positive predictor of electronic low vision aid adoption (odds ratio [OR] 1.82, 95% CI 1.31 - 2.53; P<.001). Higher institutional support (OR 1.48, 95% CI 1.12 - 1.96; P=.008) and greater awareness (OR 1.35, 95% CI 1.05 - 1.72; P=.02) were also significantly associated with increased adoption. In contrast, high device cost was a strong negative predictor (OR 0.41, 95% CI 0.27 - 0.62; P<.001), while lack of training was associated with a reduced likelihood of adoption (OR 0.65, 95% CI 0.48 - 0.88; P=.006). Limited access was not a statistically significant predictor (OR 0.78, 95% CI 0.55 - 1.08; P=.14).

Figure 1 illustrates the direction and strength of associations, showing that training exposure and institutional support exerted the most substantial positive influence, while high device cost exerted the strongest negative effect.

Figure 1. Forest plot showing adjusted odds ratios (with 95% CIs) for predictors of electronic low vision aid adoption.



Discussion

This study examined barriers to the adoption and recommendation of electronic LVAs among eye care professionals in Jordan. The main findings indicate that high device cost and lack of professional training were significantly associated with a lower likelihood of recommending electronic LVAs, whereas higher awareness and stronger institutional

XSL•FO
RenderX

support were positively associated with adoption. Although overall awareness scores were moderate, fewer than half of respondents reported current use or recommendation of these devices, indicating a persistent gap between awareness and routine clinical implementation. These findings are consistent with reports from other low- and middle-income countries, where cost and training constraints are repeatedly identified as dominant barriers to electronic assistive technology uptake [10-12,14,18,23].

Multivariable regression analysis further demonstrated that training exposure, institutional support, and awareness independently predicted adoption of electronic LVAs, whereas perceived device cost was negatively associated with adoption. Together, these findings suggest that adoption is influenced by both individual-level factors (eg, professional knowledge and perceived benefit) and system-level factors (eg, organizational support and resource availability). This dual influence is consistent with prior research indicating that isolated improvements in clinician awareness are insufficient to promote sustained technology adoption without parallel investment in training and institutional infrastructure.

International studies report similar patterns of underuse. For example, surveys of eye care professionals in high-income settings have shown that fewer than half routinely recommend electronic or digital magnification devices despite recognizing their potential clinical benefit [13]. Studies from the Asia Pacific region and the Middle East likewise highlight limited institutional resources, high device cost, and insufficient professional training as persistent barriers to adoption [2,20]. These parallels suggest that the challenges identified in this study reflect broader implementation issues rather than context-specific resistance to technology.

Interpretation of these findings should also consider broader health system context in Jordan. Reports and policy analyses indicate that low vision rehabilitation services are often fragmented, with limited availability of designated low vision clinics, variable referral pathways, and uneven access to continuing professional education in rehabilitation technologies [1,12,24,25]. Although these system-level characteristics were not directly measured in this survey, they provide important context for understanding why institutional support and training emerged as prominent determinants of adoption. Global guidance from the World Health Organization emphasizes the need for integrated national rehabilitation frameworks to support equitable access to assistive and digital technologies, particularly in resource-constrained settings [3,20]. In this study, clinicians who perceived stronger institutional facilitation and greater patient benefit were more likely to recommend electronic LVAs, consistent with core constructs of the technology acceptance model and the unified theory of acceptance and use of technology framework [21].

The findings underscore the need for coordinated strategies to support electronic LVA adoption. Structured professional training in digital rehabilitation should be incorporated into undergraduate curricula and continuing professional development programs. Health care institutions may benefit from establishing demonstration or referral pathways that allow clinicians to trial devices with patients. In addition, funding mechanisms—such as subsidy schemes or public-private partnerships—may help mitigate cost barriers, particularly for older adults and children with visual impairment. These implications are aligned with international calls to strengthen rehabilitation systems and data collection under the World Health Organization Rehabilitation 2030 initiative [3].

Several limitations warrant consideration. The cross-sectional design precludes causal inference, and adoption behavior was assessed using self-reported measures rather than objective indicators such as prescription records or device use data. In addition, although the survey instrument was informed by established technology adoption frameworks, formal construct validity testing was not undertaken. This study also did not explicitly examine potential regional or socioeconomic variation within the eye care workforce, which may influence access to training, institutional support, and technology availability. These factors may limit generalizability and should be addressed in future research.

Future studies should use longitudinal or mixed methods designs to examine how training interventions, institutional policies, and funding mechanisms influence adoption over time. Incorporating validated measurement instruments, objective adoption indicators, and stratified analyses by region or practice setting would further strengthen the evidence base and inform targeted implementation strategies.

In conclusion, adoption of digital, smart, and electronic LVAs among eye care professionals in Jordan remains limited. Cost, training exposure, and institutional support were the primary factors associated with uptake in this study. These findings suggest that strengthening professional training opportunities, improving institutional facilitation, and exploring sustainable funding mechanisms may support broader integration of digital assistive technologies within low vision rehabilitation services. Greater coordination among clinicians, policymakers, and industry stakeholders may help address persistent implementation barriers and improve access to low vision care.

Acknowledgments

The author thanks the participating optometrists, ophthalmologists, and low vision specialists across Jordan for their valuable contributions. Appreciation is also extended to the statistical support team and research assistants who facilitated data entry and analysis.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

XSL•FO
RenderX

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Subgroup analysis of mean barrier scores by professional category (N=270).
[DOCX File, 14 KB - rehab_v13i1e87685_app1.docx]

Multimedia Appendix 2
Logistic regression analysis for predictors of electronic LVA adoption.
[DOCX File, 14 KB - rehab_v13i1e87685_app2.docx]

Checklist 1
CHERRIES checklist.
[DOCX File, 15 KB - rehab_v13i1e87685_app3.docx]

References

1. World report on vision. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/world-report-on-vision [accessed 2026-02-04]
2. Burton MJ, Ramke J, Marques AP, et al. The Lancet Global Health Commission on Global Eye Health: vision beyond 2020. Lancet Glob Health 2021 Apr;9(4):e489-e551. [doi: 10.1016/S2214-109X(20)30488-5] [Medline: 33607016]
3. Rehabilitation 2030: a call for action. World Health Organization. 2017. URL: https://www.who.int/publications/m/item/rehabilitation-2030-a-call-for-action [accessed 2026-02-04]
4. Miller A, Crossland MD, Macnaughton J, Latham K. The usefulness of a wearable electronic vision enhancement system for people with age-related macular degeneration: a randomized crossover trial. Transl Vis Sci Technol 2025 Sep 2;14(9):8. [doi: 10.1167/tvst.14.9.8] [Medline: 40905748]
5. Miller A, Macnaughton J, Crossland MD, Latham K. "Such a lot of bother": qualitative results of a home trial of a wearable electronic vision enhancement system for people with age-related macular degeneration. Ophthalmic Physiol Opt 2025 May;45(3):699-712. [doi: 10.1111/opo.13453] [Medline: 39865322]
6. Crossland MD, Silva RS, Macedo AF. Smartphone, tablet computer and e-reader use by people with vision impairment. Ophthalmic Physiol Opt 2014 Sep;34(5):552-557. [doi: 10.1111/opo.12136] [Medline: 25070703]
7. Bakkar MM, Alzghoul EA, Haddad MF. Clinical characteristics and causes of visual impairment in a low vision clinic in northern Jordan. Clin Ophthalmol 2018;12:631-637. [doi: 10.2147/OPTH.S153754] [Medline: 29662299]
8. Virgili G, Acosta R, Bentley SA, Giacomelli G, Allcock C, Evans JR. Reading aids for adults with low vision. Cochrane Database Syst Rev 2018 Apr 17;4(4):CD003303. [doi: 10.1002/14651858.CD003303.pub4] [Medline: 29664159]
9. van Nispen RM, Virgili G, Hoeben M, et al. Low vision rehabilitation for better quality of life in visually impaired adults. Cochrane Database Syst Rev 2020 Jan 27;1(1):CD006543. [doi: 10.1002/14651858.CD006543.pub2] [Medline: 31985055]
10. Sivakumar P, Vedachalam R, Kannusamy V, et al. Barriers in utilisation of low vision assistive products. Eye (Lond) 2020 Feb;34(2):344-351. [doi: 10.1038/s41433-019-0545-5] [Medline: 31388131]
11. Sarika G, Venugopal D, Sailaja MVS, Evangeline S, Krishna Kumar R. Barriers and enablers to low vision care services in a tertiary eye care hospital: a mixed method study. Indian J Ophthalmol 2019 Apr;67(4):536-540. [doi: 10.4103/ijo.IJO_1215_18] [Medline: 30900589]
12. Stolwijk ML, van Nispen RMA, van der Ham AJ, Veenman E, van Rens G. Barriers and facilitators in the referral pathways to low vision services from the perspective of patients and professionals: a qualitative study. BMC Health Serv Res 2023 Jan 21;23(1):64. [doi: 10.1186/s12913-022-09003-0] [Medline: 36681848]
13. Codina CJ, Rhodes M. Low vision services provision throughout NHS trusts in the UK. Br Ir Orthopt J 2023;19(1):64-70. [doi: 10.22599/bioj.293] [Medline: 37577067]
14. Arshad M, Younas A, Amin N. Assessment of level of awareness and barriers to low vision rehabilitation among optometrists in Pakistan. J Vis Impair Blind 2024 Mar;118(2):110-117. [doi: 10.1177/0145482X241248461]
15. Qutishat Y, Shublaq S, Masoud M, Alnuman N. Low vision profile in Jordan: a vision rehabilitation center-based study. Healthcare (Basel) 2021;9(1):20. [doi: 10.3390/healthcare9010020]
16. The reality of disability: "functional difficulties" in jordan, based on the data of the General Population and Housing Census 2015. : Department of Statistics (Jordan); 2021 URL: https://dosweb.dos.gov.jo/DataBank/Analytical_Reports/Disability_2021.pdf [accessed 2026-02-24]

XSL•FO
RenderX

17. Gammoh Y, Moodley V. Situation analysis of optometric education in the Eastern Mediterranean region. East Mediterr Health J 2023 Mar 26;29(3):170-176. [doi: 10.26719/emhj.23.019] [Medline: 36987622]
18. Ovenseri-Ogbomo GO, Alghamdi W. Knowledge, attitudes, and practices of optometrists regarding low vision services in Saudi Arabia. TOOPHTJ 2021 Nov 10;15(1):217-288. [doi: 10.2174/1874364102115010217]
19. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
20. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
21. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
22. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
23. Alsaqr AM. Barriers to low vision services among optometrists in Saudi Arabia. TOOPHTJ 2021 Oct 5;15(1):178-188. [doi: 10.2174/1874364102115010178]
24. Global report on assistive technology. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240049451 [accessed 2026-02-24]
25. IAPB vision atlas. International Agency for the Prevention of Blindness. 2023. URL: https://visionatlas.iapb.org/ [accessed 2026-02-06]

Abbreviations

CHERRIES: Checklist for the Reporting of Results of Internet E-Surveys
LVA: low vision aid
OR: odds ratio

Edited by S Munce; submitted 12.Nov.2025; peer-reviewed by A Anderson, MM Hatamleh; revised version received 06.Feb.2026; accepted 10.Feb.2026; published 02.Mar.2026.

Please cite as:
Okasheh-Otoom A
Barriers to Adoption of Electronic Low Vision Aids Among Eye Care Professionals in Jordan: Descriptive Cross-Sectional Study
JMIR Rehabil Assist Technol 2026;13:e87685
URL: https://rehab.jmir.org/2026/1/e87685
doi:10.2196/87685

© Areej Okasheh-Otoom. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Can GPT-5 Support Licensing Examination Preparation? Analysis of Accuracy, Reasoning, and Semantic Similarity Across Rehabilitation Disciplines

Christy Muasher-Kerwin[1], MHA, PhD, DPT; M Courtney Hughes[2], MS, PhD; Aida Sanatizadeh[3], PhD

[1]Physical Therapy Department, North Central College, 160 E Chicago Ave, Naperville, IL, United States

[2]College of Health and Human Sciences, Northern Illinois University, Dekalb, IL, United States

[3]College of Business, Northern Illinois University, Dekalb, IL, United States

Corresponding Author:
Christy Muasher-Kerwin, MHA, PhD, DPT
Physical Therapy Department, North Central College, 160 E Chicago Ave, Naperville, IL, United States

Abstract

In this cross-sectional study of 300 board-style questions across physical therapy, occupational therapy, and speech-language pathology, we evaluated reasoning types and found high overall accuracy with variation by discipline and reasoning category; the strongest performance was in deductive and analytical reasoning and the lowest accuracy was in evaluative reasoning.

(JMIR Rehabil Assist Technol 2026;13:e91019) doi:10.2196/91019

KEYWORDS

large language model; clinical reasoning; licensing examination; rehabilitation education; GPT-5; semantic similarity; health professions training; artificial intelligence; board-preparation materials

Introduction

Large language models (LLMs) such as ChatGPT (GPT-5; OpenAI) are increasingly used by clinicians and trainees to support learning and licensing examination preparation [1]. Passing licensing examinations ensures practitioner competence and patient safety and is correlated with better patient outcomes [2,3]. However, LLMs' reasoning consistency and conceptual similarity to expert sources remain uncertain [1]. Some prior research has demonstrated that LLMs can achieve passing scores on various professional examinations, but their explanations may exhibit shallow logic or inconsistent reasoning [4]. Recent evaluations of ChatGPT on Chinese national licensing examinations similarly demonstrated failure to meet official pass thresholds, with accuracy varying across examination type and question format [5]. Our study assessed GPT-5's accuracy, reasoning patterns, and semantic similarity to verified board-preparation materials for rehabilitation examinations for physical therapy (PT), occupational therapy (OT), and speech-language pathology (SLP) to assess whether GPT-5 can serve as a reliable study assistant.

Methods

Overview

Three hundred multiple-choice questions (100 each from PT, OT, and SLP validated board-preparation sources that are not publicly available) were entered into GPT-5 (via the ChatGPT web interface) in November 2025 using default settings, without hints or chain-of-thought prompting, to mimic typical learner use. For each discipline, the first 100 sequential questions were included to create a reproducible sample. Difficulty indices and blueprint mappings were unavailable; therefore, items were not stratified.

Selected answers from the model and written rationales were compared with source materials. Accuracy was calculated across items and recorded as correct or incorrect. Rationales were coded by reasoning type: inductive (from specific examples to a general conclusion), deductive (from a general rule to a specific conclusion), analytical (examining parts and their relationships), evaluative (judging something based on criteria), or inferential (drawing a conclusion from evidence). This coding was based on the cognitive process required for the correct answer. Reasoning classifications were predefined within the board-preparation materials and aligned with established educational frameworks.

Semantic similarity was computed using cosine similarity applied to L2-normalized sentence embeddings generated by the all-MiniLM-L6-v2 Sentence-BERT model, and preprocessing was limited to minimal white space normalization, consistent with recommended Sentence-BERT use [6]. Descriptive statistics summarized performance across disciplines. Incorrect responses were qualitatively reviewed to identify recurring conceptual challenges. A single coder (CM-K) conducted inductive content analysis using an emergent coding approach [7] and grouped content by shared subject matter to identify the specific content domains that presented greater

XSL•FO
RenderX

difficulty for GPT-5. Initial codes were developed through iterative review of incorrect responses and GPT-5 rationales using constant comparison. A second coder (MCH) reviewed categories, and disagreements were discussed and resolved. Because coding focused on thematic categorization of recurring errors, formal interrater reliability statistics were not calculated. Consensus review was used to ensure interpretive agreement and consistency of category definitions prior to final theme reporting.

Ethical Considerations

This project used publicly available, non–human participant material and did not require institutional review board approval.

Results

GPT-5 demonstrated strong factual accuracy overall, although performance varied by discipline and reasoning type, with only deductive reasoning in PT receiving 100% (Table 1). Overall accuracy for GPT-5 was 91.0% for PT (95% CI 83.6%-95.8%), 79.0% for OT (95% CI 68.9%-85.5%), and 83.0% for SLP (95% CI 74.4%-89.6%).

Table . Accuracy by discipline and reasoning type.

Reasoning type	Physical therapy (n=100), n/N (%)	Occupational therapy (n=100), n/N (%)	Speech-language pathology (n=100), n/N (%)
Deductive	22/22 (100/100)	11/15 (73.3/100)	31/33 (93.9/100)
Analytical	22/23 (95.7/100)	9/11 (81.8/100)	40/47 (84.8/100)
Inductive	19/22 (86.4/100)	18/21 (85.7/100)	2/4 (50.0/100)
Inferential	22/25 (88.0/100)	26/36 (72.2/100)	6/9 (66.7/100)
Evaluative	6/8 (75/100)	15/17 (88.2/100)	4/7 (57.1/100)
Overall accuracy	(91.0/100)	(79.0/100)	(83.0/100)

Across all questions, the mean semantic similarity between GPT-5 and source-material rationales was 0.707. This indicates semantic alignment between GPT-5 rationales and expert rationales. Deductive reasoning had the highest mean similarity (0.730), followed by analytical (0.714), inductive (0.675), inferential (0.685), and evaluative (0.671) reasoning. The Kruskal-Wallis H test across all 5 reasoning types was also significant: H_4=11.7 (P=.02). This suggests that GPT-5's explanations aligned more closely with reference rationales for deductive reasoning tasks, where rules are applied to reach conclusions, than for tasks requiring generalization, inference, or evaluation.

Qualitative review of incorrect answers identified recurring conceptual challenges. Using inductive content analysis, challenges were grouped by shared conceptual features in GPT-5 rationales. Categories appearing across multiple questions within a discipline are summarized in Textbox 1.

Textbox 1. Summary of missed question-type patterns across disciplines.

Physical therapy

- Functional reasoning (3/9, 33.3%)
- Safety/priority decision-making (3/9, 33.3%)
- Technical recall (3/9, 33.3%)

Occupational therapy

- Functional reasoning (12/17, 70.6%)
- Professional scope and delegation reasoning (3/17, 17.6%)
- Safety-based decision-making (2/17, 11.8%)

Speech-language pathology

- Interpretive clinical reasoning for specialized assessment and instrumental findings (10/22, 45.5%)
- Evidence-based intervention selection (7/22, 31.8%)
- Procedural and ethical decision-making (5/22, 22.7%)

Discussion

In this evaluation of board-preparation materials, GPT-5 reproduced knowledge required for entry-level competence but continued to show deficits in higher-order reasoning. Because reasoning categories reflect examination item demands rather than GPT-5's internal reasoning, findings represent performance patterns across question types. Although semantic similarity to

XSL•FO
RenderX

expert rationales was high, reasoning lapses suggest that the model captures conceptual language patterns without consistently modeling clinicians' inferential logic. These findings reinforce that semantic similarity reflects linguistic alignment rather than equivalence in clinical reasoning. These findings extend previous work showing that LLMs can approach human-level factual performance while still lacking in cognitive depth and explainability [8].

Findings should be interpreted cautiously because board-preparation materials differ from official licensing examinations, which undergo formal psychometric validation, calibrated difficulty scaling, and secure item rotation. Thus, results do not represent performance on live examinations. Although accuracy approached commonly cited passing thresholds [9], comparisons are limited by scoring differences across professions.

Given the shift toward competency-based medical education and assessment frameworks, such as entrustable professional activities (EPAs) [10], the reasoning deficits observed here raise important questions about whether LLMs can support learners' progression toward entrustment. In structured settings, GPT-5 may be used as a formative adjunct, with educators leveraging model rationales to stimulate reflection and compare reasoning against established milestones or EPAs. Such supervised use supports competency development with oversight. While GPT-5 reproduced factual knowledge, its lapses in inferential and evaluative reasoning highlight limitations in supporting competencies related to clinical judgment, prioritization, and ethical decision-making.

This study is limited by its sample size, focus on rehabilitation disciplines only, and absence of learner outcome data. Because reasoning patterns may differ across specialties, future research should evaluate larger question sets, additional medical domains, intermodel comparisons, and performance across LLM models and prompting conditions. Evaluating LLM reasoning in interactive or case-based formats may reflect clinical decision tasks. Finally, longitudinal studies should examine how learners use LLMs during examination preparation to determine whether model reasoning errors influence learner understanding or create opportunities for instructional interventions.

Although GPT-5 reproduced substantial factual content, observed reasoning gaps support supervised integration in examination preparation. Because this study evaluated examination performance rather than clinical outcomes, any safety implications are theoretical. AI tools should complement, not replace, educator-guided reasoning development.

Funding

No funding was provided for this study.

Data Availability

Data can be made available upon reasonable request. The full set of examination questions and answer options cannot be publicly shared because they originate from commercially licensed board-preparation materials that are copyrighted and not authorized for redistribution.

Conflicts of Interest

None declared.

References

1. Abd-Alrazaq A, AlSaad R, Alhuwail D, et al. Large language models in medical education: opportunities, challenges, and future directions. JMIR Med Educ 2023 Jun 1;9:e48291. [doi: 10.2196/48291] [Medline: 37261894]
2. Norcini J, Grabovsky I, Barone MA, Anderson MB, Pandian RS, Mechaber AJ. The associations between United States Medical Licensing Examination Performance and outcomes of patient care. Acad Med 2024 Mar 1;99(3):325-330. [doi: 10.1097/ACM.0000000000005480] [Medline: 37816217]
3. Zong H, Wu R, Cha J, et al. Large language models in worldwide medical exams: platform development and comprehensive analysis. J Med Internet Res 2024 Dec 27;26:e66114. [doi: 10.2196/66114] [Medline: 39729356]
4. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023 Feb;2(2):e0000198. [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
5. Zong H, Li J, Wu E, Wu R, Lu J, Shen B. Performance of ChatGPT on Chinese national medical licensing examinations: a five-year examination evaluation study for physicians, pharmacists and nurses. BMC Med Educ 2024 Feb 14;24(1):143. [doi: 10.1186/s12909-024-05125-7] [Medline: 38355517]
6. Reimers N, Gurevych I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. 2019 Presented at: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP); Hong Kong, China p. 3982-3992 URL: https://www.aclweb.org/anthology/D19-1 [accessed 2026-05-04] [doi: 10.18653/v1/D19-1410]
7. Bingham AJ. From data management to actionable findings: a five-phase process of qualitative data analysis. Int J Qual Methods 2023 Oct;22. [doi: 10.1177/16094069231183620]

XSL•FO
RenderX

8. Brügge E, Ricchizzi S, Arenbeck M, et al. Large language models improve clinical decision making of medical students through patient simulation and structured feedback: a randomized controlled trial. BMC Med Educ 2024 Nov 28;24(1):1391. [doi: 10.1186/s12909-024-06399-7] [Medline: 39609823]
9. Typical PT. Understanding the NPTE passing score percentage: a comprehensive guide. URL: https://typicalpt.com/blogs/news/understanding-the-npte-passing-score-percentage-a-comprehensive-guide [accessed 2026-05-04]
10. Gummesson C, Alm S, Cederborg A, et al. Entrustable professional activities (EPAs) for undergraduate medical education - development and exploration of social validity. BMC Med Educ 2023 Sep 4;23(1):635. [doi: 10.1186/s12909-023-04621-6] [Medline: 37667366]

Abbreviations

EPA: entrustable professional activity
LLM: large language model
OT: occupational therapy
PT: physical therapy
SLP: speech-language pathology

Edited by S Munce; submitted 07.Jan.2026; peer-reviewed by B Shen, E Abdulwahed, F Fukuzawa; revised version received 25.Mar.2026; accepted 09.Apr.2026; published 21.May.2026.

Please cite as:
Muasher-Kerwin C, Hughes MC, Sanatizadeh A
Can GPT-5 Support Licensing Examination Preparation? Analysis of Accuracy, Reasoning, and Semantic Similarity Across Rehabilitation Disciplines
JMIR Rehabil Assist Technol 2026;13:e91019
URL: https://rehab.jmir.org/2026/1/e91019
doi:10.2196/91019

© Christy Muasher-Kerwin, M Courtney Hughes, Aida Sanatizadeh. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 21.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Remote Therapeutic Monitoring Use Among Commercially Insured and Medicare Advantage Patients

Joseph H Joo[1,2,3], MD, MS; Jessica I Billig[1,2,4], MD, MSc; Nadia Lieu[1,2], MS; Lingmei Zhou[1,2], MS; Joshua M Liao[1,2,3], MD, MSc

[1]Department of Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, United States

[2]Program on Policy Evaluation and Learning, Dallas, TX, United States

[3]Department of Health Economics, Systems, and Policy, University of Texas Southwestern Medical Center, Dallas, TX, United States

[4]Department of Plastic Surgery, University of Texas Southwestern Medical Center, Dallas, TX, United States

Corresponding Author:
Joseph H Joo, MD, MS
Department of Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, United States

Abstract

This national study found that remote therapeutic monitoring use increased rapidly between 2022 and 2024 but remained low overall, varied by insurance type and clinician specialty, and was concentrated among patients with musculoskeletal conditions.

(JMIR Rehabil Assist Technol 2026;13:e103600) doi:10.2196/103600

KEYWORDS

remote therapeutic monitoring; commercially insured; Medicare Advantage; patients with musculoskeletal conditions; RTM

Introduction

Remote therapeutic monitoring (RTM) enables the collection and monitoring of nonphysiological data, including functional status, treatment adherence, and patient-reported symptoms [1]. Through RTM, clinicians can gain insight into patients' responsiveness to and engagement with prescribed therapies outside of traditional clinical settings. This information can help identify changes in clinical status that require early intervention, particularly for patients with musculoskeletal, respiratory, and oncologic conditions [2-5]. However, despite the potential for RTM to support proactive care and improve clinical outcomes, studies regarding its use remain limited. Therefore, the aim of our nationwide study was to describe RTM use nationwide among commercially insured and Medicare Advantage (MA) patients.

Methods

Overview

We used 2022 - 2024 Merative MarketScan data, which contain longitudinal commercial and MA insurance claims in the United States, and identified RTM using Current Procedural Terminology codes (98975, 98976, 98977, 98980, 98981). Separately for commercial and MA claims, we assessed the total number of months of RTM use and median number of months of RTM use per patient. Patients were included if they had at least 1 month of RTM use and were enrolled in insurance for at least 11 out of 12 months in a calendar year.

We described the number and characteristics of patients using RTM with respect to age, sex, region, clinical complexity (ie, Charlson Comorbidity Index), and insurance plan type (eg, preferred provider organization, health maintenance organization) [6]. *International Classification of Diseases, Tenth Revision, Clinical Modification* codes were used to identify patient diagnoses for RTM use then grouped using Clinical Classification Software Refined categories (Multimedia Appendix 1) [7]. Lastly, RTM use was stratified by clinician and facility type as follows: primary care physicians (eg, internal medicine and family medicine providers), medical subspecialists (eg, cardiologists and pulmonologists), surgeons (eg, orthopedic and neurological surgeons), other physician specialists (eg, neurology and pain medicine specialists), advanced practice providers (eg, physician assistants and nurse practitioners), other clinicians (eg, therapists and nurses) and facilities (eg, rehabilitation or specialty care facilities) [8].

Ethical Considerations

As this study used deidentified data and did not involve human subjects as defined by federal regulations, it was determined to be exempt by the institutional review board of the University of Texas Southwestern Medical Center.

Results

Between 2022 and 2024, RTM was used by 16,116 patients (n=12,951 commercially insured and n=3165 MA patients; Table S1 in Multimedia Appendix 2). Use of RTM increased from 2089 months (n=1528 commercial and n=561 MA) in

XSL•FO
RenderX

2022 to 24,494 months (n=18,368 commercial and n=6126 MA) by 2024.

Commercially insured patients used RTM a total of 26,936 months, with a median of 1 (IQR 1 - 2) month of RTM use per patient. With over 48 million patients enrolled during the study period, this equated to approximately 5.6 RTM months per 10,000 commercially insured patients. Most (12,280/12,951, 94.8%) used RTM during a single calendar year, while a minority (671/12,951, 5.2%) used RTM during multiple years. The majority of the 12,951 patients were female (8061/12,951, 62.2%), enrolled in a preferred provider organization (7855/12,877, 61%), and had low clinical complexity (12,801/12,951, 98.8%). Additionally, a plurality of patients were between the ages of 55 and 64 years (4116/12,951, 31.8%) and from the US South (6112/12,951, 47.2%).

Comparatively, MA patients used a total of 9551 RTM months, with a median of 2 (IQR 1 - 3) months of RTM use per patient. With nearly 4 million patients enrolled between 2022 and 2024, there were about 24.3 RTM months per 10,000 MA patients. A large majority (2949/3165, 93.2%) used RTM during a single calendar year, and a minority (216/3165, 6.8%) used RTM across multiple years. Among the MA patients, the majority were 75 years or older (1654/3165, 52.3%), female (1982/3165, 62.6%), from the US North Central region (1775/3165, 56.1%), enrolled in a preferred provider organization (2619/2999, 87.3%), and had low clinical complexity (2768/3165, 87.5%).

Among patients with commercial insurance, 48.7% of RTM was delivered by other clinicians (eg, therapists, nurses), followed by other physician specialists (16.5%) and then advanced practice professionals (10.1%; Table 1). The predominant diagnosis category for RTM was diseases of the musculoskeletal system and connective tissue (56.3%). The next most common categories were diseases of the nervous system (17.2%), and factors influencing health status and contact with health services (5.5%).

XSL•FO
RenderX

Table . Percentage of remote therapeutic monitoring use across specialty–primary diagnosis combinations among commercial insurance beneficiaries.

	Other non-physicians, %	Other physicians, %	Advanced practice providers, %	Surgeons, %	Primary care physicians, %	Facility-based, %	Medical subspecialists, %	Total, %
Diseases of the musculoskeletal system and connective tissue	34.62	7.47	4.11	5.04	1.60	2.40	1.05	56.28
Diseases of the nervous system	0.57	7.13	4.90	0.12	1.46	2.53	0.52	17.23
Factors influencing health status and contact with health services	1.32	0.64	0.53	0.67	0.47	0.06	1.83	5.52
Symptoms, signs, and abnormal clinical and laboratory findings	3.12	0.12	0.03	0.05	0.21	0.29	0.04	3.87
Diseases of the genitourinary system	3.71	0.02	0.01	0.01	0.06	0.04	0.01	3.86
Injury, poisoning, and certain other consequences of external causes	2.13	0.14	0.11	1.09	0.07	0.16	0.05	3.75
Endocrine, nutritional, and metabolic diseases	1.91	0.11	0.06	0.18	1.09	0.14	0.13	3.62
Diseases of the circulatory system	0.27	0.14	0.07	0.03	1.19	0.05	0.10	1.84
Mental, behavioral, and neurodevelopmental disorders	0.32	0.52	0.21	0.02	0.31	0.17	0.11	1.67
Diseases of the respiratory system	0.01	0.08	0.01	0.01	0.33	0.01	0.33	0.78
Other	0.74	0.08	0.05	0.11	0.35	0.09	0.17	1.59
Total	48.72	16.45	10.11	7.32	7.15	5.94	4.32	100.00

Among MA patients, RTM was most frequently delivered by primary care physicians (37.8%), other clinicians (23.9%), and other physician specialists (21.3%; Table 2). The primary diagnosis category associated with RTM was also diseases of the musculoskeletal system and connective tissue (45.0%), followed by diseases of the circulatory system (14.2%) and nervous system (10.9%).

XSL•FO
RenderX

Table . Percentage of remote therapeutic monitoring use across specialty–primary diagnosis combinations among Medicare Advantage beneficiaries.

	Primary care physicians, %	Other non-physicians, %	Other physicians, %	Surgeons, %	Medical sub specialists, %	Facility-based, %	Advanced practice providers, %	Total, %
Diseases of the musculoskeletal system and connective tissue	8.29	19.35	8.07	4.22	1.22	2.54	1.32	45.02
Diseases of the circulatory system	9.17	0.22	4.18	0.03	0.41	0.22	0.00	14.23
Diseases of the nervous system	5.85	0.44	2.91	0.13	1.35	0.15	0.08	10.90
Endocrine, nutritional, and metabolic diseases	5.45	0.06	0.95	0.16	0.35	0.08	0.04	7.09
Symptoms, signs, and abnormal clinical and laboratory findings	1.01	2.22	1.16	0.00	0.07	0.54	0.00	5.01
Diseases of the respiratory system	2.32	0.05	1.43	0.02	0.29	0.14	0.02	4.27
Factors influencing health status and contact with health services	0.77	0.59	0.56	0.97	0.16	0.20	0.26	3.52
Diseases of the genitourinary system	0.62	0.24	0.74	0.00	0.01	0.04	0.00	1.65
Injury, poisoning, and certain other consequences of external causes	0.43	0.60	0.37	0.06	0.00	0.03	0.04	1.53
Mental, behavioral, and neurodevelopmental disorders	1.04	0.00	0.30	0.00	0.09	0.03	0.07	1.52
Other	2.86	0.11	0.61	0.05	1.54	0.04	0.05	5.26
Total	37.82	23.88	21.29	5.64	5.48	4.01	1.89	100.00

Discussion

In this nationwide study, RTM use grew during the first 3 years after the billing codes were introduced, overall use remained low, and different use patterns were observed between commercially insured and MA patients. These early findings provide several observations to build on in future studies.

First, despite a smaller sample for MA versus commercial patients during the study period, RTM use was approximately 4 times higher among MA patients (24.3 versus 5.6 months per 10,000). Because MA patients are older with more comorbidities (12.5% for MA vs 1.2% for commercially insured patients with a Charlson Comorbidity Index of ≥2), it might be expected that RTM was used more among MA patients. Given that MA's goal is to improve longitudinal outcomes through managed care, future work should assess whether more frequent RTM use among MA patients is associated with improved patient outcomes and lower spending. How such results combine with the findings of this and prior studies (eg, remote monitoring use for several months mostly among older female patients) remains within the purview of future work [9,10].

XSL•FO
RenderX

Second, RTM use among both commercial and MA patients was dominated by diseases of the musculoskeletal system and connective tissue, underscoring the role of therapeutic monitoring in supporting rehabilitation and functional monitoring. However, our results showed differences in those delivering RTM. Among commercial insurance patients, RTM was most frequently delivered by clinicians such as therapists and nurses, while among MA patients, RTM was more frequently delivered by primary care physicians. Future studies could evaluate the underlying incentives, approaches, and effectiveness of RTM when delivered through different clinician approaches.

Study limitations included the descriptive design and the inability to evaluate associations between RTM use and clinical outcomes. Although our study period included over 50 million patients, our analysis does not reflect RTM use among traditional Medicare patients. Nonetheless, we believe our study provides valuable observations of patterns in RTM use and, in turn, foundational insights for future research.

Acknowledgments

The authors attest that no AI assistance technologies were used in the writing, editing, or preparation of this manuscript.

Funding

This research protocol was funded by the University of Texas Southwestern National Institutes of Health Clinical and Translational Science Award (grant 1U54TR002361).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to data use and licensing restrictions from Merative Marketscan Commercial and MA Database but are available from the corresponding author on reasonable request.

Authors' Contributions

Study concept and design: JHJ, JML; data collection and analysis: LZ; interpretation: all authors; drafting: JHJ, NL, JML; critical revision of manuscript: all authors; supervision: JML; agreement with submission of final manuscript: all authors.

Conflicts of Interest

JML reports service on the Medicare Payment Advisory Commission, the Physician-Focused Payment Model Technical Advisory Committee, and the Health Care Payment Learning and Action Network Patient Empowerment Workgroup. The views in this work are not intended to, and do not necessarily, represent the views of these groups, nor the views of his employer or other groups with which he is affiliated. Others do not have a conflict.

Multimedia Appendix 1
Clinical Classifications Software Refined categories for *International Classification of Diseases, Tenth Revision, Clinical Modification* diagnoses.
[XLSX File, 3609 KB - rehab_v13i1e103600_app1.xlsx]

Multimedia Appendix 2
Characteristics of commercially insured and Medicare Advantage patients using remote therapeutic monitoring, as well as clinician specialties and facilities.
[DOCX File, 32 KB - rehab_v13i1e103600_app2.docx]

References

1. Telehealth & remote monitoring. Centers for Medicare and Medicaid Services. 2025. URL: https://www.cms.gov/files/document/mln901705-telehealth-remote-monitoring.pdf [accessed 2026-08-17]
2. Nowell WB, Curtis JR. Remote therapeutic monitoring in rheumatic and musculoskeletal diseases: opportunities and implementation. Med Res Arch 2023 Jul;11(7.2):3957. [doi: 10.18103/mra.v11i7.2.3957] [Medline: 38550526]
3. Flynn S, Mosher CL, Cornelison S, et al. Feasibility, usability, and pilot efficacy study of a software-enabled, virtual pulmonary rehabilitation with remote therapeutic monitoring. Int J Chron Obstruct Pulmon Dis 2025;20:231-241. [doi: 10.2147/COPD.S484558] [Medline: 39906676]
4. Derman BA, Essell JH, Ascha MS, et al. Remote therapeutic monitoring and infection-related health care resource utilization in patients with hematologic malignancies. JCO Oncol Pract 2026 Apr 10;10:OP2501188. [doi: 10.1200/OP-25-01188] [Medline: 41962066]
5. Rocque G, Patt D, Hunnicutt J, et al. Supporting reimbursement for electronic patient-reported outcomes in clinical cancer care using current procedural terminology codes for remote therapeutic monitoring. JCO Oncol Pract 2026 Apr 3;3:OP2600041. [doi: 10.1200/OP-26-00041] [Medline: 41931724]

XSL•FO
RenderX

6. Huang YQ, Gou R, Diao YS, et al. Charlson Comorbidity index helps predict the risk of mortality for patients with type 2 diabetic nephropathy. J Zhejiang Univ Sci B 2014 Jan;15(1):58-66. [doi: 10.1631/jzus.B1300109] [Medline: 24390745]
7. Clinical Classifications Software Refined (CCSR). Healthcare Cost & Utilization Project. 2025. URL: https://hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp [accessed 2026-08-17]
8. Advanced Practice Providers.: NIH Clinical Center URL: https://www.cc.nih.gov/app [accessed 2026-08-16]
9. Marshall T, Goldman A, Lyles R, et al. Retrospective case-control study on the effect of in-person physical therapy with remote therapeutic monitoring on functional outcomes and plan of care adherence amongst individuals with musculoskeletal conditions. Arch Rehabil Res Clin Transl 2025 Sep;7(3):100466. [doi: 10.1016/j.arrct.2025.100466] [Medline: 40980535]
10. Joo JH, Jiang C, Billig JI, et al. Remote patient monitoring use among commercially insured adults with cancer. JMIR Cancer 2026 Feb 12;12:e84788. [doi: 10.2196/84788] [Medline: 41678798]

Abbreviations

MA: Medicare Advantage
RTM: remote therapeutic monitoring

Edited by S Munce; submitted 04.Jun.2026; peer-reviewed by A Lakdawala, A Peton; revised version received 21.Aug.2026; accepted 27.Aug.2026; published 21.Sep.2026.

Please cite as:
Joo JH, Billig JI, Lieu N, Zhou L, Liao JM
Remote Therapeutic Monitoring Use Among Commercially Insured and Medicare Advantage Patients
JMIR Rehabil Assist Technol 2026;13:e103600
URL: https://rehab.jmir.org/2026/1/e103600
doi:10.2196/103600

© Joseph H Joo, Jessica I Billig, Nadia Lieu, Lingmei Zhou, Joshua M Liao. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 21.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Smart Textile Biofeedback Training System for Upper Limb Rehabilitation After Stroke: Co-Design Development and Evaluation Study

Maria Munoz-Novoa[1,2], PhD; Li Guo[3], PhD; Anna Björkquist[3], BSc; Morten B Kristoffersen[4,5], PhD; Peiman Khorramshahi[6], MSc; Leif Sandsjö[2,7], PhD; Margit Alt Murphy[1,8], PhD

[1]Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden

[2]Department of Work Life and Social Welfare, Faculty of Caring Science, Work Life and Social Welfare, University of Borås, Allégatan 1, Borås, Sweden

[3]Department of Textile Technology, Swedish School of Textiles, University of Borås, Borås, Sweden

[4]Chalmers Industriteknik, Gothenburg, Sweden

[5]Department of Engineering Technology, Technical University of Denmark, Ballerup, Denmark

[6]Daralabs AB, Gothenburg, Sweden

[7]Design & Human Factors, Department of Industrial and Materials Science, Chalmers University of Technology, Gothenburg, Sweden

[8]Department of Health and Rehabilitation, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden

Corresponding Author:
Leif Sandsjö, PhD
Department of Work Life and Social Welfare, Faculty of Caring Science, Work Life and Social Welfare, University of Borås, Allégatan 1, Borås, Sweden

Abstract

Background: An increasing number of rehabilitation technologies are being developed to support upper limb rehabilitation after stroke, with smart textile solutions for surface electromyography (sEMG) emerging as a promising approach. Early end-user involvement is crucial for developing user-friendly and clinically valid rehabilitation tools.

Objective: This study aims to refine and evaluate the prototype design and usability of a smart textile biofeedback system for self-administered upper limb training after stroke.

Methods: The training system includes a knitted smart textile sleeve with integrated electrodes over the forearm muscles, an sEMG unit, and tablet-based biofeedback software. An iterative co-design process was followed, including initial testing, demonstration sessions with end users (9 clinicians and 10 individuals with stroke), and a final evaluation of the co-design process. Participants' experiences were gathered through semistructured interviews, analyzed using content analysis, and the User Experience Questionnaire. The co-design team included experts in stroke rehabilitation, textile engineering, biomedical engineering, software development, and human factors, as well as a research partner with lived experience after stroke.

Results: The perspectives of the end users and the expert team were collectively integrated into prototype refinements of the sleeve and training software to meet the needs of the intended target group. The experiences of end users formed 2 main categories: "This could be an exciting new training tool for stroke rehabilitation" and "The tool works well, but some changes could enhance independent training." End users found the smart textile sleeve and biofeedback system easy to use and saw potential for integrating it into their training routines. Both end-user groups rated the system as attractive, stimulating, and novel.

Conclusions: The results of this study establish a necessary ground toward the development of a smart textile sEMG biofeedback system for self-administered upper limb training after stroke. Findings from the co-design process support the continued development and evaluation of the system as a self-administered upper limb training tool for individuals living with stroke.

(JMIR Rehabil Assist Technol 2026;13:e77999) doi:10.2196/77999

KEYWORDS

stroke; surface electromyography; sEMG; surface electrodes; textile electrodes; smart textiles; upper limb function; self-administered; rehabilitation; co-design

XSL•FO
RenderX

Introduction

Regaining upper limb function is a primary goal for individuals after a stroke [1]. In the acute phase, about 50% to 70% of individuals experience impaired arm and hand function, and in the long term, only a smaller percentage fully recover their previous level of function [2-5]. Upper limb impairment often leads to difficulties in performing daily activities and limitations in participation [6,7]. Most improvements occur early in recovery, with slower progress continuing into the chronic phase [5,8]. To follow this recovery pattern, rehabilitation programs are mostly concentrated in the acute and subacute phases, while in the chronic phase, the responsibility for continued training often falls on the individuals themselves [9,10]. To support long-term rehabilitation [1], various rehabilitation technologies have been developed to facilitate independent home-based training, ultimately aiming to improve functioning [11,12].

Surface electromyography (sEMG) biofeedback is a rehabilitation technology that provides real-time visual feedback on muscle activity, thereby supporting motor learning mechanisms for more effective movement execution [13,14]. Previous research in stroke populations indicates that sEMG biofeedback interventions can increase upper limb muscle strength, active range of motion, and awareness of the paretic arm [15-18]. sEMG biofeedback systems commonly use disposable pre-gelled Ag/AgCl electrodes [19]. As an alternative to these traditional electrodes, smart textile solutions with integrated textile-based electrodes offer improved comfort, easier application, and a reduced risk of skin irritation [20-24]. Textile solutions enable prolonged, repeated use, which makes them suitable for self-administered rehabilitation in both clinical and home settings. However, smart textile solutions for sEMG applications in upper limb stroke rehabilitation are in the early stages of development, with existing studies limited to controlled research environments and/or small sample sizes [25-27].

A challenge in developing new rehabilitation technologies is their implementation in real-world settings [12,28,29]. A failure to meet the needs of end users (patients and clinicians) and underestimating the complexity and costs of implementation in clinical practice are common reasons for limited clinical uptake [30-32]. Adopting collaborative co-design methodologies that involve end users in the development process can help identify and address specific user needs early on. This ensures better alignment among the technology, its intended use, and context and has been shown to improve the applicability and acceptance of rehabilitation technologies [30,33,34].

This study aims to refine and evaluate the prototype design and usability of a smart textile biofeedback system for self-administered upper limb training after a stroke, using a co-design approach.

Methods

Ethical Considerations

The study was approved by the Ethics Review Authority (reference 2019-00450/1074-18 and amendment 2024-01703-02) and follows the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Checklist 1) [35]. Before entering the study, all participants completed an approved informed consent form, as authorized by the Swedish Ethics Review Authority, including information about the possibility to withdraw from the study at any time without providing a reason. All collected data were anonymized and safely stored accessible only to the research team. No compensation was given for participation.

Co-Design Process

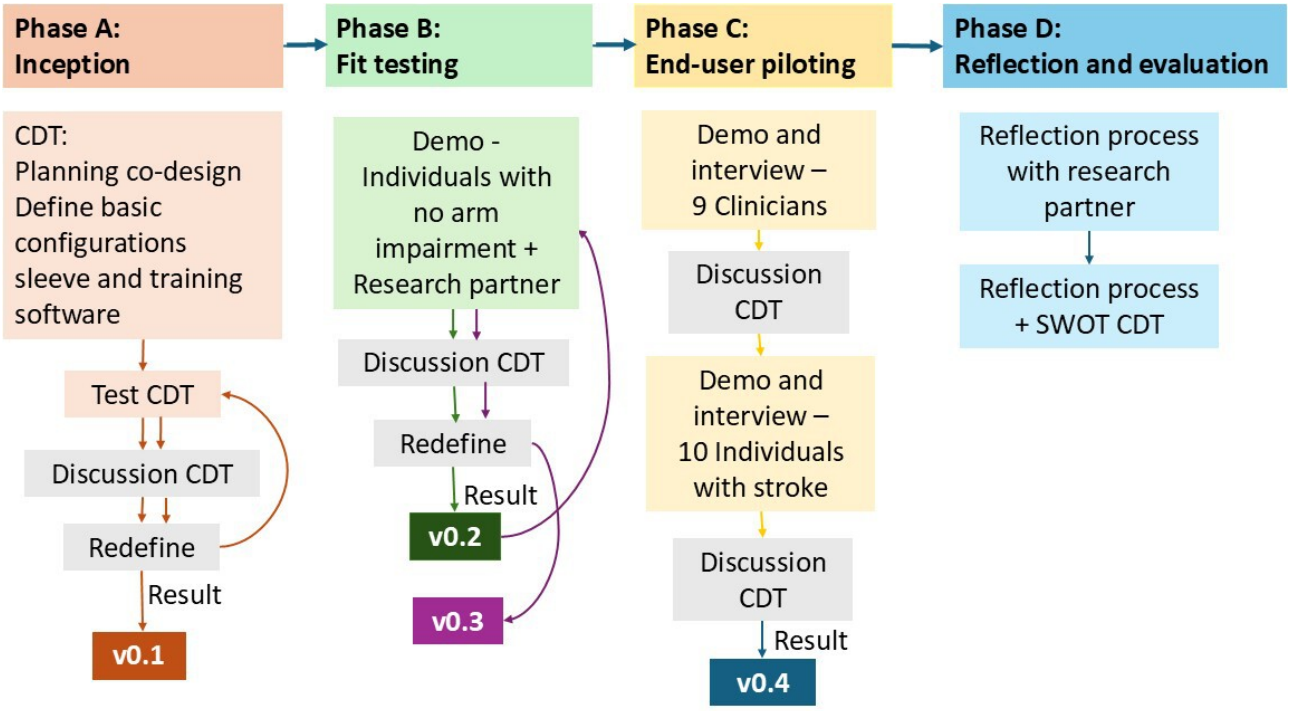
The participatory action research model [36] and the Medical Research Council framework for complex interventions [37] were used to guide the iterative and reflective co-design process, which included 4 phases (Figure 1). Each phase involved iterative cycles of problem definition, planning, acting, observing, reflecting, and redefining the problem [36]. At the end of the codesign process, a strengths, weaknesses, opportunities, and threats analysis [38] was conducted to evaluate the co-design process. The aims, procedures, and actions of each co-design phase are detailed in Table 1.

The co-design team included 7 experts from various fields, including stroke rehabilitation, biomedical engineering, textile engineering, signal processing, software development, and human factors. Most team members had over 15 years of experience in their respective areas, including 1 industry representative with a background in commercial product development. The team also included 2 junior researchers, each with approximately 5 years of experience.

A research partner with a university degree and over 5 years of lived experience with stroke was also included in the co-design team to provide end-user insights during the co-design process. Following the iterative approach of the co-design process, regular physical and digital meetings were held within the co-design team between December 2023 and February 2025.

XSL•FO
RenderX

Figure 1. The 4 phases of the co-design process, guided by the participatory action research model and the Medical Research Council framework, showing the iterations leading to gradually evolving prototype versions (v0.1-v0.4). CDT: co-design team; SWOT: strengths, weaknesses, opportunities, and threats.



XSL•FO
RenderX

Table . Co-design process and refinements of the training system.

Phase	Aim	Procedures	Actions
A: Inception	To set priorities for the co-design process and discuss basic configurations for the training system.	• The CDT[a] met several times to set priorities and define the initial configuration of the training system. • The CDT also tested the first version of the sleeve and software.	• Version 0.1: a conically shaped, circularly knitted sleeve with 2 electrode pairs (wrist flexors and extensors) and a reference electrode at the elbow was developed. The electrodes were connected via snap-button connectors to a wireless sEMG[b] acquisition unit (Shimmer3 EMG, Shimmer Sensing) with an external band. To enhance signal quality, water was applied by 3 finger strokes to each electrode prior to use. Biofeedback was delivered through a tablet using a custom-made application (iOS and Android compatible). Real-time sEMG signals from 2 channels (wrist flexors and extensors) were used during a calibration period to set training thresholds. For training, a vertical bar (presented next to the sEMG signals) turned green when the user reached a target threshold.
B: Fit testing	To test the training system with individuals without upper limb impairments and refine its design for improved fit and usability.	• Individual demonstration sessions with the research partner and with 11 participants without arm impairments. • Refinements and production of the sleeve.	• Version 0.2: sEMG signals were obtained from all participants, although the comfort of the sleeve varied across participants. Based on their feedback, medium and large sizes were produced, and a reference line following the middle finger was added to facilitate electrode alignment during donning. • Version 0.3: a second feedback round led to further refinements, including a knob for easier elbow alignment, a Velcro strap to attach the sEMG unit to the sleeve (replacing the external band), and the addition of a small sleeve size. Software updates included adding a separate window for visualizing the bar activity, tracking repetitions exceeding the threshold, and implementing calibration with a manual stop.

XSL•FO
RenderX

Phase	Aim	Procedures	Actions
C: End-user piloting	To assess usability and investigate perceptions and experiences of the training system from clinicians and individuals with stroke.	• Individual demonstration sessions with 9 stroke rehabilitation clinicians and 10 individuals with stroke. • Semistructured interviews and the UEQ[c] with end users at the end of the demonstration sessions.	• Version 0.4: on the basis of the collected feedback by end users, an additional Velcro strap was added to the sleeve to secure cables and simplify donning and doffing. The medium-sized sleeve was also downsized for a better fit. Software improvements included: fewer and larger buttons, naming of channels, adding a calibration countdown, larger size of the bars with repetition counts, higher color contrast, and a performance report.
D: Reflection and evaluation	To reflect on and to evaluate the co-design process.	• SWOT[d] analysis and reflection workshop with CDT. • Individual reflection session with research partner.	• Reflections from the CDT highlighted strengths such as strong collaboration, research partner involvement, and interdisciplinary expertise. Weaknesses included lack of an external facilitator and limited access to in-house software developers. Opportunities and threats related to future implementation were discussed. The research partner provided additional insights on her experience and valued her active role throughout the co-design process.

[a]CDT: co-design team.

[b]sEMG: surface electromyography.

[c]UEQ: User Experience Questionnaire.

[d]SWOT: strengths, weaknesses, opportunities, and threats.

Participants and Data Collection

In phase A, the co-design team defined the basic configuration of the training system (version 0.1). A textile sleeve with 2 pairs of integrated textile electrodes, placed over the wrist extensors and flexors, and a reference electrode placed on the elbow was produced. The sleeve was connected to a 2-channel wireless sEMG system (Shimmer3 EMG, Shimmer Sensing). The biofeedback software provided visual feedback of the sEMG signals displayed on a tablet from the wrist extensor muscles (eg, extensor carpi radialis brevis) and flexor muscles (eg, flexor carpi ulnaris). The sEMG signals from specific muscle activation were used to calibrate muscle activity thresholds, which were then visualized during training through a vertical bar display. This provided real-time feedback and allowed for adjustments to individualized muscle activity threshold settings.

In phase B, an initial test demonstration session was conducted with the research partner, followed by individual sessions with 11 participants without arm impairments. Convenience sampling was used, and age, gender, and anthropometric data were collected from this sample (Multimedia Appendix 1).

In phase C, individual demonstration sessions were conducted with 2 end-user groups: 9 clinicians (6 physiotherapists, 2 occupational therapists, and 1 rehabilitation doctor) and 10 individuals with stroke (Tables S1 and S2 in Multimedia Appendix 2). Participants were recruited through purposive sampling. Clinicians were recruited from stroke rehabilitation units in hospitals and primary care centers, while individuals with stroke were recruited through advertisements at rehabilitation centers in the Gothenburg urban area. Inclusion criteria for individuals with stroke were as follows: aged over 18 years, upper limb impairment after stroke, and detectable sEMG signal from the paretic arm. Exclusion criteria for all participants included open wounds, other conditions affecting arm function, uncorrected visual impairment, and inability to understand or follow instructions in Swedish or English. Demographic data, prior sEMG experience, and familiarity with using technologies (eg, smartphones, tablets, and computers) were collected from all end users. For individuals with stroke, additional information was gathered on stroke type and time since stroke, while for clinicians, data on years of experience and work setting were collected.

All demonstration sessions were conducted at Gothenburg University and led by a physiotherapist experienced with the training system (MM-N). In these sessions, the participants were first instructed on how to use the sleeve and software and

XSL•FO
RenderX

then asked to try it independently, with additional support provided as needed. At the end of the sessions, the participants were interviewed for about 30 minutes and asked to fill in the User Experience Questionnaire (UEQ) [39,40]. Additionally, in the demonstration sessions with individuals with stroke, upper limb function and spasticity were assessed using the Fugl-Meyer Assessment of Upper Extremity [41] and the Modified Ashworth Scale [42], respectively. From the demonstration sessions with end users, 1 clinician tested only the sleeve due to technical issues with the software and 1 individual with stroke provided limited feedback due to aphasia.

Evaluations and Data Analysis

Semistructured Interviews

The semistructured interviews included questions about the end users' perceptions and experiences using the sleeve and training software during the demonstration sessions, as well as their opinions on the training tool's potential for independent use in clinics or home settings. For clinicians, questions were adapted to explore how they would foresee using the training system in stroke rehabilitation (Multimedia Appendix 1). All interviews were led by the same physiotherapist who conducted the demonstration sessions (MM-N, with more than 5 years of research experience in stroke rehabilitation), unless a participant was not fluent in English, in which case the interview was conducted in Swedish by another physiotherapist (MAM, with more than 20 years of experience in stroke rehabilitation). All interviews were audio-recorded, transcribed verbatim by 2 researchers from the co-design team (MM-N and AB), and stored securely, with access restricted to the authors. The field notes and spontaneous feedback collected from participants during the demonstration sessions were also used to complement the interview data.

Content analysis using the approach described by Graneheim and Lundman [43] was used to analyze the interviews [44]. The analysis involved extracting meaning units, condensing them, coding, and sorting codes into categories and subcategories (Multimedia Appendix 1). Since the interviews were conducted following a single demonstration session, the analysis focused on manifest content, emphasizing explicit and observable elements [43,44]. Interviews conducted in Swedish were translated into English and checked for accuracy by a native Swedish speaker (AB). The analyses were conducted in English. Initial coding was carried out by a member of the co-design team (MM-N), while a second researcher (AB), a research engineer with over 5 years of experience in textile technology who was not involved in data collection but was familiar with the data, participated in the review and refinement of the codes. Interviews with clinicians and individuals with stroke were analyzed together, but coding was grouped separately for each participant type using different colors to facilitate category development while distinguishing differences between the groups. Two researchers (AB and MM-N) independently developed initial categories, followed by an iterative process of discussions with a third researcher (MAM) to refine and resolve conflicts until final categories were agreed upon. The research partner was also included in the final stages of the analysis, providing additional input on the selected categories and subcategories.

User Experience Questionnaire

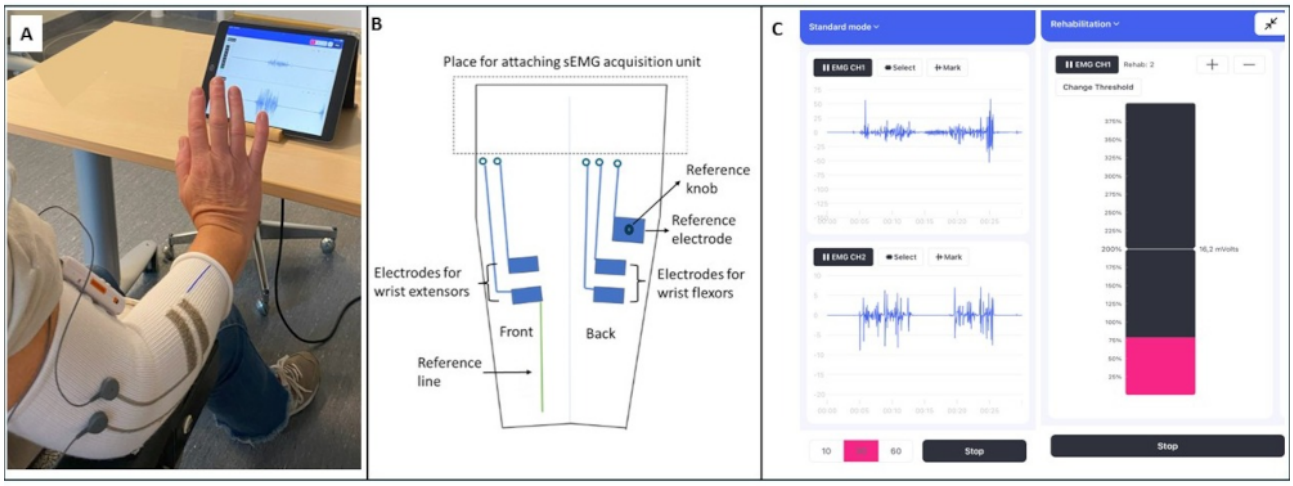
The UEQ captures a broad range of user experience aspects, helping to identify both strengths and areas for improvement [39,40]. It consists of 26 items rated on a 7-point scale, with each item presented as a pair of opposite terms. The questionnaire evaluates 6 dimensions: attractiveness (overall impression), perspicuity (ease of use), efficiency (task completion without unnecessary effort), dependability (feeling in control), stimulation (excitement and motivation), and novelty (innovation and creativity) [39,40]. The UEQ benchmark [45] was used to interpret the results on 5 levels: excellent, good, above average, below average, and bad.

Results

The specific refinements of the training system during the iterative phases (A-D) of the study are detailed in Table 1. The setup of the training system, the smart textile sleeve configuration, and the biofeedback software used in the demonstration sessions (version 0.3) are shown in Figure 2.

XSL•FO
RenderX

Figure 2. The smart textile sleeve biofeedback training system version 0.3. (A) Demonstration session setup, (B) smart textile sleeve configuration, and (C) biofeedback software (left panel: real-time sEMG signals from extensor [CH1] and flexor [CH2] channels; right panel: muscle activity bar visualization with adjustable threshold settings for training). CH1: channel 1; CH2: channel 2; sEMG: surface electromyography.



End Users' Experiences and Perceptions

From the interviews with clinicians and individuals with stroke, 2 main categories and 5 subcategories emerged, as illustrated in Textbox 1. Quotes are presented below each subcategory, where "S" indicates individuals with stroke and "C" refers to clinicians.

Textbox 1. Categories and subcategories emerging from content analysis.

Categories and subcategories

- This could be a new exciting training tool for stroke rehabilitation.
  - Rewarding and motivating to see muscle activity.
  - Use will be determined by individual needs.
  - Instructions and guidance will help independent use.
- The training tool works fine, but some changes could enhance independent training.
  - It can be difficult at first, but it is possible to find a way to make it work.
  - More is needed for independent long-term use.

This Could Be a New Exciting Training Tool for Stroke Rehabilitation

Participants with stroke and clinicians found the training system to be novel, exciting, and potentially useful for upper limb rehabilitation after stroke. Both groups expressed interest in incorporating the training system into their routines.

Rewarding and Motivating to See Muscle Activity

The real-time visual feedback on muscle activity was described as intuitive, useful [2], and rewarding by most participants. Individuals with stroke liked seeing their muscle activity, especially during training with the bar visualization, as many had thought they had little or no muscle activity due to their impairment. Some participants reported feeling more "connected" to their affected arm when seeing the activity visualized on the screen [3,4]. They described the feedback as not only giving them a sense of control over their arm but also possibly making the training more engaging, interesting, and motivating. Clinicians also highlighted the benefits of visual feedback, particularly for patients with limited muscle activity, as it could help them recognize their remaining muscle function and visualize their progress over time. Participants with stroke and clinicians also noted that visual feedback could enhance motivation and engagement and increase awareness of the affected arm, ultimately supporting and empowering independent training.

> ***S5:*** *That's, for me, [seeing the muscle activity] is something good to see, because then I see that something actually happens…I mean, it's boring to just sit like that, but if you see something happening, then I know my muscles are generating that kind of activity.*

> ***C3:*** *I like the visual feedback you get, even though we just do like a slight movement, you see the graph is changing. And if I put it on stroke patients, I think it's really good because sometimes when you tell a stroke patient, "Oh, can you do this movement?" They are quite sad, because they can't do the movement, but [with this training] you can see, "Oh, they are initiating the movement."*

XSL•FO
RenderX

Use Will Be Determined by Individual Needs

Most participants considered the training system more suitable for the chronic or subacute phases after a stroke. In the acute and subacute phases, individuals with stroke felt it could be overwhelming due to the many other challenges they faced at that stage, a concern shared by the clinicians. Some clinicians also noted that cost, time, and technical constraints could limit its use in clinical settings, while others described the tool as a complement to other training in the subacute phase. Both groups saw it as feasible and beneficial for home training in the chronic phase, especially considering that the available training opportunities are fewer compared to the subacute phase. However, they noted that some level of upper limb function in the paretic arm would be necessary, as high spasticity or limited elbow and finger control could make it difficult to don and doff the sleeve independently. Clinicians also emphasized that a certain level of cognition is required to handle and manage the training without therapist or carer support. Both groups mentioned that limited prior experience with technology could be a potential barrier to using the software.

> ***S6:*** *I think it's a very good idea [the training tool]. Eh, it's a problem for people like me [individuals with stroke], in the winter, to get out to training places when it's snowy, so it's very nice to make it possible to train home as well.*
>
> ***C7:*** *I think if you have one arm [functional arm], you could do it. You still need good cognitive abilities. Maybe not being too old. Some patients who are old are afraid of technological things because they feel uncertain. Even though they can manage… they don't think they can. It is very different from person to person.*

Instructions and Guidance Will Help Independent Use

Individuals with stroke, as well as clinicians, viewed independent home training positively, with individuals with stroke especially finding it exciting and convenient. However, both groups emphasized that clear instructions and support are essential for successful independent use. While they appreciated the simple and clear guidance provided during the demonstration session, they felt additional help with software navigation and sleeve placement would be required. Individuals with stroke generally believed they could train independently or with partial caregiver assistance at home after initial guidance. However, they desired more structured support, including written training protocols, visual and written instructions, in-software reminders, and clinician follow-ups (in person or online) to ensure success. Clinicians also stressed the need for clear, repeated instructions in multiple formats to facilitate adherence. Additionally, clinicians highlighted the importance of hands-on training for patients and caregivers, along with in-person or online follow-ups to monitor tool usage and progress.

> ***S2:*** *It was difficult [to place the sleeve independently]. I think it can be even more difficult if you, kind of, get it [the training system] home and don't get to try it on first.*
>
> ***C1:*** *...They [individuals with stroke] need guidance, as it is right now [the tested version], I think...but otherwise, I think it's quite easy to use with guidance.*

The Training Tool Works Fine, but Some Changes Could Enhance Independent Training

Overall, individuals with stroke and clinicians described the sleeve as functional and easy to use, noting that further improvements to the software were needed to promote independent training.

It Can Be Difficult at First, but It Is Possible to Find a Way to Make It Work

Most participants found the sleeve easy to use and comfortable, and they appreciated the provided instructions as well as the reference line and elbow knob for guidance. Individuals with stroke and more severe arm impairment experienced some difficulties with getting the hand inside the sleeve, rotating the sleeve into the correct position, and/or pulling the sleeve up on the arm. However, many found strategies by themselves during the demonstration session and believed that, with practice, it would not be a major problem. Another challenge was locating the reference electrode at the elbow and applying water to the reference and flexor electrodes due to a limited range of arm movement and/or sensory impairment. In the end, all participants with stroke were able to place the sleeve correctly, although some needed more time and minor assistance.

Perceptions of the software varied. Younger participants in both groups found the software easier to use, while others described it as a bit complicated and not intuitive. The buttons on the screen were too small and difficult to see for many. In general, most individuals with stroke and clinicians liked that a tablet was used for feedback, as it is larger than a phone but also portable compared to a computer screen. They emphasized that initial guidance and clear instructions on the software would be essential to ensure successful use.

> ***S7:*** *Yes, it was difficult to begin with [to put on the sleeve], but I think it's something you get used to.*
>
> ***C6:*** *To me, it was easy and fine. I don't know if a patient with a paresis would have difficulties having it drawn up on the middle of the arm …and also, placing it correctly on the elbow might be difficult.*

More Is Needed for Independent Long-Term Use

Independent home use was a possibility that many described as feasible, but they suggested improvements for better usability and long-term adherence. To simplify sleeve placement, participants recommended using a mirror or tablet webcam to better view the electrodes on the dorsal side of the arm and the reference electrode. Additional reference points, shorter electrode cables, and a pull strap for easier placement were also suggested. Regarding the software, both participants with stroke and clinicians desired a more user-friendly, engaging, and intuitive interface, with suggestions including games, videos, and motivational elements to make training more engaging. Clinicians emphasized adding functional, goal-oriented tasks and applying resistance to enhance engagement and facilitate training. Fewer and larger clickable buttons, simpler calibration

XSL•FO
RenderX

procedures, and adding more colors and contrast were other proposals for software improvements. Clinicians also suggested incorporating audio feedback to guide training and augment visual feedback. Additionally, both groups highlighted the need for customizable training parameters to facilitate sessions, such as calibration time and the number of repetitions, as well as the ability to track progress, which is particularly important for independent training. When considering home training, participants thought that a short daily training session of about 15 minutes would be feasible—short enough to avoid mental fatigue. A clinician also suggested using a bag to keep the sleeve clean between training sessions and noted that multiple sleeves might be needed if one requires cleaning or washing.

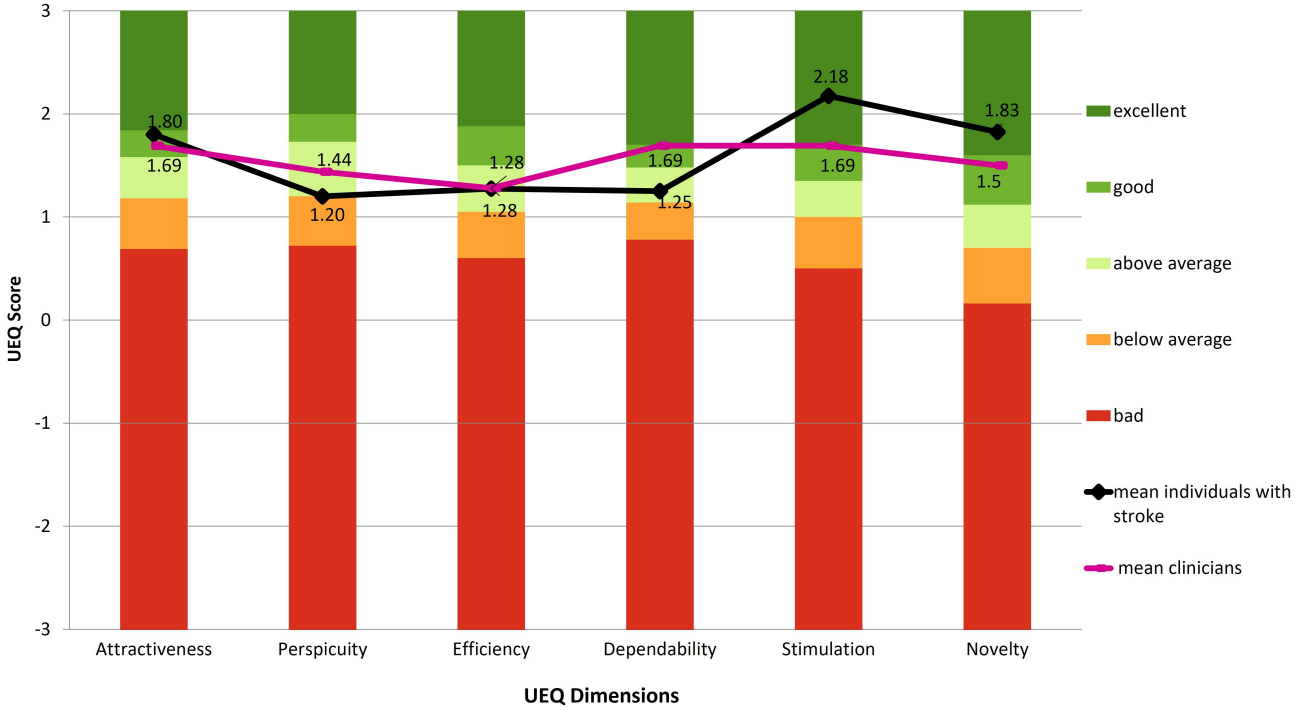
> ***S10:*** *On the tablet, I think it's too many knobs [buttons] to press on…It should be easier. One or two presses.*

> ***C2:*** *I'm thinking this [the training tool] was easy to use. But maybe adding some extra, like instructions... instructions like how fast you will do the exercise, how many repetitions, how long will I keep it. Maybe adding…like an audio, like a voice, [for example] if you have bad vision.*

User Experience Questionnaire

The UEQ was completed by all individuals with stroke and all but 1 clinician, who did not have the opportunity to try the software. Both groups rated *attractiveness* as "good" and *efficiency* as "above average." Clinicians rated *perspicuity* as "above average," while participants with stroke rated it lower, between "above average" and "below average." *Dependability* was rated "good" by clinicians and "above average" by individuals with stroke. Participants with stroke rated *stimulation* and *novelty* as "excellent," while clinicians rated them as "good" (Figure 3).

Figure 3. The results from the User Experience Questionnaire (UEQ) in relation to the standard UEQ benchmark levels. Scores range from −3 (very negative evaluation) to +3 (very positive evaluation), with 0 representing a neutral response. Line plots indicate the mean responses from individuals with stroke (black line; n=10) and clinicians (pink line; n=8).



Reflection and Evaluation of the Co-Design Process

The strengths of the co-design process identified by the co-design team were strong collaboration, effective communication, and research partner involvement, ensuring an end-user perspective. Flexibility, team commitment, and optimal facilities were also highlighted. Not having access to an external facilitator and software developers within the co-design team was identified as a weakness, with the latter contributing to delays in communication and iterative development. Opportunities included creating visual simulations before user involvement, direct contact with clinicians and patients, and a positive view of the training tool's potential value. Threats such as overdevelopment, uncertainty about real-world applicability, and challenges with commercialization and regulatory approval were also identified.

At the end of the co-design process, a reflection session with the research partner revealed that she enjoyed the experience and was pleasantly surprised by how much she liked it, as it was her first time in this role. She appreciated being involved not only in testing the training system but also in the reflective process. Overall, she found the experience fun and felt good about contributing to the research.

Discussion

Principal Findings

Supported by collaboration between experts from different fields and end users, the iterative co-design process provided valuable

XSL•FO
RenderX

insights into the technical, clinical, and user-centered aspects of the training system, highlighting its potential usability, acceptability, and areas for improvement based on end users' needs. The participants with stroke were able to independently, or with minor assistance, don and doff the sleeve and operate the training software on the tablet. Both clinicians and individuals with stroke recognized the training tool's potential for self-administered training. This study provides a starting point for the ongoing development and improvement of the training system.

The structured co-design process in this study, grounded in the participatory action research model [36] and the Medical Research Council framework [37], provided a strong foundation for the development of the training system and the shared understanding necessary for its continued refinement. This approach supported a systematic and iterative development of the training system and may serve as a useful model for future rehabilitation technology development [46,47]. The interdisciplinary nature of the co-design team enriched the process through diverse perspectives. However, differences in disciplinary knowledge and workflows, as reported in previous studies [48-50], may have prolonged the decision-making process. The co-design team suggested that adding a neutral facilitator to the team could have helped to advance the process further [33,51]. Integrating end-user feedback on design improvements with development constraints and limited available resources is a complex but well-recognized challenge in rehabilitation technology development [36,52]. This challenge was addressed through regular co-design meetings and systematic documentation, which enabled balanced prioritization of modifications based on both end-user feedback and practical constraints. Even though not all suggestions from the end users were integrated, they offered valuable insights for future improvements to the training system.

The overall positive perceptions of the smart textile biofeedback training system among individuals with stroke and clinicians underscore its potential as a possible solution for self-administered rehabilitation at home and in clinical contexts. Initial difficulties with donning and doffing the sleeve, particularly among participants with more severe upper limb impairments, highlighted the importance of user-centered design and individualized support during the development of the training system [30,33,34]. All participants were ultimately able to use the system independently or with minimal assistance, suggesting its accessibility across varying levels of motor function. The integration of textile electrodes simplified the donning process, possibly reducing barriers associated with conventional electrode placement and supporting independent use [20-24], especially for individuals with limited fine motor control. Although the system was evaluated in single sessions, the positive perceptions and willingness to incorporate it into daily training routines indicate promising initial usability and acceptability. Further longitudinal studies are needed to assess sustained engagement and rehabilitation outcomes over time.

The visual feedback, consistent with previous evidence [53-56], was perceived as intuitive and helpful in increasing awareness of muscle activity. The participants also found that being able to see and control their muscle activity increased their engagement and motivation. Future iterations, as suggested by the end users, will include more engaging interactive games and/or virtual environments, to facilitate long-term use [57-60]. Additionally, more advanced signal processing techniques, such as myoelectric pattern recognition, could be valuable for detecting specific movement patterns and enabling more intuitive control [14,61]. Ease of use and engaging designs have been identified as key factors in increasing treatment adherence [50,62,63] and will therefore be prioritized in the continued development of the training system for self-administered rehabilitation.

The participants also emphasized the need for structured support from the therapist, including regular follow-ups and adjustments to the training program. They highlighted the importance of clear instructions, in written or video formats, to support independent use and long-term adherence. Clear guidance, personalized feedback, and ongoing clinician support are essential for the successful use and sustained engagement with home-based rehabilitation technologies [50,63-65]. Both participants with stroke and clinicians found the training tool suitable for individuals with limited upper limb function but raised some concerns for individuals with limited cognitive capacity or experience with digital tools. On the basis of these insights, future iterations of the training system will focus on incorporating therapist support, clear instructional materials, the potential integration of a virtual coach, and appropriate user targeting to promote effectiveness and broader adoption in home-based rehabilitation.

Strength and Limitations

The ongoing input from the research partner was a strength of this study, offering valuable feedback that kept the co-design team focused on end users' needs. The iterative co-design process, supported by the triangulation of interviews, field notes, and questionnaires, as well as ongoing reflexive discussions within the team, strengthened the study's credibility, transparency, and overall trustworthiness [46,47]. Additionally, interviews, field notes, and a UEQ provided a broader understanding of participants' experiences.

A strength that also posed a limitation was that some participants were recruited through professional networks and, in many cases, the first author conducted both the sessions and the interviews. While this facilitated access and trust, it may have introduced sampling bias and relationship bias, as prior familiarity could have influenced participants' responses and the researcher's interpretation, potentially limiting the diversity and representativeness of perspectives.

The end users' perceptions were primarily used to inform the design and perceived usefulness of the training system, rather than to evaluate its practical application in clinical settings. The brief, single-session exposure further limited feedback to initial impressions rather than long-term use. Future studies are needed to evaluate the usability, feasibility, and potential effectiveness of the system in people with stroke after a longer training period.

Conclusions

This study developed and evaluated the usability of a smart textile sEMG biofeedback system designed for self-administered

XSL•FO
RenderX

upper limb rehabilitation after stroke. The co-design process, involving a multidisciplinary team of experts and end users, played an important role in identifying users' needs and preferences, contributing to a more targeted and user-centered design. Individuals with stroke and clinicians found the training tool useful and relevant for rehabilitation. The findings will guide ongoing development and clinical evaluation, but further research is needed to determine the system's effectiveness as a home-based rehabilitation tool for people with stroke.

Acknowledgments

The authors thank all the participants, including individuals with stroke, clinicians, researchers, and individuals with no arm impairment, for their time and effort in joining this study. The authors give special thanks to their research partner, Johanna Persson, for her invaluable guidance in the co-design process and to Elin Gustafsson for assisting with participant recruitment. The authors also acknowledge Bola AB for producing the textile prototypes used in this study and Katharina S Sunnerhagen for her support throughout the process.

Funding

This study was funded by the Promobilia Foundation (A23169); KK-Stiftelsen (20220222); Strokeförbundets Stiftelser och Fonder (S-10119402024-09-15); Norrbacka-Eugeniastiftelsen (837/24); Stiftelsen Hjalmar Svenssons (HJSV2024058); Swedish Society for Medical Research (S19-0074); and the Swedish state under the agreement between the Swedish government and the county councils, the ALF-agreement (ALF-GBG-932697, ALF-GBG 965508, and ALF-GBG 1006050).

Authors' Contributions

Conceptualization – LS (lead), LG (equal), MAM (equal), AB (supporting), MBK (supporting), MM-N (supporting), PK (supporting)
Data curation – MM-N (lead), LG (supporting), LS (supporting), MAM (supporting), PK (supporting)
Formal analysis – MM-N (lead), MAM (equal), AB (supporting)
Funding acquisition - MAM (lead), LG (equal), LS (equal), MM-N (supporting)
Investigation – MM-N (lead), MAM (equal), AB (supporting), LG (supporting), LS (supporting), MBK (supporting), PK (supporting)
Methodology – MM-N (lead), LG (equal), LS (equal), MAM (equal), AB (supporting), MBK (supporting), PK (supporting)
Project administration – MN-M (lead), LG (equal), MAM (equal), LS (supporting)
Resources – MAM (lead), LG (equal), LS (equal), MM-N (equal), PK (equal), AB (supporting), MBK (supporting)
Software – PK (lead), LG (equal), LS (equal), AB (supporting), MAM (supporting), MBK (supporting), MM-N (supporting)
Supervision – MAM (lead), LG (supporting), LS (supporting), MBK (supporting)
Validation – MAM (lead), LG (supporting), LS (supporting), MBK (supporting)
Visualization – MM-N (lead), LG (supporting), LS (supporting), MAM (supporting)
Writing – original draft – MM-N (Lead)
Writing – review & editing – MM-N (lead), LG (equal), LS (equal), MAM (equal), MBK (equal) AB (supporting), PK (supporting)

Conflicts of Interest

AB, LG, LS, and PK are founders and stakeholders of ZiniaCare, a startup company aiming to make the described biofeedback system commercially available, thereby presenting a potential conflict of interest. The company did not provide funding or support for the study.

Multimedia Appendix 1
Characteristics of individuals without arm impairment, interview templates, and an example of content analysis from end-user interviews.
[DOCX File, 55 KB - rehab_v13i1e77999_app1.docx]

Multimedia Appendix 2
Characteristics of individuals with stroke and clinicians.
[DOCX File, 47 KB - rehab_v13i1e77999_app2.docx]

Checklist 1
COREQ checklist.
[PDF File, 447 KB - rehab_v13i1e77999_app3.pdf]

XSL•FO
RenderX

References

1. Pollock A, St George B, Fenton M, Firkins L. Top 10 research priorities relating to life after stroke--consensus from stroke survivors, caregivers, and health professionals. Int J Stroke 2014 Apr;9(3):313-320. [doi: 10.1111/j.1747-4949.2012.00942.x] [Medline: 23227818]
2. Persson HC, Parziali M, Danielsson A, Sunnerhagen KS. Outcome and upper extremity function within 72 hours after first occasion of stroke in an unselected population at a stroke unit. A part of the SALGOT study. BMC Neurol 2012 Dec 29;12:162. [doi: 10.1186/1471-2377-12-162] [Medline: 23273107]
3. Nijland RHM, van Wegen EEH, Harmeling-van der Wel BC, Kwakkel G, EPOS Investigators. Presence of finger extension and shoulder abduction within 72 hours after stroke predicts functional recovery: early prediction of functional outcome after stroke: the EPOS cohort study. Stroke 2010 Apr;41(4):745-750. [doi: 10.1161/STROKEAHA.109.572065] [Medline: 20167916]
4. Harris JE, Eng JJ. Paretic upper-limb strength best explains arm activity in people with stroke. Phys Ther 2007 Jan;87(1):88-97. [doi: 10.2522/ptj.20060065] [Medline: 17179441]
5. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: a systematic review. Lancet Neurol 2009 Aug;8(8):741-754. [doi: 10.1016/S1474-4422(09)70150-4] [Medline: 19608100]
6. Raghavan P. Upper limb motor impairment after stroke. Phys Med Rehabil Clin N Am 2015 Nov;26(4):599-610. [doi: 10.1016/j.pmr.2015.06.008] [Medline: 26522900]
7. Ekstrand E, Brogårdh C. Life satisfaction after stroke and the association with upper extremity disability, sociodemographics, and participation. PM R 2022 Aug;14(8):922-930. [doi: 10.1002/pmrj.12712] [Medline: 34541828]
8. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
9. Teasell R, Mehta S, Pereira S, et al. Time to rethink long-term rehabilitation management of stroke patients. Top Stroke Rehabil 2012;19(6):457-462. [doi: 10.1310/tsr1906-457] [Medline: 23192711]
10. Boehme C, Toell T, Lang W, Knoflach M, Kiechl S. Longer term patient management following stroke: a systematic review. Int J Stroke 2021 Oct;16(8):917-926. [doi: 10.1177/17474930211016963] [Medline: 33949269]
11. Arntz A, Weber F, Handgraaf M, et al. Technologies in home-based digital rehabilitation: scoping review. JMIR Rehabil Assist Technol 2023 Jul 27;10:e43615. [doi: 10.2196/43615] [Medline: 37253381]
12. Chen Y, Abel KT, Janecek JT, Chen Y, Zheng K, Cramer SC. Home-based technologies for stroke rehabilitation: a systematic review. Int J Med Inform 2019 Mar;123:11-22. [doi: 10.1016/j.ijmedinf.2018.12.001] [Medline: 30654899]
13. Giggins OM, Persson UM, Caulfield B. Biofeedback in rehabilitation. J Neuroeng Rehabil 2013 Jun 18;10:60. [doi: 10.1186/1743-0003-10-60] [Medline: 23777436]
14. Munoz-Novoa M, Kristoffersen MB, Sunnerhagen KS, Naber A, Ortiz-Catalan M, Alt Murphy M. Myoelectric pattern recognition with virtual reality and serious gaming improves upper limb function in chronic stroke: a single case experimental design study. J Neuroeng Rehabil 2025 Jan 17;22(1):6. [doi: 10.1186/s12984-025-01541-y] [Medline: 39825410]
15. Munoz-Novoa M, Kristoffersen MB, Sunnerhagen KS, Naber A, Alt Murphy M, Ortiz-Catalan M. Upper limb stroke rehabilitation using surface electromyography: a systematic review and meta-analysis. Front Hum Neurosci 2022;16:897870. [doi: 10.3389/fnhum.2022.897870] [Medline: 35669202]
16. Nelson LA. The role of biofeedback in stroke rehabilitation: past and future directions. Top Stroke Rehabil 2007;14(4):59-66. [doi: 10.1310/tsr1404-59] [Medline: 17698458]
17. Woodford H, Price C. Emg biofeedback for the recovery of motor function after stroke. Cochrane Database Syst Rev 2007 Apr 18;2007(2):CD004585. [doi: 10.1002/14651858.CD004585.pub2] [Medline: 17443550]
18. Kim JH. The effects of training using EMG biofeedback on stroke patients upper extremity functions. J Phys Ther Sci 2017 Jun;29(6):1085-1088. [doi: 10.1589/jpts.29.1085] [Medline: 28626331]
19. Hermens H, Freriks B, Merletti R, et al. SENIAM 8: European recommendations for surface electromyography. : Roessingh Research and Development; 1999 URL: http://www.seniam.org/pdf/contents8.PDF [accessed 2026-03-13]
20. Guo L, Sandsjö L, Ortiz-Catalan M, Skrifvars M. Systematic review of textile-based electrodes for long-term and continuous surface electromyography recording. Text Res J 2020 Jan;90(2):227-244. [doi: 10.1177/0040517519858768]
21. Searle A, Kirkup L. A direct comparison of wet, dry and insulating bioelectric recording electrodes. Physiol Meas 2000 May;21(2):271-283. [doi: 10.1088/0967-3334/21/2/307] [Medline: 10847194]
22. Fleury A, Sugar M, Chau T. E-textiles in clinical rehabilitation: a scoping review. Electronics (Basel) 2015;4(1):173-203. [doi: 10.3390/electronics4010173]
23. Khan MA, Saibene M, Das R, Brunner I, Puthusserypady S. Emergence of flexible technology in developing advanced systems for post-stroke rehabilitation: a comprehensive review. J Neural Eng 2021 Dec 2;18(6). [doi: 10.1088/1741-2552/ac36aa] [Medline: 34736239]
24. Korzeniewska E, Krawczyk A, Mróz J, Wyszyńska E, Zawiślak R. Applications of smart textiles in post-stroke rehabilitation. Sensors (Basel) 2020 Apr 22;20(8):2370. [doi: 10.3390/s20082370] [Medline: 32331218]
25. Yang G, Deng J, Pang G, et al. An IoT-enabled stroke rehabilitation system based on smart wearable armband and machine learning. IEEE J Transl Eng Health Med 2018;6:2100510. [doi: 10.1109/JTEHM.2018.2822681] [Medline: 29805919]

XSL•FO
RenderX

26. De Marchis C, Santos Monteiro T, Simon-Martinez C, Conforto S, Gharabaghi A. Multi-contact functional electrical stimulation for hand opening: electrophysiologically driven identification of the optimal stimulation site. J Neuroeng Rehabil 2016 Mar 8;13:22. [doi: 10.1186/s12984-016-0129-6] [Medline: 26955873]
27. Lorussi F, Carbonaro N, De Rossi D, Paradiso R, Veltink P, Tognetti A. Wearable textile platform for assessing stroke patient treatment in daily life conditions. Front Bioeng Biotechnol 2016;4:28. [doi: 10.3389/fbioe.2016.00028] [Medline: 27047939]
28. Langan J, Subryan H, Nwogu I, Cavuoto L. Reported use of technology in stroke rehabilitation by physical and occupational therapists. Disabil Rehabil Assist Technol 2018 Oct;13(7):641-647. [doi: 10.1080/17483107.2017.1362043] [Medline: 28812386]
29. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251. [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
30. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17. [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
31. Redfern J, McKevitt C, Wolfe CDA. Development of complex interventions in stroke care: a systematic review. Stroke 2006 Sep;37(9):2410-2419. [doi: 10.1161/01.STR.0000237097.00342.a9] [Medline: 16902171]
32. Alt Murphy M, Pradhan S, Levin MF, Hancock NJ. Uptake of technology for neurorehabilitation in clinical practice: a scoping review. Phys Ther 2024 Feb 1;104(2):pzad140. [doi: 10.1093/ptj/pzad140] [Medline: 37856528]
33. Singh H, Benn N, Fung A, et al. Co-design for stroke intervention development: results of a scoping review. PLoS ONE 2024;19(2):e0297162. [doi: 10.1371/journal.pone.0297162] [Medline: 38354160]
34. Shah SGS, Robinson I. Benefits of and barriers to involving users in medical device technology development and evaluation. Int J Technol Assess Health Care 2007;23(1):131-137. [doi: 10.1017/S0266462307051677] [Medline: 17234027]
35. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): A 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
36. Cornish F, Breton N, Moreno-Tabarez U, et al. Participatory action research. Nat Rev Methods Primers 2023;3(1). [doi: 10.1038/s43586-023-00214-1]
37. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
38. Pickton DW, Wright S. What's SWOT in strategic analysis? Strategic Change 1998 Mar;7(2):101-109. [doi: 10.1002/(SICI)1099-1697(199803/04)7:2<101::AID-JSC332>3.0.CO;2-6]
39. Schrepp M, Hinderks A, Thomaschewski J. Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. 2014 Presented at: Design, User Experience, and Usability: Theories, Methods, and Tools for Designing the User Experience: 3rd International Conference, DUXU 2014; Jun 22-27, 2014; Heraklion, Crete, Greece p. 383-392. [doi: 10.1007/978-3-319-07668-3_37]
40. Schrepp M, Hinderks A, Thomaschewski J. Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). Int J Interact Multimed Artif Intell 2017;4(6):103-108. [doi: 10.9781/ijimai.2017.09.001]
41. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. Scand J Rehabil Med 1975;7(1):13-31. [Medline: 1135616]
42. Gregson JM, Leathley M, Moore AP, Sharma AK, Smith TL, Watkins CL. Reliability of the Tone Assessment Scale and the Modified Ashworth Scale as clinical tools for assessing poststroke spasticity. Arch Phys Med Rehabil 1999 Sep;80(9):1013-1016. [doi: 10.1016/s0003-9993(99)90053-9] [Medline: 10489001]
43. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
44. Lindgren BM, Lundman B, Graneheim UH. Abstraction and interpretation during the qualitative content analysis process. Int J Nurs Stud 2020 Aug;108:103632. [doi: 10.1016/j.ijnurstu.2020.103632] [Medline: 32505813]
45. Schrepp M, Hinderks A, Thomaschewski J. Construction of a benchmark for the User Experience Questionnaire (UEQ). Int J Interact Multimed Artif Intell 2017;4(4):40-44. [doi: 10.9781/ijimai.2017.445]
46. Malterud K. Qualitative research: Standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]
47. Ahmed SK. The pillars of trustworthiness in qualitative research. J Med Surg Public Health 2024 Apr;2:100051. [doi: 10.1016/j.glmedi.2024.100051]
48. Williams J, Craig TJ, Robson D. Barriers and facilitators of clinician and researcher collaborations: a qualitative study. BMC Health Serv Res 2020 Dec 5;20(1):1126. [doi: 10.1186/s12913-020-05978-w] [Medline: 33278896]
49. Domecq JP, Prutsky G, Elraiyah T, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014 Feb 26;14:89. [doi: 10.1186/1472-6963-14-89] [Medline: 24568690]
50. Mitchell J, Shirota C, Clanchy K. Factors that influence the adoption of rehabilitation technologies: a multi-disciplinary qualitative exploration. J Neuroeng Rehabil 2023 Jun 20;20(1):80. [doi: 10.1186/s12984-023-01194-9] [Medline: 37340496]

XSL•FO
RenderX

51. Eriksson S, Wallgren P, Sandsjö L, Karlsson M. Genuine co-design: An activity theory analysis involving emergency nurses in an interdisciplinary new product development project of a novel medical device. Int J Hum Factors Ergon 2021;8(4):331-369. [doi: 10.1504/IJHFE.2021.119054]
52. Barbosa IM, Alves PR, Silveira ZC. Upper limbs' assistive devices for stroke rehabilitation: a systematic review on design engineering solutions. J Braz Soc Mech Sci Eng 2021 May;43(5). [doi: 10.1007/s40430-021-02919-4]
53. Feldner HA, Papazian C, Peters K, Steele KM. "It's all sort of cool and interesting…but what do i do with it?" A qualitative study of stroke survivors' perceptions of surface electromyography. Front Neurol 2020;11:1037. [doi: 10.3389/fneur.2020.01037] [Medline: 33041981]
54. Feldner HA, Howell D, Kelly VE, McCoy SW, Steele KM. "Look, your muscles are firing!": A qualitative study of clinician perspectives on the use of surface electromyography in neurorehabilitation. Arch Phys Med Rehabil 2019 Apr;100(4):663-675. [doi: 10.1016/j.apmr.2018.09.120] [Medline: 30392855]
55. Cardoso VF, Valencia N, Loterio FA, et al. Towards an upper limb rehabilitation tool after stroke based on surface electromyography biofeedback and virtual reality. Res Biomed Eng 2022;38(3):1017-1025. [doi: 10.1007/s42600-022-00218-y]
56. Munoz-Novoa M, Andersson C, Sunnerhagen KS, Alt Murphy M. A novel intervention for upper limb rehabilitation in people with stroke combining myoelectric pattern recognition, virtual reality, and serious gaming: a qualitative study. Disabil Rehabil 2025 Jul;47(15):3930-3937. [doi: 10.1080/09638288.2024.2434643] [Medline: 39625112]
57. Lohse KR, Hilderman CGE, Cheung KL, Tatla S, Van der Loos HFM. Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. PLoS ONE 2014;9(3):e93318. [doi: 10.1371/journal.pone.0093318] [Medline: 24681826]
58. Burke JW, McNeill MDJ, Charles DK, Morrow PJ, Crosbie JH, McDonough SM. Optimising engagement for stroke rehabilitation using serious games. Vis Comput 2009 Dec;25(12):1085-1099. [doi: 10.1007/s00371-009-0387-4]
59. Bargeri S, Scalea S, Agosta F, et al. Effectiveness and safety of virtual reality rehabilitation after stroke: an overview of systematic reviews. EClinicalMedicine 2023 Oct;64:102220. [doi: 10.1016/j.eclinm.2023.102220] [Medline: 37745019]
60. Doumas I, Everard G, Dehem S, Lejeune T. Serious games for upper limb rehabilitation after stroke: a meta-analysis. J Neuroeng Rehabil 2021 Jun 15;18(1):100. [doi: 10.1186/s12984-021-00889-1] [Medline: 34130713]
61. Pregnolato G, Rimini D, Baldan F, et al. Clinical features to predict the use of a sEMG Wearable Device (REMO®) for hand motor training of stroke patients: a cross-sectional cohort study. Int J Environ Res Public Health 2023 Mar 14;20(6):5082. [doi: 10.3390/ijerph20065082] [Medline: 36981992]
62. Hochstenbach-Waelen A, Seelen HAM. Embracing change: Practical and theoretical considerations for successful implementation of technology assisting upper limb training in stroke. J Neuroeng Rehabil 2012 Aug 2;9:52. [doi: 10.1186/1743-0003-9-52] [Medline: 22856548]
63. Vourganas I, Stankovic V, Stankovic L, Kerr A. Factors that contribute to the use of stroke self-rehabilitation technologies: a review. JMIR Biomed Eng 2019;4(1):e13732. [doi: 10.2196/13732]
64. Mahmood A, Deshmukh A, Natarajan M, et al. Development of strategies to support home-based exercise adherence after stroke: a Delphi consensus. BMJ Open 2022 Jan 6;12(1):e055946. [doi: 10.1136/bmjopen-2021-055946] [Medline: 34992120]
65. Odetunde MO, Olaoye OA, Obajobi OA, Yusuf AA, Awotidebe TO. Adherence to home-based exercise programs among stroke survivors and perspectives of informal caregivers: a mixed method study. Bull Fac Phys Ther 2024;29(1). [doi: 10.1186/s43161-024-00224-4]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
sEMG: surface electromyography
UEQ: User Experience Questionnaire

Edited by S Munce; submitted 23.May.2025; peer-reviewed by A Kerr, D Hobbs; accepted 19.Feb.2026; published 13.Apr.2026.

Please cite as:
Munoz-Novoa M, Guo L, Björkquist A, Kristoffersen MB, Khorramshahi P, Sandsjö L, Alt Murphy M
A Smart Textile Biofeedback Training System for Upper Limb Rehabilitation After Stroke: Co-Design Development and Evaluation Study
JMIR Rehabil Assist Technol 2026;13:e77999
URL: https://rehab.jmir.org/2026/1/e77999
doi:10.2196/77999

XSL•FO
RenderX

© Maria Munoz-Novoa, Li Guo, Anna Björkquist, Morten B Kristoffersen, Peiman Khorramshahi, Leif Sandsjö, Margit Alt Murphy. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Voice-Assisted Technology for People With Parkinson's Disease Experiencing Speech and Voice Difficulties: Co-Designing Solutions Using Design Thinking

Jodie Mills[1], BSc, MSc; George Kernohan[2], BSc, PhD; Katy Pedlow[1], BSc, PhD; Orla Duffy[1], BSc, PhD

[1]School of Health Sciences, Faculty of Life and Health Sciences, University of Ulster, Belfast, Northern Ireland, United Kingdom

[2]School of Nursing, Faculty of Life and Health Sciences, University of Ulster, Belfast, United Kingdom

Corresponding Author:
Jodie Mills, BSc, MSc
School of Health Sciences
Faculty of Life and Health Sciences
University of Ulster
2-24 York Street
Belfast, Northern Ireland, BT15 1AP
United Kingdom
Phone: 44 2895365308
Email: mills-j12@ulster.ac.uk

Abstract

Background: While smart speakers are emerging as a novel health care technology, people with Parkinson's Disease (PwPD) and speech and language therapists (SaLTs) have reported difficulties using smart speakers with speech and voice impairments in research. To date, PwPD have identified frustration with having to repeat themselves to be understood, devices timing out before they had finished speaking, and being unable to have a conversation with smart speakers. SaLTs have reported technical and practical challenges in implementing voice-assisted technology tools. Both PwPD and SaLTs indicated a lack of knowledge about what smart speakers could do, as well as concerns about privacy and the listening nature of the devices.

Objective: This study aims to co-design solutions that support the use of smart speakers for speech and voice difficulties experienced by PwPD.

Methods: Based on the Design Thinking framework, a multistage design process was conducted, involving a lay steering group and 2 online co-design workshops. Twenty participants, including PwPD, carers, SaLTs, design and technology experts, and third-sector staff, collaborated during the co-design workshops. The ideate phase included brainstorming and ranking, and conventional content analysis was used to specify prototypes.

Results: Two main prototypes were created: (1) education and guidance, including privacy and therapeutic usage guides for PwPD and SaLTs to address troubleshooting and delivery considerations; and (2) new speech and language therapy (SLT)–specific features for smart speakers. Participants provided feedback on their experiences of co-design, highlighting feeling valued, the balance of perspectives, and making improvement suggestions. Feedback aligned with the UK standards for public involvement.

Conclusions: Smart speakers could enhance accessibility, therapy engagement, and long-term speech outcomes, offering scalable, cost-effective solutions to support SLT services, patient independence, and reduced service demand. Smart speaker solutions with a SLT focus enable PwPD to self-manage speech and voice difficulties at home and reinforce therapy gains between clinic visits. Co-designed with users, these prototypes are intended to address health disparities and relieve pressure on SLT services, offering a scalable and sustainable solution that enhances efficiency and supports ongoing rehabilitation within health care systems.

(JMIR Rehabil Assist Technol 2026;13:e84364) doi:10.2196/84364

KEYWORDS

voice-assisted technology; speech and language therapy; co-design; participatory methods; dysarthria; speech and voice; Parkinson's Disease

XSL•FO
RenderX

Introduction

Background

Voice-assisted technology (VAT) is defined as a device that uses natural language processing or automatic speech recognition (ASR) to interpret spoken language and translate it into actionable requests.

Smart speakers are commercially available VAT devices that are controlled using voice commands and are usually connected to the internet (current examples include Amazon Alexa and Google Assistant). They can feature built-in control systems for tasks on demand, including smart home automation, providing general information (not limited to weather, recipes, or health information), person-to-person calls, sending and receiving messages, and playing music. New models with screens can also support audio and video streaming. Smart speakers are readily available for purchase and use by the general public [1].

It has been reported that VAT prompts some participants with speech difficulties to modify their speech to enable interaction with VAT [2-6]. People with Parkinson's disease (PwPD) have reported adapting their speech by speaking more slowly, loudly, and clearly when interacting with a smart speaker [4]. Considering that 90% of PwPD present with reduced speech intelligibility and limited vocal loudness [7], VAT may hold potential as a therapeutic adjunct in speech and language therapy (SLT). This prior evidence indicates that VAT may enhance access to therapy [8].

Some therapists have reported using VAT to promote improved volume, clarity, and intelligibility of speech [9]. In addition to offering biofeedback on speech clarity, these tools have provided structured opportunities for home-based practice, fostering self-awareness and supporting the self-management of dysarthria and other speech difficulties [9,10]. PwPD have reported increased clarity of speech and volume when using VAT, and have used VAT as a communication partner to practice their speech and rebuild confidence in using their voice [11]. Both speech and language therapists (SaLTs) and PwPD agree that the objective nature of VAT is key to promoting interaction and providing feedback on speech. Textbox 1 presents a hypothetical vignette illustrating how a person with speech or voice difficulties may interact with smart speakers. This vignette is informed by the understanding and findings of previous research [11].

Textbox 1. Case study

1. Case

John is 65 and has had Parkinson's disease for 5 years. His phonation is impacted by poor breath support, resulting in a breathy, hoarse voice with low volume. His articulation is reduced, resulting in imprecise speech production, which reduces speech clarity and intelligibility. His speech is also hypernasal, and nasal emissions are noted. He has hypokinetic dysarthria. At home, his family can understand him, but he is frequently told that they "can't hear him" and that he "needs to speak up." This is frustrating for John, as he reports that "he feels like he is shouting," which suggests impaired self-awareness of his speech.

2. Use

John uses his smart speaker daily. When he speaks to the smart speaker, it replies with "Sorry, I didn't get that" approximately 50% of the time. As a result, John raises his volume and repeats his request. Often, he uses a loud voice, overarticulates his words, slows down, and speaks as soon as he takes a breath. The smart speaker responds when he uses these strategies. This demonstrates that smart speakers provide feedback on volume and clarity of speech in the form of an external cue: "Sorry, I didn't get that." This can encourage increased self-awareness of speech volume and intelligibility, and result in the use of LOUD, clear speech strategies. As John's smart speaker can time out before he has finished speaking, he uses *adaptive listening mode* (available on Amazon devices), which is found in the accessibility settings and gives him longer to speak.

John also plays the game "Word Tennis" on his smart speaker. He has to think of words within a semantic category quickly and remember to use a LOUD, clear voice when answering. This task focuses on a word-level activity within the speech hierarchy and adds a cognitive load to increase difficulty, which aligns with Lee Silverman Voice Treatment (LSVT) LOUD principles [12]. He also enjoys sport and cooking and uses his smart speaker to search for recipes. Common functional requests include "Add cheese and potatoes to my shopping list," "Show me my cooking library," and "What was the Man United score today?" Sometimes, he even uses his smart speaker like a diary: "Leave a sticky note for...", where he records a voice note on his smart speaker to remind someone to feed their dog.

3. Summary

Overall, interacting with his smart speaker allows John to practice a LOUD, clear voice at home, with external feedback on speech volume and clarity that may help improve his self-awareness.

Despite the facilitators discussed in Textbox 1, several barriers to the effective use of VAT among PwPD and SaLTs remain [9,11]. For example, PwPD have reported feeling frustrated by needing to repeat themselves to be understood, by devices timing out before they had finished speaking, and by being unable to have a conversation with their smart speaker [11]. SaLTs also indicated that they faced technical and practical challenges in implementing VAT tools [9]. Both PwPD and SaLTs reported a lack of knowledge about smart speaker capabilities and concerns surrounding privacy and data security.

Addressing these challenges is essential to enable the integration of VAT into SLT practice. Design Thinking is a user-centered innovation framework used to guide the development of new health care technologies, often utilizing co-design approaches [13-15]. It offers a structured approach to identifying problems and generating solutions through empathy, collaboration, and iterative prototyping and testing. This study is informed by the define, ideate, and prototyping phases of the Design Thinking process. Figure 1 outlines the Design Thinking process, and Table 1 shows the connections between the phases of the Design

XSL•FO
RenderX

Thinking framework, the specific research questions to be addressed, and the methods used.

Co-design has been used to foster collaboration that stimulates new ideas, clarifies concepts, and creates solutions that prioritize the needs and lived experiences of end users [16]. Co-design workshops have been used in SLT, health technology research, and with older adult populations [17-19], with improved outcomes for technology adoption compared with noncollaborative design processes [17,19]. Co-design is critical when developing technologies for SLT [18] and has value in engaging people with communication difficulties [20-22]. We set out to follow the co-design cycle and principles, meeting the criteria for true co-design under the ladder of co-production [23-25], through the identification and development of recommendations from participants with communication difficulties [26,27].

Figure 1. The Design Thinking stages [13] from empathize to prototyping, which can be used during the development of new health care technologies [15].

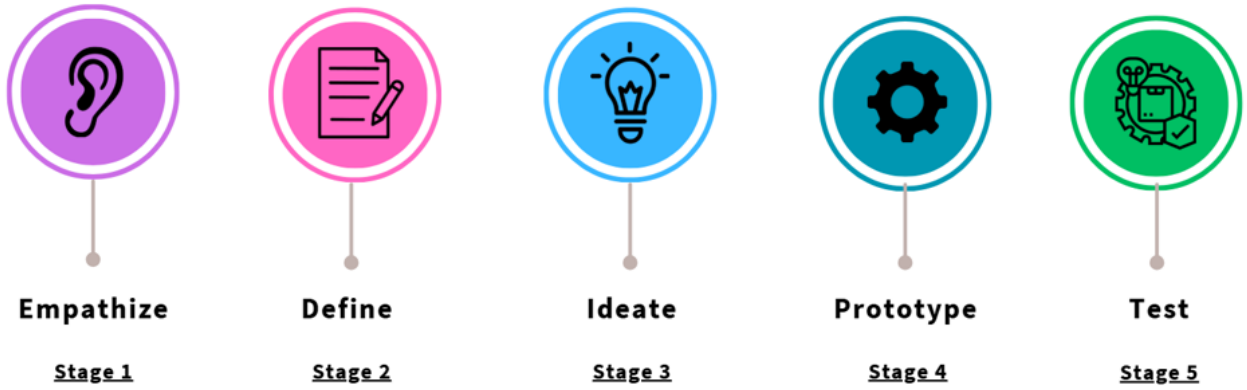


Table 1. Ways in which existing barriers to the therapeutic use of VAT[a] by people with Parkinson's can be solved.

Objective	Objective 1: To consider and select problem statements from the perspectives of experts by experience.	Objective 2: To create solutions to problem statements associated with VAT usage, alongside people with Parkinson's, carers, SaLTs[b], technology and design experts, and third-sector representatives.	Objective 3: To prioritize co-designed solutions to inform prototype VAT interventions.	Objective 3: To prioritize co-designed solutions to inform prototype VAT interventions.
Design Thinking stage	Stage 2: Define	Stage 3: Ideate	Stage 3: Ideate	Stage 4: Prototype
Method	Patient, public involvement workshop	Co-design workshop A	Co-design workshop B	Inductive content analysis

[a]VAT: voice-assisted technology.

[b]SaLT: speech and language therapist.

This research is intended to address previously noted barriers to VAT use [9,11] and aims to co-design solutions to previously identified challenges with VAT by working with PwPD, carers, SaLTs, charity representatives, and technology and design experts. This approach was taken to ensure that new technologies can be used in ways that meet end user needs. We sought to create solutions by using commercial technology, without coding or modifying VAT devices. This ensures that solutions are low cost and accessible, enhancing the potential for wider adoption of VAT in SLT contexts.

Aim

We set out to co-design solutions to support the use of smart speakers in SLT to improve volume and intelligibility for PwPD, using a Design Thinking framework. The research addressed the following question: "How can we facilitate the therapeutic use of VAT by people with Parkinson's Disease?"

Our study objectives (mapped onto Design Thinking stages) are as follows:

- To consider and select problem statements from the perspectives of experts by experience (Define).
- To co-create solutions to problem statements associated with VAT usage, alongside PwPD, carers, SaLTs, technology and design experts, and third-sector representatives (Ideate).
- To prioritize co-designed solutions to inform prototype VAT interventions (Prototype).

Methods

Participatory Co-Design Approach

Participatory methods such as co-design allow interventions to be designed around end user needs. This study co-designed solutions to previously identified barriers regarding the use of VAT when speech and voice difficulties were present. This results in technology that more readily meets user needs [28] and helps to avoid digital exclusion [29]. Workshops were held online, removing geographical and physical barriers and

XSL•FO
RenderX

enabling SaLTs from throughout the United Kingdom to share their experiences.

Ethical Considerations

Ethical approval was granted by the Ulster University Research Ethics Committee in January 2025 (approval number FCNUR-24-078-A). This study is part of a larger PhD project using Design Thinking. Previous phases of work aligned with the empathize stage, and the current co-design phase aligns with the define, ideation, and prototyping stages. All participants provided informed consent before the workshops. Participant outputs were anonymous, and ground rules were agreed upon to maintain confidentiality. Participants did not receive payment or financial incentives.

Patient and Public Involvement

A patient and public involvement (PPI) steering group was established to provide a voice for key stakeholders and ensure their active role in shaping the research. This group included a SaLT with firsthand experience using VAT in clinical settings, a person living with Parkinson's, and a caregiver. These 3 experts by experience coassessed the barriers to VAT usage identified in previous research [9,11] and co-decided the top 5 problems that reflected their experiences, in keeping with the cycle of coproduction [24].

Study Recruitment

PwPD and carers were recruited via a third-sector organization (Parkinson's UK) using advertisements on the Parkinson's UK research portal, Research Support Network monthly emails, and flyers at local Parkinson's support groups in Northern Ireland. SaLTs were recruited through the Royal College of Speech and Language Therapists, including the Parkinson's Clinical Excellence Network. In addition, Parkinson's UK staff and technology or design experts were recruited through the lead author's (JM) professional network. Some participants had established a relationship with the lead researcher through work with the local branch of Parkinson's UK in Northern Ireland.

Previous research was used to determine the number of participant collaborators invited to share their experiences during the workshops (n=20) [30,31]. Participants were asked to contact the research team to express interest in the study and were screened according to predefined criteria (Table 2). Potential participants were sent study information and consent forms by email or post, depending on preference, and were asked to indicate their availability. Once consent was obtained, participants completed a demographic survey and received links for the online workshops. This enabled interaction between diverse experiences. Participants were placed into smaller, experience-diverse groups of 4-5 participants to encourage idea generation in a safe and supportive environment.

XSL•FO
RenderX

Table 2. Inclusion and exclusion criteria for people with Parkinson's, carers, speech and language therapists, third-sector representatives, and technology or design experts.

Participant group	Inclusion	Exclusion
People with Parkinson's	• Adults over 18 years old • Mild to moderate dysarthria/voice difficulties (to include users of augmentative, alternative communication) • Diagnosis of Parkinson's disease • Current or previous use of VAT[a] • Have access to a laptop, with a camera, that facilitates videoconferencing software	• Moderate or severe cognitive impairment • History of other neurological disorders
Carers	• Adults over 18 years old • Live with or care for a PwPD[c] or both • Experience of facilitating the use of VAT with a PwPD • Have access to a laptop, with a camera, that facilitates videoconferencing software	N/A[b]
Speech and language therapists	• Adults over 18 years old • Who currently have/have had a clinical caseload of PwPD in the past 5 years • Who have used VAT in practice and have basic knowledge of the devices • Have a laptop, with a camera, that facilitates videoconferencing software	N/A
Third-sector staff	• Adults over 18 years old • Currently working in a third-sector organization for PwPD • Involvement and relationships with the local Parkinson's community • Basic knowledge of speech and voice difficulties in Parkinson's disease • Have a laptop, with a camera, that facilitates videoconferencing software	N/A
Technology/design experts	• Adults over 18 years old • Experience of VAT and detailed knowledge of its capabilities or relevant experience in designing or developing health care technologies	N/A

[a]VAT: voice-assisted technology.

[b]N/A: not applicable.

[c]PwPD: people with Parkinson's disease.

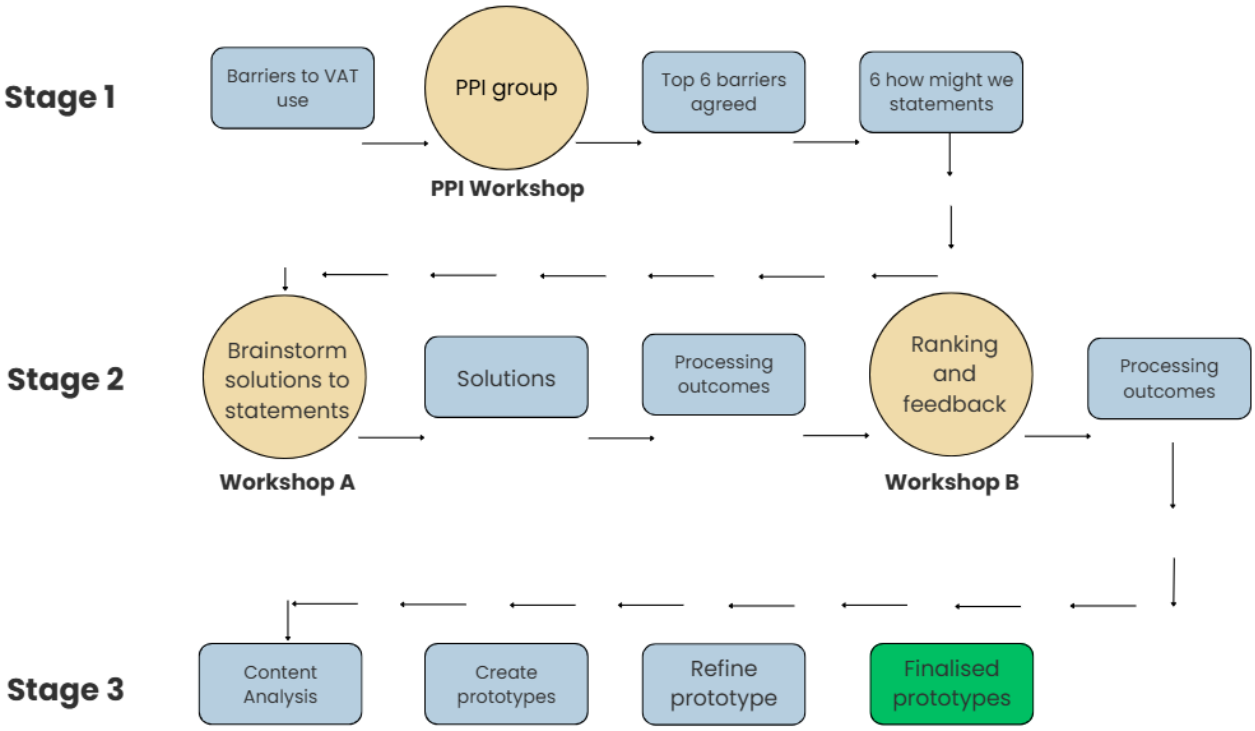
Procedure

Overview

Using principles underpinned by Design Thinking and Participatory methodology, a series of 2 co-design workshops were undertaken, informed by insights gained from previous research [30-32]. The 3-stage co-design procedure is outlined below (see Figure 2) and aligns with the define, ideate, and prototype phases of the Design Thinking framework, as presented in Figure 1.

XSL•FO
RenderX

Figure 2. The co-design process from the development of problem statements, through workshop completion, analysis, prototype creation, and refinement. PPI: patient and public involvement; VAT: voice-assisted technology.



Stage 1: Define Phase

Barriers identified by PwPD, their carers, and SaLTs during an earlier stage of the research [9,11] were reviewed by the PPI group, ensuring that the research was shaped by lived experience. Researchers JM and OD were present during the workshop. The group identified and agreed on the top 6 problems to be brought forward to workshop A, and, as per Design Thinking guidance, these were reframed into "How Might We" statements. These "How Might We" statements are used in Design Thinking to help people reframe a problem-focused perspective into solution-focused thinking [33], and previous research highlights the value of shaping research based on lived experience, as this ensures that the problems being solved are meaningful to end users [34]. The final 6 "How Might We" statements are presented in Table 3.

XSL•FO
RenderX

Table 3. Procedures and outcomes for stages 1, 2, and 3 of the co-design process.

Stage	Design Thinking framework stage	Procedure	Outcomes
1	Define	**PPI[a] workshop** • The PPI group was presented with all barriers from previous research [9,11]. • JM shared her screen via videoconferencing, using Canva PowerPoint software to present the barriers identified in previous research and to assist with live decision-making. This acted as contemporaneous notes. OD also took field notes. • PPI group identified 6 top problems—felt to be a reasonable number of barriers to brainstorm solutions to within an hour. • 6 problems reframed into "How Might We" statements.	• The following "How Might We" statements were used in workshop A: • How might we help people understand smart speaker privacy and reduce their fears? • How might we help people when smart speakers do not work? • How might we help people to have a conversation with a smart speaker? • How might we help people to know what smart speakers can do? • How could we deliver this information? • How could smart speaker technology be adapted?
2a	Ideate	**Co-design workshop A** • A document with problem statements and examples of the problems was emailed to the participants before the workshop. • Welcome and introduction PowerPoint, which presented an overview of previous research, problems experienced by people using smart speakers, and an overview of the co-design process. • Two example brainstorming activities completed. • Smaller groups discussed 6 problem statements, rotating after 10 minutes on each problem statement. All groups completed different problems at the same time. • Facilitators shaped discussions and noted contributions on a live Word document. Each problem statement had a separate Word document. • Participants were invited to share solutions to the problem statements around using VAT[b] with a speech or voice difficulty. • Participants were thanked for their time and the workshop ended.	• Solutions to each problem statement were reviewed. Similar solutions for each problem statement were grouped together and combined. No content was removed, while ensuring there was a feasible number of ideas to rank. • Ideas are presented fully in Multimedia Appendix 1, as they appeared during the workshop. • A reduced number of solutions were placed into tables for each group to rank in workshop B.
2b	Ideate	**Co-design workshop B** • Welcome and recap of solutions generated in workshop A. Participants received a document of the solutions ahead of the workshop to aid their preparation. • Each facilitator worked through 3 solutions documents with their group and ranked their top 5 priorities for each solution. • A similar procedure to workshop A was followed: each solution was captured in a separate Word document, and groups rotated after 10 minutes on each problem statement, inputting to the same solution document. • Each participant was asked a short series of questions about their experience of the co-design workshops. Questions were based on the UK standards for public involvement. • Participants were thanked for time, next steps were explained, and the workshop ended.	• Lead researcher (JM) collated solutions and rankings and removed any solutions that were not ranked. • Solutions ranked and their rankings taken forward to stage 3 (prototyping).

XSL•FO
RenderX

Stage	Design Thinking framework stage	Procedure	Outcomes
3	Prototyping	**5 stages of inductive content analysis conducted for solutions and rankings from workshop B** • Initial and final codes were discussed for refinement and agreement with the research team and broadly grouped into themes. • Themes were mapped into 2 prototypes.	• Two prototypes were created, namely, prototype 1 (education and guidance) and prototype 2 (developing new SLTc-specific features for smart speakers). • An overview of the link between workshop 2 outputs and final prototypes is available in Multimedia Appendix 2. • Participant feedback was collated.

[a]PPI: patient and public involvement.

[b]VAT: voice-assisted technology.

[c]SLT: speech and language therapy.

Stage 2: Ideate Phase

Co-design workshops were facilitated by 2 qualified SaLTs (JM and OD), 2 health care professionals (KP and RB), and an academic (GK), with participants working in smaller groups during the workshops. Facilitator guidance and training before the workshops ensured methodological consistency when participants worked within the groups (Multimedia Appendix 3). Co-design principles of valuing lived experience, sharing power, and respect were presented at the beginning of each workshop, aiming to reduce power imbalances between researchers and participants. Both workshops lasted approximately 1 hour and were conducted via videoconferencing to enable data collection across a wider geographical area [35], avoid travel, and facilitate workshops in the evenings.

Data collection was recorded as notes in live Word (Microsoft Corporation) documents. By recording content-only contributions, participant anonymity was ensured from the outset, as no identifiers were associated with the contributions. Although participants may have been known to other group members, they were asked to respect everyone's right to confidentiality by not sharing contributions outside the group setting. Additionally, workshops were not audio- or video-recorded, in keeping with co-design principles. Although this may have contributed to some data loss, the live recording of workshop contributions helped to mitigate this risk.

In workshop A, participants brainstormed solutions to the "How Might We" statements shown above. The process for this workshop is shown in Table 3. Following workshop A, these solutions were refined by combining similar ideas and removing duplicates (Multimedia Appendix 1) in preparation for workshop B. In workshop B, participants reviewed the solutions to the 6 problem statements created during workshop A and were asked to rank the top solutions for each problem statement from 1 (top priority) to 5 (lower priority). Ranking is regarded as a way to prioritize and reach an agreement [36]. To ensure priorities accurately reflected participants' lived experience, each problem statement was ranked by 2 groups; however, this meant that not every group ranked every problem statement. The workshop procedure is outlined in Table 3, and the solutions are presented in the "Results" section. At the conclusion of the workshop, participants were asked by facilitators to provide feedback on their experiences of the co-design process. Following workshop B, the lead researcher (JM) collated the rankings and removed any solutions that were not ranked.

Stage 3: Prototyping

Outcomes from the workshops were analyzed using content analysis and used to create prototypes, in line with the Design Thinking process. Conventional inductive content analysis, following 5 stages, was used to allow categories to emerge directly from the workshop outputs and to reduce the volume of information [37].

The lead researcher (JM) read through the workshop outputs and any associated field notes several times to become immersed in the data. This supported note-making on initial ideas in a reflexive journal, allowing consideration of connections, similarities, and differences within the data. This process also highlighted that participants generally lacked knowledge about smart speakers, were fearful of hackers, and wanted speech-accessible smart speakers. These insights challenged the lead researcher's confirmation bias, encouraging empathy with the experiences of PwPD and allowing the research to be shaped by user needs. This highlighted the importance of creating new features for smart speakers that better meet users' needs, as well as utilizing existing features. Reflexivity also allowed the lead researcher to reflect on her multiple roles as a SaLT, facilitator, and analyst, and the potential for these roles to introduce interpretation bias toward a clinical perspective. As a result, an audit trail was developed to demonstrate the analysis process and enhance trust in decision-making during analysis [38].

Following data immersion, the lead author created a mind map to inductively group ideas and develop initial codes for analysis. Initial codes were shared with OD, KP, and GK for discussion and refinement, enhancing credibility through investigator triangulation and peer debriefing. The final codes were both descriptive (eg, "privacy concern") and interpretative (eg, "need to increase motivation for speech practice"), and the meaning of each code was documented to ensure reliability during coding.

XSL•FO
RenderX

Subsequently, the workshop outputs were coded. This was conducted by hand, using graph paper and colored pens to assign meaning to each output.

Coding was conducted 3 times on separate days by 1 coder (JM) and was presented to the research team for discussion, redrafting, and agreement. It is acknowledged that coding by a single researcher may introduce bias, and, upon reflection, the involvement of 2 coders may have enabled data triangulation and enhanced data credibility. Despite this, peer debriefing helped to minimize potential impacts on the analysis of results. Similar codes were grouped into broader themes to capture meaning across the outputs. This process was also completed by hand, using colored pens to illustrate relationships between codes. Reflexive notes were recorded, discussed with coauthors (OD, KP, and GK), and refined accordingly. Finally, themes were conceptually mapped into 2 prototypes, in keeping with the Design Thinking Framework.

The content analysis process described above, from ranked ideas to the creation of prototypes, is available in Multimedia Appendix 2 as an audit trail, enhancing the credibility of the outputs [39]. All solutions that were ranked in workshop B were included in the prototypes to ensure that the prototypes reflected the wants and needs of participants. Furthermore, direct quotations, where available, from participant feedback are presented to provide a direct voice and to link outputs with interpretations [32]. Findings were sent to all participants for member checking to ensure that the written findings reflected their lived experiences and to enhance the rigor of the research.

Results

Study Participants

A total of 20 participants were recruited; 19 participated in co-design workshop A, and 16 in co-design workshop B. Overall, 15 participants took part in both workshops (Tables 4 and 5).

Table 4. Makeup of breakout rooms in workshop A.

Group number	Number of participants, n	Participants				
		People with Parkinson's disease	Carer	Speech and language therapist	Third sector	Technology/design
1	4	✓	✓	✓	N/A[a]	✓
2	3	✓	✓	✓	N/A	N/A
3	4	✓	✓	✓	✓	N/A
4	4	✓	✓	✓	N/A	✓
5	4	✓	N/A	✓	N/A	✓

[a]N/A: not applicable.

Table 5. Makeup of breakout rooms in workshop B.

Group number	Facilitator	Number of participants	Participants				
			People with Parkinson's disease	Carer	Speech and language therapist	Third sector	Technology/design
1	GK	4	✓	✓	✓	N/A[a]	✓
2	OD	4	✓	✓	✓	N/A	N/A
3	JM	4	✓	✓	✓	✓	N/A
4	KP	4	✓	N/A	✓	✓	✓

[a]N/A: not applicable.

The ranking of solutions aligned with the 6 problem statements is outlined in Multimedia Appendix 4. For problem statements 1, 2, 4, and 6, 3 solutions were ranked by both groups, and 4 solutions were ranked by 1 group. For problem statement 3, 2 solutions were ranked by both groups, and 5 solutions were ranked by 1 group. For problem statement 5, all 4 solutions were ranked by both groups, as only 4 options were presented. Solutions that were not given a rank by any group were removed from the results presented below. The full list of ideas available for ranking during workshop B is provided in Multimedia Appendix 1.

Stage 1: Prototyping Results

Rankings were collated into 2 main prototypes by the primary researcher and agreed upon by the team: (1) educational guidance on the therapeutic use of smart speakers, and (2) developing new SLT-specific features for smart speakers (Multimedia Appendices 5 and 6). These prototypes present the results outlined above, emphasizing cross-cutting themes. Participants' experiences of the co-design process are also presented. The process from solutions to prototypes is fully detailed in Multimedia Appendix 2.

XSL•FO
RenderX

Prototype 1: Education and Guidance

Guides for PwPD and SaLTs, detailing how to use smart speakers to improve volume, intelligibility, and clarity of speech, were unanimously agreed upon by participants. The contents of these guides are described in Multimedia Appendix 5. Participants highlighted a gap in knowledge between the traditionally available features of smart speakers and an understanding of how these features could be repurposed to benefit speech and voice in Parkinson's disease. The suggested skills catalog for therapy would create a repository of standard smart speaker features and skills that could be utilized with therapeutic intent by SaLTs and PwPD. Suggestions included integrating prompts and positive reinforcement by building routines, for example: "Could you speak louder?" or "Well done, great practice today." PwPD felt that verbal prompts to speak louder or clearer, along with positive reinforcement from smart speakers, would replicate cuing provided by SaLTs during direct therapy and motivate home practice.

Participants indicated that routines could be used to practice scripted conversations, and that these should be personalized, include prompts to help sustain conversations, and contain only personal information that users felt comfortable sharing with their smart speaker.

It was evident that not all participants were aware of these accessibility features, which are designed to maximize engagement with smart speakers, and that education in this area may help to encourage more natural conversational reciprocity. For example, the conversation mode available on Amazon Alexa devices.

Participants also indicated that education about privacy relating to smart speaker use was required. It was reported that education for both PwPD and SaLTs would help to alleviate fears regarding personal data storage and General Data Protection Regulation (GDPR) concerns.

Prototype 2: Developing New SLT-Specific Features for Smart Speakers

Participants indicated that an Alexa skill could be created to support speech therapy, as shown in Multimedia Appendix 6. Suggestions included delivering LSVT through a smart speaker or developing a speech therapy game to support speech and voice practice. Participants suggested that this could include increased feedback, such as visual cues on a screen for volume and speech clarity and live transcription of speech that repeats back what was heard, to support self-awareness in PwPD. It is acknowledged that newer Amazon Alexa models, such as the Echo Show 10, already offer subtitling features within the settings, which provide live captioning of speech or video calls.

Additionally, participants were excited about the potential for artificial intelligence (AI) integration within smart speakers and suggested that this could be used to enable more intelligent conversations with the device. Many participants indicated that current smart speakers lacked this capability. Although intelligent conversation has not yet been integrated as a core functionality across Amazon Alexa devices, skills such as ChatGPT were perceived to facilitate live, functional conversation. Furthermore, Alexa Plus, a paid feature for Amazon devices, uses generative AI to remember previous interactions and continue conversations over time. It also offers 5 personalities, which may help users feel as though they are conversing with a person rather than a device. However, Alexa Plus is not yet available in Northern Ireland, where this research was conducted. Additionally, the *follow-up* mode within Alexa accessibility settings prevents users from having to repeat the device wake word, which Amazon suggests supports a more conversational interaction with smart speakers.

Participants also indicated that extended listening time for smart speakers would prevent mid-sentence interruptions. It is acknowledged that Amazon Alexa devices currently offer an *adaptive listening* feature in the accessibility settings, which extends input time and accommodates speech differences. Although a few participants were aware of this feature, they did not indicate its impact on their smart speaker interactions.

Furthermore, enhanced privacy features were suggested. Again, it is understood that, under Alexa privacy settings, voice commands can be enabled to clear Alexa voice history; for example, "Alexa, delete everything I've ever said." Additionally, although Alexa cannot be trained to respond only to certain voices, there is an option to set up a voice profile to receive more personalized content and prevent unauthorized voice purchases.

Although participants acknowledged that adapting smart speakers to better recognize dysarthric speech could hamper their therapeutic value, they felt that this would improve accessibility for the devices more generally. They sought devices that could gradually learn their speech patterns over time, as well as deal effectively with regional accents. Notably, there is currently no research exploring the impact of improved speech recognition in smart speakers on therapy outcomes in SLT.

In addition to ranking solutions, participants were asked about their experience of co-design using questions based on the UK standards for public involvement. Overall, participants valued the online workshop format, which facilitated engagement for those with limited mobility. They felt the workshops were informal yet professional and found the tasks interesting, positively challenging them to think of solutions. Small groups were reportedly the right size for supported discussions, and participants felt this was an effective way to gather substantial information. Participants discussed their expectations and involvement in co-design, describing feeling included and respected:

> *I had some experience of delivering co-design, so I had an idea of how it should be done...I felt valued, and felt everyone has been really equally valued, no matter how you're coming at it; person with Parkinson's, speech therapist, whatever. We've all been treated equally, with respect.* [Person from Parkinson's UK]

> *I felt heard and respected throughout and you did a good job of facilitating conversations for us to feel heard.* [SaLT]

> *The carer and patient are heard. So often in NHS setting they are the last ones to be heard y'know, what*

XSL•FO
RenderX

> *would they know. But here, they were put front and centre.* [Carer]

Participants also felt that the right people were involved in the co-design process and that there was a good balance between perspectives:

> *It's involved so many stakeholders that come from that same place of making improvements for people living with Parkinson's. It was great to see various individuals are spoken to and included.* [PwPD]

> *There was a really good balance of people from different backgrounds...It absolutely worked and its so important to get everyone's view; it's mostly important to hear people with Parkinson's, carers you work alongside. You get a really holistic picture of what is the most important thing from different perspectives.* [SaLT]

> *It was useful to be able to discuss together in a group and helpful to consider all views: SaLTs, patients and tech experts.* [PwPD]

Participants provided feedback on engagement challenges and future improvements. Some PwPD or carers felt that a bridging workshop between creating and ranking solutions would be helpful. This could have included a session to discuss all brainstormed solutions and integrate them with real-world examples. Although elements of this were included in the workshops, they felt that a third workshop would have given them time to digest the large number of solutions and some more complex ideas before ranking them. One clinician who was unable to attend the first brainstorming workshop felt that this would have helped orient her more fully before the ranking task. Others suggested that color grouping or collapsing solutions for each problem statement by themes may have made it easier to rank statements. Participants also indicated that more prompts were required to remind them to think creatively and that "anything was possible."

Participants also discussed the project's focus on smart speakers, as well as their advantages and disadvantages. A few participants felt that it would be easier to create an app, as many are available for smartphones, and most people use these devices. However, most participants felt that the voice interaction of smart speakers offered advantages over smartphones, particularly for people with a tremor. Additionally, participants felt that smart speakers could remind and motivate users to practice, whereas with an app, users often have to self-motivate or remind themselves.

Discussion

Principal Findings

This study aimed to co-produce solutions to support smart speaker use for speech and voice difficulties and to inform a future intervention. PwPD, carers, SaLTs, Parkinson's UK staff, and technology and design experts collaborated during 2 online co-design workshops to brainstorm and prioritize solutions to problems identified in prior research [9,11]. Two prototypes were developed: (1) education and guidance on the therapeutic use of smart speakers and (2) the development of new speech therapy–specific features for smart speakers. By incorporating collaborators' priorities and needs, the study offers a foundation for a future smart speaker–based intervention for speech and voice therapy in PwPD.

Impact of Co-Design

This project recognizes the need to involve end users early and meaningfully when designing health care interventions [40,41], contributing to the quality and relevance of co-designed outcomes [42]. This aligns with the Design Thinking framework, specifically the ideate and prototyping phases. While co-production with people with aphasia is increasing, there is limited evidence on co-design in SLT, especially for motor speech disorders [21,32,43]. Therefore, this research continues to contribute to and develop the evidence base regarding co-design in SLT, particularly for people with dysarthria. This study is unique, as it is believed to be the first co-design study with PwPD who have speech and voice difficulties that co-designs solutions to problems experienced when using commercial VAT technology.

Participants described personal benefits of co-design, including gaining knowledge, social interaction, and feeling heard and validated, echoing previous co-production findings [32,44] and aligning with public involvement standards [45]. These benefits are particularly relevant for PwPD, who often experience reduced participation due to speech and voice issues [46], highlighting how co-design can empower participants. Power sharing and partnership can enhance engagement and lead to more patient-centered outcomes [47], and involving SaLTs may also improve future implementation of such tools into clinical practice [48]. In wider co-design research in SLT, participants with communication difficulties report improved confidence, motivation, and sense of well-being [27], and their involvement can lead to more and better-quality outcomes [43]. Overall, this demonstrates how co-production can allow participants, such as PwPD, to feel in control, empowered, and validated. For PwPD, this co-design study both physically and metaphorically provided them with a voice, building on current evidence. Despite this, wider research also acknowledges that relinquishing power in research can be challenging for researchers, requiring an active effort to make the co-design process truly collaborative [18].

Additionally, collaborator feedback highlighted the importance of skilled facilitation in enabling communication during workshops. Although evidence on co-design facilitation strategies for people with speech and voice difficulties is limited, facilitators used clinical experience and evidence-based strategies [32] to support PwPD. These included allowing preparation time before workshops, building rapport, giving extra time to speak, screen-sharing key points, regularly checking understanding, and summarizing discussions [43,49,50]. Such approaches are crucial for inclusive and accessible co-design. Some collaborators suggested improvements, such as offering more workshops and using multimedia formats to make tasks easier, which extends the evidence base on co-design with PwPD who have speech and voice difficulties. This balance of positive experiences and suggested improvements reflects the range of participants, lived experiences, and heterogeneous needs. Advantages and

XSL•FO
RenderX

disadvantages of co-design methods should be evaluated from a range of perspectives to achieve a balance between the needs of a diverse group of PwPD.

Participants indicated that training for SaLTs in the therapeutic use of smart speakers for speech and voice difficulties was a priority. Wider research supports this finding, showing that education and guidance are required to support therapeutic adoption by SaLTs and PwPD [4,9-11], and that digital health interventions for older adults should include education in effective device use, digital literacy skills, and technical support throughout [51,52]. Tailored education and guidance may contribute to PwPD and SaLTs successfully adopting and using smart speakers to support speech and voice difficulties. As such, this study begins to advance understanding of how to support VAT adoption into clinical SLT practice. While smart speaker features make them valuable tools for chronic health management among older adults [53,54], older people in particular can struggle to comprehend the full range of smart speaker functions [11,54].

Guidance should clearly link device features to SLT goals to promote understanding and demonstrate how devices can help people achieve their SLT practice and related goals [55,56]. This may positively impact digital literacy for PwPD, supporting device adoption and regular use [56], again contributing to advances in knowledge regarding the clinical adoption of VAT. Similarly, SaLTs in our earlier research made several content suggestions for guidance to empower them to use VAT [9], including sample therapy plans, scripts, goal-setting frameworks, and evidence-based practice. However, this is the first study to collate these elements into an education and guidance prototype for SaLTs. Simplified guidance is particularly important, as clinicians often discontinue technologies they perceive as overly complex for clients [57]. Similar requirements for implementation guidance have been reported in SLT research using commercial technologies, such as virtual reality (VR) [19,58,59]. These studies highlight that therapeutic usage guides should promote the ease of use and usefulness of commercial technologies to support clinical adoption and provide opportunities to trial the devices. However, it is important to note that although commercial VR technology was used, the VR program itself was specifically created by researchers. This suggests that guidance must explain how smart speakers' out-of-the-box "Alexa skills" are relevant to SLT, given that the commercial use of the technology is not intended to be therapeutic. Unlike custom VR programs, smart speakers are off-the-shelf products not originally designed for health care. Therefore, guidance must explicitly link commercial features to therapeutic aims and support clinicians in adapting features to individual client needs, ultimately contributing to the adoption of VAT into clinical practice.

Furthermore, privacy and data protection are significant barriers to the adoption of smart speakers [60]. Common concerns include the recording of conversations and data misuse, which can deter both clinicians and clients [51,61-63]. To address this, usage guides for PwPD should include clear, accessible privacy information, support informed consent, and clearly explain how devices handle user data [64]. Given SaLTs' responsibility for safeguarding client data, guidance should map VAT's GDPR compliance and potential risks to SLT governance policies, such as Data Protection Impact Assessments. This study, therefore, begins to answer questions posed by previous research [10] regarding how VAT may be implemented in accordance with clinical governance requirements. Previous findings highlight that many SaLTs lack clarity on which technologies meet governance and GDPR standards [57]. Reassuring both clinicians and PwPD about privacy may improve confidence and facilitate adoption [65]. Future evaluation of guidance acceptability and usability could apply frameworks such as the Technology Acceptance Model or Unified Theory of Acceptance and Use of Technology 2.

Delivery

Participants suggested delivering training through Royal College of Speech and Language Therapy–led webinars, live demonstrations, and group sessions led by trained SaLTs. While previous research has not identified optimal delivery formats [4,10], this study provides new insights into practical implementation and contributes to the evidence base regarding the therapeutic use of VAT in SLT clinical practice. The literature indicates that older adults often prefer hands-on, task-based learning supported by written instructions [66,67]. A training program using VAT as a tool for activities of daily living with adults with cognitive communication disorders indicated a need for written, easy-to-follow instructions, with hands-on support to overcome low technological literacy [67]. Group-based workshops can offer a supportive, low-pressure environment for exploration and skill-building with in-person support [52]. This is particularly important for users with limited experience or confidence in digital tools.

Findings indicate that SaLTs are central to introducing and supporting smart speaker use in therapy. When clinicians demonstrate relevance and ease of use, PwPD may be more likely to adopt the technology [67]. By increasing perceived usefulness and reducing concerns, training can enhance performance expectancy and digital engagement. Additionally, previous research on integrating commercial technologies in SLT has highlighted the importance of multifaceted training approaches, including device trials, workshops, clinical manuals, and information technology support [58,68]. Additional methods, such as guided observation and co-delivered interventions, may be necessary to bridge the gap between knowledge and practice [58,69]. As such, this study begins to address gaps in knowledge regarding the implementation of VAT as a therapeutic tool for speech and voice difficulties associated with Parkinson's disease.

Participants highlighted the need to develop SLT-specific features for smart speakers, designed for therapeutic use. For example, Cassano et al [70] described a SaLT building a custom skill. At the time of publication, at least three speech therapy Alexa skills existed: Speech Therapy Practice, Speech Device Practice, and Let's Talk. Additionally, 2 further speech therapy skills were identified but are no longer publicly available on the Amazon Skills store: Speech Doctor, as discussed by Makin et al [71], and Speech Therapy by Cathal Killeen. Notably, Speech Therapy Practice is a live Alexa skill developed by a SaLT that enables people with aphasia to practice very basic

XSL•FO
RenderX

words and phrases, such as colors, opposites, who/what questions, and yes/no questions. While this may potentially act as a starting point for SaLTs, the skill lacks applicability to practicing phrases and sentences and, in its current state, is unlikely to meet the speech practice needs of PwPD. To date, there are no specific Alexa skills for adults with Parkinson's or targeting dysarthria, and our research highlights the potential for future development. Future research may seek to work with developers to create an Alexa skill for this population that can be used to support home practice of speech therapy exercises. Features may include prompts for loud, clear speech; increased feedback on volume and intelligibility with suggestions for improvement; the ability to monitor progress; visual displays and biofeedback; reminders to complete therapy tasks; and LSVT-style exercises with gamification [9,11]. Such features align with wider studies integrating technology into SLT and related areas, including apps using Google Glass [72], smart speaker–based physical activity interventions [73,74], and social engagement tools for people with disabilities [75]. Development platforms like Alexa Skills Kit and Alexa Blueprint may offer scalable, cost-effective options, enabling a focus on increasing motivation, engagement, and potential adherence to intervention programs. A curated hub of Alexa skills that can be used for SLT goals may also support clinical implementation. For example, Esquivel et al [76] developed a repository of Alexa skills and recommendations for people with disabilities, by people with disabilities. Future research may explore the acceptability of a speech therapy–specific Alexa skill and its implementation within clinical practice.

Given the commercial nature of smart speakers, it may be beneficial to first assess their current therapeutic value before creating bespoke skills. As our study focused on co-design processes and did not include a formal evaluation of intervention usability, effectiveness, or acceptability, future research may consider testing the current prototypes to determine real-world clinical impact and user outcomes. This study establishes the rationale for a future feasibility study to examine the effectiveness of VAT as a therapeutic tool for speech and voice difficulties in Parkinson's disease. At the time of writing, no studies have been conducted in this area using commercial VAT. Emerging SLT research shows benefits for speech clarity in populations with intellectual disabilities and speech sound disorders [5,71], citing immediate rewards, spaced practice, enhanced autonomy, intrinsic motivation, and reduced social barriers as mechanisms of change in speech. However, these interventions do not follow established SLT intervention protocols [5,71]. Therefore, future studies should evaluate the effectiveness and usability of smart speakers for PwPD using principles of neuroplasticity and motor learning from SLT protocols, such as LSVT LOUD or Speak Out!

Despite this, the challenges surrounding the therapeutic use of smart speakers cannot be ignored. Smart speakers rely on evolving ASR models, a type of AI, which are continually being improved. ASR models can change without warning, presenting a risk to the reliability of baseline measurements and the measurement of therapy goals [77]. Furthermore, ASR errors are often higher than expected for dysarthric speech, speakers of minority languages, and those with regional accents [78-80]. Without clear and specific feedback on device or speech errors, both PwPD and SaLTs are left without information about where the "error" lies, whether it is speech- or device-related. These risks may reinforce maladaptive speech behaviors if speech practice is based on inconsistent or misleading responses from the device. It may also damage client motivation and confidence, with PwPD blaming themselves for technological errors. Research demonstrates that speakers can attribute ASR errors to themselves and link this to their sense of identity, including racial, regional, and locational identity [77]. To mitigate this lack of transparency, it is essential that SaLTs educate potential VAT users on strategies for adapting speech, managing frustration, and correctly interpreting VAT errors, as well as raising awareness of the limited ASR training on dysarthric speech and some minority or foreign languages [9,11]. This highlights the importance of a therapeutic usage manual for smart speakers for people with speech and voice difficulties and for SaLTs, as indicated in the current findings.

However, it should be noted that projects such as Voiceitt, Google Euphonia, and Project Relate aim to improve ASR accuracy in recognizing dysarthric speech, and Accessible Voice Interaction Technology for Aphasia (AVITA) aims to improve the accessibility of smart speakers [81], which may have the unintended consequence of limiting certain therapeutic applications of smart speakers in SLT. When smart speaker recognition is improved, speech difficulties no longer affect recognition, meaning all speech is easily recognized. This can be problematic, as speech that may not be intelligible in real life is recognized by devices. Consequently, this hampers therapeutic applications, because positive biofeedback provided by smart speakers does not reflect the speaker's intelligibility to unfamiliar listeners in everyday contexts. Indeed, participants in this research indicated that future adaptations of smart speakers, outside of therapeutic contexts, should aim to better recognize dysarthric speech and regional accents. Future smart speaker designs may bridge the gap between standard out-of-the-box devices and fully customized skills. For example, smart speakers could allow users to set recognition thresholds, enabling both increased accessibility for users and therapeutic usage for clinicians. Given the rapid pace of innovation, continued review of emerging literature and technologies is recommended throughout the development and implementation stages.

Limitations

This co-design study offered valuable insights into developing VAT tools for PwPD with speech and voice difficulties; however, limitations are evident.

Participants suggested an additional workshop between the ideation and prioritization phases, that is, between workshops A and B. A bridging session could have allowed more reflection and improved understanding, potentially leading to rankings that more accurately reflected lived experience. Furthermore, although recruitment was successful, there was some participant dropout between workshops A and B. This necessitated merging groups in workshop B, which may have influenced group dynamics and limited continuity of discussion.

XSL•FO
RenderX

Despite efforts to recruit a diverse group, the sample was small (n=20), and certain perspectives, such as those of people with advanced Parkinson's or severe dysarthria, were underrepresented. This may limit the generalizability of the findings.

Finally, given the rapidly evolving technology landscape in AI and ASR, some recommendations may become outdated by the time of implementation. This includes changes in smart speaker capabilities, privacy policies, and integration with large language models (eg, AI conversational agents).

Conclusions

This study highlights the value of co-designing smart speaker interventions with PwPD, carers, SaLTs, third sector representatives and technology and design experts to address challenges in using VAT for speech therapy. Using a participatory Design Thinking approach, user-centered solutions were generated to improve the accessibility, usability, and therapeutic potential of smart speakers.

Two prototypes were developed: (1) education and guidance for PwPD and SaLTs, and (2) speech therapy–specific smart speaker features.

The outputs balance commercial technology with clinical needs, focusing on privacy, troubleshooting, and feedback for home use, while reinforcing co-design as a powerful method for developing digital health tools. Co-design also ensured that interventions reflected lived experience and clinical insight, enhancing the likelihood of adoption and sustained use. This research strengthens the evidence for co-design in SLT and supports smart speakers as tools to enhance therapy access, promote self-management, and reduce pressure on SLT services. Future work should develop and evaluate these prototypes to assess their real-world impact and scalability.

Acknowledgments

We thank Rory Bradley for his contributions to the co-design workshops in this research. The authors would also like to thank Donna McGuckin, Emmet Leyden, and Dr Marc Parker for their contribution to this research as members of the patient and public involvement (PPI) group.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results condensed from workshop A.
[DOCX File , 22 KB - rehab_v13i1e84364_app1.docx]

Multimedia Appendix 2
Ranking to prototype creation.
[DOCX File , 28 KB - rehab_v13i1e84364_app2.docx]

Multimedia Appendix 3
Facilitator guidance.
[DOCX File , 21 KB - rehab_v13i1e84364_app3.docx]

Multimedia Appendix 4
Ranking results.
[DOCX File , 21 KB - rehab_v13i1e84364_app4.docx]

Multimedia Appendix 5
Prototype 1: education and guidance on the therapeutic use of smart speakers.
[DOCX File , 17 KB - rehab_v13i1e84364_app5.docx]

Multimedia Appendix 6
Prototype 2: developing new speech therapy–specific features for smart speakers.
[DOCX File , 16 KB - rehab_v13i1e84364_app6.docx]

References

1. The Infinite Dial UK. Edison Research. 2025. URL: https://www.edisonresearch.com/wp-content/uploads/2025/05/Infinite-Dial-UK-2025-webinar-presentation-for-download.pdf [accessed 2025-09-01]
2. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput-Hum Interact 2020 Sep 16;27(4):1-27. [doi: 10.1145/3373759]

XSL•FO
RenderX

3. Cave R, Bloch S. The use of speech recognition technology by people living with amyotrophic lateral sclerosis: a scoping review. Disabil Rehabil Assist Technol 2023 Oct;18(7):1043-1055 [FREE Full text] [doi: 10.1080/17483107.2021.1974961] [Medline: 34511007]
4. Duffy O, Synnott J, McNaney R, Brito Zambrano P, Kernohan WG. Attitudes toward the use of voice-assisted technologies among people with Parkinson disease: findings from a web-based survey. JMIR Rehabil Assist Technol 2021 Mar 11;8(1):e23006 [FREE Full text] [doi: 10.2196/23006] [Medline: 33704072]
5. Smith E, Sumner P, Hedge C, Powell G. Smart speaker devices can improve speech intelligibility in adults with intellectual disability. Int J Lang Commun Disord 2021 May;56(3):583-593. [doi: 10.1111/1460-6984.12615] [Medline: 33772998]
6. Bleakley A, Rough D, Roper A, Lindsay S, Porcheron M, Lee M, et al. Exploring smart speaker user experience for people who stammer. 2022 Presented at: ASSETS '22: The 24th International ACM SIGACCESS Conference on Computers and Accessibility; October 23-26, 2022; Athens, Greece p. 23-26. [doi: 10.1145/3517428.3544823]
7. Miller N. Communication changes in Parkinson's disease. Pract Neurol 2017 Aug;17(4):266-274. [doi: 10.1136/practneurol-2017-001635] [Medline: 28687681]
8. Miller N, Deane KHO, Jones D, Noble E, Gibb C. National survey of speech and language therapy provision for people with Parkinson's disease in the United Kingdom: therapists' practices. Int J Lang Commun Disord 2011;46(2):189-201. [doi: 10.3109/13682822.2010.484849] [Medline: 21401817]
9. Mills J, Duffy O, Pedlow K, Kernohan G. Exploring speech and language therapists' perspectives of voice-assisted technology as a tool for dysarthria: qualitative study. JMIR Rehabil Assist Technol 2025 Sep 02;12:e75044 [FREE Full text] [doi: 10.2196/75044] [Medline: 40896823]
10. Kulkarni P, Duffy O, Synnott J, Kernohan WG, McNaney R. Speech and language practitioners' experiences of commercially available voice-assisted technology: web-based survey study. JMIR Rehabil Assist Technol 2022 Jan 05;9(1):e29249 [FREE Full text] [doi: 10.2196/29249] [Medline: 34989694]
11. Mills J, Duffy O, Pedlow K, Kernohan G. Exploring the perceptions of voice-assisted technology as a tool for speech and voice difficulties: focus group study among people with Parkinson disease and their carers. JMIR Rehabil Assist Technol 2025 Jul 16;12:e75316 [FREE Full text] [doi: 10.2196/75316] [Medline: 40669073]
12. Ramig L, Halpern A, Spielman J, Fox C, Freeman K. Speech treatment in Parkinson's disease: randomized controlled trial (RCT). Mov Disord 2018 Nov;33(11):1777-1791 [FREE Full text] [doi: 10.1002/mds.27460] [Medline: 30264896]
13. Altman M, Huang TT, Breland JY. Design thinking in health care. Prev Chronic Dis 2018 Sep 27;15:E117 [FREE Full text] [doi: 10.5888/pcd15.180128] [Medline: 30264690]
14. Sumner J, Chong LS, Bundele A, Wei Lim Y. Co-designing technology for aging in place: a systematic review. Gerontologist 2021 Sep 13;61(7):e395-e409 [FREE Full text] [doi: 10.1093/geront/gnaa064] [Medline: 32506136]
15. Brown T. Design and thinking. Harvard Educational Review 2013 Apr 01;86(1):84-92. [doi: 10.17763/1943-5045-83.1.243]
16. Vargas C, Whelan J, Brimblecombe J, Allender S. Co-creation, co-design, co-production for public health - a perspective on definition and distinctions. Public Health Res Pract 2022 Jun 15;32(2):3222211 [FREE Full text] [doi: 10.17061/phrp3222211] [Medline: 35702744]
17. Benjamin-Thomas TE, Corrado AM, McGrath C, Rudman DL, Hand C. Working towards the promise of participatory action research: learning from ageing research exemplars. International Journal of Qualitative Methods 2018 Dec 26;17(1):e-1609406918817953. [doi: 10.1177/1609406918817953]
18. Roper A, Skeat J. Innovation through participatory design: collaborative qualitative methods in the development of speech-language pathology technology. Int J Speech Lang Pathol 2022 Oct;24(5):527-532. [doi: 10.1080/17549507.2022.2050943] [Medline: 35506478]
19. Brassel S, Brunner M, Power E, Campbell A, Togher L. Speech-language pathologists' views of using virtual reality for managing cognitive-communication disorders following traumatic brain injury. Am J Speech Lang Pathol 2023 Mar 23;32(2S):907-923. [doi: 10.1044/2022_ajslp-22-00077]
20. Wilson S, Roper A, Marshall J, Galliers J, Devane N, Booth T, et al. Codesign for people with aphasia through tangible design languages. CoDesign 2015 Jan 09;11(1):21-34. [doi: 10.1080/15710882.2014.997744]
21. Cruice M, Aujla S, Bannister J, Botting N, Boyle M, Charles N, et al. Creating a novel approach to discourse treatment through coproduction with people with aphasia and speech and language therapists. Aphasiology 2021 Jul 03;36(10):1159-1181. [doi: 10.1080/02687038.2021.1942775]
22. Shiggins C, Coe D, Gilbert L, Research Collaboration A, Mares K. Development of an “aphasia-accessible participant in research experience survey” through co-production. Aphasiology 2022 Jan 31;39(12):1659-1692. [doi: 10.1080/02687038.2021.1996532]
23. Arnstein S. A Ladder Of Citizen Participation. Journal of the American Institute of Planners 1969 Jul;35(4):216-224 [FREE Full text] [doi: 10.1080/01944366908977225]
24. Elwyn G, Nelson E, Hager A, Price A. Coproduction: when users define quality. BMJ Qual Saf 2020 Sep;29(9):711-716 [FREE Full text] [doi: 10.1136/bmjqs-2019-009830] [Medline: 31488570]
25. Penny J, Slay J. Commissioning for outcomes and co-production: a practical guide for local authorities. New Economics Foundation. London, UK: New Economics Foundation; 2014. URL: https://neweconomics.org/2014/06/commissioning-outcomes-co-production [accessed 2026-01-25]

XSL•FO
RenderX

26. Connery A, Salsberg J. Exploring participatory health research and its application to speech and language therapy research practices. Int J Lang Commun Disord 2024;59(4):1257-1268. [doi: 10.1111/1460-6984.12994] [Medline: 38130139]
27. McMenamin R, Griffin M, Grzybowska B, Pound C. Working together: experiences of people with aphasia as co-researchers in participatory health research studies. Aphasiology 2021 Jun 25;39(12):1601-1622. [doi: 10.1080/02687038.2021.1923948]
28. Cook T, Boote J, Buckley N, Vougioukalou S, Wright M. Accessing participatory research impact and legacy: developing the evidence base for participatory approaches in health research. Educational Action Research 2017 Jul 31;25(4):473-488. [doi: 10.1080/09650792.2017.1326964]
29. Roper A, Wilson S, Neate T, Marshall J. Speech and language. In: Yesilada Y, Harper S, editors. A Foundation for Research. London, UK: Springer; 2019:121-131.
30. McNaney R, Tsekleves E, Synnott J. Future opportunities for IoT to support people with Parkinson's. 2020 Presented at: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; April 25-30, 2020; Honolulu, HI p. 1-15 URL: https://dl.acm.org/doi/10.1145/3313831.3376871 [doi: 10.1145/3313831.3376871]
31. Mc Menamin R, Isaksen J, Manning M, Tierney E. Distinctions and blurred boundaries between qualitative approaches and public and patient involvement (PPI) in research. Int J Speech Lang Pathol 2022 Oct 28;24(5):515-526 [FREE Full text] [doi: 10.1080/17549507.2022.2075465] [Medline: 35762365]
32. Clay P, Walton T, Malin E, Hutchinson M, Levitt K, Williams C, et al. Better conversations with Parkinson's: co-production of a novel speech and language therapy intervention with people living with Parkinson's. Research for All 2024;8(1):1-17. [doi: 10.14324/RFA.08.1.07]
33. Beckman SL. To frame or reframe: where might design thinking research go next? California Management Review 2020 Mar 02;62(2):144-162. [doi: 10.1177/0008125620906620]
34. Mah H, Dobson R, Thomson A. The importance of lived experience: a scoping review on the value of patient and public involvement in health research. Health Expect 2025 Apr 26;28(2):e70205-e70205 [FREE Full text] [doi: 10.1111/hex.70205] [Medline: 40135803]
35. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. International Journal of Qualitative Methods 2019 Sep 11;18(1):1-8. [doi: 10.1177/1609406919874596]
36. Larkman CS, Lanyon L, Rose ML. Co-designing solutions to the challenges speech pathologists and interpreters encounter when they collaborate to deliver aphasia therapy. Disabil Rehabil 2025 Dec;47(24):6468-6482 [FREE Full text] [doi: 10.1080/09638288.2025.2486459] [Medline: 40188384]
37. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
38. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
39. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
40. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
41. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
42. National Institute for Health and Care Research (NIHR). Briefing notes for researchers in public involvement in NHS, health and social care research. NIHR. 2021. URL: https://www.nihr.ac.uk/documents/briefing-notes-for [accessed 2025-05-26]
43. Jayes M, Moulam L, Meredith S, Whittle H, Lynch Y, Goldbart J, et al. Making public involvement in research more inclusive of people with complex speech and motor disorders: the I-ASC project. Qual Health Res 2021 Jun 28;31(7):1260-1274 [FREE Full text] [doi: 10.1177/1049732321994791] [Medline: 33645331]
44. Bird M, Ouellette C, Whitmore C, Li L, Nair K, McGillion MH, et al. Preparing for patient partnership: a scoping review of patient partner engagement and evaluation in research. Health Expect 2020 Jun 10;23(3):523-539 [FREE Full text] [doi: 10.1111/hex.13040] [Medline: 32157777]
45. UK standards for public involvement. UK Public Involvement Standards Development Partnership/National Institute for Health and Care Research (NIHR). URL: https://sites.google.com/nihr.ac.uk/pi-standards/standards [accessed 2025-05-26]
46. Gillivan-Murphy P, Miller N, Carding P. Voice treatment in Parkinson's disease: patient perspectives. JPRLS 2019 Jul;Volume 9(1):29-42. [doi: 10.2147/jprls.s180183]
47. Carman KL, Dardess P, Maurer M, Sofaer S, Adams K, Bechtel C, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. Health Aff (Millwood) 2013 Feb;32(2):223-231. [doi: 10.1377/hlthaff.2012.1133] [Medline: 23381514]

XSL•FO
RenderX

48. Singer I, Klatte IS, de Vries R, van der Lugt R, Gerrits E. Using co-design to develop a tool for shared goal-setting with parents in speech and language therapy. Int J Lang Commun Disord 2022 Nov 20;57(6):1281-1303 [FREE Full text] [doi: 10.1111/1460-6984.12753] [Medline: 35859264]
49. Burton A, Ogden M, Cooper C. Planning and enabling meaningful patient and public involvement in dementia research. Curr Opin Psychiatry 2019 Nov;32(6):557-562. [doi: 10.1097/YCO.0000000000000548] [Medline: 31306247]
50. Swinburn K. Working alongside people with stroke and aphasia. In: Volkmer A, Broomfield K, editors. Seldom Heard Voices in Service User Involvement: The How and Why of Meaningful Collaboration. Havant, UK: J&R Press; 2022:129-150.
51. McCloud R, Perez C, Bekalu MA, Viswanath K. Using smart speaker technology for health and well-being in an older adult population: pre-post feasibility study. JMIR Aging 2022 May 09;5(2):e33498 [FREE Full text] [doi: 10.2196/33498] [Medline: 35532979]
52. Masters K, Correia R, Nemethy K, Benjamin J, Carver T, MacNeill H. Online learning in health professions education. Part 2: Tools and practical application: AMEE Guide No. 163. Medical Teacher 2023 Sep 23;46(1):18-33. [doi: 10.1080/0142159x.2023.2259069]
53. Kim S, Choudhury A. Exploring older adults' perception and use of smart speaker-based voice assistants: a longitudinal study. Computers in Human Behavior 2021 Nov 18;124(3):106914-106924. [doi: 10.1016/j.chb.2021.106914]
54. Chang F, Sheng L, Gu Z. Investigating the integration and the long-term use of smart speakers in older adults' daily practices: qualitative study. JMIR Mhealth Uhealth 2024 Feb 12;12:e47472 [FREE Full text] [doi: 10.2196/47472] [Medline: 38345844]
55. Arnold A, Kolody S, Comeau A, Miguel Cruz A. What does the literature say about the use of personal voice assistants in older adults? A scoping review. Disabil Rehabil Assist Technol 2024 Jan;19(1):100-111. [doi: 10.1080/17483107.2022.2065369] [Medline: 35459429]
56. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019 Apr 08;21(4):e11604 [FREE Full text] [doi: 10.2196/11604] [Medline: 30958272]
57. Tazi F, Nandakumar A, Dykstra J, Rajivan P, Das S. Privacy, security, and usability tradeoffs of telehealth from practitioners' perspectives. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2023 Oct 25;67(1):1862-1867. [doi: 10.1177/21695067231199689]
58. Vaezipour A, Aldridge D, Koenig S, Theodoros D, Russell T. A mixed-method investigation of clinician acceptance, barriers and enablers of virtual reality technology in communication rehabilitation. Disabil Rehabil 2022 Jul 14;44(15):3946-3958. [doi: 10.1080/09638288.2021.1895333] [Medline: 33715566]
59. Mills J, Duffy O. Speech and language therapists' perspectives of virtual reality as a clinical tool for autism: cross-sectional survey. JMIR Rehabil Assist Technol 2025 Feb 27;12:e63235-e63235 [FREE Full text] [doi: 10.2196/63235] [Medline: 40014826]
60. Maccario G, Naldi M. Alexa, is my data safe? The (ir)relevance of privacy in smart speakers reviews. International Journal of Human–Computer Interaction 2022 Apr 20;39(6):1244-1256. [doi: 10.1080/10447318.2022.2058780]
61. Lustgarten SD, Garrison YL, Sinnard MT, Flynn AW. Digital privacy in mental healthcare: current issues and recommendations for technology use. Curr Opin Psychol 2020 Dec;36:25-31 [FREE Full text] [doi: 10.1016/j.copsyc.2020.03.012] [Medline: 32361651]
62. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220 [FREE Full text] [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
63. Quinn K, Leiser Ransom S, O'Connell C, Muramatsu N, Marquez DX, Chin J. Assessing the feasibility and acceptability of smart speakers in behavioral intervention research with older adults: mixed methods study. J Med Internet Res 2024 Aug 30;26:e54800 [FREE Full text] [doi: 10.2196/54800] [Medline: 39213034]
64. Zhang R, Li H, Chen A, Liu Z, Lee Y. AI privacy in context: a comparative study of public and institutional discourse on conversational AI privacy in the US and Chinese social media. Social Media + Society 2024 Oct 28;10(4):1-19. [doi: 10.1177/20563051241290845]
65. Saripalle R, Patel R. From command to care: a scoping review on utilization of smart speakers by patients and providers. Mayo Clin Proc Digit Health 2024 Jun;2(2):207-220 [FREE Full text] [doi: 10.1016/j.mcpdig.2024.03.002] [Medline: 40207172]
66. Mitzner TL, Fausset CB, Boron JB, Adams AE, Dijkstra K, Lee CC, et al. Older adults' training preferences for learning to use technology. Proc Hum Factors Ergon Soc Annu Meet 2008 Sep;52(26):2047-2051 [FREE Full text] [doi: 10.1177/154193120805202603] [Medline: 25309139]
67. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. Front Public Health 2024;12:1447747 [FREE Full text] [doi: 10.3389/fpubh.2024.1447747] [Medline: 39555039]
68. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251.e1 [FREE Full text] [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]

XSL•FO
RenderX

69. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
70. Cassano F, Pagano A, Piccinno A. Supporting speech therapies at (smart) home through voice assistance. 2021 Presented at: International Symposium on Ambient Intelligence; October 6-8, 2021; Salamanca, Spain. [doi: 10.1007/978-3-031-06894-2_10]
71. Makin L, Smith E, Hedge C, Sumner P, Powell G. Smart speakers are an acceptable and feasible speech practice tool for children with speech difficulties. Disabil Rehabil Assist Technol 2025 Oct 17;20(7):2510-2521 [FREE Full text] [doi: 10.1080/17483107.2025.2491636] [Medline: 40244154]
72. McNaney R, Vines J, Roggen D, Balaam M, Zhang P, Poliakov I, et al. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Apr 26 Presented at: CHI '14: The SIGCHI Conference on Human Factors in Computing Systems; May 1, 2014; Toronto, Canada p. 2551-2554. [doi: 10.1145/2556288.2557092]
73. Jansons P, Fyfe J, Via JD, Daly RM, Gvozdenko E, Scott D. Barriers and enablers for older adults participating in a home-based pragmatic exercise program delivered and monitored by Amazon Alexa: a qualitative study. BMC Geriatr 2022 Mar 25;22(1):248-258 [FREE Full text] [doi: 10.1186/s12877-022-02963-2] [Medline: 35337284]
74. Carlin A, Logue C, Flynn J, Murphy MH, Gallagher AM. Development and feasibility of a family-based health behavior intervention using intelligent personal assistants: randomized controlled trial. JMIR Form Res 2021 Jan 28;5(1):e17501 [FREE Full text] [doi: 10.2196/17501] [Medline: 33507155]
75. Greuter S, Balandin S, Watson J. Social games are fun: exploring social interactions on smart speaker platforms for people with disabilities. In: CHI PLAY '19 Extended Abstracts: Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2019 Presented at: CHI PLAY '19: The Annual Symposium on Computer-Human Interaction in Play; October 22-25, 2019; Barcelona, Spain p. 429-435. [doi: 10.1145/3341215.3356308]
76. Esquivel P, Gill K, Chung C, Ding D, Faieta J. Smart speakers and skill use: what do we know? Disabil Rehabil Assist Technol 2025 Feb 08;20(2):488-497. [doi: 10.1080/17483107.2024.2387801] [Medline: 39115270]
77. Cave R. How people living with amyotrophic lateral sclerosis use personalized automatic speech recognition technology to support communication. J Speech Lang Hear Res 2024 Nov 07;67(11):4186-4202. [doi: 10.1044/2024_jslhr-24-00097]
78. Mengesha Z, Heldreth C, Lahav M, Sublewski J, Tuennerman E. "I don't think these devices are very culturally sensitive."—Impact of automated speech recognition errors on African Americans. Front Artif Intell 2021;4:725911 [FREE Full text] [doi: 10.3389/frai.2021.725911] [Medline: 34901836]
79. Green J, MacDonald RL, Jiang PP, Cattiau J, Heywood R, Cave R. Automatic speech recognition of disordered speech: personalized models outperforming human listeners on short phrases. 2021 Presented at: The 22nd Annual Conference of the International Speech Communication Association; August 30 to September 3, 2021; Brno, Czechia p. 4778-4782 URL: https://www.isca-archive.org/interspeech_2021/green21_interspeech.pdf [doi: 10.21437/interspeech.2021-1384]
80. Tobin J, Nelson P, MacDonald B, Heywood R, Cave R, Seaver K, et al. Automatic speech recognition of conversational speech in individuals with disordered speech. J Speech Lang Hear Res 2024 Nov 07;67(11):4176-4185. [doi: 10.1044/2024_jslhr-24-00045]
81. Improving access to Alexa for people with aphasia. City of St George's, University of London. 2025. URL: https://tinyurl.com/45fcp8k6 [accessed 2025-09-10]

Abbreviations

AI: artificial intelligence
ASR: automatic speech recognition
GDPR: General Data Protection Regulation
LSVT: Lee Silverman Voice Treatment
PPI: patient and public involvement
PwPD: people with Parkinson’s disease
SaLT: speech and language therapist
SLT: speech and language therapy
VAT: voice-assisted technology
VR: virtual reality

XSL•FO
RenderX

Edited by S Munce; submitted 22.Sep.2025; peer-reviewed by D Singhal, Y Lin; comments to author 04.Nov.2025; revised version received 18.Dec.2025; accepted 13.Jan.2026; published 04.Feb.2026.

Please cite as:
Mills J, Kernohan G, Pedlow K, Duffy O
Voice-Assisted Technology for People With Parkinson's Disease Experiencing Speech and Voice Difficulties: Co-Designing Solutions Using Design Thinking
JMIR Rehabil Assist Technol 2026;13:e84364
URL: https://rehab.jmir.org/2026/1/e84364
doi:10.2196/84364
PMID:

©Jodie Mills, George Kernohan, Katy Pedlow, Orla Duffy. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 04.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Delivering Virtual Group Interventions for Youth With Persisting Symptoms After Concussion (Move&Connect-Youth) and Their Caregivers (Move&Connect-Caregivers): Fidelity and Engagement Evaluation

Andrea Hickling[1,2], MScOT, OT Reg. (Ont.); Hiba Al-Hakeem[1,3], MA; Andrew Lovell[1], MSc; Sophie Madge-Nyren[1], HBSc; Kylie D Mallory[1,2], PhD, OT Reg. (Ont.); Jennifer L Ryan[1], PhD, PT; Christine F Provvidenza[1], MSc; Ruth Wilcock[4], RP; Patricia Knapp[1]; Brenda Knapp[1], BScN; Shari L Wade[5,6], PhD; Keith Owen Yeates[7], PhD; Shannon E Scratch[1,8,9], C.Psych, PhD

[1]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, 150 Kilgour Rd, Toronto, ON, Canada

[2]Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada

[3]Department of Psychology, University of Windsor, Windsor, ON, Canada

[4]Ontario Brain Injury Association, Thorold, ON, Canada

[5]Division of Rehabilitation Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[6]Department of Pediatrics, University of Cincinnati, Cincinnati, OH, United States

[7]Department of Psychology, Alberta Children's Hospital Research Institute, University of Calgary, Calgary, AB, Canada

[8]Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[9]Department of Pediatrics, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Shannon E Scratch, C.Psych, PhD
Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, 150 Kilgour Rd, Toronto, ON, Canada

Abstract

Background: Youth with persisting symptoms after concussion (PSaC) can experience ongoing physical, cognitive, and emotional challenges, and their caregivers experience increased stress. There are currently limited interventions that support youth with PSaC or their caregivers. *Move&Connect* is a virtual group program with 2 parallel interventions: *Move&Connect-Youth* (*M&C-Y*) and *Move&Connect-Caregivers* (*M&C-C*). *M&C-Y* includes concussion education, active rehabilitation, and goal setting. *M&C-C* uses psychoeducation and practical tools to promote family communication and problem solving.

Objective: Within the context of a broader pilot study of *M&C-Y* and *M&C-C*, we completed a fidelity and engagement evaluation with the objectives of (1) evaluating intervention fidelity and (2) exploring facilitators and barriers to youth and caregiver engagement.

Methods: *M&C-Y* and *M&C-C* were delivered virtually using Zoom Healthcare and ran in parallel. Youth with PSaC and caregivers were recruited from a children's rehabilitation hospital in Ontario, Canada. Intervention fidelity was evaluated according to the 5 principles of fidelity from Bellg et al: design, training, delivery, receipt, and enactment. Intervention delivery, attendance, and engagement ratings were monitored through weekly logs and field notes. Participants completed exit satisfaction surveys, and a subset of youth and caregivers completed semistructured interviews to understand facilitators and barriers to engagement. Descriptive statistics were used to analyze logs and surveys. Directed content analysis, guided by a motivational framework, and inductive coding to identify barriers and facilitators to engagement were applied to qualitative data.

Results: A total of 34 youth (age: mean 15.1, SD 2.28 years) and 36 caregivers (age: mean 48.9, SD 5.92 years) participated in 10 *M&C-Y* and 9 *M&C-C* groups, respectively. Most youth identified as girls (24/34, 71%), and most caregivers identified as mothers (30/36, 83%). Clinician facilitators adhered to the components of the intervention protocols between 92% and 97% of the time in *M&C-Y* and 82% and 93% of the time in *M&C-C*. The majority of youth (22/34, 65%) and caregivers (24/36, 67%) attended 5 or more sessions. In addition, 82% (28/34) of youth and 94% (34/36) of caregivers agreed they would recommend the interventions to others. Qualitative findings revealed youth engagement was supported by motivation, comfort in the virtual setting, interactive intervention components, feelings of trust, and diverse and modifiable exercises. Barriers to youth engagement included ongoing symptoms, discrepancies between exercises and physical abilities, and previous knowledge of education. Caregiver engagement was facilitated by perceived benefits of *M&C-C*, emotional connection and validation, and comfort and

XSL•FO
RenderX

convenience of the virtual setting. Barriers to caregiver engagement included stress, competing scheduling demands, and limited reminders or access to intervention materials.

Conclusions: The pilot study of *M&C-Y* and *M&C-C* was conducted with high delivery adherence. Exit satisfaction surveys combined with qualitative data revealed high levels of caregiver satisfaction and engagement, whereas youth satisfaction and engagement had some variability, particularly related to the exercise. Findings demonstrate important implications for delivering virtual group interventions for youth with PSaC and their caregivers.

Trial Registration: ClinicalTrials.gov NCT05631301; https://clinicaltrials.gov/study/NCT05631301

(JMIR Rehabil Assist Technol 2026;13:e88771) doi:10.2196/88771

KEYWORDS

concussion; brain injury; telehealth; telemedicine; virtual care; rehabilitation; pediatrics; parents; group intervention; qualitative

Introduction

Background

After a concussion, youth can experience an array of physical, cognitive, and emotional and behavioral symptoms [1]. Although most youth recover relatively quickly from concussion, up to 30% continue to experience persisting symptoms after concussion (PSaC) for months or even years [2,3]. PSaC can negatively impact youths' quality of life [4,5] and lead to ongoing mental health challenges [6,7]. Youth with PSaC also experience feelings of isolation and a lack of social support and understanding from peers [8].

An interdisciplinary approach is recommended to support youth with PSaC [9,10], and active rehabilitation (ie, subthreshold aerobic exercise, coordination exercises) [11] and psychoeducation are considered standard best practices in the treatment of PSaC [12]. Active rehabilitation has been shown to reduce symptoms and improve mental health outcomes in youth with PSaC [13-16]. Psychoeducational intervention about recovery expectations and treatment recommendations is important for youth with PSaC and their caregivers [17,18]. In addition to these approaches, emerging research has explored the feasibility and impact of providing concussion care in a virtual model [19-21] and virtual active rehabilitation to youth with PSaC [22-25]. These virtual approaches are safe and tolerable for youth with PSaC and provide comfort and convenience for families [19,21-23,26].

Caregivers play a critical role in managing their child's PSaC and describe burden [27] and elevated stress regarding their child's recovery [28]. Caregiver and family functioning may have an important bidirectional association with PSaC after pediatric concussion [29]. Providing caregivers with psychoeducation and therapeutic support can help to navigate challenges after concussion [30]. Caregivers report a need for clear and reliable information about concussion recovery [31], and caregiver education on concussion signs and symptoms has been linked to reduced recurrence rates [32]. Despite this, there is a dearth of interventions tailored to caregivers of youth with PSaC and a clear need to develop support that targets caregivers' well-being during concussion recovery [33,34].

Move&Connect

The *Move&Connect* program was designed to support youth with PSaC and their caregivers and consists of 2 parallel intervention arms: *Move&Connect-Youth* (*M&C-Y*: for youth with PSaC) and *Move&Connect-Caregivers* (*M&C-C*: for caregivers of youth with PSaC). *M&C-Y* and *M&C-C* are 6-week virtual group interventions with 1-hour sessions that were developed using the intervention mapping framework [35] and codesigned with youth, caregivers, and clinicians. The detailed design and development processes of these interventions are described elsewhere [36,37].

Delivering rehabilitation interventions such as *M&C-Y* and *M&C-C* via virtual platforms (eg, Zoom Healthcare) can enhance access to services, support participation, and promote continuity of care [38,39]. The virtual group model in *Move&Connect* is particularly appropriate for youth with PSaC and their caregivers due to geographic inequities accessing specialized concussion care [40-42], family burden navigating appointments [43,44], and engagement of individuals who may become symptomatic with travel or in-person appointments [42]. The *Move&Connect* program differs from existing virtual concussion and group rehabilitation programs [25,45-48] due to the parallel youth and caregiver intervention design, theory driven development, and interdisciplinary delivery.

M&C-Y is guided by the biopsychosocial model [49] and includes psychoeducation, active rehabilitation, and goal-setting. Each week, the intervention begins with an ice breaker followed by an education or discussion topic (eg, headache management, managing stress, advocacy). Youth then complete exercises (ie, cardiovascular, balance, coordination, strength) with the goal of improving their confidence and tolerance to different physical demands. All exercises are designed to be completed at home with no equipment required. In-session music during exercise and weekly polls are used to engage participants. At the end of the session, youth set a weekly goal, often related to the psychoeducation topic or exercise, which is reviewed and discussed in a solution-focused manner [50] the following week [36].

M&C-C is guided by the family-directed approach to brain injury (FAB) model [51] and includes psychoeducation and active skill development to promote family communication and problem-solving. Psychoeducation, discussion, and activities are completed as a group and are related to weekly themes: the ripple effect, school advocacy, considering your child's well-being, family and you, stress and daily challenges, and parenting is hard. Weekly take-home reflection activities are

XSL•FO
RenderX

introduced at the end of each session, and caregivers share their insights during the group discussion the following week [37].

Initial testing of *M&C-Y* [36] (n=10) and *M&C-C* [37] (n=9) revealed that the interventions were feasible based on the key domain areas (recruitment, acceptability, and evaluation of participant responses) from Orsmond and Cohn [52] and our a priori criteria. Youth indicated that *M&C-Y* was enjoyable and they were able to form meaningful peer connections [36]. Caregivers indicated that *M&C-C* was a safe space to share personal experiences and they valued receiving advice from facilitators and their peers [37].

Intervention Fidelity and Engagement

Following establishment of feasibility, it is important to consider intervention fidelity, which is the ongoing assessment, monitoring, and enactment of the reliability and internal validity of a study [53]. Intervention or treatment fidelity (herein referred to as intervention fidelity) is a critical part of intervention research as it increases confidence that changes in the study outcome of interest (dependent variable) are due to the intervention (independent variable) [54]. A rigorous fidelity plan should include study design, training, and quality and consistency of delivery [54,55]. Importantly, recent studies have emphasized moving beyond only assessing intervention delivery to including a greater focus on the understanding of participants' receipt and enactment (collectively called engagement) of intervention activities [53,56,57]. Understanding intervention engagement is especially important as it is multifaceted and may be influenced by numerous contextual factors [57] (eg, virtual delivery, group setting, multiple intervention components). Engagement has previously been described by King and colleagues [58] as the affective (expecting positive outcomes), cognitive (believing that therapy will be effective), and behavioral (actively participating in therapy tasks) involvement or investment in a therapy session [58]. Using engagement-promoting strategies is recommended to improve adherence to interventions [59-62]. Additionally, family members being actively engaged in their child's rehabilitation journey is associated with better outcomes [63]. Thus, exploring intervention fidelity and participant engagement are vital steps in delivering an intervention study.

Objectives

Based on the aforementioned promising feasibility findings, we conducted a pilot study of *M&C-Y* and *M&C-C*. Within the context of the broader pilot study, the goal of this paper was to (1) evaluate intervention fidelity and (2) explore facilitators and barriers to youth and caregiver engagement.

Methods

Study Design

This paper describes a multimethod fidelity and engagement evaluation of *M&C-Y* and *M&C-C*. Between November 2022 and June 2025, 10 *M&C-Y* and 9 *M&C-C* groups were delivered. Group interventions were delivered virtually using Zoom Healthcare. Two trained clinician facilitators delivered each *M&C-Y* (occupational therapists and/or physiotherapists) and *M&C-C* (neuropsychologist, neuropsychology trainee, and/or social worker) session. A research staff observer attended all *M&C-Y* and *M&C-C* sessions to record observations and support any technological challenges. Youth were allocated to groups by age/grade, when possible (ie, elementary and high school), to account for developmental differences. Caregivers were also allocated to groups for those with elementary or high school–aged children. Participants were provided with weekly reminder emails of upcoming sessions including any supplies and weekly handouts or materials in advance of each session.

To ensure safety and personalization during *M&C-Y*, modifications to make each exercise more or less challenging were provided, breaks were encouraged, and clinicians monitored participant symptom changes. Consistent with concussion best practice guidelines [12], if youth experienced >2-point change in their symptoms on a symptom severity scale ranging from 0 to 10, they were instructed to stop the activity until their symptoms had returned to their baseline level. Virtual breakout rooms were utilized for 1-on-1 conversations regarding symptom monitoring. Reminders about participant privacy and group confidentiality were provided at the beginning of each session. The study's principal investigator randomly audited intervention sessions. No adverse events occurred during the study.

Youth and caregiver participants also completed outcome measures before (T_0), immediately following (T_1), and 3 months following (T_2) intervention completion. Clinical outcome results will be described in subsequent publications, as the primary contribution of this paper is the evaluation of intervention fidelity and engagement. Intervention fidelity evaluation involved data collected during intervention delivery and at the T_1 follow-up. Engagement evaluation involved optional youth and caregiver exit interviews completed at T_1. The qualitative components of the study follow guidelines outlined in the Consolidated Criteria for Reporting Qualitative Studies (COREQ) checklist [64] (see Checklist 1).

Recruitment and Eligibility Criteria

Youth were included if they were between the ages of 8 years and 21 years, had a diagnosis of concussion, and had been experiencing PSaC for >4 weeks at intervention onset. Youth were excluded if they had a physical condition that would impact participation (eg, requiring a mobility device), a disorder of social cognition or communication (eg, intellectual disability, autism spectrum disorder), a functional neurological disorder (eg, conversion disorder, somatization), or an acute psychiatric condition that resulted in a recent hospital admission. Given the pilot nature of the study, youth with these conditions were excluded as they may have unique support and safety needs. Caregivers were included if they had a child (aged 8 - 21 years) who was diagnosed with a concussion and had been experiencing PSaC for >4 weeks at intervention onset. Youth and caregivers had to be able to read and understand English, be willing to engage in weekly sessions, and have access to a reliable internet connection. Youth and caregivers could enroll as a dyad (ie, both youth and caregiver enroll) or independently (ie, youth could enroll without caregiver; caregiver could enroll without youth). Recruitment predominantly occurred via convenience sampling through concussion services at Holland

XSL•FO
RenderX

Bloorview Kids Rehabilitation Hospital (HBKRH) [10], but community participants were also eligible.

Ethical Considerations

Ethics approval was obtained from HBKRH (REB #0513). All procedures complied with the ethical principles outlined in the Declaration of Helsinki. Youth and caregivers were informed that their decision to participate in the study would not affect the care that they received at HBKRH. A capacity assessment (ie, a formal assessment of decision-making capacity) [65] was performed for all youth participants, and all participants (youth and caregivers) provided informed written consent.

Study Measures

Intervention Fidelity Assessment

This pilot study used the framework by Bellg et al [55] for assessing the fidelity of the *M&C-Y* and *M&C-C* interventions, which describes 5 key domains to ensure optimal intervention fidelity: (1) intervention design, (2) provider training, (3) intervention delivery, (4) intervention receipt, and (5) intervention enactment. Rather than define strict a priori fidelity thresholds, we used triangulation to evaluate intervention delivery, receipt, and enactment. Multiple data sources and approaches were used for each domain to allow for triangulation and reduction of potential bias [57] (Table 1).

Table . Intervention fidelity components and assessment approaches. Table adapted from Ginsburg and colleagues [57].

Intervention fidelity component [55]	Our approach to fidelity assessment
Fidelity of design	
Intervention is defined and operationalized consistent with its underlying theory	• Intervention approach and components designed with incorporation of underlying theory; see Scratch et al [36] and Al-Hakeem et al [37] for details
Fidelity of training	
Clinician facilitator training	• Research and institutional orientation • Clinicians provided with intervention manuals and materials • Train-the-trainer approach including session shadowing and discipline specific peer debriefing • Feedback provided by trainees to inform train-the-trainer approach
Fidelity of delivery	
Adherence to protocol (components, timing) and consistent delivery	• Clinician delivery checklist completed after each session • Debriefing after each session with clinicians and research staff observer • Bi-weekly research team meetings • Clinician consultation with study principal investigator, as needed
Fidelity of receipt	
Participants understood/can use intervention components	• Participant attendance tracking • Research staff observer field notes
Fidelity of enactment	
Engagement in Move&Connect activities	• Participant engagement rating • Exit satisfaction survey • Research staff observer field notes • Semistructured interviews[a]

[a]Used in objective 2 to further explore barriers and facilitators to engagement.

Intervention Fidelity Tools

All fidelity tools, with the exception of the semistructured interviews, were captured using REDcap [66].

Clinician Delivery Checklist

To monitor adherence to the components and timing of the intervention delivery, 2 clinician facilitators completed the clinician delivery checklist weekly after each session. The *M&C-Y* clinician delivery checklist asked if the education, exercise, and goal setting components of the program were delivered as intended (3 yes/no questions). The *M&C-C* clinician delivery checklist asked if the education, activity, and take-home reflections were delivered as intended (3 yes/no questions). An open-ended text box allowed clinicians to explain any deviations from intervention delivery or additional contextual information.

Attendance Tracking Log

Youth and caregiver attendance for each session was monitored in an attendance tracking log completed by the research staff observer. The attendance tracking log contained information regarding session number (1-6), attendance (yes/no), and if a participant was late for the session (yes/no). An open-ended text box allowed for any additional contextual information related to attendance (eg, if a participant emailed that they were unable to attend).

XSL•FO
RenderX

Research Staff Observer Field Notes

A structured field note template was used by the research staff observer in all *M&C-Y* and *M&C-C* sessions to capture behavioral observations, participant goals, and any feedback on the intervention provided by participants throughout the sessions.

Participant Engagement Rating

The research staff observer completed the participant engagement rating for each participant that attended each *M&C-Y* and *M&C-C* session. The participant engagement rating was developed by the research team to provide an exploratory measure of participation and engagement and was based on insights and observations from initial feasibility testing [36]. The 1 - 3 rating asked, "how would you rate the participant's overall engagement and participation during this session?" For *M&C-Y*: 1-Fair (Participant demonstrated reduced engagement; did not share thoughts, experiences, or goals without considerable prompting; and did not engage or attempt exercises.), 2-Good (Participant was engaged for the duration of the session; shared their thoughts, experiences, and goals with prompting; and participated in the exercise activities.), or 3-Excellent (Participant was engaged during the session and shared their thoughts, experiences, and goals in detail with minimal prompting and completed and engaged with exercises.). For *M&C-C*, the same 3-point rating structure was used, with the omission of the exercise component. An average engagement score was calculated for each participant (sum of scores/number of sessions attended) and at the group level (sum of participant scores/number of participants in the group). The *M&C-Y* and *M&C-C* groups were observed by 3 different research staff members. The primary study coordinator (AL) conducted 1-on-1 observer training on the participant engagement rating, incorporating didactic instruction and case review from the initial feasibility study [36,37]. Questions or uncertainties from alternate observers were discussed with the main study coordinator (AL).

Exit Satisfaction Survey

Youth and caregivers were invited to complete an exit satisfaction survey after the interventions were complete (T_1). The survey asked about satisfaction with intervention components, the group setting, group size and timing, clinician facilitators, and future intent to implement strategies (5-point Likert scale; youth: 20 items; caregivers: 17 items). It also included 3 open-ended questions about the structure and delivery of the intervention, the best feature of the intervention, and what participants would like to see changed or improved.

Semistructured Interviews

Youth and caregivers had the opportunity to participate in an optional semistructured interview after the interventions were complete (T_1) to better understand participants' experiences with the interventions. All youth and caregivers who participated in the interventions were invited to participate in the interview through email. Toward the end of recruitment, the team focused on inviting families from underrepresented backgrounds, including fathers, boys, and participants from diverse ethnic backgrounds. Consistent with the concept of information power, sample adequacy was considered in relation to the study aims and the range of participant perspectives represented in the qualitative sample [67]. Interviews were completed virtually using Zoom Healthcare by a member of the research team who did not facilitate the intervention. Interviews were audio-recorded, transcribed verbatim, and checked for accuracy. Of note, the quotes presented in this paper were lightly edited to remove dysfluencies (eg, um) and repeated words for clarity and conciseness [68] and were labeled using data codes. See Multimedia Appendix 1 for the complete interview guides. Note that this interview guide was used in the broader pilot study and not all questions align with the scope of the research questions addressed in this manuscript. Additional qualitative findings will be presented in subsequent publications.

Data Analysis

Overview

Results of the clinician delivery checklist, attendance tracking logs, and participant engagement ratings were reported using frequencies and descriptive statistics. Exit satisfaction survey Likert-scale question responses were collapsed into 3 groups: (1) favorable response (strongly agree, agree), (2) neutral response (neither agree nor disagree), and (3) unfavorable response (strongly disagree, disagree). They were then reported descriptively. Responses to open-ended exit satisfaction survey questions were grouped into initial codes by 2 members of the research team (AH, AL). Codes were then grouped into broader categories, and the frequency of responses in each category was reported. Two researchers (AH, AL) coded observer field notes descriptively, grouped related codes into categories, and presented categories related to engagement. Any discrepancies in the coding process were resolved through discussion. Findings from checklists, logs, field notes, surveys, and interviews were triangulated by assessing consistency and identifying complementary insights (eg, did observations align with participants' reported experiences?) [57].

Interviews

Content analysis was used to examine the interview data, following both directed and inductive approaches. Pre-identified codes were derived from the multifaceted motivational framework of engagement [58]. Inductive coding was used to generate additional codes from the interview data to capture facilitators and barriers to engagement across both interventions. Facilitators were defined as factors that supported participants' ability to engage, whereas barriers reflected factors that hindered their engagement. An initial codebook informed by the multifaceted motivational framework of engagement [58] and the concepts of facilitators and barriers to engagement was developed collaboratively by 2 coders (HAH, SMN) prior to coding. The 2 coders then independently applied the codebook to the interview transcripts and met regularly to discuss coding decisions, resolve discrepancies through consensus, and iteratively refine the codebook by clarifying existing codes and incorporating additional inductively derived codes. The analytic process involved (1) identifying topics or patterns, (2) gathering examples of topics from the transcripts, and (3) examining connections and relationships among codes to refine categories. This process included organizing examples of engagement using

XSL•FO
RenderX

the predefined codebook and developing categories that reflected key facilitators and barriers. Quirkos software was used to apply codes and organize relevant excerpts [69].

Reflexivity

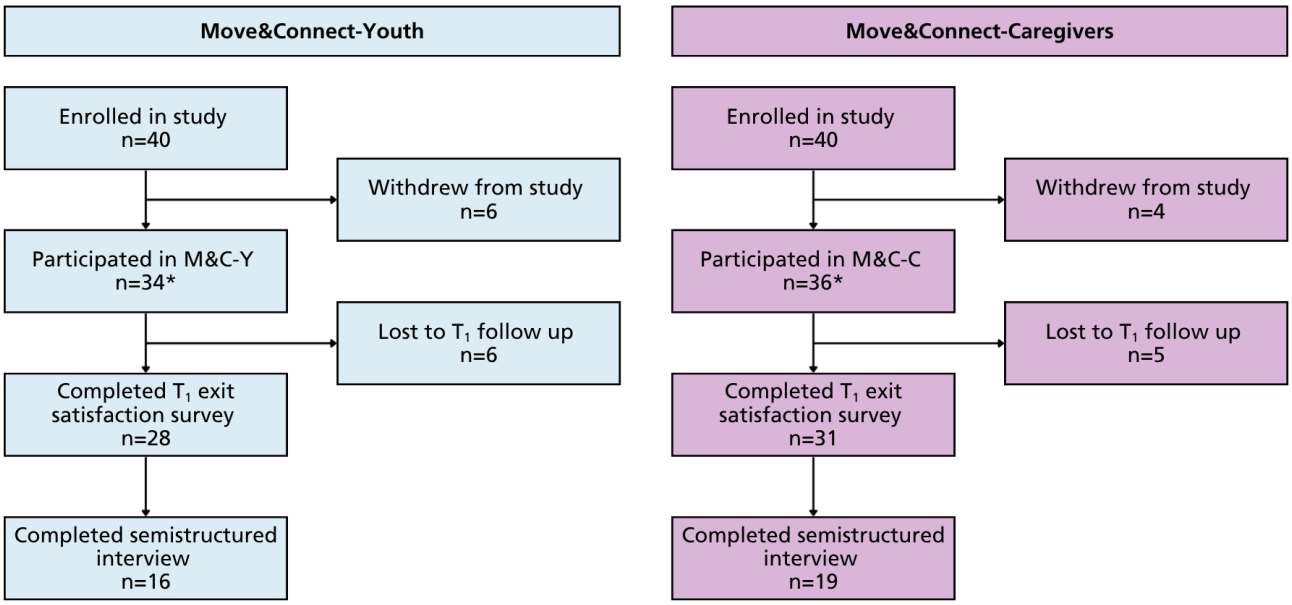
Interviews were conducted by AL (male research coordinator with a master's degree in human health and nutritional sciences), HAH (female graduate student in clinical neuropsychology), and SMN (female research assistant with an undergraduate degree in neuroscience and psychology). Some interviewers had limited prior contact with participants through onboarding procedures. Interviewers did not conduct interviews with participants from a group if they acted as a clinician facilitator for that group. The broader study team has expertise in brain injury rehabilitation and experience with the *M&C-Y* and *M&C-C* interventions, which may have influenced the interpretation of the findings. To support reflexive practice and enhance trustworthiness, data analysis was completed by 2 individuals (HAH, SMN), and analytic memos documenting coding decisions and interpretations were maintained throughout the analytic process. Interpretations were also shared with the broader research team to encourage reflection and dialogue [70].

Results

Participants

Youth and caregiver participant progression throughout the study is outlined in Figure 1.

Figure 1. Youth and caregiver participant flow throughout the study, including participant numbers for enrollment, withdrawals, intervention participants, lost to T_1 (immediately following invervention completion) follow-up, T_1 exit satisfaction surveys, and semistructured interviews. *Demographics, attendance, and engagement ratings are reported for all participants who participated in the interventions.



Youth Participants

A total of 40 youth enrolled in the study (32 as dyads with parent/caregiver participants; 8 independents): 4 youth withdrew before the intervention began, 2 youth withdrew from the intervention due to scheduling demands, and 6 youth participants attended group sessions but were lost to T_1 follow up.T_1 follow-up procedures were completed by 28 youth. At the time of enrollment, youth participants completed a demographics form (see Table 2 for youth characteristics). All youth who attended T_1 follow-up (n=28) were invited to an optional semistructured interview: 16 youth completed an interview (10 identified as girls, 5 identified as boys, and 1 identified as gender fluid), and 12 youth declined. All 6 *M&C-Y* sessions were attended by 7 of the 16 youth who were interviewed, and interviews lasted from 4 minutes to 27 minutes (mean 18 minutes). For clarity, this leaves the following sample sizes: demographics, attendance, and engagement ratings: n=34, 27 dyads, 7 independents; exit satisfaction survey data for fidelity of enactment: n=28, 21 dyads, 7 independents; and interviews: n=16, 11 dyads, 5 independents).

XSL•FO
RenderX

Table . Youth participant demographic information (n=34).

Demographic characteristics	Results
Age (years), mean (SD)	15.1 (2.28)
Age (years), range	9 - 18
Sex, n (%)	
Female	24 (71)
Male	9 (27)
Prefer not to specify	1 (3)
Gender, n (%)	
Girl/woman	24 (71)
Boy/man	9 (27)
Gender fluid	1 (3)
Ethnicity, n (%)	
Black	3 (9)
Chinese	1 (3)
South Asian	3 (9)
West Asian	1 (3)
White	21 (62)
Multiple ethnic identities	5 (15)[a]
Health history, n (%)	
Learning disability[b]	9 (27)
Attention deficit hyperactivity disorder (ADHD)	8 (24)
Headache history prior to concussion	4 (12)
Anxiety diagnosis	8 (24)
Depression diagnosis	3 (9)
Other developmental disorders[b]	1 (3)
Other psychological/psychiatric conditions	4 (12)
Education history, n (%)	
Special education	5 (15)
Gifted program[b]	5 (15)
Individual education plan (IEP)	16 (47)
Previous concussion history[b], n (%)	
0	9 (27)
1	11 (33)
2	3 (9)
3	5 (15)
4	3 (9)
≥5	2 (6)
Days since most recent concussion (at T_0), mean (SD)	326.5 (380.97)
Days since most recent concussion (at T_0), median (IQR)	198.5 (198.75)
Days since most recent concussion (at T_0), range[c]	22 - 1905
Setting/mechanism of injury, n (%)	

XSL•FO
RenderX

Demographic characteristics	Results
Sport	17 (50)
School	6 (18)
Motor vehicle collision	2 (6)
Fall	2 (6)
Assault	2 (6)
Accident (eg, from general clumsiness)	4 (12)
Other	1 (3)
Currently receiving therapy support for concussion, n (%)[b]	
Yes	18 (55)
No	15 (46)
Previous concussion support group attendance (in person or virtual), n (%)	
Yes	0 (0)
No	34 (100)
Current school involvement, n (%)	
Not going to school because of injury	3 (9)
Going to school, but in a less demanding way (eg, reduced hours/classes) compared to before injury	10 (29)
Going to school, just as I was before my injury	18 (53)
Other (eg, co-op, homeschool, graduated)	3 (9)
Job history, n (%)	
I am not working because of my injury	1 (3)
I am working, but in a less demanding way (eg, reduced hours/responsibility) compared to before my injury	1 (3)
I was not working before my injury, and I am not working now	26 (77)
I am working, just as I was before my injury	6 (18)

[a]White/Chinese, n=2; Greek/Arab/English/German, n=1; Asian/African, n=1; Japanese/Chinese/Trinidadian/White, n=1.

[b]1 missing response (n=33).

[c]1 participant was 22 days postinjury at T_0 but >4 weeks postinjury at intervention onset.

Caregiver Participants

A total of 40 caregivers enrolled in the study (32 as dyads with their child as youth participants; 8 independents): 3 caregivers withdrew before the intervention began, 1 caregiver withdrew during the intervention, and 5 caregivers attended group sessions but were lost to T_1 follow-up. T_1 follow-up procedures were completed by 31 caregivers. At the time of enrollment, caregiver participants completed a demographics form (see Table 3 for caregiver characteristics). All caregivers who attended T_1 follow-up (n=31) were invited to an optional semistructured interview: 19 caregivers completed an interview (16 identified as women/mothers, and 3 identified as men/fathers), and 12 caregivers declined. All 6 *M&C-C* sessions were attended by 14 of the 19 caregivers who were interviewed, and interviews lasted between 13 minutes and 46 minutes (mean 26 minutes). For clarity, this leaves the following sample sizes: demographics, attendance, and engagement ratings: n=36, 27 dyads, 9 independents; exit survey data for fidelity of enactment: n=31, 21 dyads, 10 independents; and interviews: n=19, 11 dyads, 8 independents.

XSL•FO
RenderX

Table . Caregiver participant demographic information (n=36).

Demographic characteristics	Results
Age (years)[a], mean (SD)	48.9 (5.92)
Age (years), range	35 - 63
Sex, n (%)[b]	
Female	29 (83)
Male	6 (17)
Gender, n (%)[c]	
Girl/woman	28 (82)
Boy/man	6 (18)
Relationship to child, n (%)	
Mother	30 (83)
Father	6 (17)
Ethnicity, n (%)[b]	
Arab	2 (6)
Black	2 (6)
Chinese	1 (3)
Latin American	1 (3)
South Asian	1 (3)
Southeast Asian	1 (3)
West Asian	1 (3)
White	22 (63)
Multiple ethnic identities	4 (11)[d]
Education, n (%)	
High school	2 (6)
College	4 (11)
University	23 (64)
Graduate school	7 (19)
Household income (CAD), n (%)[c]	
<30,000	3 (9)
31,000 - 50,000	1 (3)
51,000 - 70,000	3 (9)
71,000 - 90,000	2 (6)
91,000 - 100,000	2 (6)
>100,000	23 (68)
Support group involvement, n (%)	
Currently attending any support group	2 (6)
Ever accessed a support group	11 (31)
Mental health/therapy support, n (%)	
Currently attending therapy/mental health support[b]	6 (17)
Ever accessed therapy support/mental health support	12 (33)

[a]3 missing responses (n=33).

[b]1 missing response (n=35).

XSL•FO
RenderX

[c]2 missing responses (n=34).

[d]n=1 each: Chinese/Japanese, Metis/Ukrainian/Scottish, White/South Asian, Caribbean/German/Asian.

Objective 1: Intervention Fidelity Assessment

This section describes the results of the intervention fidelity assessment for *M&C-Y* and *M&C-C*.

Fidelity of Design

An intervention mapping framework [35] was applied to the intervention design process for *M&C-Y* and *M&C-C*. Scratch and colleagues [36] and Al-Hakeem and colleagues [37] described the detailed intervention development process, alignment with theory, and feasibility testing results. The perspectives of youth, caregivers, and clinicians were incorporated throughout the development process. To ensure fidelity of the intervention design, an intervention manual was created that outlined the intervention purpose, format, and protocols required to complete *M&C-Y* and *M&C-C*. This was provided to clinician facilitators and research staff.

Fidelity of Training

Four occupational therapists and 2 physiotherapists were trained to deliver *M&C-Y*, and 2 social workers, 2 neuropsychologists, and 2 neuropsychology trainees were trained to deliver *M&C-C*. Our train-the-trainer model was underpinned by best practices and frameworks within the training field [71,72]. Train-the-trainer frameworks are commonly used to update health care professionals' knowledge and skills and implement evidence-based interventions [71]. The training protocol used a blended learning approach, consisting of didactic and interactive methods, and was intended to ensure that all clinician facilitators delivered *M&C-Y* and *M&C-C* correctly, consistently, and safely to youth and caregivers. Feedback was also provided by the trainees (ie, individuals who were trained) about the training, which also informed the train-the-trainer approach. All clinician facilitators received institutional research training (eg, Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans [TCPS-2], Collaborative Institutional Training Initiative [CITI]), orientation to the study protocols and procedures, and a study manual with all intervention materials (eg, slides, REDcap access) [66]. All clinician facilitators observed group sessions prior to facilitating the intervention independently and participated in 1-on-1, discipline-specific, peer debriefing throughout this process. Weekly debrief meetings were held between the clinician facilitators and the research staff observer to resolve any logistical or operational issues. Individual consultation with the study principal investigator occurred, as needed (eg, to review the mental health safety protocol).

Fidelity of Delivery

The *M&C-Y* intervention protocol was consistently delivered as intended across all 10 groups (discussion: 58/60, 97% sessions; exercise: 55/60, 92% sessions; goal setting: 57/60, 95% sessions). Reasons for modifying or deviating from the *M&C-Y* intervention protocol included sessions being canceled or modified due to only 1 participant attending (n=2), youth declining to participate in the exercises due to illness or being symptomatic (n=2), and discussions that were so productive that the exercise or goals setting components of the program needed to be shortened to accommodate session timing (n=2). The *M&C-C* intervention protocol was also consistently delivered as intended across all 9 groups (psychoeducation: 50/54, 93% sessions; activity: 50/54, 93% sessions; take home reflection [weeks 2 - 6 only]: 37/45, 82% sessions). Reasons for modifying or deviating from the *M&C-C* intervention protocol included shortening the activity component because caregivers had questions and thoughts to share about the psychoeducation (n=3) or moving directly to the psychoeducation if caregivers did not complete the take home reflection (n=5). See Multimedia Appendix 2 for the detailed group breakdown of the clinician delivery checklist for *M&C-Y* and *M&C-C*.

Fidelity of Receipt

Attendance rates were as follows: all 6 sessions: 14/34 youth, 41% and 17/36 caregivers, 47%; 5 sessions: 8/34 youth, 24% and 7/36 caregivers, 19%; 4 sessions: 6/34 youth, 18% and 5/36 caregivers, 14%; 3 sessions: 4/34 youth, 12% and 4/36 caregivers, 11%; 2 sessions: 2/34 youth, 6% and 2/36 caregivers, 6%; 1 session: 1/36 caregivers, 3%. Participants frequently emailed the research team in advance when they were unable to attend a session (ie, due to scheduling conflicts or illness). In terms of fidelity receipt (ie, that participants understood and could use the intervention content), the research staff observer field notes revealed that participants frequently asked questions related to weekly topics (verbally and using the meeting chat function). Additionally, youth and caregivers were observed attending sessions with previously circulated educational materials and referring to them during group discussions. Participants completed weekly polls and discussions related to goal setting (youth) and take-home reflections (caregivers). Youth also frequently came to the sessions with the required supplies for the exercises (ie, running shoes, water bottles) that were provided in the weekly email reminders.

Fidelity of Enactment

For *M&C-Y*, the average exploratory participant engagement rating score was 2.72 (SD 0.36; range 1.5 - 3.0), and the overall group engagement scores ranged from 2.37 to 3.0 (mean 2.73, SD 0.21; see Multimedia Appendix 2). Overall, youth participants were satisfied with the *M&C-Y* intervention, indicated by favorable responses across the exit satisfaction survey (Table 4). In addition, youth responses on open-ended questions in this survey revealed that youth perceived the best feature of *M&C-Y* to be connection with other participants (10/34, 36%), the weekly exercises (7/34, 25%), the weekly goal setting (6/34, 21%), and learning strategies that they could apply outside the sessions (2/34, 7%). In terms of changes or opportunities to improve *M&C-Y*, some youth recommended moving the intervention from a virtual to an in-person format (4/34, 14%), providing more opportunities for connection between participants (4/34, 14%), making the exercises easier or harder (4/34, 14%), and including different educational topics (4/34, 14%).

XSL•FO
RenderX

Table . Youth (n=28) and caregiver (n=31) exit satisfaction survey Likert-scale question results.

Questions	Frequency of favorable responses, n (%)	Frequency of neutral responses, n (%)	Frequency of unfavorable responses, n (%)
Youth exit satisfaction survey			
I felt that this program was a good use of my time	24 (86)	2 (7)	2 (7)
I felt that the topics covered were relevant to my personal experiences.	20 (71)	7 (25)	1 (4)
The exercise component of Move&Connect was beneficial for me.	20 (71)	1 (4)	7 (25)
The exercises were the right amount of difficulty for me.	16 (57)	5 (18)	7 (25)
The exercise component was an appropriate length of time.	20 (71)	6 (21)	2 (7)
The educational component of the Move&Connect program was beneficial for me.	23 (82)	3 (11)	2 (7)
The program content (slides/handouts) complimented the discussions.	24 (86)	4 (14)	0 (0)
The discussions lasted an appropriate amount of time.	24 (86)	2 (7)	2 (7)
The goal setting component of the Move&Connect program was helpful for me.[a]	22 (81)	4 (15)	1 (4)
I will start/continue to set goals related to my recovery.	20 (71)	4 (14)	4 (14)
The group size was appropriate.	22 (78)	3 (11)	3 (11)
The sessions lasted an appropriate amount of time.	23 (82)	2 (7)	3 (11)
Meeting other youth with similar experiences was important for me.	23 (82)	4 (14)	1 (4)
I was comfortable sharing my experiences with other youth in a group setting.	22 (78)	3 (11)	3 (11)
I was comfortable completing the exercise portion of the program in a group setting.	23 (82)	4 (14)	1 (4)
I would recommend this program to other young people or friends who have had a concussion.	23 (82)	3 (11)	2 (7)
I felt supported by the group facilitators.	27 (96)	1 (4)	0 (0)
I felt supported by other members of the group.	23 (82)	3 (11)	2 (7)
I felt that the other members of the group understood my experiences.	19 (67)	8 (29)	1 (4)
I felt connected to the other members of the group.	15 (54)	7 (25)	6 (21)
Caregiver exit satisfaction survey			
I felt that this program was a good use of my time.	31 (100)	0 (0)	0 (0)

XSL•FO
RenderX

Questions	Frequency of favorable responses, n (%)	Frequency of neutral responses, n (%)	Frequency of unfavorable responses, n (%)
I felt that the topics covered were relevant to my personal experiences	31 (100)	0 (0)	0 (0)
I felt that the take-home activities were beneficial to my experience with the program.	23 (74)	6 (20)	2 (6)
The program handouts/resources were relevant and supplemented the program sessions well.	23 (74)	7 (23)	1 (3)
The educational component of the Move&Connect program was beneficial.[b]	29 (97)	1 (3)	0 (0)
The activities/strategies taught in the program were applicable.	28 (90)	3 (10)	0 (0)
I will start/continue to apply the activities/strategies taught during the program.	27 (87)	4 (13)	0 (0)
I found the weekly discussions to be helpful.	30 (97)	1 (3)	0 (0)
The group size was appropriate.	27 (87)	2 (6)	2 (6)
The sessions lasted an appropriate amount of time.[b]	28 (93)	2 (7)	0 (0)
Meeting other parents with similar experiences was important for me.	30 (97)	1 (3)	0 (0)
I was comfortable sharing my experiences in a group setting.	31 (100)	0 (0)	0 (0)
I would recommend this program to other caregivers or friends.	29 (94)	2 (6)	0 (0)
I felt supported by the group facilitators.	31 (100)	0 (0)	0 (0)
I felt supported by other members of the group	29 (94)	2 (6)	0 (0)
I felt that the other members of the group understood my experiences.	30 (97)	1 (3)	0 (0)
I felt connected to the other members of the group.	26 (84)	4 (13)	1 (3)

[a]1 missing youth response (n=27).

[b]1 missing caregiver response (n=30).

For *M&C-C*, the average exploratory participant engagement rating score was 2.79 (SD 0.27; range 2.0 - 3.0), and the overall group engagement score ranged from 2.5 to 3.0 (mean 2.8, SD 0.17; see Multimedia Appendix 2). Caregiver participants were also satisfied with the *M&C-C* intervention, with favorable responses indicated across the exit satisfaction survey (Table 4). Through open-ended questions on this survey, caregiver participants also indicated that they enjoyed the convenience of the virtual format (7/31, 23%) and that the best features of *M&C-C* were connecting with other caregivers with similar experiences (19/31, 61%) and the education and strategies provided (11/31, 36%). Suggestions to improve *M&C-C* included providing different educational topics and resources (6/31, 16%), additional reminder emails with program content (5/31, 16%), allowing more opportunities for connection between participants (4/31, 13%), modifying the timing of sessions (3/31, 10%), and increasing the number of participants in each group (2/31, 7%).

Examples of how youth and caregivers (Table 5) engaged with the interventions are outlined using the motivational framework from King et al [58], followed by an examination of the individual, intervention-related, and contextual factors that influenced these experiences of engagement (Objective 2). Behavioral engagement in both interventions centered around taking part in intervention components (eg, discussions and group activities) and participants' confidence in their ability to perform intervention tasks. Affective engagement involved the

XSL•FO
RenderX

positive feelings and attitudes participants had toward the interventions. Cognitive engagement included participants' interests, ways they felt inspired, and an appreciation of the structure and utility of *M&C-Y* and *M&C-C*.

Table . How youth and caregivers engaged with *Move&Connect-Youth* (*M&C-Y*) and *Move&Connect-Caregivers* (*M&C-C*), with each example of behavioral, affective, and cognitive engagement presented with an illustrative participant quote.

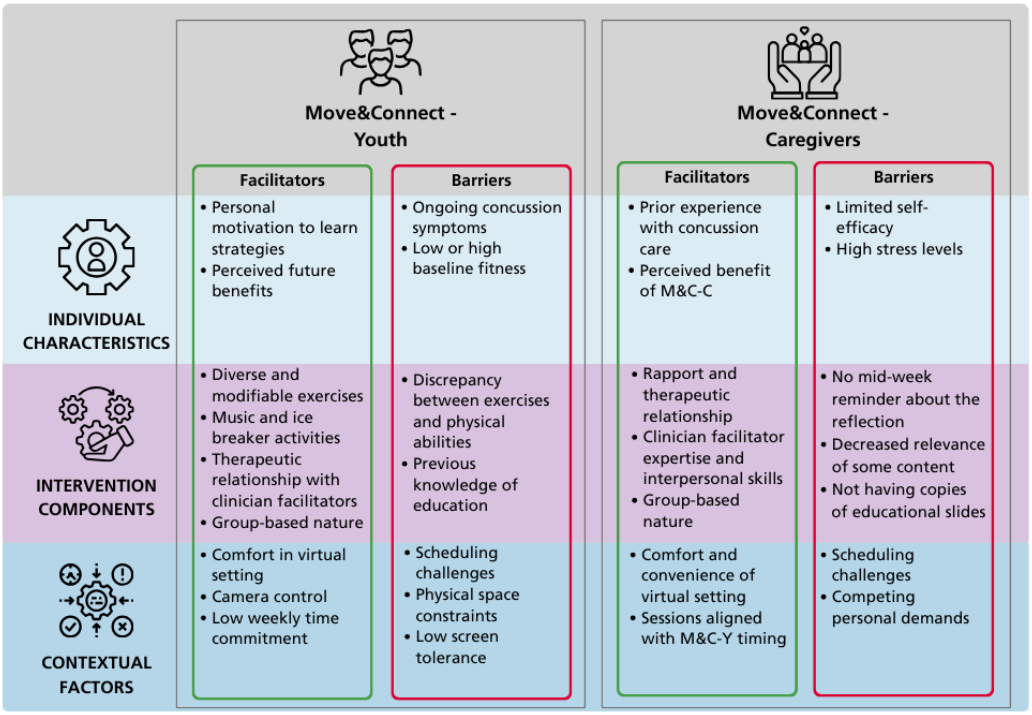
Examples	Quotes
M&C-Y	
Behavioral (participation/confidence in ability to do intervention tasks)	
Group discussions	We would talk about topics, and it would be nice to hear people's opinions and ideas on how to help […] it helped sometimes, it is not bad at all (Y01, age 11 years).
Goal setting	I was sharing my goals with everyone and they were also, they also had goals for themselves, kind of was a motivation to continue out for that, that goal. And... so I can come back next week and say that I accomplished it (Y09, age 17 years).
Exercising	I thought it was really cool to do all the exercises and have good discussions (Y10, age 16 years).
Beliefs about self-efficacy	I think I'll be able to. Uh, if you know a goal comes up that I think I will be able to follow it and accomplish it pretty well (Y16, age 17 years).
Affective (emotional investment and attitude towards intervention)	
Comfort and keenness	It made me feel like if I can do it, they can do it. If they can do it, I can (Y01, age 11 years).
Active and energized	I think it made me feel less tired, uh, just like yeah just less tired and more felt better, yeah, just like more active (Y11, age 16 years).
Cognitive (conviction of intervention appropriateness)	
Inspired by peers	It was inspirational to see them [the other participants] getting better throughout the weeks (Y14, age 17 years).
Belief in utility and structure	It was really educational, I think, and it was really good to talk to other people that were like experiencing the same thing as me. I think that was the most beneficial part for me (Y12, age 16 years).
M&C-C	
Behavioral (participation/confidence in ability to do intervention tasks)	
Emotionally vulnerable discussions	We were all a part of a group that was sharing information and relating to each other (C17).
Active listening and verbal participation	When I would make a comment and the other parent would respond by saying, "yeah, I get that, I get that I have that too." And then I, you know, the reciprocal was there (C02).
Application of strategies outside of sessions	It brought a heightened awareness about that might be what the children have been going through, right, and, and just being more aware and understanding about everything that is going on (C12).
Affective (emotional investment and attitude towards intervention)	
Validation and connection	It was great connection and kind of, it's nice to be able to talk about it with people who are going through similar experiences (C03).
Relief	I don't know how else to say it but it was it really was umm liberating. I don't know if that's the right word, but I just felt like a huge relief of saying, okay, I'm not the only one. I'm not going crazy when my kid is actually experiencing this, and you know somebody else actually finally understands me (C04).
Cognitive (conviction of intervention appropriateness)	
Relevance and benefit of *M&C-C*	I'm actually learning something, the presentation is good, you know? This is knowledge coming to me that I didn't have before, so yes (C11).
Positive perceptions of *M&C-C* structure	I thought it was just, really well run. It was just enough of everything like do you know what I mean like, um, education versus, um, chatting, um, yeah, I just thought it was done really well (C12).

XSL•FO
RenderX

Objective 2: Facilitators and Barriers to Engagement

This section describes factors that influenced participant engagement within *M&C-Y* and *M&C-C*. Facilitators and barriers to intervention engagement are organized across 3 overarching categories: (1) individual characteristics, (2) intervention-related components, and (3) contextual considerations (Figure 2). Findings for *M&C-Y* are presented first, followed by those for *M&C-C*.

Figure 2. Visual representation of key barriers and facilitators to engagement in *Move&Connect-Youth* (*M&C-Y*) and *Move&Connect-Caregivers* (*M&C-C*), organized by individual characteristics, intervention components, and contextual factors.



Facilitators of Engagement in M&C-Y

Individual Characteristics

Individual motivation and personal interests facilitated engagement. With respect to motivation, youth described joining *M&C-Y* to improve health, manage symptoms, and return to valued activities. Some also noted an interest in contributing to research and helping others.

Intervention Components

A lighthearted, comfortable, and nonjudgmental environment was reported to drive engagement in *M&C-Y*. Youth described *M&C-Y* sessions as friendly and positive, with a lighthearted tone that included laughter and playful moments among participants and clinician facilitators:

> *We have like those like ball throwing or sock throwing things, and things would just fly everywhere. [We would] like poke fun at each other but like in a positive way that made you feel like you can like laugh at yourself and not feel so bad if you're not good at something.* [Y15, age 17 years]

Creative icebreaker activities and conversation prompts were reported to ease nerves and encourage engagement. Youth noted that the personalized music played in the session supported the active rehabilitation exercises. Youth also emphasized that seeing their peers being open and transparent regarding their symptoms and challenges made it easier to participate in discussions and relate the intervention content to their own situations:

> *They [the other participants] were really open to sharing, which I really appreciated. And they shared some similar things. And so it was kind of like, I can share something like this. I felt comfortable sharing, some of the stuff that I was dealing with, because they were also sharing some of the stuff they were dealing with and it was similar.* [Y12, age 16 years]

Clinician facilitators were consistently described as supporting youth engagement. Youth commented that the clinician facilitators' style felt relatable, and they offered encouragement and validation. Clear reminders about confidentiality or "what's said on the call stays on the call" (Y01, age 11 years) reinforced trust and made disclosure feel appropriate within the group. Further, youth emphasized that the clinician facilitators taking part in activities and exercises themselves, alongside participants, enhanced engagement and comfort.

Most youth described the educational content as helpful and relatable when it linked concussion information to their own symptoms, daily routines, and recovery goals. They identified the variety of exercises and flexibility as supportive and appreciated that the exercises were tailored to different abilities and broken down into manageable blocks:

> *There were a lot of options, so even if I did not want to do one, if I was doing an exercise for a long time*

XSL•FO
RenderX

> *and could not do it anymore, I could move to an easier version.* [Y15, age 17 years]

Youth also valued that the exercises were challenging enough to feel meaningful without being overwhelming:

> *They [the exercises] were not too hard where I would be sore the next day, but they were hard enough that I would be tired in the end. That helped my physical and mental health, and they were fun to do.* [Y11, age 16 years]

Youth noted that watching peers share openly encouraged their own participation and improved feelings of connection. More specifically, having others in the group of similar age with comparable concussion experiences was helpful:

> *Everyone was always smiling, um, like there was never any criticism or anything, any questions were answered like there was kind of always a solution, like yeah it was just a nice environment. And again, with the age like being so close and similar in age and experiences too that it just it was a comfortable environment.* [Y16, age 17 years]

Importantly, the group format supported accountability and follow-through with setting and achieving goals. Youth indicated that the clear expectation to set goals, return the following week, and report on progress to the group promoted accountability. Others described how hearing peers' goals provided ideas for their own targets and a sense of shared effort:

> *I did set goals [before M&C-Y], and I tried to stick to them, but I just, I found I really wasn't sticking to them at all. And then within the group, it really helped me to stick to them because […] I was sharing my goals with everyone and they were also, they also had goals for themselves, kind of was a motivation to continue […] so I can come back next week and say that I accomplished it.* [Y09, age 17 years]

A small group size was also a facilitator to engagement. Youth described that a smaller group made it easier to speak, reduced stress levels, and increased the sense that individual contributions would be heard. It also allowed for more direct check-ins and time from the clinician facilitators supporting the sessions:

> *There weren't many people, and I felt like my voice was heard a lot of the time.* [Y04, age 16 years]

Contextual Factors

Broader environmental and technological factors also facilitated *M&C-Y* engagement. Youth were able to attend sessions more regularly when they had fewer extracurricular and school-related commitments. The virtual format supported comfort and engagement for some youth, particularly in their ability to control the camera and position themselves at a comfortable distance from the camera:

> *You don't actually have to see people like super close up. You can kind of move away from your camera and just focus on yourself instead of, oh, what are other people doing? Or what are other people thinking about watching me exercise.* [Y06, age 16 years]

Barriers to Engagement in M&C-Y

Individual Characteristics

Ongoing concussion symptoms (eg, headaches, fatigue, dizziness) were barriers to full participation in the exercises or made it harder to attend sessions. One youth described that where they were in their concussion recovery affected their participation and did not feel that the exercise component aligned with their needs:

> *I don't think exercises are really on par with my recovery. Because I've noticed that exercise is actually in fact making my symptoms worse. Even if it's just light exercises like walking. So, I wouldn't say that it was bad. Once again, I would say it just wasn't for me per se.* [Y05, age 17 years]

Intervention Components

Having exercises that felt too easy or too difficult was a barrier to engage fully in the session, particularly when youths perceived a mismatch between their current abilities and the session of intensity. For example, 1 youth who competed in track and field wanted "more challenging exercises," noting that even the harder options felt easy at times (Y08, age 15 years). However, the availability of modifications often mitigated some of these barriers:

> *I liked how they had very different variations depending on whether you could do it or not. I cannot really jump a lot because I get dizzy, and they have different alterations to some of the exercises.* [Y06, age 16 years]

Several youth reported that portions of the education content did not feel new nor relevant to their stage of recovery and were described as repetitive. For some, prior exposure to concussion resources or longer recoveries meant they had previously received education, which impacted how they engaged with the material. Others noted that, although the education was acceptable, it did not provide additional ideas for self-management beyond what the participant was already doing:

> *If people who are starting to realize what [concussion] is and what to do. And to me, I feel like I was a little bit of a step ahead. So I had to learn all this stuff already by myself, but it was still a good kind of process.* [Y03, age 15 years]

Youth offered suggestions to improve relevance and depth of education materials, including expanding topics such as sleep hygiene (Y01, age 11 years), brain anatomy, and nutrition (Y05, age 17 years). They also suggested integrating interactive elements or games into the education to increase engagement with younger youth. Some requested additional time for open-ended and peer-to-peer discussion to share experiences more fully:

> *I do wish I had more time to kind of have like more open-ended discussions like between like everyone. Just a bit more like time like specifically for having*

XSL•FO
RenderX

> *the discussions between all the kids. And like how they feel and their experiences…* [Y15, age 17 years]

A barrier to group participation involved social comparison, with 1 participant commenting that they evaluated themselves against others in terms of their recovery:

> *It was kind of like, ughh, you know, 'wow I'm so much worse than everyone else here.' It was kind of like a little not nice but yeah, whatever, like again it doesn't really affect me whether they're better than I or not as good, so yeah.* [Y16, age 17 years]

Contextual Factors

Scheduling demands, physical space, and screen tolerance were noted as barriers to engagement. Scheduling commitments related to schoolwork and extracurricular activities impacted attendance. One youth also commented that physical space in their home was a barrier to participating in the exercises:

> *I was doing it in my room, and I'm tall, my ceiling's a bit low, so it was a bit difficult at times.* [Y04, age 16 years]

Some youth noted limited screen tolerance made paying attention harder during the session. One youth, however, noted gradual improvement in screen tolerance over the course of the intervention:

> *At first, I was like 'how can I get on a Zoom call because screens hurt my eyes?' [yeah] but I tried it the first time. I wasn't feeling well that first week, but, I felt better as we got into the exercises. And then I also wasn't as afraid of going on to screens anymore, so yeah.* [Y14, age 17 years]

Facilitators of Engagement in M&C-C

Individual Characteristics

Individual characteristics that facilitated caregivers' engagement in *M&C-C* were motivation, perceived benefit of the intervention, and prior experiences with concussion care. Participants were motivated to join *M&C-C* to gain practical strategies to support their child's recovery, enhance their own well-being, and increase their understanding of concussion. Many caregivers were also interested in connecting with others in similar situations. Their frustration with previous concussion care also motivated them to join *M&C-C*:

> *I think we were at our wit's end right? And […] no matter where we were going, the doctors… we, we just felt like we weren't being heard, right?* [C04]

Intervention Components

Elements of the intervention that promoted engagement in *M&C-C* were having a dedicated space to connect and receive validation from other group members; clinician facilitator expertise; and a small, cohesive group format. Caregivers shared that the group provided a unique opportunity for open discussion with others who understood the challenges of supporting a child through prolonged concussion recovery, as a caregiver stated "No one judged me here" (C17). Connecting with other caregivers, exchanging stories, and hearing about others' coping strategies were among the most helpful aspects of the sessions. Similarities among members also supported engagement, with caregivers highlighting that their group happened to include parents in similar caregiving roles with similar aged children as them, which appeared to ease sharing among one another:

> *Just recognizing things that I was like oh yes I, I've got that too, and, and so, you know, recognizing the commonalities that are there and kind of then kind of what they're going through and how they deal with it was really nice.* [C18]

Caregivers reported feeling listened to and encouraged to participate, noting that the facilitators maintained a steady pace, balanced education with support, and created space for discussion. Sessions were described as well-structured and smoothly run, with facilitators who were personable, responsive to questions, experts in their fields, and clear in presenting material:

> *[The clinician facilitators] were just phenomenal. Besides the fact of being very personable, I think just the ability to ask some questions and the way they took us along the program. You know it, they were just really well-handled, well-structured sessions.* [C04]

Small group size was consistently mentioned as an important facilitator of engagement in *M&C-C*. Participants reported that fewer members increased connection, made communication easier, and fostered a sense of responsibility to attend and contribute. One caregiver shared:

> *This was the smallest group I've been involved with, and I think that the benefits of these small numbers really… they can't be ignored. […] I found, like I said, the level of intimacy was very beneficial and not only did it drive engagement, it really helped me make a full commitment to the program.* [C02]

Contextual Factors

Caregiver engagement was also facilitated by environmental and technological considerations. Caregivers highlighted that it was helpful to have both sessions (*M&C-Y* and *M&C-C*) run at the same time, as their children were occupied and not listening in on the caregiver group. The virtual format was widely viewed as convenient by caregivers. They emphasized that the virtual format reduced travel time and fit more easily into busy schedules. The ability to join from home was described as feasible and manageable:

> *I wouldn't have been able to do those [the sessions] if it wouldn't have been over Zoom because being so far away there, there's no way I would have been able to go for an hour and spend 3 hours getting there and back, wouldn't have made it work.* [C17]

Barriers to Engagement in M&C-C

Individual Characteristics

Some caregivers had initial doubts about their ability to understand the intervention content:

XSL•FO
RenderX

> *You're thinking, oh god, you know, is this going to be, uh, a really science-y thing, and I don't understand science.* [C19]

Others noted that stressful life demands made it challenging to fully engage in sessions or implement strategies between sessions:

> *It's still been hard to implement some of this only because I'm being pulled in so many different directions.* [C04]

Intervention Components

Although information shared was generally viewed positively, caregivers identified practical barriers in how materials were delivered by facilitators. Some requested receiving the full slide deck in advance to review and save for later use, as well as accessibility improvements such as larger font to make slides easier to read. Several caregivers also suggested sending additional reminders to complete the optional reflection activities:

> *I was remiss at not doing my first week's homework. And the reason why is because I couldn't remember what it was. And you know, had too many distractions on…* [C06]

Caregivers acknowledged that a limitation of the group model is that not all content resonated equally with all participants, with some material feeling less applicable to their circumstances:

> *Some weeks are maybe like a little bit longer or more dull if you want to say but that's always the way, right. You're going to latch on to like, some of the other parents latched on to things that I was like, I don't see myself or my kid in any of this. And then, so those weeks seem a little bit less connecting but that is the nature of different people being different.* [C07]

Contextual Factors

Time and logical constraints were common barriers to engagement. Caregivers described feeling stressed and overwhelmed with other life demands, which made it difficult to attend consistently, complete between-session activities, or implement new strategies. Scheduling the sessions during the evening (ie, 5 PM or 6 PM) was described as challenging due to appointments and family routines. Caregivers proposed alternative scheduling to improve engagement, such as later evenings, weekend options, or condensed workshop formats to broaden access. Some suggested that weekend sessions might attract caregivers who were underrepresented in the groups:

> *So, if it was a weekend workshop, maybe you'd get some of those dads out. Maybe they could go for a weekend. I could see them being like, okay, let's get those done in a one-er, you know?* [C06]

Despite the virtual convenience, a few caregivers also expressed a preference for at least one in-person touchpoint, such as a closing session or celebration, to strengthen personal connections.

Discussion

Principal Findings

Within the context of a broader pilot study, we used a multimethod approach to assess the fidelity of the *M&C-Y* and *M&C-C* interventions in terms of intervention development, training, delivery, receipt, and enactment. Second, we explored barriers and facilitators to youth and caregiver engagement with the interventions at the individual, intervention, and contextual levels. A rigorous approach to fidelity monitoring ensured high adherence to our theory-driven intervention protocols, quality of virtual delivery, and participant satisfaction and engagement. Caregivers showed consistent and high levels of satisfaction with *M&C-C*, whereas youth satisfaction with *M&C-Y* had some variability, particularly related to exercise difficulty. Both youth and caregivers expressed that clinician facilitators established a supportive, nonjudgmental environment that was central to building feelings of comfort and supported engagement. The group format provided important validation of youth and caregiver experiences and supported accountability to apply strategies outside of the interventions. Barriers to youth engagement included concussion symptoms, baseline fitness, and previous knowledge of education, whereas barriers to caregiver engagement included stress, scheduling demands, and limited access to intervention materials. Results provided important implications for delivering virtual group interventions for youth with PSaC and their caregivers.

Comparison With Prior Work

Our approach to fidelity assessment aligned with the framework from Bellg and colleagues [55] and best practices for evaluating intervention fidelity [53,54] by using multiple methods to allow for triangulation and reducing the chance of biases associated with any one approach [73]. Triangulation recognizes that each individual will have a unique perspective on a shared phenomenon [57]. We used clinician facilitator checklists and research staff observer logs and field notes and provided multiple mediums for youth and caregivers to provide their feedback. There is heterogeneity in what constitutes "good" attendance of an intervention [74]; however, it is promising that the majority of participants in this study attended at least 5 of 6 sessions. As recommended by Toomey and colleagues [75], our approach extended beyond intervention delivery to include an evaluation of fidelity receipt and enactment. Opportunities to assess fidelity receipt and enactment were also built directly into the interventions to reduce administrative burden [57]. A rigorous approach to fidelity assessment is especially important in complex interventions like *M&C-Y* and *M&C-C* that likely have multiple potential active components, and this fidelity evaluation can enhance replicability and help ensure that future outcomes can more confidently be attributed to the interventions [53,57,75].

Virtual videoconferencing technology (eg, Zoom Healthcare) enables access and scalability of pediatric rehabilitation interventions [76,77], and families report that virtual care offers benefits such as convenience, flexibility, and comfort in the home environment [78]. Similar to findings reported by Shore and colleagues [26], where youth with PSaC and parents

XSL•FO
RenderX

appreciated the convenience and comfort of engaging in individual virtual rehabilitation from home, participants in this study described the virtual setting as a facilitator of attendance and felt comfortable completing the interventions from home. Although participants appreciated the benefits provided by the virtual setting and almost all participants felt supported by the other members of the group, some felt that they did not connect with other members of the group. The discontinuous nature of cyberspace or the "loss of shared space" in virtual therapy has been identified as a threat to group engagement and relational processes in virtual groups [79,80]. To increase connection and engagement, some participants requested an in-person component or more time for unstructured or open-ended conversations between participants. The virtual nature of *M&C-Y* and *M&C-C* may lack the informal opportunities for connection that happen in the waiting room or parking lot before and after the weekly sessions. Indeed, there seems to be a trade-off between convenience and connection in the virtual group setting for youth with PSaC and caregivers.

The findings from this study have important implications for providing virtual group-based education and active rehabilitation to youth with PSaC. Active rehabilitation and education are recommended in the treatment of PSaC [12], and virtual active rehabilitation interventions are emerging as feasible approach to care [22,23]. The educational and exercise components of *M&C-Y* were delivered as intended in the majority of sessions, which is especially important as intervention personalization and deviations have been poorly reported in previous exercise interventions for concussion [81]. Although most youth agreed the education was beneficial and relevant to their personal experiences, some youth described that the education was repetitive, with information that they had already received from other health care providers or strategies that they had to previously learn on their own. Importantly, youth in this study had many risk factors for prolonged recovery [2,82] and were an average of almost 1 year postinjury (mean 327 days). As such, perspectives of these youth and their caregivers, as well as their responses to the interventions, could be different than youth injured more recently. Consistent with current clinical care guidelines [12], this suggests that education should be provided earlier in the recovery process.

The *M&C-Y* program aligns with a recent call-to-action to develop concussion interventions that do not require expensive gym-based equipment (eg, stationary bike) and to implement home exercise programs via virtual care that only require accessible and affordable materials [83]. Interestingly, both high and low levels of baseline fitness were identified as a barrier to engagement in the exercise portion of *M&C-Y*. The intervention uses a predetermined set of exercises, and clinician facilitators provided modifications for each exercise to make it easier or more challenging, which was viewed as a facilitator to engagement for participants. Even with these modifications, some youth felt that the exercise component was not the right level of difficulty. This is consistent with previously identified challenges with personalizing exercise in the virtual group setting, including decreased direct observation [84] and not being able to use your hands to correct exercise techniques [85]. Future delivery of *M&C-Y* could mitigate this by providing an individual assessment before the program begins or further individualize the exercises by providing additional opportunities for youth to provide feedback on the exercises throughout the intervention. This aligns with the principles of client and family-centered care, which emphasizes the importance of working with youth with PSaC to align aspects of the intervention to their needs [44].

There are limited educational interventions designed specifically for caregivers of youth with PSaC [34]. Unlike youth in this study who described receiving education about managing their concussion from multiple health care providers, *M&C-C* was the first time that the caregivers received education designed specifically for them. Caregivers were overwhelmingly satisfied with *M&C-C*, and virtually all caregivers agreed that the facilitators were supportive, meeting other caregivers with similar experiences was important, and the discussions were helpful and relevant to their personal experiences. These findings reinforce the importance of developing and evaluating caregiver-centered interventions and support [34]. Caregivers of youth with PSaC report experiencing high levels of stress [28] and being overwhelmed managing multiple health care appointments to support their child [44]. Similarly, caregivers in this study also identified high levels of stress and busy conflicting schedules as barriers to attendance and engagement in *M&C-C*. This was apparent, as the most common modification to the *M&C-C* intervention protocol was omitting the take-home reflection review due to caregivers not completing the reflection between sessions. Strategies like customizing mid-week reminders about reflections were suggested to support caregivers to more fully engage in the intervention.

Limitations

This study has limitations worth noting. First, as is the case with most intervention studies, results from the T_1 satisfaction surveys and interviews only include the individuals that attended the T_1 data collection session and may represent the perspectives of participants who were most engaged with the interventions. To ensure a balanced approach to intervention receipt and enactment, data from individuals who were lost to T_1 follow-up were included in the demographics, attendance, field notes, and engagement ratings. This is important, as individuals with concussion who do not attend follow-up are vastly underrepresented in the literature [86]. As participation in the interviews was optional, the qualitative findings may also be subject to self-selection and positivity bias and may not fully represent the experiences of all intervention participants. There is also potential of nonindependence of observations from members of the same family. In addition, fidelity receipt and enactment were assessed primarily through indirect observer field notes, engagement ratings, and self-report. The participant engagement rating was developed by the research team, as we were unable to identify a validated, observer-rated measure of in-session engagement that reflected the components of the interventions and was feasible for the virtual group context. Although the scale provided useful exploratory insights, its reliability and validity have not been established and therefore may not capture the intended construct of engagement. The scale may also be subject to observer bias, as sessions were not

XSL•FO
RenderX

recorded and observed by a single rater. To mitigate this, our engagement evaluation was multifaceted, and triangulation occurred with surveys and interviews. Additionally, recruitment for this study primarily took place in a hospital-based clinic in a large urban center where individuals may have had increased access to concussion care in general. Although the virtual format of *M&C-Y* and *M&C-C* makes the interventions more accessible to individuals living in rural areas, future efforts should be made to provide access to the interventions and evaluate it in this group. In addition, most participants in this study identified as White and were from high socioeconomic status households. There is a dearth of information about the association between socioeconomic status and concussion; however, recent Canadian studies have identified that individuals from households with lower socioeconomic status are less likely to access care for concussion [41,87,88]. Understanding how individuals from diverse backgrounds engage with the *Move&Connect* program should be a priority for future research. Strategies to engage these underrepresented groups could include providing alternative scheduling, offering translation services, and a focus on community-based recruitment. Last, most participants in this study identified as girls/women. This is consistent with findings from initial feasibility testing of *M&C-Y* [36] and *M&C-C* [37] and may indicate a preference by girls/women toward group-based interventions. Future efforts should be made to understand the experiences of boys/men with the interventions and explore barriers they may experience when enrolling in this type of intervention.

Conclusions

This paper aimed to evaluate the intervention fidelity of *M&C-Y* and *M&C-C* and explore facilitators and barriers to youth and caregiver engagement with the virtual group interventions. The theory-driven fidelity evaluation, including intervention development, training, delivery, receipt, and enactment, demonstrated that the pilot study of *M&C-Y* and *M&C-C* was administered with high delivery adherence. Caregivers reported high levels of satisfaction and engagement, whereas youth satisfaction and engagement had some variability, particularly related to the exercise component. Engagement with the interventions was shaped by individual, intervention, and contextual factors. Important considerations for facilitating virtual group interventions for youth with PSaC and their caregivers are presented. Future research should examine the impact of *M&C-Y* and *M&C-C* on outcomes such as concussion symptoms, goal attainment, self-efficacy, mental health, and family functioning.

Acknowledgments

Thank you to the clients and families that participated in this study and made this research possible. We would like to acknowledge the clinician facilitators on the *Move&Connect* team, specifically Lindsay Brazill, Cynthia Campos, Michelle Tanel, Rashan Edwards, Tess Bardikoff, Sabrina Freud, Emilee Morin, and Kassandra Korcsog. We also acknowledge members of the *Move&Connect* Canadian Institutes of Health Research project grant team, specifically Katie Mah, Nick Reed, William Montelpare, and Jessica Brian for their ongoing support. We would also like to acknowledge the efforts of the members of the NOvEL Lab team (Bloorview Research Institute), specifically Willow Barton, Alexa Irvin, and Christine Muscat.

Funding

This work was supported by the Canadian Institutes of Health Research (#463185) and the Holland Family Professorship in Acquired Brain Injury. Open access publishing costs were generously supported by Dr. Peter Rumney. We are deeply grateful for his support and for his remarkable legacy of advancing care for families affected by pediatric brain injury. HAH received graduate funding support from the Canadian Institutes of Health Research and the Ontario Graduate Scholarship.

Data Availability

Data will be made available upon reasonable request to the corresponding author and in line with ethical standards of practice and institutional data transfer agreements.

Authors' Contributions

Conceptualization: AH, HAH, AL, SES, KDM
Data curation: AH, HAH, AL, SMN, KDM
Formal analysis: AH, HAH, AL, SMN
Funding acquisition: AH, CFP, RW, PK, BK, SLW, KOS, SES
Investigation: AH, HAH, AL, SMN, KDM, SES
Methodology: AH, HAH, AL, KDM, CFP, SLW, KOY, SES
Project administration: AH, HAH, AL, SES
Resources: SES
Supervision: AH, KDM, SES
Visualization: AH, HAH, AL, SMN, SES
Writing – original draft: AH, HAH, AL, SMN, SES
Writing – review & editing: HAH, AL, SMN, KDM, JLR, CFP, RW, PK, BK, SLW, KOY, SES

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Youth and caregiver interview guides.
[DOCX File, 32 KB - rehab_v13i1e88771_app1.docx]

Multimedia Appendix 2
Clinician delivery checklist results and group engagement scores.
[DOCX File, 18 KB - rehab_v13i1e88771_app2.docx]

Checklist 1
COREQ (Consolidated Criteria for Reporting Qualitative Studies) checklist.
[PDF File, 439 KB - rehab_v13i1e88771_app3.pdf]

References

1. Patricios JS, Schneider KJ, Dvorak J, et al. Consensus statement on concussion in sport: the 6th International Conference on Concussion in Sport-Amsterdam, October 2022. Br J Sports Med 2023 Jun;57(11):695-711. [doi: 10.1136/bjsports-2023-106898] [Medline: 37316210]
2. Zemek R, Barrowman N, Freedman SB, et al. Clinical risk score for persistent postconcussion symptoms among children with acute concussion in the ED. JAMA 2016 Mar 8;315(10):1014-1025. [doi: 10.1001/jama.2016.1203] [Medline: 26954410]
3. Babcock L, Byczkowski T, Wade SL, Ho M, Mookerjee S, Bazarian JJ. Predicting postconcussion syndrome after mild traumatic brain injury in children and adolescents who present to the emergency department. JAMA Pediatr 2013 Feb;167(2):156-161. [doi: 10.1001/jamapediatrics.2013.434] [Medline: 23247384]
4. Fineblit S, Selci E, Loewen H, Ellis M, Russell K. Health-related quality of life after pediatric mild traumatic brain injury/concussion: a systematic review. J Neurotrauma 2016 Sep 1;33(17):1561-1568. [doi: 10.1089/neu.2015.4292] [Medline: 26916876]
5. Novak Z, Aglipay M, Barrowman N, et al. Association of persistent postconcussion symptoms with pediatric quality of life. JAMA Pediatr 2016 Dec 5;170(12):e162900. [doi: 10.1001/jamapediatrics.2016.2900] [Medline: 27775762]
6. Conder R, Conder AA. Neuropsychological and psychological rehabilitation interventions in refractory sport-related post-concussive syndrome. Brain Inj 2015;29(2):249-262. [doi: 10.3109/02699052.2014.965209] [Medline: 25291459]
7. Sheldrake E, Al-Hakeem H, Lam B, et al. Mental health outcomes across the lifespan in individuals with persistent post-concussion symptoms: a scoping review. Front Neurol 2022;13:850590. [doi: 10.3389/fneur.2022.850590] [Medline: 35481264]
8. Kita H, Mallory KD, Hickling A, Wilson KE, Kroshus E, Reed N. Social support during youth concussion recovery. Brain Inj 2020 May 11;34(6):782-790. [doi: 10.1080/02699052.2020.1753243] [Medline: 32315217]
9. Kapadia M, Scheid A, Fine E, Zoffness R. Review of the management of pediatric post-concussion syndrome-a multi-disciplinary, individualized approach. Curr Rev Musculoskelet Med 2019 Mar;12(1):57-66. [doi: 10.1007/s12178-019-09533-x] [Medline: 30758705]
10. Scratch SE, Rumney P, Agnihotri S, Reed N. Pediatric concussion: managing persistent symptoms with an interdisciplinary approach. J Head Trauma Rehabil 2019;34(6):385-393. [doi: 10.1097/HTR.0000000000000542] [Medline: 31688376]
11. Gagnon I, Galli C, Friedman D, Grilli L, Iverson GL. Active rehabilitation for children who are slow to recover following sport-related concussion. Brain Inj 2009 Nov;23(12):956-964. [doi: 10.3109/02699050903373477] [Medline: 19831492]
12. Reed N, Zemek R, Dawson J, Ledoux AA, Provvidenza C, Paniccia M. Living Guideline for Pediatric Concussion. OSF. 2024. URL: https://osf.io/3vwn9/overview [accessed 2026-09-04]
13. Gauvin-Lepage J, Friedman D, Grilli L, et al. Effectiveness of an exercise-based active rehabilitation intervention for youth who are slow to recover after concussion. Clin J Sport Med 2020 Sep;30(5):423-432. [doi: 10.1097/JSM.0000000000000634] [Medline: 30095507]
14. Hunt AW, Agnihotri S, Sack L, et al. Mood-related changes in children and adolescents with persistent concussion symptoms following a six-week active rehabilitation program. Brain Inj 2020 Jul 2;34(8):1068-1073. [doi: 10.1080/02699052.2020.1776396] [Medline: 32529851]
15. Anderson V, Manikas V, Babl FE, Hearps S, Dooley J. Impact of moderate exercise on post-concussive symptoms and cognitive function after concussion in children and adolescents compared to healthy controls. Int J Sports Med 2018 Sep;39(9):696-703. [doi: 10.1055/a-0592-7512] [Medline: 30068005]
16. Gagnon I, Grilli L, Friedman D, Iverson GL. A pilot study of active rehabilitation for adolescents who are slow to recover from sport-related concussion. Scand J Med Sci Sports 2016 Mar;26(3):299-306. [doi: 10.1111/sms.12441] [Medline: 25735821]

XSL•FO
RenderX

17. Alfonso D, Gerst E. Psychoeducation as intervention in complex concussion cases: profile of an adolescent patient. Journal of Pediatric Neuropsychology 2024 Dec;10(4):306-313. [doi: 10.1007/s40817-024-00175-7]
18. Belanger HG, Barwick F, Silva MA, Kretzmer T, Kip KE, Vanderploeg RD. Web-based psychoeducational intervention for postconcussion symptoms: a randomized trial. Mil Med 2015 Feb;180(2):192-200. [doi: 10.7205/MILMED-D-14-00388] [Medline: 25643387]
19. Bales NJ, Perera DC, Foerster R, Poirier L, Ducis K. Analysis of a novel virtual pediatric concussion clinic in a rural setting. Childs Nerv Syst 2024 Apr;40(4):1199-1205. [doi: 10.1007/s00381-023-06231-5] [Medline: 38015249]
20. Ellis MJ, Mendez I, Russell K. Preliminary clinical algorithm to optimise remote delivery of paediatric concussion care in Canada's North. Int J Circumpolar Health 2020 Dec;79(1):1832390. [doi: 10.1080/22423982.2020.1832390] [Medline: 33089768]
21. Ellis MJ, Boles S, Derksen V, et al. Evaluation of a pilot paediatric concussion telemedicine programme for northern communities in Manitoba. Int J Circumpolar Health 2019 Dec;78(1):1573163. [doi: 10.1080/22423982.2019.1573163] [Medline: 30714513]
22. Shore J, Hutchison MG, Nalder E, Reed N, Hunt A. Tele-active rehabilitation for adolescents with concussion: a feasibility study. BMJ Open Sport Exerc Med 2022;8(1):e001277. [doi: 10.1136/bmjsem-2021-001277] [Medline: 35309373]
23. Teel E, Alarie C, Swaine B, Cook NE, Iverson GL, Gagnon I. An at-home, virtually administered graded exertion protocol for use in concussion management: preliminary evaluation of safety and feasibility for determining clearance to return to high-intensity exercise in healthy youth and children with subacute concussion. J Neurotrauma 2023 Aug;40(15-16):1730-1742. [doi: 10.1089/neu.2022.0370] [Medline: 37212272]
24. Chrisman SPD, Bollinger BJ, Mendoza JA, et al. Mobile Subthreshold Exercise Program (MSTEP) for concussion: study protocol for a randomized controlled trial. Trials 2022 Apr 26;23(1):355. [doi: 10.1186/s13063-022-06239-3] [Medline: 35473570]
25. Doucet C, Gagnon IJ, Beaulieu C, Grilli L, Teel EF. Physiotherapy-based concussion management outcomes are similar for children whether care is provided in-person or via telehealth. Brain Inj 2026;40(9):763-771. [doi: 10.1080/02699052.2026.2669601] [Medline: 42095598]
26. Shore J, Bernick A, Nalder E, Hutchison M, Reed N, Hunt A. Adolescent and parent experiences with tele-active rehabilitation for concussion: an exploratory qualitative study. Brain Inj 2022 Jul 29;36(9):1140-1148. [doi: 10.1080/02699052.2022.2114610] [Medline: 35993317]
27. Ganesalingam K, Yeates KO, Ginn MS, et al. Family burden and parental distress following mild traumatic brain injury in children and its relationship to post-concussive symptoms. J Pediatr Psychol 2008 Jul;33(6):621-629. [doi: 10.1093/jpepsy/jsm133] [Medline: 18227110]
28. Teel EF, Caron JG, Gagnon IJ. Developing a conceptual framework to identify and classify sources of parental stress following pediatric concussion. Brain Inj 2022 Jul 29;36(9):1149-1157. [doi: 10.1080/02699052.2022.2110283] [Medline: 35946143]
29. Chadwick L, Marbil MG, Madigan S, Callahan BL, Yeates KO. The relationship between parental and family functioning and post-concussive symptoms after pediatric mild traumatic brain injury: a scoping review. J Neurotrauma 2024 Feb;41(3-4):305-318. [doi: 10.1089/neu.2023.0201] [Medline: 37565282]
30. Miley AE, Fisher AP, Moscato EL, et al. A mixed-methods analysis examining child and family needs following early brain injury. Disabil Rehabil 2022 Jul;44(14):3566-3576. [doi: 10.1080/09638288.2020.1870757] [Medline: 33459078]
31. Minney MJ, Roberts RM, Mathias JL, Raftos J, Kochar A. Service and support needs following pediatric brain injury: perspectives of children with mild traumatic brain injury and their parents. Brain Inj 2019;33(2):168-182. [doi: 10.1080/02699052.2018.1540794] [Medline: 30396278]
32. Eliason PH, Galarneau JM, Kolstad AT, et al. Prevention strategies and modifiable risk factors for sport-related concussions and head impacts: a systematic review and meta-analysis. Br J Sports Med 2023 Jun;57(12):749-761. [doi: 10.1136/bjsports-2022-106656] [Medline: 37316182]
33. Lucas S, Bolter ND, Petranek LJ, Nilsson K, Pardue K, Flint H. Parents'/guardians' experiences with their adolescents' prolonged recovery from a sport-related concussion. Journal of Concussion 2023 Jan;7. [doi: 10.1177/20597002231160946]
34. Kamba G, Plourde V. Psychoeducational interventions and postconcussive recovery in children and adolescents: a rapid systematic review. Arch Clin Neuropsychol 2022 Feb 22;37(3):568-582. [doi: 10.1093/arclin/acac011] [Medline: 35262665]
35. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct;25(5):545-563. [doi: 10.1177/109019819802500502] [Medline: 9768376]
36. Scratch SE, Mallory KD, Al-Hakeem H, et al. Move&Connect-Youth: a virtual group intervention for youth experiencing persisting symptoms after concussion. Dev Neurorehabil 2023 Nov;26(8):471-482. [doi: 10.1080/17518423.2024.2331455] [Medline: 38531782]
37. Al-Hakeem H, Hickling A, Mallory KD, et al. Move&Connect-Caregivers: a virtual group intervention for caregivers of youth experiencing persisting symptoms after concussion. Dev Neurorehabil 2024 Oct;27(7):217-227. [doi: 10.1080/17518423.2024.2398161] [Medline: 39313848]

XSL•FO
RenderX

38. Brigo E, Rintala A, Kossi O, et al. Using telehealth to guarantee the continuity of rehabilitation during the COVID-19 pandemic: a systematic review. Int J Environ Res Public Health 2022 Aug 19;19(16):10325. [doi: 10.3390/ijerph191610325] [Medline: 36011959]
39. Goldman JG, Merkitch D, Brewington D, et al. Patient experiences receiving rehabilitation care via telehealth: identifying opportunities for remote care. Front Rehabil Sci 2023;4:1049554. [doi: 10.3389/fresc.2023.1049554] [Medline: 36817717]
40. Ellis MJ, Russell K. The potential of telemedicine to improve pediatric concussion care in rural and remote communities in Canada. Front Neurol 2019;10:840. [doi: 10.3389/fneur.2019.00840] [Medline: 31428043]
41. Wittevrongel K, Barrett O, Hagel BE, et al. Factors associated with follow-up care after pediatric concussion: a longitudinal population-based study in Alberta, Canada. Front Pediatr 2022;10:1035909. [doi: 10.3389/fped.2022.1035909] [Medline: 36699293]
42. Womble MN, Reynolds E, Kissinger-Knox A, et al. The emerging role of telehealth for concussion clinical care during the coronavirus (COVID-19) pandemic. J Head Trauma Rehabil 2022;37(2):E49-E54. [doi: 10.1097/HTR.0000000000000713] [Medline: 34320559]
43. Ganesalingam K, Yeates KO, Ginn MS, et al. Family burden and parental distress following mild traumatic brain injury in children and its relationship to post-concussive symptoms. J Pediatr Psychol 2007 Oct 23;33(6):621-629. [doi: 10.1093/jpepsy/jsm133]
44. Hickling A, Mah K, Al-Hakeem H, Scratch SE. Exploring the experiences of youth with persistent post-concussion symptoms and their families with an interprofessional team-based assessment. J Interprof Care 2023;37(4):558-567. [doi: 10.1080/13561820.2022.2137482] [Medline: 36373207]
45. Chrisman SPD, Mendoza JA, Zhou C, et al. Pilot study of telehealth delivered rehabilitative exercise for youth with concussion: the Mobile Subthreshold Exercise Program (MSTEP). Front Pediatr 2021;9:645814. [doi: 10.3389/fped.2021.645814] [Medline: 34123963]
46. Macheras S, Ponsford J, Peverill T, et al. Feasibility and preliminary efficacy of telehealth interdisciplinary rehabilitation for concussion recovery: a single-case experimental design. Ann Phys Rehabil Med 2026 May;69(4):102126. [doi: 10.1016/j.rehab.2026.102126] [Medline: 42019192]
47. Callahan CE, Beisecker L, Zeller S, Donnelly KZ. LoveYourBrain mindset: feasibility, acceptability, usability, and effectiveness of an online yoga, mindfulness, and psychoeducation intervention for people with traumatic brain injury. Brain Inj 2023 Apr 16;37(5):373-382. [doi: 10.1080/02699052.2023.2168062] [Medline: 36692090]
48. Callahan CE, Donnelly KZ, Gaylord SA, et al. Feasibility and preliminary effectiveness of an online meditation intervention in young adults with concussion history. J Sport Rehabil 2024 Jul 1;33(5):346-355. [doi: 10.1123/jsr.2023-0329] [Medline: 38843862]
49. Engel GL. The need for a new medical model: a challenge for biomedicine. Science 1977 Apr 8;196(4286):129-136. [doi: 10.1126/science.847460] [Medline: 847460]
50. Baldwin P, King G, Evans J, McDougall S, Tucker MA, Servais M. Solution-focused coaching in pediatric rehabilitation: an integrated model for practice. Phys Occup Ther Pediatr 2013 Nov;33(4):467-483. [doi: 10.3109/01942638.2013.784718] [Medline: 23611310]
51. Fisher A, Bellon M, Lawn S, Lennon S, Sohlberg M. Family-directed approach to brain injury (FAB) model: a preliminary framework to guide family-directed intervention for individuals with brain injury. Disabil Rehabil 2019 Apr;41(7):854-860. [doi: 10.1080/09638288.2017.1407966] [Medline: 29171308]
52. Orsmond GI, Cohn ES. The distinctive features of a feasibility study: objectives and guiding questions. OTJR (Thorofare N J) 2015 Jul;35(3):169-177. [doi: 10.1177/1539449215578649] [Medline: 26594739]
53. Borrelli B, Sepinwall D, Ernst D, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. J Consult Clin Psychol 2005 Oct;73(5):852-860. [doi: 10.1037/0022-006X.73.5.852] [Medline: 16287385]
54. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. J Public Health Dent 2011;71(s1):S52-S63. [doi: 10.1111/j.1752-7325.2011.00233.x] [Medline: 21499543]
55. Bellg AJ, Borrelli B, Resnick B, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 2004 Sep;23(5):443-451. [doi: 10.1037/0278-6133.23.5.443] [Medline: 15367063]
56. Walton H, Spector A, Williamson M, Tombor I, Michie S. Developing quality fidelity and engagement measures for complex health interventions. Br J Health Psychol 2020 Feb;25(1):39-60. [doi: 10.1111/bjhp.12394] [Medline: 31693797]
57. Ginsburg LR, Hoben M, Easterbrook A, Anderson RA, Estabrooks CA, Norton PG. Fidelity is not easy! Challenges and guidelines for assessing fidelity in complex interventions. Trials 2021 May 29;22(1):372. [doi: 10.1186/s13063-021-05322-5] [Medline: 34051830]
58. King G, Currie M, Petersen P. Child and parent engagement in the mental health intervention process: a motivational framework. Child Adolesc Ment Health 2014 Feb;19(1):2-8. [doi: 10.1111/camh.12015] [Medline: 32878365]
59. Bolster EAM, van Gessel C, Welten M, et al. Using a co-design approach to create tools to facilitate physical activity in children with physical disabilities. Front Rehabil Sci 2021;2:707612. [doi: 10.3389/fresc.2021.707612] [Medline: 36188842]

XSL•FO
RenderX

60. Elbers S, van Gessel C, Renes RJ, van der Lugt R, Wittink H, Hermsen S. Innovation in pain rehabilitation using co-design methods during the development of a relapse prevention intervention: case study. J Med Internet Res 2021 Jan 20;23(1):e18462. [doi: 10.2196/18462] [Medline: 33470937]
61. Antoniadou M, Granlund M, Andersson AK. Strategies used by professionals in pediatric rehabilitation to engage the child in the intervention process: a scoping review. Phys Occup Ther Pediatr 2024;44(4):461-488. [doi: 10.1080/01942638.2023.2290038] [Medline: 38213190]
62. Danzl MM, Etter NM, Andreatta RD, Kitzman PH. Facilitating neurorehabilitation through principles of engagement. J Allied Health 2012;41(1):35-41. [Medline: 22544406]
63. Foster AM, Armstrong J, Buckley A, et al. Encouraging family engagement in the rehabilitation process: a rehabilitation provider's development of support strategies for family members of people with traumatic brain injury. Disabil Rehabil 2012;34(22):1855-1862. [doi: 10.3109/09638288.2012.670028] [Medline: 22489631]
64. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care 2007 Sep 16;19(6):349-357. [doi: 10.1093/intqhc/mzm042]
65. Biros M. Capacity, vulnerability, and informed consent for research. J Law Med Ethics 2018 Mar;46(1):72-78. [doi: 10.1177/1073110518766021] [Medline: 29991882]
66. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
67. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444]
68. Collins H, Leonard-Clarke W, O'Mahoney H. 'Um, er': how meaning varies between speech and its typed transcript. Qualitative Research 2019 Dec;19(6):653-668. [doi: 10.1177/1468794118816615]
69. Turner D. Quirkos 3 is released: our best offline qualitative analysis software. Quirkos. 2025. URL: https://www.quirkos.com/blog/post/quirkos-3-is-released-best-qualitative-software/ [accessed 2026-09-04]
70. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. Int J Qual Methods 2017 Dec 1;16(1). [doi: 10.1177/1609406917733847]
71. Nexø MA, Kingod NR, Eshøj SH, Kjærulff EM, Nørgaard O, Andersen TH. The impact of train-the-trainer programs on the continued professional development of nurses: a systematic review. BMC Med Educ 2024 Jan 4;24(1):30. [doi: 10.1186/s12909-023-04998-4] [Medline: 38178050]
72. Poitras ME, Couturier Y, Doucet E, et al. Enhanced train-the-trainer program for registered nurses and social workers to apply the founding principles of primary care in their practice: a pre-post study. BMC Prim Care 2024 Sep 3;25(1):322. [doi: 10.1186/s12875-024-02574-3] [Medline: 39227782]
73. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity criteria: development, measurement, and validation. American Journal of Evaluation 2003 Sep;24(3):315-340. [doi: 10.1177/109821400302400303]
74. Peycheva AA, Guevara AF, Jordans MJD. Systematic review of strategies for improving attendance in health and mental health interventions for children and adolescents in LMICs: implications for mental health interventions. SSM - Mental Health 2023 Dec;4:100284. [doi: 10.1016/j.ssmmh.2023.100284]
75. Toomey E, Hardeman W, Hankonen N, et al. Focusing on fidelity: narrative review and recommendations for improving intervention fidelity within trials of health behaviour change interventions. Health Psychol Behav Med 2020 Mar 12;8(1):132-151. [doi: 10.1080/21642850.2020.1738935] [Medline: 34040865]
76. Hsu N, Monasterio E, Rolin O. Telehealth in pediatric rehabilitation. Phys Med Rehabil Clin N Am 2021 May;32(2):307-317. [doi: 10.1016/j.pmr.2020.12.010] [Medline: 33814060]
77. Tanner K, Bican R, Boster J, et al. Feasibility and acceptability of clinical pediatric telerehabilitation services. Int J Telerehab 2020 Dec 8;12(2):43-52. [doi: 10.5195/ijt.2020.6336]
78. Lindsay S, Ragunathan S, Kingsnorth S, et al. Understanding the benefits and challenges of outpatient virtual care during the COVID-19 pandemic in a Canadian pediatric rehabilitation hospital. Disabil Rehabil 2024 Jun;46(11):2396-2404. [doi: 10.1080/09638288.2023.2221902] [Medline: 37306595]
79. Ben-David Y, Ickeson T, Kaye-Tzadok A. Lost in the matrix: dialectical tensions in facilitating virtual video groups during COVID-19 pandemic. Internet Interv 2021 Dec;26:100445. [doi: 10.1016/j.invent.2021.100445] [Medline: 34485095]
80. Kashyap H, Chandur J, Reddy RP. The loss of the shared space: process issues in telepsychotherapy. Indian J Psychol Med 2020 Sep;42(5):469-472. [doi: 10.1177/0253717620952313] [Medline: 33414595]
81. Tracey I, Gomes E, Signal N, Alder G, Forch K, Olsen S. Highlighting gaps in the reporting of aerobic exercise interventions for mild traumatic brain injury: a systematic review using the TIDieR-Rehab checklist. JSAMS Plus 2025 Dec;6:100120. [doi: 10.1016/j.jsampl.2025.100120] [Medline: 41878248]
82. Yeates KO, Tang K, Zhang C, Pediatric Emergency Research Canada (PERC) A-CAP study team, et al. Symptom trajectories and their biopsychosocial correlates in pediatric concussion: an A-CAP study. J Neurotrauma 2025 Nov;42(21–22). [doi: 10.1177/08977151251365530] [Medline: 40779427]

XSL•FO
RenderX

83. Cook NE, Kissinger-Knox A, Iverson IA, et al. Social determinants of health and health equity in the treatment and rehabilitation of sport-related concussion: a content analysis of intervention research and call-to-action. J Neurotrauma 2024 Oct;41(19-20):2201-2218. [doi: 10.1089/neu.2023.0550] [Medline: 38753708]
84. Cronin E, McCallion M, Monaghan K. "The best of a bad situation?" A mixed methods survey exploring patients' perspectives on physiotherapy-led online group exercise programmes. Ir J Med Sci 2023 Dec;192(6):2595-2606. [doi: 10.1007/s11845-023-03386-7] [Medline: 37154996]
85. Paul SS, Hubbard A, Johnson J, Dennis SM. Transition to a virtual model of physiotherapy and exercise physiology in response to COVID-19 for people in a rural Australia: is it a viable solution to increase access to allied health for rural populations? PLoS ONE 2023;18(1):e0280876. [doi: 10.1371/journal.pone.0280876] [Medline: 36662817]
86. Carroll LJ, Cassidy JD, Holm L, Kraus J, Coronado VG, WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. Methodological issues and research recommendations for mild traumatic brain injury: the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. J Rehabil Med 2004 Feb;36(43 Suppl):113-125. [doi: 10.1080/16501960410023877] [Medline: 15083875]
87. Macpherson AK, Harkins J, Sergio LE, Sadrmanesh O, Emery C, Rothman L. Emergency department visits for concussion in children and youth by age, sex and material deprivation in Ontario, Canada, 2010-2020: a population-based study. Inj Prev 2025 Jul 21;31(4):298-303. [doi: 10.1136/ip-2024-045556] [Medline: 40645644]
88. Wittevrongel K, Barrett O, Couloigner I, et al. Longitudinal trends in incidence and health care use for pediatric concussion in Alberta, Canada. Pediatr Res 2023 May;93(6):1752-1764. [doi: 10.1038/s41390-022-02214-5] [Medline: 36085365]

Abbreviations

CITI: Collaborative Institutional Training Initiative
COREQ: Consolidated Criteria for Reporting Qualitative Studies
FAB: family-directed approach to brain injury
HBKRH: Holland Bloorview Kids Rehabilitation Hospital
***M&C-C*:** Move&Connect-Caregivers
***M&C-Y*:** Move&Connect-Youth
PSaC: persisting symptoms after concussion
TCPS-2: Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans

Edited by J Sarvestan; submitted 01.Dec.2025; peer-reviewed by S Huang, TA Wani; revised version received 30.Jun.2026; accepted 01.Jul.2026; published 21.Sep.2026.

Please cite as:
Hickling A, Al-Hakeem H, Lovell A, Madge-Nyren S, Mallory KD, Ryan JL, Provvidenza CF, Wilcock R, Knapp P, Knapp B, Wade SL, Yeates KO, Scratch SE
Delivering Virtual Group Interventions for Youth With Persisting Symptoms After Concussion (Move&Connect-Youth) and Their Caregivers (Move&Connect-Caregivers): Fidelity and Engagement Evaluation
JMIR Rehabil Assist Technol 2026;13:e88771
URL: https://rehab.jmir.org/2026/1/e88771
doi:10.2196/88771

© Andrea Hickling, Hiba Al-Hakeem, Andrew Lovell, Sophie Madge-Nyren, Kylie D Mallory, Jennifer L Ryan, Christine F Provvidenza, Ruth Wilcock, Patricia Knapp, Brenda Knapp, Shari L Wade, Keith Owen Yeates, Shannon E Scratch. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 21.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Asynchronous Broadcasting of Audiovisual Content as a Telerehabilitation Strategy for Patients in Rural Areas: Development and Usability Study

Lilia Aparicio Pico, PhD; Roberto Ferro Escobar, PhD; Paulo César Coronado Sánchez, MSc

Faculty of Engineering, Universidad Distrital Francisco José De Caldas, Carrera 7 # 40B - 53, Bogotá DC, Colombia

Corresponding Author:
Lilia Aparicio Pico, PhD
Faculty of Engineering, Universidad Distrital Francisco José De Caldas, Carrera 7 # 40B - 53, Bogotá DC, Colombia

Abstract

Background: Geographical and economic barriers limit access to health care services in rural regions of Colombia. In San Vicente del Caguán, the lack of infrastructure and rehabilitation professionals forces patients to travel long distances. Asynchronous telerehabilitation using video broadcasting is a viable strategy to address these challenges.

Objective: This study aims to design and validate a telerehabilitation model using asynchronous audiovisual content broadcasting for rural patients, evaluating functionality, usability, and clinical effectiveness.

Methods: A 4-stage case study developed and validated the model in San Vicente del Caguán: (1) analysis of telemedicine experiences and video-based therapy; (2) solution design including telecommunications infrastructure (radio links and Wi-Fi), mobile app (HSRehabiAPP), and web platform (HSRehabiWEB); (3) fieldwork with 7 patients receiving physical, occupational, or speech therapy, evaluating functionality (11 criteria), usability (8 criteria), and content quality (5 criteria); and (4) results analysis. The infrastructure connected San Rafael Hospital with remote centers in Los Pozos and Tres Esquinas. Participants (aged 7-68 years) from urban and rural areas had conditions including stroke, shoulder injuries, knee pathologies, hypertension, and attention-deficit hyperactivity disorder.

Results: All 7 patients achieved 100% compliance across functional, usability, and audiovisual content criteria. Functional evaluation covered login, navigation, therapy access, session viewing, exercise execution, pain assessment, therapist communication, and satisfaction surveys. Usability assessment evaluated initial access, content location, navigation comfort, instructional guidance, session organization, video playback, instruction clarity, and interface intuitiveness. Content criteria included exercise clarity, step-by-step instructions, visual quality, audio quality, and correct posture demonstration. Patients reported high satisfaction, noting reduced travel costs and time, family convenience, and effective outcomes. Offline functionality proved essential in areas with limited internet connectivity.

Conclusions: The asynchronous audiovisual telerehabilitation model is an effective solution for improving access to rehabilitation services in rural areas. It successfully addressed geographical barriers and infrastructure limitations while maintaining clinical effectiveness across therapies. Implementation requires adequate technological infrastructure, user-friendly platforms with offline capabilities, and quality therapeutic content. Future work demands inclusive health policies, professional training, and research with larger sample sizes to assess long-term sustainability in diverse rural contexts.

(JMIR Rehabil Assist Technol 2026;13:e66343) doi:10.2196/66343

KEYWORDS

telerehabilitation; asynchronous; rehabilitation; telemedicine; telehealth; telebroadcasting; virtual therapy; video broadcasting; audiovisual; broadcasting; broadcast; case study; physical therapy; physiotherapy; occupational therapy; speech therapy; model; rural; access; accessibility; service; delivery

Introduction

Geographical barriers, population dispersion, and economic limitations hinder timely and continuous access to health care, especially in rural areas [1]. In addition, the lack of technological infrastructure and health care professionals forces patients to travel long distances to receive treatment [2]. In this context, telemedicine and telerehabilitation [3], which involve the use of information and communication technologies (ICT) to provide remote rehabilitation services [4], become particularly relevant in rural and hard-to-reach areas where health care resources are scarce [5].

A critical case in Colombia can be observed in San Vicente del Caguán, a municipality in the Department of Caquetá, where social conflict, coupled with difficult access and limited availability of physical, occupational, and speech therapy

XSL•FO
RenderX

rehabilitation services, is exacerbated by the region's geographical and topographical conditions. Although the San Rafael Hospital in the municipality offers rehabilitation services, it does not meet the demand from patients in rural areas. Consequently, the implementation of a telerehabilitation model based on the asynchronous broadcasting of audiovisual content emerges as a viable strategy to improve access to rehabilitation services and support the comprehensive recovery of patients facing geographical barriers. This proposal includes the use of interactive audiovisual content to facilitate patient education and remote monitoring as a solution for physical rehabilitation [6], while also training patients and their families, promoting self-management of health, and reducing dependence on in-person health care services [7]. The proposed infrastructure consists of a radio link and a mobile app containing interactive videos to asynchronously guide telerehabilitation patients.

Methods

Overview

The development of the model consists of 4 stages. The first stage is a case study analysis, aimed at identifying successful experiences to gather references on key aspects that impact the target groups served through telemedicine and telerehabilitation modalities. These aspects include patient satisfaction with the services provided, opportunities to improve access, cases of technological innovation, and technology-centered approaches with coverage analysis. In addition, important considerations regarding the creation and use of videos in therapy are taken into account. This stage also includes a specific case study focused on the regional conditions and services offered by San Rafael Hospital in San Vicente del Caguán and its surrounding areas, addressing telerehabilitation for physical therapy, occupational therapy, and speech therapy. The second stage focuses on designing the proposed solution to support the strategy's model. The third stage involves fieldwork, including functional testing, usability testing, and audiovisual content evaluation, for which evaluation instruments were applied. Finally, the fourth stage involves analyzing the results to establish the final conclusions.

Case Study Analysis

Case Analysis Focused on Service and Patients

Successful cases of telerehabilitation processes and medical consultations carried out through telemedicine platforms are presented in the study by Finak et al [8], highlighting the application of remote communication technologies for physical rehabilitation, occupational therapy, and chronic pain management [9], particularly for rural populations [10,11]. In addition, the following aspects were analyzed:

1. Patient satisfaction: the reviewed studies show high patient satisfaction with telemedicine and telerehabilitation interventions [12]. Similarly, rural patients reported high satisfaction, mainly due to reduced travel costs and time [13].
2. Access to care: telemedicine was found to be not only a viable but also a necessary solution to ensure equitable access to health care [14]. Other studies highlight improvements in the management of chronic diseases in rural areas, such as stroke management in China [15].
3. Technological innovation: the studies emphasize the innovative use of technologies such as mobile apps, eHealth platforms, and remote monitoring devices [16]. These tools enhance treatment adherence and chronic disease management. For example, improvements were reported in participation and self-efficacy among wheelchair users through an eHealth program [17].
4. Consistency measures: consistency across findings suggests that telemedicine and telerehabilitation technologies are effective across a variety of contexts and populations. Although the magnitude of the effects may vary, the general trend indicates improved accessibility, satisfaction, and clinical effectiveness [18].

Case Analysis Focused on Technologies

In this aspect, a review was conducted to understand how the type of technology used and the characteristics of the population influence the effectiveness of telemedicine and telerehabilitation interventions [19]. Effective components in different clinical contexts were analyzed, aiming to optimize technologies to maximize their benefits in medical practice [20].

Regarding access and coverage of telemedicine services in rural and remote areas, significant challenges were identified, which have been addressed through various telecommunication technologies. Key factors in choosing appropriate technology in contexts with limited communication infrastructure, such as rural areas [21], include the scarcity of terrestrial networks and the need for sustainable solutions. The use of technologies such as Very Small Aperture Terminals (VSAT) [22] is justified by their availability and adaptability to local needs, highlighting cost-effectiveness and broad coverage, especially in rural areas. Furthermore, in extremely remote locations such as Maewo Island [23], satellite technology emerges as the most viable solution [24].

For mobile networks in telemedicine, efforts focus on integrating patients and health care professionals through advanced ICT such as 5G, which offers high bandwidth and the capacity to quickly transmit large medical files [25], enabling timely and informed decision-making about patient health [26]. In addition, the future scalability of 5G and its ability to easily upgrade to new technologies are key advantages. The adaptation to available communication infrastructure, along with diverse technological options such as Code Division Multiple Access (CDMA 2000 1X) [27] and Global System for Mobile Communications (GSMC) [28,29], enhances availability and connection stability in both urban and rural environments, complemented by the use of ad hoc mobile networks such as Mobile Ad Hoc Networks (MANET) and Mobile Internet Protocol (Mobile IP) [30].

Worldwide Interoperability for Microwave Access (WiMAX) is another useful technology for covering large distances while providing adequate bandwidth for telemedicine applications, such as the transmission of high-resolution medical images and video [31].

XSL•FO
RenderX

Fiber optic infrastructure and radio networks also enhance access. In the study by Baena and Bríñez [32], a high-speed backbone network is shown to enable the creation of a faster and more robust telemedicine network, compared to other technologies such as satellite connections or radio links. In another case, radiofrequency-based solutions [33] connect a hospital with rural health posts in remote areas using a radio link.

In summary, the technologies used in telemedicine are diverse and vary according to the specific needs of each environment. WiMAX stands out for its high bandwidth and its ability to prioritize critical data. Studies presented in Niyato et al [34], Tchao et al [35], and Aliyu et al [36] further highlight these advantages. The use of portable devices and long-range communication protocols is also important. In the study by Pramanik et al [37], Long Range (LoRa) technology is selected for its broadband wireless access capabilities and security features, facilitating effective transmission of medical data.

Wireless Body Area Networks (WBAN) combined with cloud computing enable remote patient monitoring and real-time health data analysis. An example of remote health monitoring can be found in Suleiman and Adinoyi [38].

A description of how 5G technology, artificial intelligence, machine learning, and the Internet of Medical Things (IoMT) are transforming health care delivery is provided in the study by Latha et al [39]. In addition, WBAN technology [40] and cloud computing are used to create a telemedicine framework, as shown in the study by Park et al [40]. Finally, LTE-Advanced (LTE-A) can also be selected for its higher bandwidth, lower latency, better coverage, increased capacity, and improved quality of service, as highlighted in Kress and Leeuwen [41].

Application Case in San Vicente Del Caguán

To analyze the health care access issues in this region of Colombia, it is essential to consider the years of armed conflict experienced in the Department of Caquetá. This context has hindered progress in various sectors, such as infrastructure, education, and health, limiting access to essential basic services. These deficiencies, recognized as unsatisfied basic needs, not only contradict the Sustainable Development Goals but also widen socioeconomic and territorial inequality gaps in the region.

In the Department of Caquetá, medium-complexity health services are centralized in Florencia, the department's capital. In some municipalities, specialist services (such as gynecology, cardiology, internal medicine, and pediatrics) are only offered through mobile health brigades. Outside the capital, municipalities only provide low-complexity health services, which results in an uneven distribution of specialists across the department. In addition, there is insufficient supply, high demand, long waiting times for appointments managed by health entities, and geographical limitations.

The San Rafael Hospital in San Vicente del Caguán has recently opened a Comprehensive Rehabilitation Center, offering physical therapy, occupational therapy, and speech therapy, making it the only municipality outside Florencia with these 3 rehabilitation services available.

San Vicente del Caguán covers an area of 28,300 km² and has a population of approximately 54,575 inhabitants, where 51.3% of the population lives in rural areas, according to 2024 projections by Departamento Administrativo Nacional de Estadística [42]. The municipality is made up of 287 rural settlements distributed across 14 police inspection zones within the municipal territory. The primary access routes are dirt roads (unpaved), which significantly complicate the journey of rural patients traveling from their villages to San Rafael Hospital for various health care procedures.

In addition, there is no internet access at health service points in the region. However, a significant number of inhabitants own mobile phones, which they use for telephone calls and, when possible, access mobile data services in communal zones where coverage is available.

Considering these regional conditions, the proposed telerehabilitation system, based on broadcasting therapy content (physical therapy, occupational therapy, and speech therapy) through an asynchronous computer-based format, offers a viable solution to expand health care service coverage to rural areas served by San Rafael Hospital. This model would also overcome geographical barriers and internet access limitations, improving health care delivery in the region.

Design and Production of Audiovisual Content for Therapy

A video is an audiovisual content format that combines visual and audio elements to facilitate the understanding of a message and enhance information retention [43-45]. Videos can vary in format, duration, and purpose and are widely used across different industries, including education, entertainment, advertising, and health care, particularly in telerehabilitation where digital platforms have shown effectiveness [46,47].

Based on these criteria, teletherapy is developed as a care modality that uses ICTs, including mobile apps and video-based content, to deliver remote therapeutic rehabilitation services [46].

Considerations for the Use of Videos in Telerehabilitation

When supporting the use of videos in teletherapy, it is necessary to consider the following aspects: the positive impacts that are reflected in the improvement of the understanding of exercises, by providing a visual representation, allow patients to observe precise movements, postures, and appropriate techniques, which reduces the likelihood of errors and increases the accuracy in the execution of exercises [46]. In turn, adherence to rehabilitation treatment is another aspect that is strengthened with the use of videos.

Another significant impact is the reduction of the patients' dependence on the therapist. Video-based interventions and visual feedback allow patients to perform exercises more independently, providing a constant guide that they can review as many times as necessary [48]. In Latin America, several studies have also demonstrated significant positive impacts. Video-based therapeutic interventions for rehabilitation resulted in significant improvements in mobility and pain reduction [49].

XSL•FO
RenderX

Technology-based rehabilitation interventions for poststroke patients helped improve motor coordination and reduce recovery time [50]. Technology-assisted rehabilitation programs have reported improvements in aerobic capacity and reduction of fatigue in patients with chronic conditions [51]. Rehabilitation programs have shown faster recoveries and pain reduction in postsurgical hip replacement patients [52].

The design of rehabilitation videos requires a methodical approach to ensure that audiovisual content is accurate, understandable, and accessible to all patients. Studies have highlighted [53] the importance of a systematic approach to creating educational videos for rehabilitation.

The need for a structured process for the design of audiovisual content has been emphasized in rehabilitation research [53], highlighting the importance of conducting pilot tests and collecting patient feedback to improve the quality of the videos, an iterative approach that makes it possible to adjust the content to make it more effective and accessible.

Finally, it is important to take into account that the methods or processes that guide the design of videos in telerehabilitation should consider the following recommendations: having the collaboration of multidisciplinary professionals [53], the duration of the video should be short enough to maintain the patient's attention, but long enough to cover all important aspects of the exercise or technique [46], including clear and descriptive visual elements, as effective visual communication is essential in digital health interventions [54], to address the diversity of patients' needs, the videos must be adaptable and customizable [51], the creation of videos must be an iterative process that includes evaluation and continuous feedback from patients and health care professionals, as demonstrated in recent systematic reviews [55], and it is important to consider options for downloading videos in areas with limited internet access [42].

Development of the Telerehabilitation Model for the Implementation of the Content Broadcasting Strategy

Overview

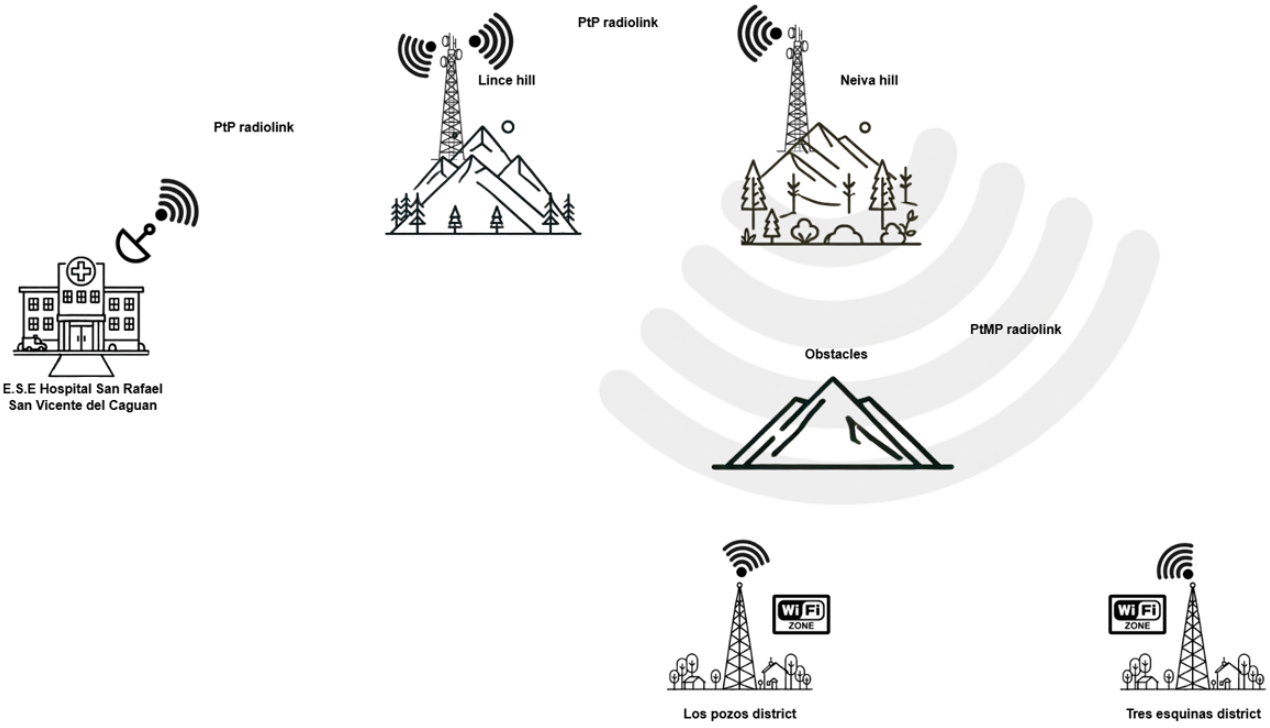
The following are the 3 components of the model, consisting of the telecommunications infrastructure, the mobile app, and the design of audiovisual content with its mode of operation. The teletherapy services implemented support the telerehabilitation services for physiotherapy, occupational therapy, and speech therapy.

The solution has 3 components: a telecommunications component supported by a wireless network based on radiofrequency solutions (long-range and high-bandwidth radio links) and Wi-Fi technologies (802.11a, 802.11b, 800.11n, 802.11ac, and 802ax), which enable the interconnection of a hospital with prioritized and remote health posts far from the municipal or urban center. A software platform (Progressive Web Application [PWA]; Google), user-friendly and interactive, is complemented by the creation of audiovisual content (videos and animations) that are stored in a cloud repository or are synchronously and programmatically loaded with a rehabilitation therapy plan onto the patient's mobile device. This solution can be available for direct access to the cloud or located at remote or rural centers for loading onto the patient's mobile device to be used later at home. The platform allows the administration and configuration of therapies that will be scheduled for patients according to their medical history, needs, and clinical records.

The Implemented Network

Figure 1 shows the general configuration of the telecommunications system.

Figure 1. General configuration of the telecommunications system. (Own elaboration).



XSL•FO
RenderX

The network is made up of a radio link that guarantees access to rehabilitation services in the telemedicine modality to connect the E.S.E. Hospital San Rafael with remote health centers with difficult access to serve the communities of Los Pozos and Tres Esquinas. The network is composed of several components described below.

E.S.E. Hospital San Rafael in San Vicente del Caguán: from this point, the telerehabilitation service is coordinated, bringing together specialists, training patients referred to teletherapy, and providing follow-up through tele-expertise to patients by connecting doctors and nurses through the radio link. These professionals are responsible for maintaining direct contact with patients once they start follow-up when they come to the center for their check-ups. Finally, monitoring is carried out via telemonitoring and synchronous and asynchronous teleconsultation.

The system has 2 signal relay hops at Lince Hill and Cerro Neiva. This configuration allows communication coverage to reach the remote areas.

Figure 2 shows the installed infrastructure at each of the described points. Figure 2A shows the antenna installed at Hospital San Rafael. Figure 2B presents the configuration at Lince Hill, which maintains a direct line with the hospital. Figure 2C shows the configuration at Cerro Neiva, which maintains a link with Lince Hill. From Cerro Neiva, the highest point in the configuration, communication is established with Tres Esquinas and Los Pozos, as shown in the configuration in Figure 2D for each point.

Figure 2. Installation of the system antennas.



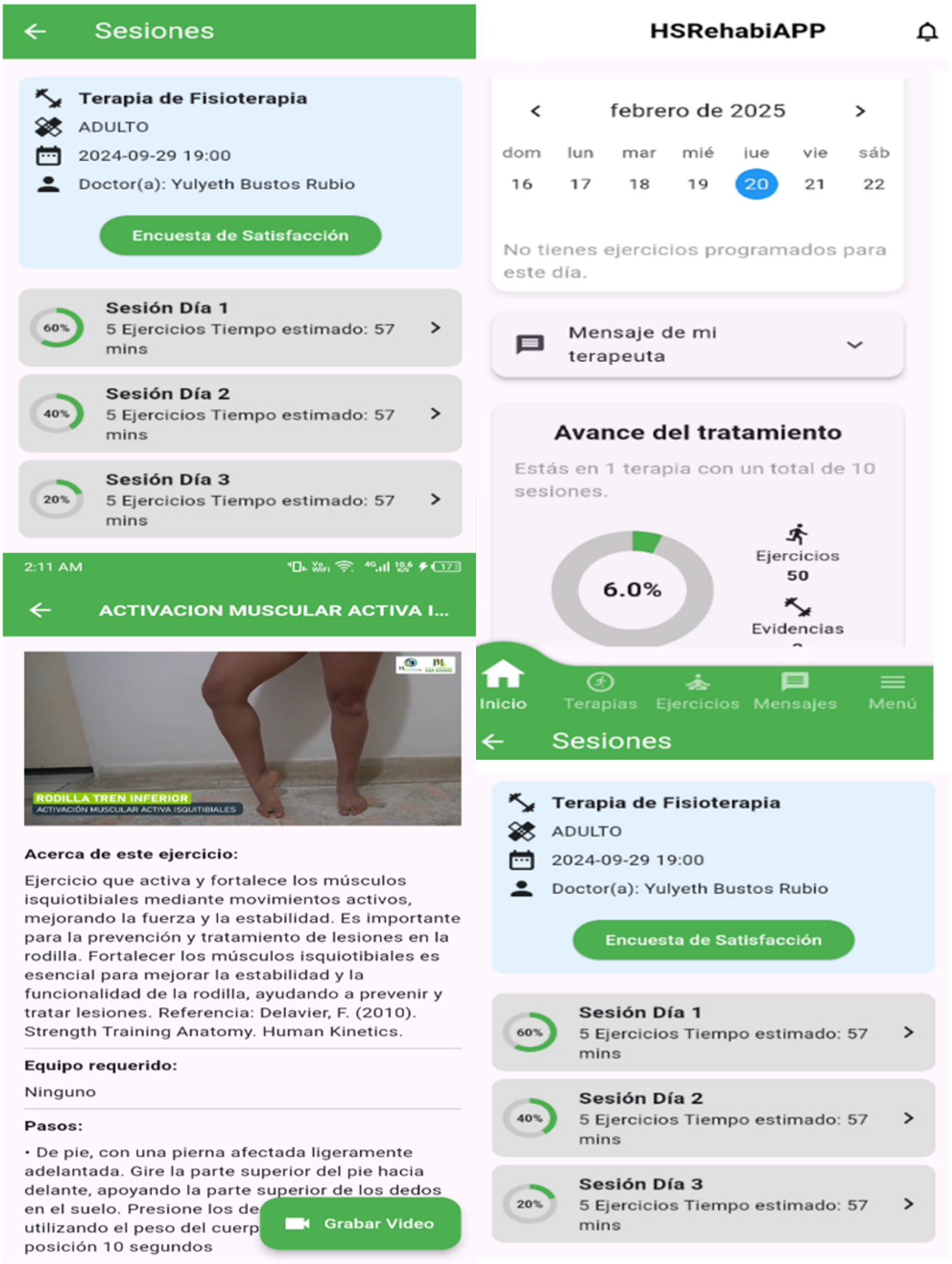
XSL•FO
RenderX

Implemented Software System

The software system of the solution follows a layered architecture that ensures separation of responsibilities. The presentation layer consists of 2 apps. The first one, HSRehabiAPP, shown in Figure 3, is a mobile app designed for online and computer-based use, where users can access their assigned therapies, obtain exercise instructions, prepare evidence, monitor their progress, and securely communicate with the assigned health care professional.

XSL•FO
RenderX

Figure 3. HSReHabiAPP mobile app.



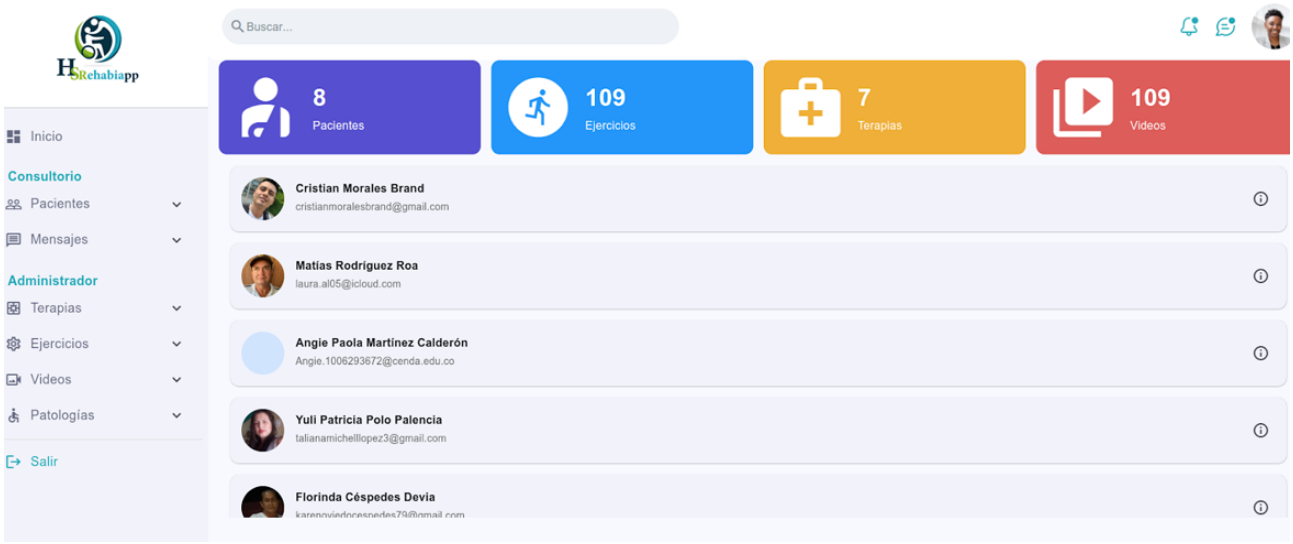
A critical aspect in the design of the mobile app is its offline functionality. This is particularly important considering that most patients are in rural areas with little or no internet connectivity. Data synchronization is fully automated and takes place whenever patients have a connection, using a publish-subscribe model that optimizes data transfer.

The second app, HSRehabiWEB (Figure 4), is a web app where health care professionals manage the catalog of videos and

XSL•FO
RenderX

exercises, configure therapies, register their patients, and monitor their progress. This app serves as the control center of the solution.

Figure 4. Main page of HSRehabiWEB.



The application layer consists of an application programming interface, which provides access points to the application logic (Backend). The application layer is deployed on the AWS platform (Amazon) to ensure adequate quality of service. The data layer, also deployed on AWS, implements a relational model in a PostgreSQL database (Amazon).

Field Work

Overview

A pilot test was carried out between the San Rafael Hospital in San Vicente del Caguán and the centers of Pocos and Tres Esquinas in rural areas, to evaluate the implementation and effectiveness of the telerehabilitation model based on asynchronous transmission of audiovisual content in the municipality of San Vicente del Caguán, Colombia. The test included 7 patients in the telerehabilitation program, which covered physical, occupational, and speech therapies. The criteria observed considered functional tests, usability criteria, and content evaluation.

Description of Participants

Table 1 summarizes the demographic, clinical, and geographic characteristics of the patients in the trial.

Table . Demographic, clinical, and geographic characteristics of the patients in the pilot test of the telerehabilitation model.

ID	Name	Area	Pathology	Life cycle	Age (y)	Zone	Disability	Surgeries
P-1	Cristian Iván Morales Brand	Physical therapy	Left shoulder dislocation	Adult	25	Urban	No	Shoulder surgery
P-2	Matías Rodríguez Roa	Physical therapy	Stroke with left hemiplegia	Adult	55	Rural	Physical	Stroke
P-3	Yuli Patricia Polo Palencia	Physical therapy	Rotator cuff syndrome (right shoulder)	Adult	Not specified	Urban	No	No
P-4	Florinda Céspedes Debia	Speech therapy	Hypertension	Older adult	68	Rural	No	Hypertension
P-5	Emmanuel Lozano Ríos	Occupational therapy	Attention-deficit hyperactivity disorder	Early childhood	7	Rural	No	No
P-6	Jheisson Aguilera Betancourt	Physical therapy	Patella dislocation and right meniscus tear	Adult	31	Rural	No	Right knee surgery
P-7	Cristian Camilo Castro Cuéllar	Physical therapy	Knee sprain	Adult	27	Urban	No	No

XSL•FO
RenderX

Evaluation Instruments

Table 2 presents the criteria evaluated during the tests of the telerehabilitation model and its app, organized by number, specific criterion, and corresponding description. The evaluation covered 3 main areas: functional testing, usability testing, and audiovisual content assessment. Functional testing verified key aspects related to the usability, accessibility, and functionality of the app. Usability testing focused on identifying strengths and opportunities for improvement in the user experience by analyzing different aspects of user interaction with the app. Additionally, the audiovisual content evaluation examined the quality and effectiveness of the multimedia materials used in the telerehabilitation model to ensure that patients could perform the exercises properly and safely.

XSL•FO
RenderX

Table . Evaluation for CF[a], CU[b], and CA[c] in the telerehabilitation model.

Number	Criterion	Description
CF-1	App login	Were you able to successfully log in using the email and password provided during registration?
CF-2	Navigation through the patient's main page	Were you able to easily access and navigate through the patient's main page, including the calendar, exercises, and messages sections?
CF-3	Access to the therapy module	Were you able to enter the therapy module and view the assigned therapies and their progress?
CF-4	Viewing scheduled therapy sessions	Were you able to access the scheduled sessions for a specific therapy and monitor their progress?
CF-5	Viewing exercises scheduled per session	Were you able to access the exercises defined for each session and track their progress?
CF-6	Execution of scheduled exercises	Were you able to see the instructions, illustrative videos, and other necessary details to correctly perform the exercises?
CF-7	Recording exercise execution	Does the app allow you to properly record and save the execution of an exercise?
CF-8	Pain assessment after the session	Were you able to enter and evaluate your pain level at the end of the session using a 0 to 10 pain scale?
CF-9	Communication with the therapist	Were you able to send messages to your therapist and receive responses through the message module?
CF-10	Therapy satisfaction assessment	Were you able to access and complete the satisfaction survey after completing the therapy sessions?
CF-11	Navigation through the side menu	Were you able to access additional features such as "My therapies," "Messages," and "Settings" through the side menu?
CU-1	Initial access	Was it easy for you to log into the app for the first time?
CU-2	Assigned content	Were you able to quickly find the videos of your assigned exercises?
CU-3	Navigation	Do you feel comfortable navigating through the app menus?
CU-4	Instructional guide	Did you find the step-by-step guide provided by the app at startup helpful?
CU-5	Session organization	Do you find the way your rehabilitation sessions are displayed clear?
CU-6	Video playback	Were you able to play the exercise videos without any issues?
CU-7	Therapist instructions	Was it easy for you to follow the therapist's instructions through the app?
CU-8	Intuitive interface	Do you feel that the app's interface is intuitive and easy to understand?
CA-1	Clarity in the presentation of the exercise	Do the videos clearly present the name and objective of the exercise?
CA-2	Step-by-step instructions	Do the videos adequately explain each step of the exercise in a way that is understandable for the patient?
CA-3	Visual quality	Is the video image quality sufficient to clearly observe the exercise movements?

XSL•FO
RenderX

Number	Criterion	Description
CA-4	Audio quality	Is the audio in the videos clear, and is the verbal explanation of the exercise understandable without interference?
CA-5	Correct postures movements	Does the video highlight the correct postures and proper movements, showing how to perform the exercise safely?

[a]CF: functional criteria.

[b]CU: usability criteria.

[c]CA: audiovisual content criteria.

Ethical Considerations

Ethics Approval

This study was approved by the Research Ethics Committee of Universidad Distrital Francisco José de Caldas. All procedures performed in this study involving human participants were conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its subsequent amendments. Written informed consent was obtained from all participants included in the study (or from their legal guardians in the case of minors) prior to their participation. Participants were provided with detailed information regarding the nature and objectives of the telerehabilitation model, the procedures involved, the potential risks and benefits, and their right to withdraw from the study at any time without penalty. All participants were given sufficient time to consider their participation and had the opportunity to ask questions before signing the informed consent forms.

Privacy and Confidentiality

Patient data, including demographic information, medical records, therapy videos, and evaluation results, were securely stored in encrypted digital formats with access restricted exclusively to authorized members of the research team. The mobile app (HSRehabiAPP) and the web platform (HSRehabiWEB) were designed with security measures aimed at protecting personal health information, in compliance with Colombian data protection regulations (Law 1581 of 2012 on personal data protection). Video recordings used for therapy sessions were stored on secure servers and were accessible only to the assigned health care professionals and the corresponding patient. Participants were informed about data storage, access, and usage procedures through the informed consent process.

Vulnerable Populations

Additional ethical considerations were adopted to ensure adequate protection of these participants. In the case of the older adult participant, cognitive capacity to provide informed consent was assessed prior to enrollment in the study. For the minor participant, both parental informed consent and child assent were obtained, and therapy sessions were adapted according to the child's age and conducted under parental supervision. Given the rural and potentially socioeconomically vulnerable condition of participants from Los Pozos and Tres Esquinas, measures were taken to ensure that participation was entirely voluntary and that individuals did not feel coerced to participate due to limited access to alternative rehabilitation services.

Data Sharing and Dissemination

Anonymized participant data will not be made publicly available due to the small sample size and the potential risk of reidentification, particularly considering the specific rural and clinical context of the participants. Aggregate results and findings are presented in this publication in a manner that protects the individual confidentiality of participants.

Results

Table 3 presents the results obtained from the evaluation of the telerehabilitation model, including functional tests, usability tests, and audiovisual content assessment conducted with 7 patients (P-1 to P-7). The evaluated criteria are organized in rows, while the patients are presented in columns, indicating whether each criterion was met. Functional test results describe the performance of the telerehabilitation app according to the established functional criteria. Usability test results analyze the interaction of patients with the app, identifying the fulfillment of usability criteria (CU-1 to CU-8). Additionally, the audiovisual content evaluation presents the results associated with 5 assessment criteria (CA-1 to CA-5), focused on the quality and effectiveness of the multimedia materials used to support the correct and safe execution of rehabilitation exercises within the telerehabilitation model.

XSL•FO
RenderX

Table . Results for CF[a], CU[b], and CA[c] by patient in the telerehabilitation model[d].

Patient \ Criterion	P-1	P-2	P-3	P-4	P-5	P-6	P-7
CF-1	✔	✔	✔	✔	✔	✔	✔
CF-2	✔	✔	✔	✔	✔	✔	✔
CF-3	✔	✔	✔	✔	✔	✔	✔
CF-4	✔	✔	✔	✔	✔	✔	✔
CF-5	✔	✔	✔	✔	✔	✔	✔
CF-6	✔	✔	✔	✔	✔	✔	✔
CF-7	✔	✔	✔	✔	✔	✔	✔
CF-8	✔	✔	✔	✔	✔	✔	✔
CF-9	✔	✔	✔	✔	✔	✔	✔
CF-10	✔	✔	✔	✔	✔	✔	✔
CF-11	✔	✔	✔	✔	✔	✔	✔
CU-1	✔	✔	✔	✔	✔	✔	✔
CU-2	✔	✔	✔	✔	✔	✔	✔
CU-3	✔	✔	✔	✔	✔	✔	✔
CU-4	✔	✔	✔	✔	✔	✔	✔
CU-5	✔	✔	✔	✔	✔	✔	✔
CU-6	✔	✔	✔	✔	✔	✔	✔
CU-7	✔	✔	✔	✔	✔	✔	✔
CU-8	✔	✔	✔	✔	✔	✔	✔
CA-1	✔	✔	✔	✔	✔	✔	✔
CA-2	✔	✔	✔	✔	✔	✔	✔
CA-3	✔	✔	✔	✔	✔	✔	✔
CA-4	✔	✔	✔	✔	✔	✔	✔
CA-5	✔	✔	✔	✔	✔	✔	✔

[a]CF: functional criteria.

[b]CU: usability criteria.

[c]CA: audiovisual content criteria.

[d]Convention: ✔: Compliant

Discussion

Principal Findings

The considerations made based on the literature analysis to establish the system solution and the telerehabilitation strategy using audiovisual content broadcasting facilitated the validation of the proposal and verification through the case study in San Vicente del Caguán.

Regarding improved access to care, it was observed that telemedicine and telerehabilitation facilitate access to health care services, especially in regions with limited infrastructure.

In terms of patient satisfaction, participants expressed high satisfaction with these services, highlighting the timeliness of care and the convenience for families during recovery processes. In addition, patients noted a reduction in travel costs (both in time and money) and satisfactory levels of treatment effectiveness, achieving acceptable recovery outcomes across the 3 therapy types.

On the other hand, regarding the feasibility in rural health care centers, the technologies implemented proved effective in rural areas, where technology helps address the shortage of specialists to meet service demand. The ease of extending coverage was noted, improving the hospital's service offering, while patients readily accepted and engaged with asynchronous computer-based therapies.

It was necessary to develop training and continuous education tools for all community members involved in the process. In addition, this initiative opened the opportunity to formulate inclusive health policies, recognizing the need to cover diverse conditions and populations, particularly those affected by economic and technological limitations.

XSL•FO
RenderX

Finally, the need for ongoing evaluation and monitoring was established to ensure the impact and effectiveness of technology in supporting patient recovery cases.

Conclusions

The implementation of the telerehabilitation model based on the broadcasting of audiovisual content has proven to be a highly effective strategy for improving access to rehabilitation services in rural areas.

In the case study conducted in San Vicente del Caguán, patients showed high levels of satisfaction with the telerehabilitation interventions. The use of interactive audiovisual content not only facilitated rehabilitation in physical therapy, occupational therapy, and speech therapy, but also empowered patients and their families, promoting greater self-management of their health.

Findings from the literature review highlight the need to integrate telerehabilitation into public health policies, especially in rural contexts. Health authorities should develop appropriate technological infrastructure and promote continuous training for health care professionals to maximize the benefits of these technologies. Furthermore, inclusive policies should be implemented to ensure equitable access to telemedicine and telerehabilitation services for all populations, regardless of economic or technological limitations.

Several areas for future research were identified, including studies with larger sample sizes and long-term follow-up to assess the sustainability of the observed effects. It is also essential to explore the implementation of these technologies in low- and middle-income settings, as well as to investigate new technologies and innovative approaches that could further enhance the effectiveness and accessibility of telerehabilitation.

The positive effects of telemedicine and telerehabilitation interventions remain consistent across studies, regardless of the specific technology used or the target population. This indicates that the observed benefits are not specific to a particular context but are generalizable to various situations and populations.

Despite the variability in the evaluation methods used in the studies, including satisfaction surveys, clinical measures, and cost analyses, the overall findings remained robust. This suggests that telemedicine and telerehabilitation technologies are consistently effective, regardless of how outcomes are measured.

Acknowledgments

Special thanks to all the staff at Hospital San Rafael de San Vicente del Caguán for welcoming and supporting this solution, and to all the personnel at Universidad Distrital Francisco José de Caldas for their work in the execution of the project in San Vicente del Caguán. We also extend our gratitude to the Ministry of Science, Technology and Innovation of Colombia, the Telecommunications Regulation Commission (CRC) of Colombia, and Universidad Distrital Francisco José de Caldas, who founded the research and application project.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Loy E, Scheidler A, Packham T, Dow H, Winston P. Patient satisfaction with virtual evaluation, diagnosis, and treatment of CRPS. Can J Pain 2022;6(1):77-84. [doi: 10.1080/24740527.2022.2063113] [Medline: 35694140]
2. Wajapey PB, Shenoy AB. Application of telemedicine in pediatric physiotherapy. Physiother J Indian Assoc Physiother 2021;15(1):5-11. [doi: 10.4103/PJIAP.PJIAP_49_20]
3. Ouendi N, Avril E, Dervaux B, Pudlo P, Wallard L. Effectiveness of telerehabilitation programs in elderly with hip or knee arthroplasty: a systematic review. Telemed J E Health 2024 Jun;30(6):1507-1521. [doi: 10.1089/tmj.2023.0622] [Medline: 38574249]
4. Frenzel JE, Porter AL. Design and assessment of telepharmacy and telehealth training in two pharmacy programs. Am J Pharm Educ 2023 Mar;87(2):ajpe8800. [doi: 10.5688/ajpe8800] [Medline: 35260413]
5. Hung Kn G, Fong KN. Effects of telerehabilitation in occupational therapy practice: a systematic review. Hong Kong J Occup Ther 2019 Jun;32(1):3-21. [doi: 10.1177/1569186119849119] [Medline: 31217758]
6. Bailey JE, Gurgol C, Pan E, et al. Early patient-centered outcomes research experience with the use of telehealth to address disparities: scoping review. J Med Internet Res 2021 Dec 7;23(12):e28503. [doi: 10.2196/28503] [Medline: 34878986]
7. Dubois A, Sielker F. Digitalization in sparsely populated areas: between place-based practices and the smart region agenda. Reg Stud 2022 Oct 3;56(10):1771-1782. [doi: 10.1080/00343404.2022.2035707]
8. Finak P, Hastings-Truelove A, Fecica A, Batorowicz B. Therapy services for children and youth living in rural areas of high-income countries: a scoping review. Disabil Rehabil 2023 Jun;45(11):1893-1915. [doi: 10.1080/09638288.2022.2074552] [Medline: 35611460]

XSL•FO
RenderX

9. Kulkarni K, Sevakula RK, Kassab MB, et al. Ambulatory monitoring promises equitable personalized healthcare delivery in underrepresented patients. Eur Heart J Digit Health 2021 Sep;2(3):494-510. [doi: 10.1093/ehjdh/ztab047] [Medline: 34604759]
10. Naeemabadi M, Fazlali H, Najafi S, Dinesen B, Hansen J. Telerehabilitation for patients with knee osteoarthritis: a focused review of technologies and teleservices. JMIR Biomed Eng 2020;5(1):e16991. [doi: 10.2196/16991]
11. Salsabilla A, Azzahra AB, Syafitri RIP, Supadmi W, Suwantika AA. Cost-effectiveness of telemedicine in Asia: a scoping review. J Multidiscip Healthc 2021;14:3587-3596. [doi: 10.2147/JMDH.S332579] [Medline: 35002248]
12. Main F, Zubala A, Gorman J, et al. Technology-enabled remote management of diabetes foot disease and potential for reduction in associated health costs: a pilot study. J Foot Ankle Res 2021 Jan 13;14(1):7. [doi: 10.1186/s13047-020-00444-6] [Medline: 33436028]
13. Harkey LC, Jung SM, Newton ER, Patterson A. Patient satisfaction with telehealth in rural settings: a systematic review. Int J Telerehabil 2020 Dec 8;12(2):53-64. [doi: 10.5195/ijt.2020.6303] [Medline: 33520095]
14. Tommasino N, Megna M, Cacciapuoti S, et al. The past, the present and the future of teledermatology: a narrative review. Clin Cosmet Investig Dermatol 2024;17:717-723. [doi: 10.2147/CCID.S462799] [Medline: 38529172]
15. Pan Y, Xie R, Yan Q, Zhou T. Telemedicine assessment for the mental health of rural residents based on the safety degree of housing in seismically active regions. Front Public Health 2021;9:604298. [doi: 10.3389/fpubh.2021.604298] [Medline: 34408998]
16. Stanimirovic A, Francis T, Shahid N, et al. Tele-retina screening of diabetic retinopathy among at-risk populations: an economic analysis. Can J Ophthalmol 2020 Feb;55(1 Suppl 1):8-13. [doi: 10.1016/j.jcjo.2019.06.008] [Medline: 31712029]
17. Giesbrecht E, Faieta J, Best K, Routhier F, Miller WC, Laberge M. Impact of the TEAM Wheels eHealth manual wheelchair training program: study protocol for a randomized controlled trial. PLoS One 2021;16(10):e0258509. [doi: 10.1371/journal.pone.0258509] [Medline: 34644350]
18. Smith AH, Touchett H, Chen P, et al. Patient satisfaction with a coach-guided, technology-based mental health treatment: qualitative interview study and theme analysis. JMIR Ment Health 2024 Feb 2;11:e50977. [doi: 10.2196/50977] [Medline: 38306167]
19. Khalid UB, Naeem M, Stasolla F, Syed MH, Abbas M, Coronato A. Impact of AI-powered solutions in rehabilitation process: recent improvements and future trends. Int J Gen Med 2024;17:943-969. [doi: 10.2147/IJGM.S453903] [Medline: 38495919]
20. Bisu AA, Gallant A, Sun H, Brigham K, Purvis A. Telemedicine via satellite: improving access to healthcare for remote rural communities in Africa. Presented at: 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC); Dec 6-8, 2018; Malambe, Sri Lanka p. 1-6. [doi: 10.1109/R10-HTC.2018.8629855]
21. Maria MS, Silvia AN, Beatriz DG, Andrew D, Guillermo PF. Health care in rural areas: proposal of a new telemedicine program assisted from the reference health centers, for a sustainable digitization and its contribution to the carbon footprint reduction. Heliyon 2022 Jul;8(7):e09812. [doi: 10.1016/j.heliyon.2022.e09812] [Medline: 35800252]
22. Srinivasan S, Yoo CS. Deploying telemedicine in capacity constrained contexts: lessons from the vanuatu inter-island telemedicine and learning project. Presented at: TPRC48: The 48th Research Conference on Communication, Information and Internet Policy; Feb 17, 2021 URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3763578 [accessed 2026-05-24]
23. Kheyrati L, Keshvari H, Fatehi F, et al. Presenting a model for telemedicine in earthquake for Iran. Health Emerg Disasters Q 2019;4:157-164. [doi: 10.32598/hdq.4.3.157]
24. Amuomo N. The evolution of GSM Technologies into 5G and the imminent emergence of transformative telemedicine applications: a review. East Afr J inf technol 2020;2(1):8-17 [FREE Full text] [doi: 10.37284/eajit.2.1.131]
25. Dubgorn A, Kalinina O, Lyovina A, Rotar O. Foundation architecture of telemedicine system services based on Health 4.0 concept. SHS Web Conf 2018;44. [doi: 10.1051/shsconf/20184400032]
26. Meethal SP, J J. A low cost connectivity solution for rural mobile telemedicine. Presented at: GHTC '11: Proceedings of the 2011 IEEE Global Humanitarian Technology Conference; Oct 30 to Nov 1, 2011; Seattle, WA p. 506-511. [doi: 10.1109/GHTC.2011.55]
27. Sutjiredjeki E, et al. Development of a mobile telemedicine system with multi-communication links for urban and rural areas in indonesia. : Proceedings of the IEEE; 2007 Presented at: p. 660-663. [doi: 10.1007/978-3-540-68017-8_166]
28. Pathak S, Kumar B. Wireless teleophthalmology: a novel, low-cost, flexible network architecture and its performance evaluation for remote eye care solutions. Telemed J E Health 2017 Sep;23(9):753-762. [doi: 10.1089/tmj.2016.0196] [Medline: 28304241]
29. Viswacheda DV, Barukang L, Hamid MY, Arifianto MS. Performance evaluation of mobile ad hoc network-based communications for future mobile tele-emergency systems. J Appl Sci 2007 Jul 15;7(15):2111-2119. [doi: 10.3923/jas.2007.2111.2119]
30. Chaudhari K, Karule PT. WiMAX network based E health service and telemedicine applications for rural and remote populations in india. Presented at: 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom); Nov 7-8, 2014; Greater Noida, India p. 398-406. [doi: 10.1109/MedCom.2014.7006041]
31. Chowdhury M. A telecommunication network architecture for telemedicine in Bangladesh and its applicability. Int J Digit Content Technol Appl 2009;3(3). [doi: 10.4156/jdcta.vol3.issue3.20]

XSL•FO
RenderX

32. Baena JET, Bríñez RC. Caracterización de radioenlaces para interconectar el nuevo hospital de La Candelaria con sus puestos de salud en Purificación para el desarrollo de la telemedicina. Cuaderno Activa 2019;11(1):121-136. [doi: 10.53995/20278101.579]
33. El Khaddar MA, Harroud H, Boulmalf M, Elkoutbi M, Habbani A. Emerging wireless technologies in e-health trends, challenges, and framework design issues. Presented at: 2012 International Conference on Multimedia Computing and Systems; May 10-12, 2012; Tangiers, Morocco p. 440-445. [doi: 10.1109/ICMCS.2012.6320276]
34. Niyato D, Hossain E, Diamond J. IEEE 802.16/WiMAX-based broadband wireless access and its application for telemedicine/e-health services. IEEE Wireless Commun 2007;14(1):72-83. [doi: 10.1109/MWC.2007.314553]
35. Tchao ET, Diawuo K, Ofosu WK. Mobile telemedicine implementation with WiMAX technology: a case study of Ghana. J Med Syst 2017 Jan;41(1):17. [doi: 10.1007/s10916-016-0661-8] [Medline: 27975206]
36. Aliyu B, Alenoghena CO, Zubair S, Salawu N. Wearable device for telemedicine: an architecture and prototype implementation for remote medical diagnosis using long-range communication protocols. : Telecommunication Engineering Department, Federal University of Technology Presented at: 4th International Engineering Conference (IEC 2022); Minna, Nigeria URL: http://irepo.futminna.edu.ng:8080/jspui/bitstream/123456789/18465/1/IEC%20CONFERENCE%202023.pdf [accessed 2026-05-27]
37. Pramanik PKD, Nayyar A, Pareek G. WBAN: driving e-healthcare beyond telemedicine to remote health monitoring: architecture and protocols. In: Telemedicine Technologies: Academic Press; 2019:89-119. [doi: 10.1016/B978-0-12-816948-3.00007-6]
38. Suleiman TA, Adinoyi A. Telemedicine and smart healthcare—the role of artificial intelligence, 5G, cloud services, and other enabling technologies. Int J Commun Netw Syst Sci 2023;16(3):31-51. [doi: 10.4236/ijcns.2023.163003]
39. Latha R, Vetrivelan P, Geetha S. Telemedicine setup using wireless body area network over cloud. Procedia Comput Sci 2019;165:285-291. [doi: 10.1016/j.procs.2020.01.074]
40. Park RC, Jung H, Shin DK, Cho YH, Lee KD. Telemedicine health service using LTE-Advanced relay antenna. Pers Ubiquit Comput 2014 Aug;18(6):1325-1335. [doi: 10.1007/s00779-013-0744-1]
41. Kress G, Leeuwen T. Multimodal Discourse: The Modes and Media of Contemporary Communication: Arnold; 2001. URL: https://www.scirp.org/reference/referencespapers?referenceid=538574 [accessed 2026-05-24]
42. Departamento Administrativo Nacional de Estadística (DANE). (28 de agosto de 2023). proyecciones de población municipal por área y pertenencia étnicoracial. : Recuperado de; 2023 URL: https://www.dane.gov.co/files/censo2018/proyecciones-de-poblacion/Nacional/anex-DCD-Proypoblacion-PerteneniaEtnicoRacialmun.xlsx [accessed 2026-05-24]
43. Flew T. New Media: An Introduction: Oxford University Press; 2008. URL: https://www.researchgate.net/publication/200026356_New_Media_An_Introduction [accessed 2026-05-25]
44. Bordwell D, Thompson K. Film Art: An Introduction: McGraw-Hill; 2010. URL: https://www.scirp.org/(S(y23rubfvg45z345vbrepxrl))/reference/referencespapers?referenceid=4064446 [accessed 2026-05-24]
45. Mayer RE. Multimedia Learning, 3rd edition: Cambridge University Press; 2021. [doi: 10.1017/9781108894333.003]
46. Ryan S, Ní Chasaide N, O' Hanrahan S, Corcoran D, Caulfield B, Argent R. mHealth apps for musculoskeletal rehabilitation: systematic search in app stores and content analysis. JMIR Rehabil Assist Technol 2022 Aug 1;9(3):e34355. [doi: 10.2196/34355] [Medline: 35916688]
47. Völter C, Stöckmann C, Schirmer C, Dazert S. Tablet-based telerehabilitation versus conventional face-to-face rehabilitation after cochlear implantation: prospective intervention pilot study. JMIR Rehabil Assist Technol 2021 Mar 12;8(1):e20405. [doi: 10.2196/20405] [Medline: 33709934]
48. Cortés-Pérez I, Desdentado-Guillem JM, Camacho-Delgado MS, Del Rocío Ibancos-Losada M, Obrero-Gaitán E, Lomas-Vega R. Virtual reality-based therapy after anterior cruciate ligament injury effectively reduces pain and improves knee function, movement patterns, and dynamic balance: a systematic review and meta-analysis. Knee Surg Sports Traumatol Arthrosc 2025 May;33(5):1736-1753. [doi: 10.1002/ksa.12477] [Medline: 39302094]
49. Cano-Mañas MJ, Collado-Vázquez S, Rodríguez Hernández J, Muñoz Villena AJ, Cano-de-la-Cuerda R. Effects of video-game based therapy on balance, postural control, functionality, and quality of life of patients with subacute stroke: a randomized controlled trial. J Healthc Eng 2020;2020:5480315. [doi: 10.1155/2020/5480315] [Medline: 32148744]
50. El-Kafy EMA, Alshehri MA, El-Fiky AAR, Guermazi MA. The effect of virtual reality-based therapy on improving upper limb functions in individuals with stroke: a randomized control trial. Front Aging Neurosci 2021;13:731343. [doi: 10.3389/fnagi.2021.731343] [Medline: 34795574]
51. Osterloh J, Knaack F, Bader R, et al. The effect of a digital-assisted group rehabilitation on clinical and functional outcomes after total hip and knee arthroplasty-a prospective randomized controlled pilot study. BMC Musculoskelet Disord 2023 Mar 14;24(1):190. [doi: 10.1186/s12891-023-06270-8] [Medline: 36915066]
52. Papalia R, Campi S, Vorini F, et al. The role of physical activity and rehabilitation following hip and knee arthroplasty in the elderly. J Clin Med ;9(5):1401. [doi: 10.3390/jcm9051401]
53. Heisinger S, Huber D, Matzner MP, et al. Online videos as a source of physiotherapy exercise tutorials for patients with lumbar disc herniation-a quality assessment. Int J Environ Res Public Health 2021 May 28;18(11):5815. [doi: 10.3390/ijerph18115815] [Medline: 34071461]

XSL•FO
RenderX

54. Lichtenauer N, Schlachetzki F, Eberl I, Meussling-Sentpali A. Differences in teletherapy and telecare?-Experiences of health professionals and patients with video communication in nursing, physiotherapy, occupational therapy, and speech therapy. Digit Health 2024;10:20552076241301963. [doi: 10.1177/20552076241301963] [Medline: 39640968]
55. Vezér M, Gresits O, Engh MA, et al. Effectiveness of video-game-based therapy to improve hand function in children with cerebral palsy: a systematic review and meta-analysis. J Clin Med 2024 Dec 11;13(24):7524. [doi: 10.3390/jcm13247524] [Medline: 39768447]

Abbreviations

CDMA: Code Division Multiple Access
GSMC: Global System for Mobile Communications
ICT: information and communication technologies
IoMT: Internet of Medical Things
LoRa: Long Range
LTE-A: LTE-Advanced
MANET: Mobile Ad Hoc Networks
Mobile IP: Mobile Internet Protocol
PWA: Progressive Web Application
VSAT: Very Small Aperture Terminals
WBAN: Wireless Body Area Networks
WiMAX: Worldwide Interoperability for Microwave Access

Edited by JJ Zhang; submitted 10.Sep.2024; peer-reviewed by D Chong, IT Ring; revised version received 15.Mar.2025; accepted 26.Mar.2025; published 29.May.2026.

Please cite as:
Aparicio Pico L, Ferro Escobar R, Coronado Sánchez PC
Asynchronous Broadcasting of Audiovisual Content as a Telerehabilitation Strategy for Patients in Rural Areas: Development and Usability Study
JMIR Rehabil Assist Technol 2026;13:e66343
URL: https://rehab.jmir.org/2026/1/e66343
doi:10.2196/66343

© Lilia Aparicio Pico, Roberto Ferro Escobar, Paulo César Coronado Sánchez. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 29.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Influence of Digitalization on Multidisciplinary Poststroke Rehabilitation Practice: Qualitative Study

Ann Marie Hestetun-Mandrup[1,2], MHS; Charlotta Hamre[2], MHS, PhD; Anne Lund[1], Prof Dr; Anne Catrine Trægde Martinsen[1,2], Prof Dr; Hong-Gu He[3,4], Prof Dr; Minna Pikkarainen[1], Prof Dr

[1]Department of Rehabilitation Science and Health Technology, Faculty of Health, OsloMet – Oslo Metropolitan University, Pb. 4 St. Olavs Plass, Oslo, Norway

[2]Department for Research and Education, Sunnaas Rehabilitation Hospital, Oslo, Norway

[3]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[4]National University Health System, Singapore, Singapore

Corresponding Author:
Ann Marie Hestetun-Mandrup, MHS
Department of Rehabilitation Science and Health Technology, Faculty of Health, OsloMet – Oslo Metropolitan University, Pb. 4 St. Olavs Plass, Oslo, Norway

Abstract

Background: Leveraging digital technologies in health care is recognized as essential for effective and efficient services. However, significant challenges remain in implementing these technologies in stroke rehabilitation practice, and research on their influence is limited.

Objective: This study aimed to explore the current influence of digital technologies on stroke rehabilitation practices and consider how these technologies could shape the future landscape of rehabilitation for multidisciplinary health care professionals in poststroke rehabilitation.

Methods: A qualitative, exploratory design was used. Data were collected from 12 experienced multidisciplinary health care professionals at 2 Norwegian rehabilitation settings via semistructured interviews, and the data were analyzed using reflexive thematic analysis. Data analysis was guided by social practice theory.

Results: The 12 participants included experienced physiotherapists, occupational therapists, speech therapists, nurses, physicians, and social workers. The following three main themes were generated: (1) Outsourcing information about and to stroke survivors: coordination and continuity within and across services (subthemes on follow-up and interservice collaboration, and user-centered approaches); (2) Navigating the ambivalence of remaining human relations in digital psychosocial support conversations (highlighting multidisciplinary challenges in building relational depth and addressing sensitive topics); and (3) Enhancing digital supplements for assessment and engagement in motor rehabilitation (subthemes on progress monitoring and motor skills exercises). Overall, the use of digital technologies in specialized stroke rehabilitation practices was seen as an adjunct to practices. While digital technologies influenced rehabilitation practices, ambivalence and challenges were noted, particularly in digitalizing multidisciplinary psychological support and exercise programs. Systems for sharing medical records and goal-setting apps, which enhance coordination and involve stroke survivors, were emphasized as future digital technologies that can shape stroke rehabilitation.

Conclusions: Health care professionals used various technologies in their daily specialist practices, as well as for the coordination and follow-up of stroke survivors after referral to community services. This study identified several organizational processes, roles, standards, and rules that can act as barriers or drivers to implementing digital technologies in practice. Viewing familiar digital technology as a supplement to existing practices, rather than as a singular solution for all areas of specialized stroke rehabilitation, offers significant potential for quality improvement. These findings provide valuable insights for technology developers, health care personnel, and user groups in specialized neurological rehabilitation settings.

(JMIR Rehabil Assist Technol 2026;13:e77753) doi:10.2196/77753

KEYWORDS

digital therapeutics; digital technology; goals; practice theory; rehabilitation; stroke

XSL•FO
RenderX

Introduction

The World Health Organization (WHO) has promoted digital health to enhance the quality and accessibility of health care services [1]. Leveraging digital technologies in health care is recognized as essential for effective and efficient services [2]. Limited resources have led to earlier hospital discharge and outsourced rehabilitation services aided by digital technologies, which are areas where politicians, clinicians, stroke survivors, and researchers have turned their attention to [3]. WHO's initiative of Digital and Assistive Technologies for Aging (DATA) [4] further underscores the importance of using accessible digital technologies in developing and sustaining new stroke care pathways [5]. The term "digital technologies" refers to electronically driven devices or applications [6,7]. Digital technologies encompass various information and communication technologies (ICTs) used in medicine and health care for disease management and wellness enhancement [6]. These technologies include devices, such as computers, smartphones, and tablets, along with their associated applications and internet connectivity [7]. Studies have shown that digital technologies improve activity and function [8-11], social participation [12], and medical adherence among individuals with stroke [13-15]. Technologies, such as gaming, videoconferencing, online health websites, apps, and wearables, have been especially explored for their potential to enhance exercise programs in both clinic and home settings in the stroke population [16]. Despite the rapid adaptation of digital technologies after the COVID-19 pandemic, scaling up remains slow [17].

Effective neurorehabilitation and tailored digital interventions are crucial for future stroke home rehabilitation [18], as they are cost-effective, portable, and motivational solutions to support daily activities [19]. However, challenges persist in home implementation [19], which might explain the lack of superior effects on supervised digital motor rehabilitation at home over conventional training for stroke survivors [18]. Although studies have found overall positive attitudes toward digital technologies in poststroke rehabilitation from the perspective of individuals with stroke, due to increased accessibility and convenience, there is an advocated need for personal contact, shared responsibility, and collaboration with health care professionals (HCPs) [20].

Building on these findings, rehabilitation specialists have introduced many digital solutions for home-based rehabilitation [21], yet significant challenges remain in adapting these technologies in practice, and research on their influence is limited. Despite guidelines and research recommending technology use in stroke rehabilitation, actual implementation and clinical validation are inconsistent [16,22-24]. A systematic review by Borg et al [25] focused on older people and people with disabilities and highlighted the need for research on technology adoption in practice, particularly from a theoretical, multidisciplinary approach. A longitudinal study across inpatient, outpatient, and community rehabilitation settings found that the type and amount of technology use varied between neurological populations, with the most frequent and varying use in stroke rehabilitation [22]. Contrary to this, Langan et al [16] found that physical and occupational therapists do not widely incorporate digital technologies into their practices, often prescribing written, paper-based, and nontechnological home exercises after hospital discharge. The primary barriers often include financial constraints, lack of technological experience, and limited time [26,27]. Despite inconsistencies in technological uptake, researchers assert that organizational readiness for eHealth reveals a dichotomy, with rehabilitation staff feeling ready, but practice preparedness being slower [2].

A broader and multidisciplinary perspective is essential to fully capture the complexity of rehabilitation practices [2,27,28]. This study addresses these opportunities for further inquiry. Existing research argues that the challenge in technology application often lies in social behavior around technology use, rather than the technology itself [29,30]. A second critical challenge lies in understanding organizational readiness for the adaptation of digital technologies in rehabilitation settings [2]. This includes the preparedness of both the organization and its members to adapt and integrate these technologies in their daily practices [2]. A third challenge is that existing studies often overlook the broader social context in which these technologies are used in practice [31]. Therefore, there is a need to consider the interconnectedness of health and social care when evaluating how digital technologies shape practices, where digital solutions should aim to lower barriers and facilitate drivers for access to rehabilitation [32].

We applied social practice theory as a theoretical framework to understand how digitalization affects stroke rehabilitation practices at the individual and social levels. Social practice theory, a behavioral theory, explores how people engage with the social world and how it impacts human behavior [33,34]. This applied practice theory is based on work by Reckwitz and incorporates elements from Bourdieu, Giddens, Foucault, and Schatzki [35]. Instead of focusing only on concepts, such as intersubjectivity, embodiment, language, and power within organizations, it incorporates all these concepts and takes a broader perspective by examining what people do in practice and the interconnected elements of individual actions and institutional structures [33]. Practice entails routine behavior involving linked key concepts, such as agents' skills and meanings needed for practices, the things that shape and enable practices, the rules and social norms guiding practices, and the process structure [35]. This theory, which has gained recognition in various fields, provides a useful lens to address the challenges and solutions related to digital inclusivity [36]. It has been used in environmental research [37], education [33], digital information [38], and public and clinical health [31,34,39,40]. The application of this theory in poststroke settings is novel, and this study adds to the limited qualitative research on the changes in practice through digitalization [40]. This study applies the concept of digitalized stroke rehabilitation as a social practice to explore how digitalization influences rehabilitation within a local context. Therefore, this study aims to examine the current influence of digital technologies on stroke rehabilitation practices and consider how these technologies could shape the future landscape of rehabilitation. This is performed through the following research question: How do digital technologies shape the current and perceived future

XSL•FO
RenderX

everyday practices of multidisciplinary HCPs in poststroke rehabilitation?

Methods

Design

In this descriptive qualitative study, a theoretical framework of social practice theory (henceforth "practice theory") [35] was applied as a lens to explore and interpret the structures shaping the use of digital technologies in 2 Norwegian stroke rehabilitation settings. Inductively generated thematic themes were mapped onto practice theory concepts. This design allowed for an in-depth exploration and assessment of the relationship between various theoretical key components of the challenges and opportunities of digitalization in stroke rehabilitation, providing insights that could shape the future landscape of rehabilitation.

Settings and Participants

Multidisciplinary HCPs from 2 rehabilitation wards in Norway were invited to participate via email and staff meetings. Recruitment at the first hospital (June 2022 to January 2023) and the second hospital (April to May 2024) involved multidisciplinary HCPs, facilitated by team coordinators. Purposeful sampling was used to recruit appropriate and information-rich data for the in-depth qualitative study. Participants were selected for their direct experience with multidisciplinary specialized stroke rehabilitation. There were no a priori guidelines for the amount or the specific data required, such as regular engagement with new technologies. The study started broadly and narrowed naturally as we became interested in what qualifies as specialized stroke rehabilitation and how technologies shape practices, justifying the theoretical delineation of the sample [41]. Both settings specialize in poststroke rehabilitation, regularly engaging with new technologies. Some technologies have been implemented in practice, while others are still being tested or are under consideration.

The first rehabilitation hospital is a large specialist facility with subacute and poststroke wards, allowing for both short- and long-term rehabilitation. Recruitment occurred in the poststroke ward, with an average of 10 days poststroke at admission and an annual average of 279 stroke survivors admitted. The second ward is a community-based clinic within a hospital and focuses mainly on subacute stroke rehabilitation, with an average of 32 days poststroke at admission and an annual range of 150 - 200 stroke survivors admitted. Despite differences in organization, both settings were chosen for their high staffing levels for specialized and intensive multidisciplinary stroke rehabilitation. The first hospital qualifies as performing complex rehabilitation due to the involvement of at least six different professions. Maximum variation purposive sampling [42] was used to recruit 12 experienced HCPs, ensuring diversity in age, sex, and work experience (Table 1). Participants had at least 2 years of experience in poststroke rehabilitation, and both female and male professionals were included. Interviews were conducted on-site at the hospital during appropriate times to minimize interference with daily practices.

XSL•FO
RenderX

Table . Demographic information of the health care professionals who participated in this study (N=12).

ID	Profession	Age (years)	Time working in stroke rehabilitation (years)	Workplace and rehabilitation practice	Current digital technologies used in rehabilitation
1.1	Speech therapist	31	2 - 5	Specialized stroke rehabilitation hospital	Videoconference, digital speech training, video fluoroscopy, and electronic medical records
1.2	Occupational therapist	35	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, pilot project on stroke survivor access to medical records, QR codes to health-related webpages, robotic technology training, digital group motor rehabilitation, and electronic medical records
1.3	Physiotherapist	30	2 - 5	Specialized stroke rehabilitation hospital	Videoconference, virtual reality lab, and electronic medical records
1.4	Occupational therapist	38	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, webinars, virtual reality, and electronic medical records
1.5	Nurse/head of the ward	42	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, virtual reality lab, robotic technology training, and electronic medical records
1.6	Nurse	53	2 - 5	Specialized stroke rehabilitation hospital	Videoconference and electronic medical records
1.7	Physiotherapist	52	25 - 30	Specialized stroke rehabilitation hospital	Videoconference, web-based exercise programs, and electronic medical records
1.8	Social worker	40	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, national health webpages, electronic individual plan, and electronic medical records
1.9	Physician/head physician	57	15 - 20	Specialized stroke rehabilitation hospital	Videoconference, webinars, and electronic medical records
2.1	Physiotherapist	54	<30	Rehabilitation clinic within a hospital	Videoconference, pulse monitoring, digital blood pressure equipment, digital exercise program, and electronic medical records
2.2	Physiotherapist	35	5 - 10	Rehabilitation clinic within a hospital	Virtual reality, videoconference, pulse monitoring, step counter, iPad, dynamometer, and electronic medical records

XSL•FO
RenderX

ID	Profession	Age (years)	Time working in stroke rehabilitation (years)	Workplace and rehabilitation practice	Current digital technologies used in rehabilitation
2.3	Physiotherapist	36	5 - 10	Rehabilitation clinic within a hospital	Web-based exercises, pulse monitoring, dynamometer, biometric equipment, Nintendo Wii, step counter, and electronic medical records

Theory

In this study, practice theory offered a lens through which to examine how different HCPs talked about their actions. Grounded in constructionism, the theory posits that learning and knowledge are collectively produced through shared understandings, practices, and language [33]. Practice theory situated digitalization in poststroke rehabilitation within 2 practice settings, and its terminology served as an analytical advancement for understanding how digital technologies influence poststroke practices. Its key concepts are mapped in Table 2 to clarify the description of each concept as it relates to specialized poststroke rehabilitation settings.

XSL•FO
RenderX

Table . Glossary of practice theory concepts related to poststroke rehabilitation practice.

Concept	Practice theory concepts	Example related to poststroke rehabilitation practice
Agent (including skills)	An agent is a carrier of actions with the necessary skills, and this encompasses both bodily and mental routines. It involves ways of understanding the social world and implicit and explicit know-how [35].	Health care professional agents are connected through specialized hospital processes [43] and are norm-following individuals shaped by professional know-how, who adhere to stroke guidelines and hospital policies relative to their hierarchy at the hospital [40]. They use their minds and bodies to perform routine actions during stroke rehabilitation, such as instructing stroke survivors on walking while encouraging progress. In both contexts, they interact with medically stable stroke survivors past the acute phase, when they are cleared for intense rehabilitation. Occupational therapists, physiotherapists, nurses, medical social workers, and speech and language therapists are central to practices when motor rehabilitation is the main focus [43].
Things	To carry out a practice, agents use things in a certain way, such as physical entities or material conditions, where actions take place. However, writing, printing, and electronic media are also a part of things [35].	Things represent all therapeutic devices, equipment, and health solutions that are necessary to carry out the practice of stroke rehabilitation. Physical elements, such as conversation rooms and rooms for exercise or written material, encompass the things that the agents use or interact with [35].
Rules	Structural rules construct a practice around communication chains and signs. Agents endorse cultural norms and meanings to objects, determining actions through routinized nonsubjective understanding [35].	Health care professionals in poststroke rehabilitation follow recommended interventions from recent rehabilitation guidelines and national clinical practice guidelines for stroke rehabilitation [43]. Communication between professionals often occurs via health information systems, electronic medical records, and IT platforms where rehabilitation plans, progress notes, and patient outcomes are documented and coded according to standardized protocols. Communication with patients or their families typically happens through face-to-face consultations, telephone follow-ups, or digital platforms for remote monitoring, ensuring adherence to evidence-based practices [40].
Process structure	Routinized social practices and processes occur in a sequence over time and through repetition, forming the core of the social structure. Changes in practices arise internally, as practitioners challenge or resist established routines, and externally, as intersecting practices influence each other. Practitioners improvise new actions in novel situations, prompting internal shifts, while external interactions among diverse practices foster adaptation and transformation [35].	In poststroke rehabilitation, health care professionals facilitate a process-centered rehabilitation model for stroke survivors. Although structured, it starts flexibly with initial consultations, followed by planning, a discharge meeting, and ultimately self-care. The process includes multidisciplinary subprocesses like therapeutic examinations and collaboration with other institutions [44].

Data Collection

Individual semistructured interviews were conducted over 2 years, from 2022 to 2024, each lasting 40 to 60 minutes. Information power guided the inclusion of 12 interviews, based on data relevance and richness [45] and when sufficient data "told a coherent, convincing, and comprehensive story" [41]. The interview guide, collaboratively developed by the study's authors (Multimedia Appendix 1), explored participants' stroke rehabilitation practices, shared decision-making, and the use of technology and future digital innovations. The inclusion of decision-making as a part of the stroke rehabilitation process was inspired by its person-centered foundation, drawing on literature by scholars, such as Rose et al [46] and Bomhof-Roordink et al [47], in formulating interview guide questions. Interviews were conducted in Norwegian, allowing for probing questions and guided conversations. Audio recordings were securely stored using services for sensitive data. A transcription software assisted with transcription to Norwegian, which was later manually edited for accuracy. Deidentified selected quotations were then translated into English and checked for accuracy by the research team and a secure university-based artificial intelligence (AI) tool

XSL•FO
RenderX

(ChatGPT-4, OpenAI). The research team retained responsibility for final wording and meaning.

Data Analysis

Thematic analysis was chosen to identify, analyze, and interpret patterns across varied meanings [48]. This approach is well-suited as a method for exploring complex phenomena, as it allows for a rich and nuanced understanding of the HCPs' perspectives on how they construct and navigate their practices. The flexibility of thematic analysis enables it to be effectively applied across diverse theoretical frameworks and research questions, making it appropriate for this study's aim of understanding multidisciplinary HCPs' practices. We used reflexive thematic analysis (RTA), specifically developed by Braun and Clarke [48], because it aligns with the interpretative paradigm, emphasizing the active role between researchers and interviewees in generating knowledge and the subjective, situated nature of data interpretation. The in-depth engagement with the data enabled us to explore the social, cultural, and contextual complexities of digitalization in stroke rehabilitation practices, as well as the iterative and reflective process of capturing the underlying meanings and nuances of HCPs' perspectives [48]. The 6 phases of RTA—familiarization, coding, theme generation, review, refinement, and write up—were applied in a back-and-forth manner to accommodate curiosity and interpretation. Familiarization and coding were performed in NVivo (Lumivero), while theme generation, refinement, and write up were conducted in Word (Microsoft Corp). We conducted 2 analyses. The first analysis followed RTA, and the second aligned the findings according to practice theory concepts using Excel (Microsoft Corp). Both descriptive ("semantic," according to Braun and Clarke [48]) and interpretative ("latent") codes were utilized [49] with emphasis on capturing the empirical data during the coding stage. After several rounds of coding, the codes were manually organized into main themes, with practice theory guiding further analysis. The themes from the RTA did not change in meaning, but further descriptions and interpretations related to themes and subthemes were inspired by practice theory as a lens. During theme refinement, an additional analysis mapping qualitative data according to practice theory's 4 concepts was conducted [35] (Multimedia Appendix 2). Practice tasks, forming the final subthemes, were defined. An example of the six phases of RTA, including analyses from both RTA and practice theory, is presented in Multimedia Appendix 3. Key concepts are written in italics in the Results section. Each concept was examined from the points of both the current and future states to understand how digital technologies shape HCPs' practices.

Reflexivity

The researchers' backgrounds acknowledge how the authors' positionality influenced data analysis [48], with several having clinical experience with stroke survivors (AMH-M, CH, AL, and HGH). All authors have academic backgrounds, experience in digital solutions, and multidisciplinary expertise in the fields of rehabilitation within occupational therapy, physiotherapy, nursing, computer science, and medical physics and technology. The first author, a doctoral student and experienced physiotherapy specialist, conducted all interviews, considering subjectivity to be a resource in conducting and analyzing data; maintained a listener's role; and emphasized her role as a researcher rather than an HCP. Still, her clinical background allowed her own experiences to resonate with participants, enabling her to listen and repeat participants' recognizable stories. Reflexivity was considered early in the analysis process by recording initial assumptions after each interview and writing memos and annotations in NVivo during coding to ensure the quality of analytical points and the possibility of reviewing interpretations. A description related to all codes also supported analytical interpretation. To ensure rigor, the authors collaborated in regular meetings and workshops to discuss and enhance understanding, develop comprehensive descriptions, and refine the analysis, strengthening confirmability, coherence, and clear theme boundaries [48,50].

Ethical Considerations

As the study's focus on health service research fell outside the scope of the Health Research Act § 2, ethical approval from The Norwegian National Research Ethics Committee for Medical and Health Research was not required (reference number: 279418). Approval was granted by the Norwegian Agency for Shared Services in Education and Research (reference number: 668899) to ensure privacy protection and adherence to ethical standards. Participants received oral and written information about the study, and written informed consent was obtained prior to the interviews. The first author reiterated consent form details, emphasizing the voluntary nature of participation, confidentiality, and the participants' right to withdraw from the study. Participants did not receive any compensation for their involvement in the study.

Results

Themes and Subthemes

The results are presented in themes and subthemes and exemplified with selected quotes extracted from the analysis based on practice theory.

Theme 1: Outsourcing Information About and to Stroke Survivors: Coordination and Continuity Within and Across Services

Overview

HCPs actively leveraged digital technologies to enable better coordination and continuity of care across services. Digital tools facilitated collaboration not only between HCPs in cross-service collaboration but also with stroke survivors, relatives, and interpreters, forming the first subtheme. Through their practices, they sought to harness digital health to streamline communication and improve rehabilitation outcomes, despite the challenges of sustaining structured coordination. Maintaining continuity after referrals to community services was challenging, requiring constant goal adjustments and balancing person-centered approaches. The rule of involving stroke survivors in their rehabilitation was viewed as a necessary but complex procedure, reflected in the second subtheme "navigating user-centered approaches to rehabilitation for stroke survivors in the digital era."

XSL•FO
RenderX

Subtheme 1.1: Follow-Up and Continuity in Rehabilitation in Collaboration With Stroke Survivors and Between HCPs in Cross-Service Collaboration

In their current practices, HCPs use digital technologies, primarily videoconferencing, for stroke follow-up after discharge, involving additional *agents* such as stroke survivors, community HCPs, interpreters, Norwegian Labor and Welfare Administration (NAV) advisors, and relatives. The importance of HCPs with ICT knowledge as an added skill was highlighted. Coordination and continuity were perceived as crucial owing to the challenges in maintaining structure after referrals to community services, but they often relied on the agency of stroke survivors.

> *Keeping track of who to contact and when is challenging. We've faced logistical issues regarding when the reablement team is supposed to come, and when it conflicts with orthopedic appointments. We lack an overview of these things when the stroke survivors aren't hospitalized and rely on information from stroke survivors or other responsible persons. We do have good contact with some teams, but the organization varies greatly between different districts, making it very inconsistent.* [Informant 2.2]

Increased coordination and continuity aligned with the explicit *rules* of increasing outpatient activity and outsourcing information about stroke survivors. For future practices, an HCP suggested creating a unified individual digital plan between specialist services and communities to address uncertainty about digital progress. This system would provide an overview and ensure that individuals needing long-term, coordinated services receive appropriate care and benefit from effective collaboration to meet their goals. The following comment was made:

> *It would have been quite easy to get started with creating the document and then pass it on to the community. I know that there is no common solution, because it is somewhat up to the community whether they have a digital individual plan, and at least it was a few years back, so I don't know if there is a common solution that all communities can use, but if there is, then there is great potential to look into it.* [Informant 1.8]

Others proposed sharing medical records reporting standardized test results across services and using apps with push-warnings for goal-setting visible to both the HCP and the stroke survivor. Future solutions could involve extended digital therapy skills to include other digital platforms, such as NAV's screening tools, forming a new type of *agent* bridging HCPs and NAV advisors. They also proposed a system between the stroke survivor and health care provider to navigate in-hospital appointments and communicate with staff.

> *Maybe an overview for the patient, in an app, or some kind of guide, showing you this way, you can follow this path to get there. Your physiotherapist is sick today, so your appointment is cancelled. This is somewhat similar to what has already started in other wards. We will also implement it here eventually. It just needs to be adapted to our ward.* [Informant 1.6]

There were societal encouragements for increasing the outpatient activity and outsourcing information to the stroke survivor, enabled by more videoconferencing. Closer follow-up and a more standardized program under the direction of either a general practitioner (GP) or hospitals were mentioned. Enhanced outpatient activity demanded restructuring practices, creating better access to stroke-related information, and introducing a joint communication system between the stroke survivor and the HCP. Digitalizing content on stroke exercise programs, fatigue, rehabilitation experience forums, and cognitive or physical constraints was suggested for improved coordination. However, the complexity arising from multiple medical record systems was highlighted, showing that these systems were used exclusively between HCPs and GPs, leaving out stroke survivors. This suggests the need for integration through a unified platform for visibility and collaboration.

Even though collaboration between HCPs across health services is essential, they expressed a lack of knowledge about available resources for stroke survivors outside their own settings, underscoring the role of stroke survivors in informing their own GPs about rehabilitation opportunities. Paper-based reports were still the main communication tool. Clarifying realistic expectations and outcomes for the stroke survivor's condition helped in establishing routines and maintaining standards and rehabilitation plans. To enable this, HCPs routinely used SMART (specific, measurable, achievable, relevant, and time-bound) goals and educated stroke survivors in SMART goals [51] based on the understanding of being *specific or measurable, achievable or attractive, realistic, and time-bound* (informant 1.5), which gave direction for the future rehabilitation process. However, HCPs expressed that stroke survivors' goals, as revealed in goal-setting conversations and documented in written reports, were often forgotten or changed. The flow of information heavily depended on the stroke survivor's effort to communicate with other HCPs. As a future extension of implementing SMART goal principles in home-based rehabilitation, some HCPs perceived that goals could shift to being more focused on daily routines rather than solely the therapeutic understanding of goals for stroke survivors. For future development, HCPs valued apps with push notifications for goal-setting visible to both HCPs and stroke survivors, and indicated that future practice could consist of increased health-related data and goals. The following comment was made:

> *I wish we could have a bit more follow-up along the way, mostly for the stroke survivor's sake, because it's quite easy to put that binder in a drawer when you get home. Then you have, in a sense, "free time" until they come back. Since we say that goal setting is so important, I wish we could have communication along the way....If they could almost get a push notification through an app.* [Informant 1.2]

Collaboration faced time constraints in cross-service communication, and HCPs highlighted the importance of "setting boundaries for oneself" (informant 1.4) to balance their workload and reduce personal-professional responsibilities.

XSL•FO
RenderX

This perspective was contrasted by another participant, who highlighted differences between specialist health care services and communities in care provision and time allocation, noting the challenges in aligning social expectations.

> *I imagine that in the communities, you might get a couple of assessment visits and then wait a few months for some assistive devices, and you might not get them. You certainly don't get anything weekly and can only address one thing at a time. If we propose ten measures, it might be that you can either spend ten minutes per measure spread over a few weeks, or half of the measures are left out, is the impression I have.* [Informant 1.2]

Conversely, an HCP made the following comment: "We can't really demand that the community follow up on everything we do either" (informant 1.5). Some HCPs desired more follow-up time in their practices due to a known hesitancy in working with stroke survivors in communities, suggesting more knowledge on stroke rehabilitation and restructuring practices for better access to information and logistics. The potential benefits of a joint communication system include a less fragmented goal process and user involvement in rehabilitation goals for stroke survivors and their relatives.

Subtheme 1.2: Navigating User-Centered Approaches to Rehabilitation for Stroke Survivors in the Digital Era

Stroke survivors were routinely involved in their rehabilitation through discussions at several rehabilitation meetings, aiming to foster independence and proactive behavior. Digital preassessment conversations encouraged participation of stroke survivors in decision-making. The *rule* of involving stroke survivors in rehabilitation was to address their existential needs and was echoed by HCPs as fundamental to the nature of rehabilitation. They emphasized constant goal adjustments and the balance of person-centered care as a common but complex procedure. For instance, cognitive (eg, cognitive fatigue) or cultural determinants could undermine their accountability.

> *But of course, there are quite a few cognitive difficulties that cause significant disruption, which also means that if there are psychiatric and other conditions, stroke survivors may resist some of the services offered or do not take the initiative. It is often said that they are not motivated, as many lose due to cognitive impairment or other conditions, the ability to motivate themselves or make decisions and take initiative. And they often fall behind.* [Informant 1.9]

Relatives were involved, and they were invited to meetings either in person or remotely. HCPs supported stroke survivors' relatives by providing written material for support, owing to their role of being both a close relative and a caregiver. The challenge of maintaining this balance was highlighted as follows:

> *It's a challenge for relatives to live with someone who has had a severe stroke and to be both the closest relative, personal trainer, and everything else. They may need to be a home helper, assist with meals, and serve. It's a significant burden. It is beneficial for them to have someone come in who can take on some of the burden.* [Informant 2.1]

HCPs initiated local opportunities to continue rehabilitation, often referring stroke survivors to other services, acknowledging the complexity of stroke rehabilitation where "most need more rehabilitation" (informant 1.7). Many stroke survivors returned to the same rehabilitation hospital owing to gaps in service availability, such as occupational therapy lacking in some communities. Future digital rehabilitation would require stroke survivors to be more involved in discussions about rehabilitation opportunities throughout the process and to take ownership, which are often omitted in current practices. The following comment was made:

> *[In digital communication with stroke survivors it's important to] provide alternatives because there's not just one option; there are often several factors to consider. So, it's also about involving the stroke survivor more actively in the discussions about their future as well.* [Informant 1.6]

Concerns were expressed about current routines that excluded stroke survivors from discussions regarding alternatives in rehabilitation. HCPs viewed passivity after stroke and a lack of insight and comprehension as barriers to increased agency. However, HCPs observed a growing trend of stroke survivors taking an active role in their treatment. Although this was problematic for most owing to a lack of comprehensive medical and rehabilitation knowledge, there was a clear need for enhanced stroke survivor education. The potential for increased use of digitalized materials for relatives was also suggested. However, some HCPs criticized the assumption that most people are digitally proficient, noting that this responsibility often falls on not only stroke survivors but also their relatives. In situations where stroke survivors struggle to learn new skills, the assistance of relatives was seen as essential. The following comment was made:

> *Not all stroke survivors can use digital solutions, and then [the responsibility] might fall to their relatives. However, not all stroke survivors have resourceful relatives either, creating ambivalence about whether this user group can manage it. I believe that a large proportion can, whether it's the stroke survivors or their relatives who take responsibility. As shown in videoconference meetings here, where relatives have participated, most manage if they get a brief introduction to what they need to do, how to do it, and when. Many will be able to be part of such an arrangement in the long term, but it is probably a long way to get there.* [Informant 1.3]

Education was needed for stroke survivors with cognitive difficulties or limited digital skills. High demands were placed on those facilitating digital interactions. There were divided opinions on whether the rehabilitation hospital or the community held the main responsibility for digital home exercises and follow-ups. Some claimed that districts were responsible for digital home rehabilitation and follow-up, while others believed that the hospital's main future task was to spread and utilize their specialist expertise. Nonetheless, the user's capabilities to

XSL•FO
RenderX

become digitally self-sufficient were questioned, and an individualized approach was the top priority. The following comment was made:

> *So, when we say that we aim to provide equitable services to our users and stroke survivors, we must also remember that equitable can mean that some need in-person appointments while others can manage with digital solutions. There are many ahead of us in digital communication, including users and partners, and we are somewhat behind. But many are also far behind us and will never reach where we want to be, so we must remember to cater to the full range of our users. We cannot demand that everyone go digital. If we do, it must be very much based on collaboration. Maybe they can manage an app, but not video. We need to find the right levels.* [Informant 1.2]

Theme 2: Navigating the Ambivalence of Remaining Human Relations in Digital Psychosocial Support Conversations

Multidisciplinary HCPs played a crucial role in providing psychological support through communication, but grappled with the dual promise and limitation of digital technologies in providing support. Psychosocial support remains a multidisciplinary task as a person-centered rehabilitation approach often frames these practices. They engaged in both structured and unstructured conversations to assess existential struggles, self-perception changes following stroke, and quality of life, especially for stroke survivors with multiple cognitive and psychiatric challenges. These conversations often focused on enhancing well-being, and the rehabilitation *process* was about gaining knowledge about the difficulties and the approaches to tackle them.

While digital communication enabled flexibility and accessibility, providing effective support still required building trust and relational depth. HCPs emphasized therapeutic communication skills involving active listening, openness, and sharing, capturing sensitive information in both digital and in-person meetings. Conversations sometimes included topics like depression or suicidality, but often, these were viewed as expressions of an unwanted situation rather than an active plan. The significance of building a relationship with stroke survivors digitally was highlighted through practical strategies to achieve it.

> *Just giving them the time and space to express themselves, joking a lot. I use a lot of humor with the stroke survivors and try to show some understanding of their situation. But I find that if you can mention relevant life experiences, I think it's a good way to build relationships, to dare to open up a bit, so we're not just sitting there as rigid authority figures, making it a one-sided conversation.* [Informant 1.1]

Some HCPs felt that building relations is harder to achieve digitally and raised concerns about digital health potentially diminishing their therapeutic skills, emphasizing the importance of human relationships over relying solely on technology.

> *I think one should be a bit cautious about saying that all therapists are now redundant. Because now they can just train with an app. I don't believe in that. It's about human relationships.* [Informant 2.1]

This ambivalence was further shaped by the need to balance technological efficiency with the irreplaceable value of human connection. As professionals sought to navigate the tension between digital and in-person care, building relationships digitally was seen as a prerequisite for further effective support, particularly for follow-up in the later rehabilitation stages.

> *These conversations, right, which many experience as being more difficult- yes, life can become challenging when you return home. The challenges where perhaps there could also have been more follow-up conversations via video, but also something in a group setting. We currently have groups for those who come here, but the number of participants is somewhat limited Many live far away, and it's difficult for them to travel from home. So, perhaps there should be a type of service specifically directed at stroke survivors and their relatives at a later stage.* [Informant 1.9]

Theme 3: Enhancing Digital Supplements for Assessment and Engagement in Motor Rehabilitation

Overview

HCPs embraced digital tools to enhance assessment, treatment, and patient engagement in motor rehabilitation. These tools provided detailed insights into physical progress, fostering participation and curiosity about their health metrics, as shown in the first subtheme "assessment and monitoring of progress." HCPs emphasized individualized approaches and explored creative solutions, such as using virtual reality (VR) and augmented reality (AR), to simulate real-life scenarios or design exercises based on videos in home environments, as illustrated in the second subtheme "exercises and rehabilitation programs targeting motor skills." However, digital tools also presented challenges, such as balancing safety concerns for patients with poor balance and ensuring engagement across generations. While digital tools offered significant potential for self-managed and supported rehabilitation, they were best suited for less complex cases.

Subtheme 3.1: Assessment and Monitoring of Progress

HCPs used various digital tools to enhance the assessment of stroke survivors' progress and engagement in therapeutic activities. *Things* included pulse monitors, dynamometers with apps to measure and track physical strength and progress, video-based assessments like video fluoroscopy and endoscopic swallowing evaluations, VR tools like "Job Simulator" for cognitive assessments involving kitchen tasks, and step counters that display progress graphs. This fostered curiosity and provided detailed insights into treatment intensity. For instance, pulse monitors, which have been valued as indispensable over the past years, ensured that the pulse rate of stroke survivors remained within a target range.

XSL•FO
RenderX

> *We measure the pulse and check that we are in the zone as much as possible....And after using it for a few years now, I kind of don't understand how I managed without it. It says a lot about the patient's intensity.* [Informant 2.1]

Multidisciplinary assessment was integral to the *processes* of specialized stroke rehabilitation, focusing on holistic understanding and ongoing evaluation, which directly influenced HCPs' replicated behavior and their ability to act on it.

HCPs proposed using digital tools for preadmission consultations to evaluate engagement and proposed involving more community therapists, with local resources aiding preassessment. They suggested video-based visualizations of stroke survivors' home environments and the storage and processing of stroke survivor assessments. Monitoring was perceived as identifying stroke survivors in need of more interventions, but motivation and performance remained challenges. The following comment was made:

> *Monitoring by itself will not be enough, because when you look at the studies that have been done, at least in my experience, stroke survivors, even when they improve in arm function and actually have a higher capacity, don't use their arm as much when they get home. They revert to their usual patterns. Even though capacity increases, performance does not necessarily increase. This is a challenge that the monitor can detect, but the monitor itself will not address the stroke survivor's motivation, incentives, or desire to use the arm, or their understanding that they probably can walk more or use the arm more than they currently do.* [Informant 2.3]

Integrated feedback systems within the monitoring process were seen as potential solutions to motivation challenges. Nonetheless, the question of whether the encouragement of therapists alone is sufficient to enhance arm use or gait remains unresolved, and HCPs are yet to find effective solutions.

Subtheme 3.2: Exercises and Rehabilitation Programs Targeting Motor Skills

Both hospitals implemented gamified rehabilitation. In one hospital, VR was used for intensive balance training or as part of constraint-induced movement therapy for arm training, in addition to being utilized during digital group sessions. The other hospital used biometric equipment with force plates or the Nintendo Wii device in practice and VR for projects. One HCP had experience with asynchronous virtual individual exercises for stroke survivors living at home, when working in the community.

HCPs provided assistance either in person or remotely during the digital programs, but VR sessions required their presence. Language training, in which stroke survivors underwent naming therapy and practiced word recall, was more effective with an HCP present. *Rules* on therapy dose varied, but quantity over quality was mainly emphasized, with the intention of promoting cerebral blood flow and brain reorganization.

Despite efforts to fully utilize each stroke survivor's rehabilitation potential, challenges in maximizing their potential persisted. The following comment was made:

> *You really want that person to have the interventions implemented so they can achieve good functionality. It's very, very sad if there's potential that isn't fulfilled. Then our efforts here feel somewhat wasted. We can figure out and write down what is needed, but if it's not followed up, it feels a bit pointless. I just want people to get as well as they possibly can.* [Informant 1.7]

To enable greater potential for rehabilitation, HCPs proposed future practices involving goal-oriented digital rehabilitation opportunities and motivating exercises. However, they highlighted variations in the types of stroke survivors suitable for undergoing digital motor rehabilitation. As a result, the tasks of future HCPs would include designing motivating exercises while ensuring environmental safety and space. HCPs noted that poor balance and fall risk complicated independent exercise, making approaches to individualized adaptation and treatment effectiveness critical. Simple, self-managed technologies were seen as valuable for catering to different generations. For example, sitting exercises for arm rehabilitation were favored among HCPs owing to safety concerns, especially given that arm rehabilitation tends to receive less focus in stroke rehabilitation across services. Rehabilitation programs could be self-managed or supported by HCPs, depending on the complexity of the stroke survivor's condition. Moreover, they suggested the use of digital apps for low-threshold language training, incorporating features like animations, facial expressions, and tongue position guides. Other HCPs saw potential in creating videos of stroke survivors' homes to design VR or AR rehabilitation programs that simulate real-life challenges and create virtual training apartments, enhancing the transferability of rehabilitation practices to the home setting. Despite some doubts, many saw the potential of digital technologies to improve rehabilitation practices. The following comment was made:

> *Stroke survivors tend to walk very little. Many of them move very little. When they return home, they might train or have follow-up sessions with a physiotherapist maybe two, three days a week, at most. Some might attend a day center, but their overall movement remains limited. In that sense, if there were a low-threshold alternative to encourage them to get up and engage in physical activity would be beneficial.* [Informant 2.2]

Discussion

Principal Findings

This study explored how digital technologies shape the current and future practices of multidisciplinary HCPs in poststroke rehabilitation. According to the results and as linked to practice theory concepts illustrated in Figure 1, the primary agents (HCPs) collaborated with various other agents, including stroke survivors, relatives, interpreters, community HCPs, and NAV advisors, during their practices.

XSL•FO
RenderX

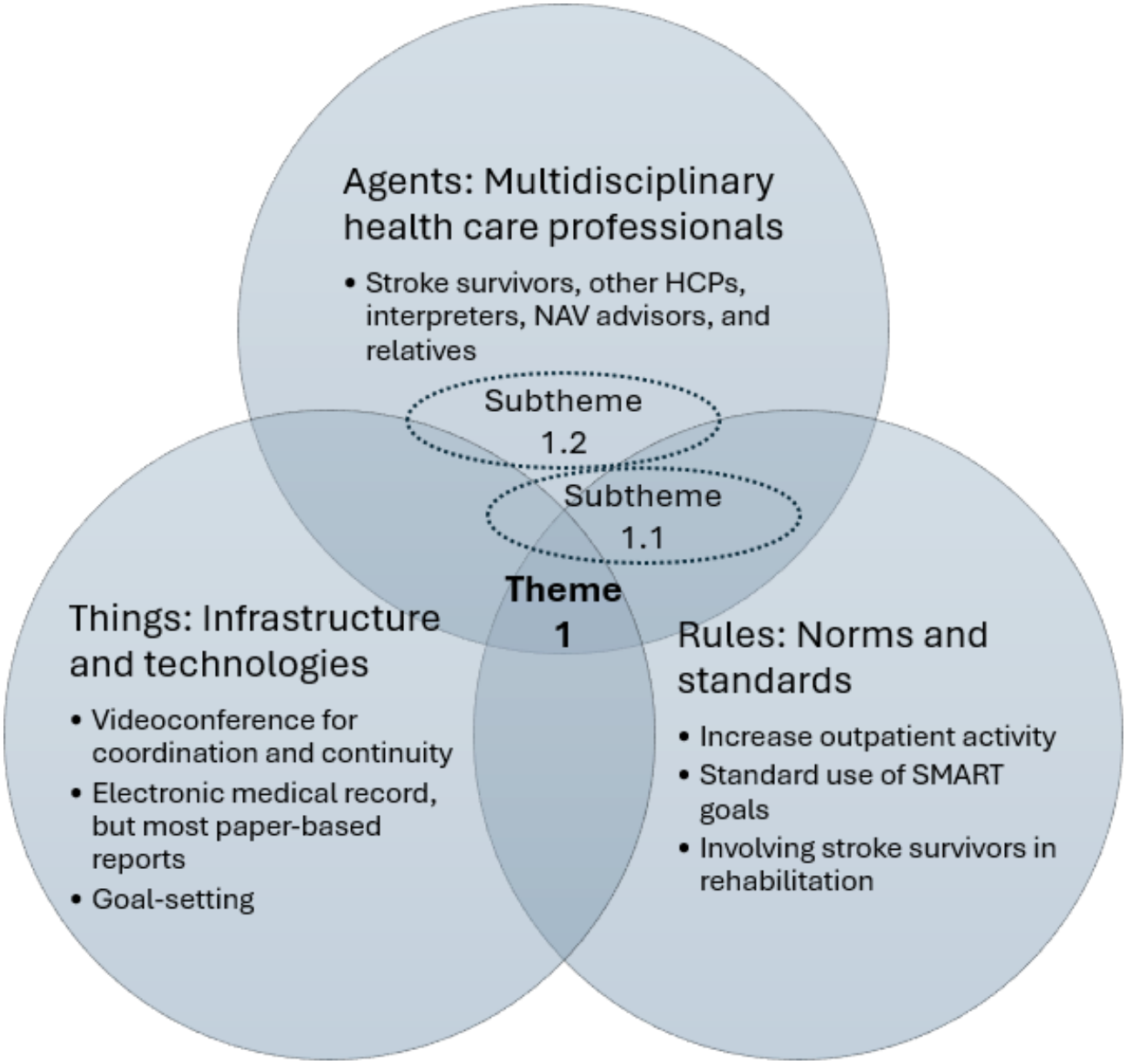
HCPs were influenced by a range of digital technologies, particularly videoconferencing, for rehabilitation coordination and continuity, although paper-based reports remained the primary communication tool across services and with stroke survivors. They used digital instruments for assessment and engagement, including gamified exercises. However, there was ambivalence in using digital technologies for multidisciplinary psychosocial support owing to concerns about their impact on human relationships.

Researchers have highlighted that contextual factors shape technology uptake, including professional discipline, rehabilitation settings, and device features [52]. The use of social practice theory helped us realize and identify how digitalization intersects with particular contextual factors that have not been emphasized in earlier stroke research, which can influence the digitalization of specialized poststroke rehabilitation practices. For example, this theory illuminated the relational dynamics between HCPs and stroke survivors by directing our attention to the required skills and know-how, as outlined in Table 2. Specifically, with regard to theme 2, this perspective highlighted how trust and relational depth are integral to the integration of digital tools. Additionally, this lens revealed how institutional structures, such as differing medical record systems and the digitalization of support conversations, altered established practice routines and were sometimes challenged or resisted by HCPs. Drawing from these core concepts and the results, this study highlights that the primary agents need digital therapy skills, potentially expanding to include knowledge of other digital platforms in cross-service collaboration. Previous research has shown that digital tools, such as apps, can give HCPs a better overview of stroke survivors' rehabilitation processes and can help in follow-up after discharge [53], particularly by using existing digital technologies in transitions between services. However, transitions between specialist health care and community-based services were found to be challenging in this study, requiring familiarization with different practices. Extended use of digital technologies could bridge this gap by enabling therapists to familiarize themselves with different practices during the coordination and continuity of rehabilitation. HCPs highlighted future unified digital systems for sharing medical records and test results across health care services, with goal-setting apps visible to both HCPs and stroke survivors. These could enhance coordination, user involvement in rehabilitation goals, and more active stroke survivor participation and agency in digital follow-up. Despite claims of low technological uptake among therapists [16], other longitudinal studies have found that in a multidisciplinary team, occupational therapists and physiotherapists used digital technologies more frequently across inpatient, outpatient, and community rehabilitation settings. However, studies agree that advancements are needed to match the rehabilitation context, including specific technology training and guidance of implementation [22]. Related responsibilities regarding whether the rehabilitation hospital or the community should lead this practice reorganization were found to be confusing among HCPs in this study. Additionally, they proposed future digital rehabilitation innovations visualizing stroke survivors' home environments and involving more community therapists while addressing safety concerns in stroke rehabilitation exercises. The rehabilitation process usually comprises the following 4 steps: assessment, goal-setting, intervention, and evaluation [54-56]. The results of this study relate to all aspects of rehabilitation and illustrate the complexity of how digitalization is shaping specialized stroke rehabilitation practices.

Consistent with our findings, other qualitative research has indicated that using digital technologies adds value in coordination and continuity, especially where user involvement is lacking. Comprehensive care coordination programs, including home telehealth, can aid stroke survivors and their relatives in managing stroke recovery through weekly real-time video calls and in-home messaging [57], enhancing accountability and self-management [58]. Moreover, tablet-based therapeutic programs can provide alternative access to rehabilitation services and follow-up. Studies have found positive attitudes among other HCPs toward digital apps for sharing stroke rehabilitation information with stroke survivors and their families [53]. However, researchers have highlighted the importance of infrastructure, equipment, space, and support in adapting telerehabilitation [17]. While ICT platforms facilitate communication and effective goal-setting among stakeholders in the rehabilitation process [32], they should also provide feedback to support the interest and progress of individuals with stroke [58,59], which is often limited or not adequately provided in current practices, as observed in this study. However, a range of digital technologies have been found to be integral to self-management for stroke survivors in the rehabilitation process, suggesting that these technologies can prevent therapists from taking over before users are able to do things themselves [20]. Nevertheless, both our findings and the findings of existing research have revealed a distinction between the therapeutic understanding of goal effectiveness and actual digital rehabilitation activities. Resistance to trivial digital technologies offering only engagement has been noted, and there are opportunities to target stroke survivor–specific goals and functional outcomes [53].

XSL•FO
RenderX

Figure 1. Diagram illustrating an example of the first overarching theme (with subthemes) and its relationship with social practice theory concepts. HCP: health care professional; NAV: Norwegian Labor and Welfare Administration; SMART: specific, measurable, achievable, relevant, and time-bound.



The exploration of how digital practices manifest within written documents and in community practices could be valuable, as discharge summaries are important for communication between health care levels and for ensuring continuity of rehabilitation [60]. There is potential for digitalizing written reports, and future research should be performed in this area. Although written reports are widely used, which is consistent with the findings of other studies, they lack the ability to facilitate both adherence and feedback [53]. While this was highlighted by the HCPs, it remains unclear whether therapists' decision-making regarding new technological infrastructure is constrained by institutional requirements of productivity, and if so, to what extent [61]. In this study, paper-based reports contrasted with the incentives of outsourcing digital information. Studies have found that stroke discharge summaries often report on activities of daily living and sensorimotor and general cognitive functions, but they frequently omit stroke survivors' needs and goals [60], and this is likely to increase with further digitalization and user feedback.

Our results contribute to the debate on whether technologies conflict with care and relationship building, emphasizing the need to transfer traditional communication skills, such as empathy, to digital platforms [62]. Claims about jeopardizing interpersonal connections due to being lost in data or a lack of in-person interactions have been made [63-66]. One reason technology may conflict with care is the limited sensory input it provides, highlighting the distinction between the critical social dimensions of humans as social sensors and some technologies as environmental sensors [67]. Despite this limitation, it has been found that using technologies can balance the roles toward a more symmetric relationship compared with the classic asymmetric stroke survivor-caregiver and HCP relationship often found in hospitals [62]. Digital platforms for collaboration, shared decision-making, and data storage can add value to a person-centered rehabilitation process [32], which, despite technology uptake, still drives rehabilitation and remains the rule of practice [52]. Video-based technologies enhance the ability to strengthen confidential and honest communication

XSL•FO
RenderX

[62], but they present ethical [24] and privacy considerations for family members, involve sharing sensitive information, and may invade stroke survivors' homes [68]. However, it is crucial not to lose or diminish face-to-face contact due to increased use of ICT, especially during the initial stage of establishing a relationship with a stroke survivor [12]. In our study, digital technologies interfered with some therapeutic communication skills, especially during communication with individuals having depressive or cognitive impairments, such as aphasia. Nonetheless, being reflexive about the technological environment in which therapists work and its impact on clinical practice has been advocated [69], despite disagreements in rating the importance of factors related to therapists' decision-making in implementing new technologies [61].

This study presented examples of therapists being unable to practice without certain technologies and concerns about losing therapeutic skills or positions owing to technology takeover. Concepts like technological determinism and disruption indicate that technologies shape societal structures and practices [70,71]. Similarly, the idea that new technologies disrupt social practices, norms, and even moral concepts has been revisited in the context of 21st-century technologies, particularly AI [71]. These concepts imply a reduced role for agents in changing practices, which we do not acknowledge, and this rather points back to the distribution of responsibility. In this study, we identified examples of technology that could shift responsibility and change routines. Therefore, new therapeutic communication skills are essential to adapt technology to people [70,72].

Assessments of stroke survivors' functional levels are needed not just in the initial phase but throughout rehabilitation to evaluate necessary adjustments. However, consideration of assessment accuracy must be taken into account [53]. Research supports that assessment is an important feature from the perspective of HCPs when using digital technologies in stroke rehabilitation [57,73]. In this study, digital monitoring enabled closer follow-up on progress and increased stroke survivors' self-management through collaboration with therapists. Shared understanding and priorities can promote technology implementation, including understanding its role relative to conventional therapy, training models, and evidence-based practice [52].

In motor rehabilitation, digital technologies appeared more complicated owing to the principles of intensity and repetition, which often require in-person assistance. While stroke rehabilitation may be too complex for full digitalization, partial implementation is feasible, especially combined with in-person support. Studies have found significant effectiveness of VR for varied outcomes [74], particularly for improving cognitive function compared with conventional care in a hospital setting [75]. Digital technologies for motor rehabilitation might be effective in a hospital setting but might face limitations in home environments [74]. In a previous study, participants continued the same types of exercises at home, transferring skills and preserving performance [75], which HCPs in our study noted as challenging. Additionally, studies have cautioned against the misfit between participants' needs and available technology, emphasizing the importance of coherent content and use to avoid misunderstandings and misuse in practice [32,53]. This approach depends on the stroke trajectory and the practice setting, as individuals may experience different recovery paths and require tailored interventions based on the stroke onset and the environment in which they receive rehabilitation [53]. The diverse perceptions between the 2 practices in this study appear to be related to engaging with people at different stages of stroke onset in the cross-section of hospitalization and home rehabilitation. The consideration of a continuous digital transition that follows stroke trajectories and becomes more self-managed in the chronic phase seems appropriate.

Researching rehabilitation practices raises questions about what is involved in stroke rehabilitation. Motor rehabilitation, as described by Kwakkel et al [43], involves a process that enhances the motor function, activity capacity, and daily performance of stroke survivors, aligning with the International Classification of Functioning, Disability and Health (ICF) framework [76]. Rehabilitation extends beyond restoring function, aiming to maximize independence and facilitate participation in home and community life [77]. Our study highlights varied perceptions of rehabilitation, underlining its multidisciplinary nature and the importance of coordinated and cross-professional contribution [32]. A digitalized rehabilitation service will require changes in the workflows of professionals, fostering cross-disciplinary collaboration to build capacity and ensure greater availability across time and settings. This shift will challenge traditional professional boundaries while breaking down silos between researchers and clinicians [78].

Strengths and Limitations

The data and findings in this study are unique as we applied social practice theory, which allowed us to comprehensively analyze the current and potential future influence of digital technologies for HCPs in poststroke rehabilitation and increased our awareness and reflexive interpretation of the contextual factors influencing these practices. It is acknowledged that the successful implementation of digital technologies depends on not only their inherent qualities but also how they are adopted and integrated into existing practices by HCPs. Maintaining an explorative approach, the findings are broad and consistent with the complexity and comprehensiveness of rehabilitation. A potential limitation of this study is the extended data collection period, which spanned 2 years but was divided into specific recruitment timepoints. While this allowed for a broader capture of experiences and perspectives, shifts in practice or external factors during this period may have influenced the findings. Some digital technologies, such as videoconference systems, VR tools, and webinar tools, were more frequently integrated into HCPs' daily practices at the first site during an earlier recruitment period (2022 - 2023), while step counters were more regularly used at the second site. These differences may reflect evolving practices influenced by project-based initiatives and specific time periods. These temporal variations should be considered when interpreting the results. Representatives of most multidisciplinary HCPs in specialized stroke rehabilitation participated in this study, with the majority recruited from a hospital rather than a community-based clinic. The inclusion of specialist stroke rehabilitation hospital wards provides a more specific context. As the participants emphasized cross-service collaboration and community collaboration, reflections from

XSL•FO
RenderX

community therapists could have added nuances to the findings. No psychologists participated as individual professionals, and only 1 HCP with a background in psychology took part. Future research could investigate the perspectives of psychologists and mental health specialists to expand the theme of psychosocial support and better tailor digital technologies to the psychological and emotional needs of stroke survivors. This study focused on the statements of HCPs regarding how they performed practices, while an ethnographic study would have provided detailed information about the implicit social structures in a specialized stroke rehabilitation practice. The use of RTA allowed for deep engagement with the data, enhancing transparency in the research process. Moreover, the collaboration between experienced researchers and clinicians ensured a robust and informed analysis, strengthening the study's credibility. Although this study covered current and proposed future technologies, it might have missed emerging innovations or future novel digital tools that the participants might not have been aware of, such as generative AI and robotics, which could potentially revolutionize practices. The study acknowledged challenges in stroke survivor follow-up continuity after referral, suggesting that future research is needed to explore detailed solutions and develop strategies for effective digital integration in rehabilitation.

Conclusion

This study found that HCPs used different technologies during their daily specialist care practices, as well as for the coordination and follow-up of stroke survivors after referral to community services. While all practice themes influenced rehabilitation practices, ambivalence and challenges were noted, particularly in digitalizing multidisciplinary psychosocial support and exercise programs. Despite ambivalence, digital technologies can serve as valuable supplements for stroke recovery, enhancing accessibility, competence, information flow, and quality care. This study identified several factors, such as organizational processes, roles, standards, and rules, that can act as barriers or drivers to implementing digital technologies in practices, emphasizing coherence among content (thing), use (agent), and aim (rules). Systems for sharing medical records and goal-setting apps, which are aimed at enhancing coordination and stroke survivor involvement, were emphasized as future digital technologies that can shape stroke rehabilitation. Viewing technology as a supplement to existing practices, rather than a singular solution for all areas of specialized stroke rehabilitation, offers significant potential for quality improvement. By focusing on leveraging already familiar technologies, specialized stroke rehabilitation integrates agent skills, physical tools, and practice guidelines, forming strategies to enhance the use and effectiveness of digital technologies in the field. Our findings may inform technology developers, health care personnel, and user groups in specialized neurological rehabilitation settings.

Acknowledgments

We would like to thank all the health care professionals who participated in the individual interviews in this study.

Funding

No funding was received for this study.

Data Availability

Complete data sets are not publicly available, as this was not outlined in the ethical approval and informed consent form. The data are available from the corresponding author upon reasonable request.

Authors' Contributions

All authors contributed to this study, gave final approval for the version submitted for publication, and agreed to be accountable for the work as presented.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 17 KB - rehab_v13i1e77753_app1.docx]

Multimedia Appendix 2
Mapping of qualitative data.
[PDF File, 75 KB - rehab_v13i1e77753_app2.pdf]

Multimedia Appendix 3

XSL•FO
RenderX

Six phases of reflexive thematic analysis.
[DOCX File, 21 KB - rehab_v13i1e77753_app3.docx]

References

1. Digital Health Platform Handbook: Building a Digital Information Infrastructure (Infostructure) for Health: World Health Organization; 2020.
2. Touré M, Poissant L, Swaine BR. Assessment of organizational readiness for e-health in a rehabilitation centre. Disabil Rehabil 2012;34(2):167-173. [doi: 10.3109/09638288.2011.591885] [Medline: 21936712]
3. DiCarlo JA, Gheihman G, Lin DJ, 2019 Northeast Cerebrovascular Consortium Conference Stroke Recovery Workshop Participants. Reimagining stroke rehabilitation and recovery across the care continuum: results from a design-thinking workshop to identify challenges and propose solutions. Arch Phys Med Rehabil 2021 Aug;102(8):1645-1657. [doi: 10.1016/j.apmr.2021.01.074] [Medline: 33556351]
4. Khasnabis C, Holloway C, MacLachlan M. The Digital and Assistive Technologies for Ageing initiative: learning from the GATE initiative. Lancet Healthy Longev 2020 Dec;1(3):e94-e95. [doi: 10.1016/S2666-7568(20)30049-0] [Medline: 36094191]
5. Haji Mukhti MI, Ibrahim MI, Tengku Ismail TA, et al. Exploring the need for mobile application in stroke management by informal caregivers: a qualitative study. Int J Environ Res Public Health 2022 Oct 10;19(19):12959. [doi: 10.3390/ijerph191912959] [Medline: 36232257]
6. Yew SQ, Trivedi D, Adanan NIH, Chew BH. Facilitators and barriers to the implementation of digital health technologies in hospital settings in lower- and middle-income countries since the onset of the COVID-19 pandemic: scoping review. J Med Internet Res 2025 Mar 6;27:e63482. [doi: 10.2196/63482] [Medline: 40053793]
7. Fischl C, Asaba E, Nilsson I. Exploring potential in participation mediated by digital technology among older adults. J Occup Sci 2017 Jul 3;24(3):314-326. [doi: 10.1080/14427591.2017.1340905]
8. Veerbeek JM, Langbroek-Amersfoort AC, van Wegen EEH, Meskers CGM, Kwakkel G. Effects of robot-assisted therapy for the upper limb after stroke: a systematic review and meta-analysis. Neurorehabil Neural Repair 2017;31(2):107-121. [doi: 10.1177/1545968316666957] [Medline: 27597165]
9. Lin JJ, Mamykina L, Lindtner S, Delajoux G, Strub HB. Fish'n'Steps: encouraging physical activity with an interactive computer game. Presented at: UbiComp'06: Proceedings of the 8th International Conference on Ubiquitous Computing; Sep 17-21, 2006. [doi: 10.1007/11853565_16]
10. Paul L, Wyke S, Brewster S, et al. Increasing physical activity in stroke survivors using STARFISH, an interactive mobile phone application: a pilot study. Top Stroke Rehabil 2016 Jun;23(3):170-177. [doi: 10.1080/10749357.2015.1122266] [Medline: 27077973]
11. English C, Ceravolo MG, Dorsch S, et al. Telehealth for rehabilitation and recovery after stroke: state of the evidence and future directions. Int J Stroke 2022 Jun;17(5):487-493. [doi: 10.1177/17474930211062480] [Medline: 34983266]
12. Gustavsson M, Ytterberg C, Guidetti S. Exploring future possibilities of using information and communication technology in multidisciplinary rehabilitation after stroke - a grounded theory study. Scand J Occup Ther 2020 Apr;27(3):223-230. [doi: 10.1080/11038128.2019.1666918] [Medline: 31545665]
13. White J, Janssen H, Jordan L, Pollack M. Tablet technology during stroke recovery: a survivor's perspective. Disabil Rehabil 2015;37(13):1186-1192. [doi: 10.3109/09638288.2014.958620] [Medline: 25212736]
14. Fruhwirth V, Berger L, Gattringer T, et al. Evaluation of a newly developed smartphone app for risk factor management in young patients with ischemic stroke: a pilot study. Front Neurol 2021;12:791545. [doi: 10.3389/fneur.2021.791545] [Medline: 35069420]
15. Taylor E, Fusari G, Darzi A, Jones F. Is a novel digital system for arm and hand rehabilitation suitable for stroke survivors? A qualitative process evaluation of OnTrack. BMJ Open 2023 Aug 18;13(8):e062119. [doi: 10.1136/bmjopen-2022-062119] [Medline: 37597873]
16. Langan J, Subryan H, Nwogu I, Cavuoto L. Reported use of technology in stroke rehabilitation by physical and occupational therapists. Disabil Rehabil Assist Technol 2018 Oct;13(7):641-647. [doi: 10.1080/17483107.2017.1362043] [Medline: 28812386]
17. Munce S, Andreoli A, Bayley M, et al. Clinicians' experiences of implementing a telerehabilitation toolkit during the COVID-19 pandemic: qualitative descriptive study. JMIR Rehabil Assist Technol 2023 Mar 10;10:e44591. [doi: 10.2196/44591] [Medline: 36897634]
18. Hestetun-Mandrup AM, Toh ZA, Oh HX, He HG, Martinsen ACT, Pikkarainen M. Effectiveness of digital home rehabilitation and supervision for stroke survivors: a systematic review and meta-analysis. Digit Health 2024;10:20552076241256861. [doi: 10.1177/20552076241256861] [Medline: 38832099]
19. Vourganas I, Stankovic V, Stankovic L, Kerr A. Factors that contribute to the use of stroke self-rehabilitation technologies: a review. JMIR Biomed Eng 2019;4(1):e13732. [doi: 10.2196/13732]
20. Hestetun-Mandrup AM, Hamre C, Lund A, Martinsen ACT, He HG, Pikkarainen M. Exploring people with stroke's perceptions of digital technologies in post-stroke rehabilitation - a qualitative study. Disabil Rehabil 2026 Jan;48(2):359-368. [doi: 10.1080/09638288.2025.2504615] [Medline: 40382688]

XSL•FO
RenderX

21. Willems EMG, Vermeulen J, van Haastregt JCM, Zijlstra GAR. Technologies to improve the participation of stroke patients in their home environment. Disabil Rehabil 2022 Nov;44(23):7116-7126. [doi: 10.1080/09638288.2021.1983041] [Medline: 34607474]
22. Pearce LMN, Pryor J, Redhead J, Sherrington C, Hassett L. Advanced technology in a real-world rehabilitation setting: longitudinal observational study on clinician adoption and implementation. J Med Internet Res 2024 Dec 30;26:e60374. [doi: 10.2196/60374] [Medline: 39753210]
23. Jarvis K, Thetford C, Turck E, Ogley K, Stockley RC. Understanding the barriers and facilitators of digital health technology (DHT) implementation in neurological rehabilitation: an integrative systematic review. Health Serv Insights 2024;17:11786329241229917. [doi: 10.1177/11786329241229917] [Medline: 38690403]
24. Silva GS, Andrade JD. Digital health in stroke: a narrative review. Arq Neuropsiquiatr 2024 Aug;82(8):1-10. [doi: 10.1055/s-0044-1789201] [Medline: 39187259]
25. Borg J, Gustafsson C, Landerdahl Stridsberg S, Zander V. Implementation of welfare technology: a state-of-the-art review of knowledge gaps and research needs. Disabil Rehabil Assist Technol 2023 Feb;18(2):227-239. [doi: 10.1080/17483107.2022.2120104] [Medline: 36103349]
26. Jette DU, Halbert J, Iverson C, Miceli E, Shah P. Use of standardized outcome measures in physical therapist practice: perceptions and applications. Phys Ther 2009 Feb;89(2):125-135. [doi: 10.2522/ptj.20080234] [Medline: 19074618]
27. Brouns B, Meesters JJL, Wentink MM, et al. Why the uptake of eRehabilitation programs in stroke care is so difficult-a focus group study in the Netherlands. Implement Sci 2018 Oct 29;13(1):133. [doi: 10.1186/s13012-018-0827-5] [Medline: 30373611]
28. Weiner BJ, Amick H, Lee SYD. Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields. Med Care Res Rev 2008 Aug;65(4):379-436. [doi: 10.1177/1077558708317802] [Medline: 18511812]
29. Wears RL, Berg M. Computer technology and clinical work: still waiting for Godot. JAMA 2005 Mar 9;293(10):1261-1263. [doi: 10.1001/jama.293.10.1261] [Medline: 15755949]
30. Dyb K, Berntsen GR, Kvam L. Adopt, adapt, or abandon technology-supported person-centred care initiatives: healthcare providers' beliefs matter. BMC Health Serv Res 2021 Mar 17;21(1):240. [doi: 10.1186/s12913-021-06262-1] [Medline: 33731078]
31. Frost J, Wingham J, Britten N, et al. The value of social practice theory for implementation science: learning from a theory-based mixed methods process evaluation of a randomised controlled trial. BMC Med Res Methodol 2020 Jul 6;20(1):181. [doi: 10.1186/s12874-020-01060-5] [Medline: 32631324]
32. Gåsvær JI, Jepsen R, Heldal I, Sudmann T. Supporting collaboration in rehabilitation trajectories with information and communication technologies: scoping review. JMIR Rehabil Assist Technol 2023 Jul 11;10(1):e46408. [doi: 10.2196/46408] [Medline: 37432715]
33. Lloyd A. Information Literacy Landscapes: Information Literacy in Education, Workplace and Everyday Contexts: Elsevier; 2010.
34. Harries T, Rettie R, Gabe J. Shedding new light on the (in)compatibility of chronic disease management with everyday life - social practice theory, mobile technologies and the interwoven time-spaces of teenage life. Sociol Health Illn 2019 Sep;41(7):1396-1409. [doi: 10.1111/1467-9566.12952] [Medline: 31124176]
35. Reckwitz A. Toward a theory of social practices: a development in culturalist theorizing. Eur J Soc Theory 2002;5(2):243-263. [doi: 10.1177/13684310222225432]
36. Tan KSY, Chan CML. Unequal access: applying Bourdieu's practice theory to illuminate the challenges of ICT use among senior citizens in Singapore. J Aging Stud 2018 Dec;47:123-131. [doi: 10.1016/j.jaging.2018.04.002] [Medline: 30447865]
37. Galvin R. Humans and stuff: interweaving social and physical science in energy policy research. Energy Res Soc Sci 2017 Apr;26:98-102. [doi: 10.1016/j.erss.2017.01.012]
38. Halford S, Savage M. Reconceptualizing digital social inequality. Information, Communication & Society 2010 Oct;13(7):937-955. [doi: 10.1080/1369118X.2010.499956]
39. Guni JN, Wanjala SW, Manguro G, et al. Using social practice theory in measuring perceived stigma among female sex workers in Mombasa, Kenya. BMC Public Health 2023 May 26;23(1):972. [doi: 10.1186/s12889-023-15809-2] [Medline: 37237349]
40. Niemelä R, Pikkarainen M, Ervasti M, Reponen J. The change of pediatric surgery practice due to the emergence of connected health technologies. Technol Forecast Soc Change 2019 Sep;146:352-365. [doi: 10.1016/j.techfore.2019.06.001]
41. Morse JM. The significance of saturation. Qual Health Res 1995 May;5(2):147-149. [doi: 10.1177/104973239500500201]
42. Gill SL. Qualitative sampling methods. J Hum Lact 2020 Nov;36(4):579-581. [doi: 10.1177/0890334420949218] [Medline: 32813616]
43. Kwakkel G, Stinear C, Essers B, et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. Eur Stroke J 2023 Dec;8(4):880-894. [doi: 10.1177/23969873231191304] [Medline: 37548025]
44. Langhorne PF, Coupar FB, Pollock AP. Motor recovery after stroke: a systematic review. Lancet Neurol 2009 Aug;8(8):741-754. [doi: 10.1016/S1474-4422(09)70150-4] [Medline: 19608100]

XSL•FO
RenderX

45. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
46. Rose A, Rosewilliam S, Soundy A. Shared decision making within goal setting in rehabilitation settings: a systematic review. Patient Educ Couns 2017 Jan;100(1):65-75. [doi: 10.1016/j.pec.2016.07.030] [Medline: 27486052]
47. Bomhof-Roordink H, Gärtner FR, Stiggelbout AM, Pieterse AH. Key components of shared decision making models: a systematic review. BMJ Open 2019 Dec 17;9(12):e031763. [doi: 10.1136/bmjopen-2019-031763] [Medline: 31852700]
48. Braun V, Clarke V. Thematic analysis: a practical guide. QMiP Bulletin 2022;1(33):46-50. [doi: 10.53841/bpsqmip.2022.1.33.46]
49. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qualitative Psychology 2022;9(1):3-26. [doi: 10.1037/qup0000196]
50. Braun V, Clarke V. Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: a review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). Palliat Med 2024 Jun;38(6):608-616. [doi: 10.1177/02692163241234800] [Medline: 38469804]
51. Bovend'Eerdt TJH, Botell RE, Wade DT. Writing SMART rehabilitation goals and achieving goal attainment scaling: a practical guide. Clin Rehabil 2009 Apr;23(4):352-361. [doi: 10.1177/0269215508101741] [Medline: 19237435]
52. Pearce LMN, Hassett L, Sherrington C, Pryor J. Human interactions remain at the heart of rehabilitation with advanced technology: a practice-embedded longitudinal qualitative study with allied health clinicians. J Neuroeng Rehabil 2025 Mar 7;22(1):52. [doi: 10.1186/s12984-025-01576-1] [Medline: 40050924]
53. Epalte K, Grjadovojs A, Bērziņa G. Use of the digital assistant Vigo in the home environment for stroke recovery: focus group discussion with specialists working in neurorehabilitation. JMIR Rehabil Assist Technol 2023 Apr 14;10(1):e44285. [doi: 10.2196/44285] [Medline: 37058334]
54. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
55. Wade DT. Describing rehabilitation interventions. Clin Rehabil 2005 Dec;19(8):811-818. [doi: 10.1191/0269215505cr923ed] [Medline: 16323380]
56. Lexell J, Brogårdh C. The use of ICF in the neurorehabilitation process. NeuroRehabilitation 2015;36(1):5-9. [doi: 10.3233/NRE-141184] [Medline: 25547759]
57. Lutz BJ, Chumbler NR, Roland K. Care coordination/home-telehealth for veterans with stroke and their caregivers: addressing an unmet need. Top Stroke Rehabil 2007;14(2):32-42. [doi: 10.1310/tsr1402-32] [Medline: 17517572]
58. Jayasree-Krishnan V, Ghosh S, Palumbo A, Kapila V, Raghavan P. Developing a framework for designing and deploying technology-assisted rehabilitation after stroke: a qualitative study. Am J Phys Med Rehabil 2021 Aug 1;100(8):774-779. [doi: 10.1097/PHM.0000000000001634] [Medline: 33141773]
59. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS. Telerehabilitation: review of the state-of-the-art and areas of application. JMIR Rehabil Assist Technol 2017 Jul 21;4(2):e7. [doi: 10.2196/rehab.7511] [Medline: 28733271]
60. Solbakken LM, Sundseth A, Langhammer B, Brovold T. Are physiotherapists and occupational therapists following the guidelines for discharge summary?-an analysis of the content of physiotherapists' and occupational therapists' discharge summaries and their adherence to stroke guideline recommendations. PLoS ONE 2024;19(9):e0308039. [doi: 10.1371/journal.pone.0308039] [Medline: 39226253]
61. Chen CC, Bode RK. Factors influencing therapists' decision-making in the acceptance of new technology devices in stroke rehabilitation. Am J Phys Med Rehabil 2011 May;90(5):415-425. [doi: 10.1097/PHM.0b013e318214f5d8] [Medline: 21765257]
62. Solli H, Hvalvik S. Nurses striving to provide caregiver with excellent support and care at a distance: a qualitative study. BMC Health Serv Res 2019 Nov 27;19(1):893. [doi: 10.1186/s12913-019-4740-7] [Medline: 31771566]
63. Narbutaitienė J, Björklund Carlstedt A, Fischl C. Stroke survivors' experiences and meaning of digital technology in daily life: a phenomenological study. Disabil Rehabil Assist Technol 2024 May;19(4):1334-1342. [doi: 10.1080/17483107.2023.2174605] [Medline: 36740734]
64. Walker RC, Tong A, Howard K, Palmer SC. Patient expectations and experiences of remote monitoring for chronic diseases: systematic review and thematic synthesis of qualitative studies. Int J Med Inform 2019 Apr;124:78-85. [doi: 10.1016/j.ijmedinf.2019.01.013] [Medline: 30784430]
65. Fiske A, Buyx A, Prainsack B. The double-edged sword of digital self-care: physician perspectives from Northern Germany. Soc Sci Med 2020 Sep;260:113174. [doi: 10.1016/j.socscimed.2020.113174] [Medline: 32659512]
66. Lucivero F. Lessons about so-called "difficult" patients from the UK controversy over patient access to electronic health records. AMA J Ethics 2017 Apr 1;19(4):374-380. [doi: 10.1001/journalofethics.2017.19.4.stas1-1704] [Medline: 28430572]
67. Nectoux S, Magee L, Soldatic K. Sensing technologies, digital inclusion, and disability diversity. J Comput Mediat Commun 2023 Aug 4;28(5). [doi: 10.1093/jcmc/zmad026]
68. Wherton J, Greenhalgh T, Hughes G, Shaw SE. The role of information infrastructures in scaling up video consultations during COVID-19: mixed methods case study into opportunity, disruption, and exposure. J Med Internet Res 2022 Nov 10;24(11):e42431. [doi: 10.2196/42431] [Medline: 36282978]

XSL•FO
RenderX

69. Killingback C, Naylor J. Emerging technologies in healthcare: interpersonal and client-based perspectives. In: Hayre CM, Muller D, Scherer M, Hackett PMW, Gordley-Smith A, editors. Emerging Technologies in Healthcare: CRC Press; 2024:1-14. [doi: 10.1201/9781003272786]
70. Hallström J. Embodying the past, designing the future: technological determinism reconsidered in technology education. Int J Technol Des Educ 2022 Mar;32(1):17-31. [doi: 10.1007/s10798-020-09600-2]
71. Löhr G. Conceptual disruption and 21st century technologies: a framework. Technol Soc 2023 Aug;74:102327. [doi: 10.1016/j.techsoc.2023.102327]
72. Lindberg J, Kvist E, Lindgren S. The ongoing and collective character of digital care for older people: moving beyond techno-determinism in government policy. J Technol Hum Serv 2022 Oct 2;40(4):357-378. [doi: 10.1080/15228835.2022.2144588]
73. Marwaa MN, Kristensen HK, Guidetti S, Ytterberg C. Physiotherapists' and occupational therapists' perspectives on information and communication technology in stroke rehabilitation. PLoS ONE 2020;15(8):e0236831. [doi: 10.1371/journal.pone.0236831] [Medline: 32857781]
74. Demeco A, Zola L, Frizziero A, et al. Immersive virtual reality in post-stroke rehabilitation: a systematic review. Sensors (Basel) 2023 Feb 3;23(3):1712. [doi: 10.3390/s23031712] [Medline: 36772757]
75. Torrisi M, Maresca G, De Cola MC, et al. Using telerehabilitation to improve cognitive function in post-stroke survivors: is this the time for the continuity of care? Int J Rehabil Res 2019 Dec;42(4):344-351. [doi: 10.1097/MRR.0000000000000369] [Medline: 31464812]
76. McDougall J, Wright V, Rosenbaum P. The ICF model of functioning and disability: incorporating quality of life and human development. Dev Neurorehabil 2010;13(3):204-211. [doi: 10.3109/17518421003620525] [Medline: 20450470]
77. Stinear CM, Lang CE, Zeiler S, Byblow WD. Advances and challenges in stroke rehabilitation. Lancet Neurol 2020 Apr;19(4):348-360. [doi: 10.1016/S1474-4422(19)30415-6] [Medline: 32004440]
78. Bernhardt J, Hayward KS, Kwakkel G, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the stroke recovery and rehabilitation roundtable taskforce. Neurorehabil Neural Repair 2017 Sep;31(9):793-799. [doi: 10.1177/1545968317732668] [Medline: 28934920]

Abbreviations

AI: artificial intelligence
AR: augmented reality
GP: general practitioner
HCP: health care professional
ICT: information and communication technology
NAV: Norwegian Labor and Welfare Administration
RTA: reflexive thematic analysis
SMART: specific, measurable, achievable, relevant, and time-bound
VR: virtual reality
WHO: World Health Organization

Edited by A Mastropietro; submitted 19.May.2025; peer-reviewed by H Sadeghsalehi, J Hao, NA Merriman; revised version received 16.Dec.2025; accepted 17.Dec.2025; published 17.Feb.2026.

Please cite as:
Hestetun-Mandrup AM, Hamre C, Lund A, Martinsen ACT, He HG, Pikkarainen M
Exploring the Influence of Digitalization on Multidisciplinary Poststroke Rehabilitation Practice: Qualitative Study
JMIR Rehabil Assist Technol 2026;13:e77753
URL: https://rehab.jmir.org/2026/1/e77753
doi:10.2196/77753

© Ann Marie Hestetun-Mandrup, Charlotta Hamre, Anne Lund, Anne Catrine Trægde Martinsen, Hong-Gu He, Minna Pikkarainen. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Experiences With Technology Among Adults Aging With HIV Engaged in an Online Community–Based Exercise Intervention Study: Longitudinal Qualitative Descriptive Study and Secondary Data Analysis

Julia Mucha[1*], MSc, MScPT; Rana Hamdy[1*], MScPT; Melina Marini[1], MScPT; Reda Aasem[1], MScPT; Chung Duong[1], MScPT; Tai-Te Su[1,2], PhD; Soo Chan Carusone[1,3], PhD; Kelly K O'Brien[1,4,5], BScPT, PhD

[1]Department of Physical Therapy, Temerty Faculty of Medicine, University of Toronto, 160-500 University Avenue, Toronto, ON, Canada

[2]School and Graduate Institute of Physical Therapy, College of Medicine, National Taiwan University, Taipei, Taiwan

[3]McMaster Collaborative for Health and Aging, McMaster University, Hamilton, ON, Canada

[4]Institute of Health Policy, Management and Evaluation (IHPME), Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[5]Rehabilitation Sciences Institute (RSI), University of Toronto, Toronto, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Kelly K O'Brien, BScPT, PhD
Department of Physical Therapy, Temerty Faculty of Medicine, University of Toronto, 160-500 University Avenue, Toronto, ON, Canada

Abstract

Background: As individuals with HIV live longer, many now face the health consequences of aging and multimorbidity, known as disability. Exercise can mitigate disability; however, engagement in exercise among adults living with HIV varies. Technology-based interventions, such as telerehabilitation, may help mitigate geographical, financial, and time barriers to community-based exercise (CBE). However, little is known about the experiences with technology uptake and usage among adults living with HIV. Understanding these experiences is essential to inform the design of inclusive, accessible, and sustainable online interventions.

Objective: This study aimed to describe experiences with technology uptake and usage among adults aging with HIV participating in a 6-month online CBE intervention and explore how these experiences changed over time, from baseline to postintervention.

Methods: We conducted a longitudinal qualitative descriptive study and secondary analysis using interview data from adults living with HIV who were engaged in a CBE intervention study in Toronto, Canada. Participants engaged in a 6-month online CBE intervention consisting of thrice-weekly exercise supervised biweekly through online personal coaching sessions, weekly group exercise classes, and monthly self-management education sessions (via Zoom). The technology used included Zoom software and a webcam, as well as the Sweat for Good YMCA app and the YMCA Virtuagym website; participants wore a wireless physical activity monitor (Fitbit Inspire 2) throughout. Participants completed interviews at baseline and postintervention. We conducted a group-based content analysis of interview transcripts, focusing on digital access, setup, usage, and perceptions of technology. Questionnaire data describing digital literacy and access to technology provided additional context to the interview data.

Results: Eleven participants completed at least one interview. We analyzed 19 interview transcripts from 11 participants (women: n=6, 55%; men: n=5, 45%; median age 52, IQR 45-60 y). Experiences with technology uptake and usage among adults aging with HIV were characterized by four components: (1) preparations for technology (technology setup), (2) interactions with technology (preferences for different types of technology, preferences for mode of delivery, and ease of usage), (3) facilitators and satisfaction with technology (facilitators to technology uptake and usage and satisfaction with technology), and (4) challenges and frustrations with technology (barriers to technology uptake and usage and frustrations with technology). Experiences with technology across participants were influenced by intrinsic contextual factors (prior exposure to technology) and extrinsic contextual factors (COVID-19 pandemic and technological and social support).

Conclusions: Experiences with technology among adults aging with HIV engaging in an online CBE intervention varied from increasing ease of use to increasingly burdensome over time. Results highlight the need to incorporate personal preferences and ongoing technological support when implementing online CBE with adults aging with HIV.

(JMIR Rehabil Assist Technol 2026;13:e86785) doi:10.2196/86785

XSL•FO
RenderX

KEYWORDS

HIV; exercise; telerehabilitation; telehealth; digital health; wearable technology; online coaching

Introduction

Approximately 65,000 people in Canada are living with HIV as of 2022, with an estimated 40.8 million people globally living with HIV at the end of 2024 [1,2]. Antiretroviral therapy has enabled people with HIV to live longer; however, many may also be living with associated multimorbidity affecting cardiovascular, autoimmune, musculoskeletal, and neurodegenerative systems [3,4]. These health challenges constitute disability that may manifest as episodic over time [5].

The Episodic Disability Framework conceptualizes disability as multidimensional and episodic, characterized by periods of wellness and illness [5,6]. The framework includes 6 dimensions of disability experienced by adults aging with HIV: physical, cognitive, and mental-emotional symptoms and impairments; difficulties with day-to-day activities; challenges to social inclusion; and uncertainty or worry about the future [5,6]. Intrinsic contextual factors (eg, personal attributes and coping strategies) and extrinsic contextual factors (eg, social support and stigma) can exacerbate or alleviate disability over time [5]. This framework provides a foundation to understand the health challenges experienced by adults living with HIV and the strategies to mitigate episodic disability.

Exercise and physical activity are rehabilitation strategies that can mitigate episodic disability among adults aging with HIV. Physical activity is defined as any type of bodily movement that results in energy expenditure, while exercise is a subset of physical activity that is planned, structured, and repetitive, with a goal of improving or maintaining physical fitness [7]. Systematic reviews indicate that adults living with HIV who engage in exercise at least three times weekly can experience improvements in cardiopulmonary fitness, strength, and quality of life [8-11]. The World Health Organization (WHO) recommends that adults, including persons living with chronic conditions, such as HIV, should engage in at least 150 minutes of moderate to vigorous intensity aerobic physical activity per week alongside muscle-strengthening activities on 2 or more days [12]. Despite the benefits, engagement in exercise can vary among adults living with HIV.

Systematic review evidence indicates that only 51% of adults living with HIV met exercise guidelines of 150 minutes of moderate to vigorous aerobic exercise weekly [13,14]. Barriers to exercise engagement were multifactorial, encompassing personal factors (lack of motivation, time constraints, and health concerns), social barriers (competing priorities and HIV-related stigma), and environmental challenges (transportation and financial constraints) [15-18]. Community-based exercise (CBE) programs can facilitate engagement in exercise among adults living with HIV, fostering social interaction and self-management strategies through organized exercise programs under supervision in community-based facilities [19,20]. In-person CBE can improve cardiovascular health, strength, and flexibility outcomes among adults living with HIV [21]. However, the complex and episodic nature of HIV, coupled with stigma, body image concerns, geographical barriers, and scheduling conflicts, can limit the engagement of persons living with HIV with in-person CBE programs [17].

Telerehabilitation, the delivery of rehabilitation services remotely using technologies, offers promising solutions to geographical, financial, and time barriers limiting access to exercise interventions [22-26]. Telecoaching, a component of telerehabilitation, uses technology for remote supervision and guidance to deliver an exercise program to improve physical activity levels and symptom management in chronic diseases [27-29]. Technology is an integral part of telecoaching CBE interventions to facilitate engagement in exercise; however, experiences with the uptake and usage of technology among adults aging with HIV engaged in such an intervention are unknown.

Technology encompasses digital tools designed to streamline care delivery, increase exercise engagement, and improve health outcomes [30]. This includes telecommunication platforms, web applications, and wireless physical activity monitors (WPAMs). Technology uptake refers to the initial adoption of technology by users, whereas usage refers to how individuals apply these tools, considering the frequency, functionality, effectiveness, and adherence of use [31]. Among adults aging with HIV, technology can support health-promoting behaviors and exercise engagement [25,26,32]. Systematic reviews highlight the role of technology in enhancing group exercise participation through social support and can help to alleviate barriers by addressing geographical, financial, and time constraints [33,34]. Technology-based exercise interventions are supported for use in chronic conditions such as Parkinson disease, chronic obstructive pulmonary disease, and stroke [27,29,35]. Despite positive trends in technology-facilitated exercise interventions, it is unclear how technology is taken up and used among adults aging with HIV.

Several factors can influence technology uptake and usage among adults living with chronic conditions, including prior exposure to, and awareness of, types of technologies, digital literacy, accessibility, and required support for use [36-39]. In HIV-specific contexts, there is growing interest in using technology to support engagement in exercise [32,40-42]. The purpose of this study was to describe experiences with technology among adults aging with HIV engaged in an online CBE intervention. The specific objectives were to describe (1) experiences with technology uptake and usage and (2) how experiences with technology may change over time among adults aging with HIV engaging in an online 6-month CBE intervention.

Methods

Study Design

We conducted a longitudinal qualitative descriptive study and secondary analysis using interviews and questionnaire data from

XSL•FO
RenderX

adults aging with HIV who participated in a 6-month online CBE intervention (Tele-Coaching CBE Study) [42]. Our qualitative design was supported by quantitative (questionnaire) data collected at baseline, month 2 (after initial exposure to the intervention), and month 6 (post intervention).

Context: Telecoaching CBE Study

The telecoaching CBE study was a prospective longitudinal mixed-method intervention study designed to evaluate the implementation of a 6-month online telecoaching CBE program for adults aging with HIV, delivered in collaboration with the Toronto YMCA [42]. Participants were community-dwelling adults living with HIV who self-identified as safe to engage in exercise based on the Physical Activity Readiness Questionnaire. Recruitment occurred through the Ontario HIV Treatment Network Cohort Study and community-based organizations in Toronto [42,43].

In phase 1 (intervention phase), participants were expected to engage in 60 minutes of exercise thrice weekly. Sessions were supervised biweekly by a YMCA personal trainer via Zoom, and participants were asked to attend monthly online self-management sessions for 6 months. In phase 2 (independent phase), participants continued with unsupervised exercise thrice weekly for 6 months. Participants were provided with a WPAM (Fitbit Inspire 2; Fitbit Inc) to track steps, distance, calories burned, and active minutes. Throughout both phases, participants received a YMCA online membership that provided access to synchronous group exercise classes via Zoom and recorded workouts (also called Avatar Workouts), YMCA Virtuagym website, and the Sweat for Good YMCA app. An engagement coordinator assisted participants with the setup of the technology and was available to provide support as needed throughout the intervention.

Investigators of the online CBE study conducted qualitative interviews via Zoom with a subset of participants at baseline (time point 1 [T1]) and postintervention (time point 2 [T2]) to explore their experiences and perceived impact of the intervention [44]. Full protocol details for the online CBE study have been published [42]. In our study, we focus on interview data pertaining to participants' experiences with technology during phase 1 of the telecoaching CBE study.

Technology in Online CBE Intervention

Technology was integral to the online CBE study, serving roles in (1) intervention delivery, (2) facilitating engagement in physical activity, and (3) measuring outcomes. For the purpose of this study, we defined technology as the *tools* (hardware, software, networks, and web applications) and *processes* (methods and strategies used for instruction, assessment, and physical activity tracking) used in the telecoaching CBE study. Key technological components included Zoom software (Zoom Communications, Inc) and a webcam for online personal training sessions and self-management education sessions; Sweat for Good YMCA app and YMCA Virtuagym website for live group exercise classes, prerecorded workouts, and recorded personalized workouts created by a personal trainer; WPAM (Fitbit Inspire 2) for motivation and measurement of physical activity; and technology-supported assessments, including online fitness assessments and interviews via Zoom, administration of patient-reported outcome measures via Qualtrics web-based survey software, body composition scale, and Fitbit app for syncing activity data [42]. In this study, we focused on experiences with technology pertaining to components of the intervention delivery.

Participants and Recruitment

Eligible participants for the telecoaching CBE study were required to have access to (1) a cell phone, tablet, laptop, or desktop computer; (2) internet access via Wi-Fi or a wired connection; (3) a webcam; and (4) a personal space for exercise. A subset of the participants in the telecoaching CBE study was invited to participate in the qualitative component of the study. Women living with HIV were purposively recruited to the qualitative study to enhance gender diversity within the sample.

In our study, we included community-dwelling adults aging with HIV (≥18 y) in Toronto, Canada, who considered themselves medically stable and who were safe to participate in the telecoaching CBE study and completed at least one qualitative interview at either baseline (T1) or postintervention (T2).

Patient and Public Involvement

We consulted with 3 community experts who self-identified as living with HIV. These consultations provided contextual insights into exercise and technology experiences within both in-person CBE and online CBE interventions. We met with each community expert via Zoom to discuss their experiences related to exercise engagement, intervention process, technology usability, and feasibility. These consultations provided foundational information for understanding the intervention and technology use and ensured our analysis and interpretation of interview data were informed by persons living with HIV.

Data Sources

Interviews

We used interview transcripts from semistructured interviews conducted at baseline (T1) and postintervention (T2) within the online CBE study to explore participants' experiences with technology uptake and usage and how experiences changed over time. The interviews were completed from October 8, 2021, to July 28, 2022, by members of the research team. We reviewed full transcripts, focusing our analysis on data that pertained to technology experiences. Interview questions at baseline about technology uptake and usage included: "Do you have any concerns or perceived challenges of participating in the upcoming CBE intervention?" and "How do you feel about adopting the technology necessary for the upcoming online CBE telecoaching intervention?" Interview questions at T2 focused on technology experiences, such as: "How would you describe your experience using the technology as part of the CBE telecoaching intervention?" and "Have there been any strengths or challenges associated with the translation of the telecoaching CBE intervention into the HIV community?" Interviewers used open-ended general and probing questions to explore technology uptake, learnability, usability, and change in experiences over time. Relevant interview guide questions

XSL•FO
RenderX

are highlighted in Supplemental File 1 in Multimedia Appendix 1. Participants received a $30 CAD (US $23) electronic gift card for their participation in each interview.

Questionnaires

We used 3 questionnaires to capture participant demographics and describe technology access characteristics, digital literacy, and telehealth experiences. The demographic questionnaire (administered at baseline) was used to describe participants' personal, health, and exercise behavior characteristics. The Information and Communications Technologies Development Index Questionnaire (administered at baseline) included 11 items to describe participants' digital access, connectivity, and literacy [45,46]. The Telehealth Indicator Questionnaire (administered at months 2 and 6) included 22 items with the aim to assess technologies used as part of the online CBE intervention across key telehealth domains, including usability (ease of use and navigation; 7 items), helpfulness (6 items), satisfaction (6 items), and reliability (3 items) [45,47].

Data Analysis

Interview Data

We conducted a collaborative group–based qualitative content analysis using a 6-step approach informed by the DEPICT model [48]. Five team members (JM, RH, MM, CD, and RA) independently read all transcripts, followed by group discussions to identify emerging codes related to technology uptake and usage. A preliminary codebook was developed based on the study objectives and operationalized technology experiences from the online CBE intervention study. We piloted the codebook on 3 transcripts with the most technology-related content. The team met 6 times to refine code definitions, incorporate barriers and facilitators to technology uptake and usage, and eliminate redundancies of codes. After 3 iterative revisions, 2 team members independently coded each transcript [48]. We used NVivo 15 software for data management [49]. For each participant, we wrote brief summaries of their technology experiences and interpretations of change over the intervention with the aim to identify patterns in how participants interpreted technology use and its impact over time. Our analytical process included team discussions for clarification, critical reflection, and consensus building [48]. To identify technology experiences over time, we coded each T1 and T2 transcript concurrently. We summarized individual experiences, documented reflections on changes in experiences, and identified data reflecting temporal changes. We discussed interview findings with questionnaire data to provide contextual interpretations.

Questionnaire Data

We analyzed the questionnaire data using Microsoft Excel [50]. For the Information and Communications Technologies Development Index Questionnaire, Telehealth Indicator Questionnaire, and demographic questionnaire, we calculated frequencies and percentages for categorical variables, and medians with 25th and 75th percentiles for continuous variables. For the Telehealth Indicator Questionnaire, we assessed the direction of change in responses between months 2 and 6 when responses differed between time points. For example, in the "satisfaction" domain, a change from "excellent" to "poor" indicated a negative change in participants' experiences with technology. Supplemental File 2 in Multimedia Appendix 2 provides the Telehealth Indicator Questionnaire. We included all questionnaire data where available at each time point.

We analyzed the interview and questionnaire data concurrently using descriptive approaches. We used the questionnaire data to inform our content analysis of the interview data, specifically by providing additional context to participant personal and health characteristics and their access to and familiarity with technology to enhance our interpretations and deepen our understanding of experiences with uptake and usage of the technology.

Sample Size Justification

We had a fixed sample of 11 participants from the online CBE study who completed at least one interview at T1 or T2. Of these, 8 participants completed both T1 and T2 interviews, 2 participants completed the T1 interview only, and 1 participant completed the T2 interview only. This sample size is comparable to those of prior qualitative studies examining technology experiences in telehealth exercise interventions among individuals with chronic conditions [51,52].

Ethical Considerations

Ethics approval for this study (protocol 47552) and the broader telecoaching CBE study (protocol 40410) was obtained from the University of Toronto Health Sciences Research Ethics Board. All participants in the telecoaching CBE study who completed an interview provided written or verbal informed consent. Data sources including transcripts and questionnaires were deidentified, meaning participant data were identified using a participant number to ensure confidentiality of participants. Participants received a CAD $30 (US $23) electronic gift card for their participation in each interview.

Results

Characteristics of Participants

Of the 11 participants (median age 52, IQR 45-60 y), 6 (55%) identified as women, 5 (45%) identified as heterosexual, and 6 (55%) identified as Black. The majority of participants (n=9, 82%) had an undetectable viral load. Five (45%) participants reported living with ≥3 concurrent health conditions in addition to HIV. Exercise behavior among participants prior to the intervention varied. Three (27%) participants did not engage in exercise in the past week, 5 (45%) engaged in 1 to 2 days, and 3 (27%) engaged in 5 or 7 days of ≥30 minutes of moderate to vigorous intensity exercise in the past week. Table 1 provides participant characteristics.

XSL•FO
RenderX

Table . Characteristics of participants (n=11).

Characteristic	Participants (n=11)
Personal characteristics	
Age (y), median (25th, 75th percentile)	52 (45, 60)
Sex, n (%)	
Female	6 (55)
Male	5 (45)
Gender, n (%)	
Woman, cis-woman	6 (55)
Man, cis-man	5 (45)
Sexual orientation, n (%)	
Heterosexual	5 (45)
Gay	3 (27)
Queer	1 (9)
Bisexual	1 (9)
Other	1 (9)
Living arrangement, n (%)	
Living with others	7 (64)
Living alone	4 (36)
Average gross yearly income (CAD), n (%)	
<$20,000	4 (36)
$20,000 to <$50,000	5 (45)
>$50,000	2 (18)
Primary source of income, n (%)	
Ontario Disability Support Program	4 (36)
Full-time employment	3 (27)
Pension plan	2 (18)
Part-time employment	1 (9)
Casual work	1 (9)
Education level, n (%)	
Completed college (received degree or diploma)	5 (45)
Completed university (received degree)	3 (27)
Postgraduate education	3 (27)
Ethnicity[a], n (%)	
Black	6 (55)
South Asian	3 (27)
Latin American, Hispanic, or Latino	1 (9)
Southeast Asian	1 (9)
White	1 (9)
Clinical characteristics	
Year of HIV diagnosis, median (25th, 75th percentile)	2002 (1995, 2006)
Taking antiretroviral medication, n (%)	11 (100)
Number of missed doses in past 7 d, n (%)	
No doses missed	7 (64)

XSL•FO
RenderX

Characteristic	Participants (n=11)
1 dose missed	4 (36)
Undetectable viral load (<50 copies/mL)	9 (82)
Recent CD4 count, n (%)	
>500 cells/mm^3	9 (82)
<500 cells/mm^3	2 (18)
Number of participants with ≥3 comorbidities, n (%)	5 (45)
Self-reported health status, n (%)	
Very good	7 (63)
Good	2 (18)
Fair	1 (9)
Poor	1 (9)
Exercise behavior characteristics, n (%)	
Days per week physically active ≥30 min at a moderate or vigorous intensity in the past week prior to intervention, n (%)	
0 (none)	3 (27)
1 d	2 (18)
2 d	3 (27)
5 d	1 (9)
7 d	2 (18)
Engaged in >300 min of moderate intensity activity or >150 of vigorous intensity activity in the past week prior to intervention, n (%)	2 (18)
Impact of COVID-19 pandemic on physical activity, n (%)	
Exercises at a greater frequency or intensity	5 (45)
Exercises at a lesser frequency or intensity	5 (45)
No impact on exercise	1 (9)

[a]One participant listed 2 ethnicities.

Access To, Skills in, and Use of Technology at Baseline

At baseline, most participants had access to a cell phone (n=10, 91%), a laptop with a webcam (n=9, 82%), and reliable Wi-Fi access (n=9, 82%). All participants (100%) reported they were comfortable using web browsers, online meeting tools, WPAMs, and fitness apps at baseline. Table 2 provides participant characteristics regarding access to, skills in, and use of technology.

XSL•FO
RenderX

Table . Access to, skills in, and use of technology at baseline (N=11).

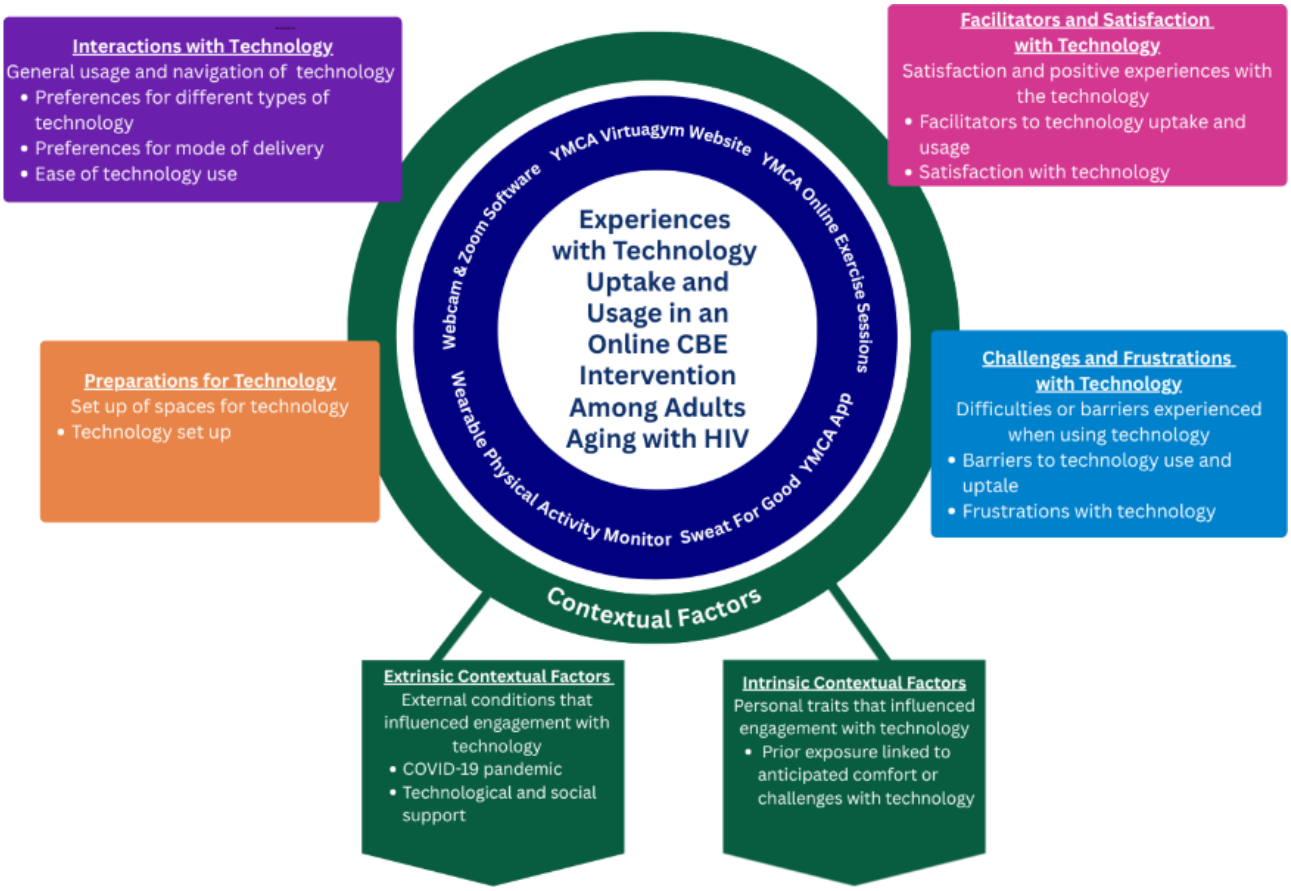
Information and communications technologies questionnaire variable	Participants, n (%)
Access to technology	
Access to a type of technology	
Cell phone	10 (91)
Laptop computer with webcam	9 (82)
Tablet	5 (45)
Desktop computer with webcam	4 (36)
Desktop computer without webcam	0 (0)
Most reliable access to the internet	
Through Wi-Fi	9 (82)
Through a wired connection	2 (18)
Skills and use of technology	
Comfortable using the World Wide Web with a browser	11 (100)
Comfortable sending and receiving emails	11 (100)
Comfortable using online meeting tools	11 (100)
Comfortable using wireless physical activity monitor	11 (100)
Comfortable using webcam	11 (100)
Comfortable using physical activity apps to track physical activity	11 (100)
Confirmed skills and use of technology	11 (100)

Experiences With Technology Uptake and Usage

Participants' experiences with technology during the online CBE intervention varied. We organized our findings into four interrelated components: (1) preparations for technology, (2) interactions with technology, (3) facilitators and satisfaction with technology, and (4) challenges and frustrations with technology. Participants' experiences with technology were shaped by both intrinsic contextual factors, such as their prior exposure and anticipated comfort with technology, and extrinsic contextual factors, including the COVID-19 pandemic and availability of technological and social support. These factors affected how participants engaged with the technology (uptake and usage) in the study, as well as the benefits and challenges they experienced (Figure 1). Below, we present each component with supportive quotations identified by participant number and interview time point.

XSL•FO
RenderX

Figure 1. Experiences with technology uptake and usage in an online community-based exercise (CBE) intervention among adults aging with HIV.



Preparations for Technology

This component captures how participants prepared their physical spaces and devices to engage with the study. Preparation was not limited to technical setup; it also involved adapting the home environments to accommodate space for exercise and facilitate active participation. These preparatory steps often required creativity and problem-solving. Most participants reported having adequate space for exercise, although some required modifications, as illustrated below:

> *Other than the sense to figure out where my laptop would be so that the instructor could see me and figure out how my apartment would work. It's a small apartment... I wish I didn't have to use my box [exercise equipment] because it would be great to put my laptop on to get the whole view of my workout area. But no. I think I'll figure it out. I've just got to move some things, just change the angle of some things.* [P2, T1]

The following participant described their experiences trying to set up their Zoom to the TV for optimal viewing of their personal trainer during the online session:

> *I'm hoping I can run a Zoom meeting off my iPhone. But I think what I'm going to need is get a cable to connect to the HDMI cable on the TV. So it's just a bit of technical stuff to get all that set up. I've got a little tripod that I can set the phone on. The microphone will work through the phone. It will just be a question of running the display out of the phone with a cable into the HDMI connection.* [P5, T1]

Interactions With Technology

Participants described a variety of experiences interacting with technology, reflecting differences in preferences, confidence, and adaptability. Participants' descriptions of their experiences interacting with technology focused primarily on preferences for different types of technology, preferences for mode of delivery, and ease of technology use.

The preferences for types of technology participants used included Zoom, Fitbit, computer, laptop, webcam, Fitbit app, and Sweat for Good YMCA app. Zoom was the primary platform for online telecoaching sessions with the personal trainer, with most participants accessing Zoom via a laptop or desktop with a webcam, while a smaller number preferred mobile phones or connecting Zoom to their TVs to enhance convenience or visibility, respectively.

The Fitbit device and the Fitbit app were largely used to track step counts by most participants. The following participant explained the use of the Fitbit:

> *I like wearing it [Fitbit] but not for the reason for exercise because I don't understand anything on it. I like it for two reasons; number one, I can see steps and number two I can see time [...]. Besides that, because I don't know technology, I didn't really benefit from the usefulness of having a Fitbit that it*

XSL•FO
RenderX

> *can tell you more stuff. But for these two reasons it was good enough for me.* [P11, T2]

The Sweat for Good YMCA app was used to attend online group exercise classes, prerecorded workouts, and the Avatar workouts, offering participants flexibility with both synchronous and asynchronous options. Preferences for mode of delivery varied between synchronous and asynchronous modes of online exercise delivery. Some found the synchronous Zoom training sessions straightforward, while others found it stressful due to technical challenges and preferred the convenience of the asynchronous classes of the Sweat for Good YMCA app. The following participant struggled with navigating the Sweat for Good YMCA app stating:

> *However, first of all trying to get everything set up for a synchronous event I just found it incredibly stressful again because of trying to navigate the technical challenges. Then when I saw the instructor on the screen they were doing some warmup things at a pace that I could barely keep up with.* [P5, T2]

This same participant (P5) specifically struggled to maintain pace during the synchronous exercise class through the Sweat for Good YMCA app, stating:

> *So the synchronous thing, it just wasn't for me. Really the thing that has most worked for me is just knowing routines that [name of personal trainer] has given me and writing them down and keeping up with that on my own. The less I have to interface with technology in this kind of remote setup, the better.* [P5, T2]

Ease of technology use referred to how easy participants found the technology to be. Many participants anticipated positively adopting the new technology at T1. However, their actual ease of use varied once they began engaging with the technology. This participant describes their experiences with their ease-of-use learning about the technology:

> *With me in trying to figure out technology, it was straightforward. Like once I plugged in to charge it, it told me to download the App and then once I did that, it was a step by step. It's more still figuring out the features. Like, I had no idea that it already recorded my sleep overnight until my roommate and bestie looked it up and was like okay.* [P2, T1]

Another participant was surprised to find the use of the technology to be challenging and required additional support to navigate:

> *Well initially… and this kind of took me by surprise. I didn't realize how difficult the technical part of things would be. It was horrendously difficult, the technical part. In fact, AM, we gave up an entire hour of training because AM got her daughter involved on her daughter's laptop, and we went through setting by setting to try to eliminate the terrible echo. You know I'm a big advocate of clear communication and Zoom has done a terrible job and most software providers do a terrible job of explaining things.* [P5, T2]

Facilitators and Satisfaction With Technology

This component explores what helped participants successfully adopt and use technology and how these factors contributed to their satisfaction. Facilitators ranged from supported features such as activity tracking and sharing their progress in social circles to practical benefits such as convenience and cost savings. These elements often reinforced participants' motivation and sense of accountability.

One participant described using the Fitbit to stay accountable:

> *So my mom has a Fitbit too and I added her as a family member or friend. So now she gets to see how much I am doing for the day. There will be days that she'll talk to me or whatever or like I'll send her my badges that I've won* [P2,T2]

The following participant appreciated not having to commute to a gym facility. They highlighted the convenience and cost-saving benefits of using Zoom rather than having an in-person personal training session:

> *The advantage of just having to get up and not have to drive anywhere. Just get up and you know have your Zoom on. So that was very convenient, very convenient. Yeah. I'm not spending gas. Like gas prices are ridiculous right now.* [P1, T2]

Satisfaction with technology referred to participants' approval with the online platforms they engaged with through the CBE intervention. Some participants indicated that they were satisfied with having access to the tools that supported their goals. Several participants described the Fitbit as helpful for monitoring their progress with physical activity. Many participants found that the weekly summary reports from Fitbit were rewarding and motivated their continued engagement:

> *A part of the study that worked best for me. For me I would say like I have best... okay definitely Fitbit. Fitbit is like... not only having the Fitbit. It helped me to track all my things right; the steps, those all sleeping patterns and even like weekly we have to something that report. Those are all in the Fitbit Apps. So those things that, seeing those all tracked, So those things that, seeing those all tracked, it satisfied me. It motivates me. I want to see like in the number or in the papers and these apps really help give me all those things what I'm looking* [P8, T2]

One participant talked about how features of the Fitbit as well as the online platforms encouraged their participation:

> *It's sort of like an encouragement to do more. These watches are good because I've sort of got used now to whenever I'm out walking to enter the time. The classes, the sessions or lectures that you have are very good.* [P3, T2]

Challenges and Frustrations With Technology

Participants also experienced challenges and frustrations with the study technology. At times, these challenges hindered engagement in components of the intervention.

XSL•FO
RenderX

Zoom was generally perceived as a useful platform; however, some participants reported technical challenges, including audio delays, echoing, and unstable internet connections. At time point 1, one participant shared their concerns about the stability of their internet:

> *In the neighborhood I live in, the internet has gone down I don't know maybe two handfuls of a time within the past six months. So you never really know.* [P2, T1]

Another participant talked about the lack of accessibility of the Fitbit features:

> *The second is if you've got any issues around sight, any, I mean even if you don't have issues around sight, you can't see these things. You go into the... you go outside, it's... so like you do your walk and you're finished and you come back and you go am I at the screen where it's turned off? You can't tell because you can't read it. And of course the writing is so small that people that have impaired vision, even if you can see it, it's difficult to read.* [P10, T2]

Frustrations with technology were tied to the reliability of the technology and to design limitations. This included Zoom audio delays, sound issues, echoing, and connectivity problems. Participants noted that Zoom provided insufficient support, requiring them to troubleshoot issues independently, as per the following participant:

> *It's just technical stuff, those little quarks about my particular setup. There was a time when I was having audio issues, and it was downright stressful to be there [self-management sessions] for the hour and a half-* [P5, T2]

Another participant shared their frustrations related to the Fitbit, including device setup, blank display screen, and concerns about inaccurate tracking:

> *The biggest demotivator for this whole thing is that stupid Fitbit that they gave us, utterly useless, good for nothing, utterly useless. The only thing I use it for is to count whether it's counting the number of steps. That's it. Every month since I got it there's been some issue with it. For some reason it went into some other country time on its own and then at some point it just completely the screen was blank and it dimmed on its own at some point. Every time it does something on its own.* [P7, T2]

Contextual Factors

Participants' experiences with technology were shaped by a range of contextual factors that influenced both uptake and ongoing use of technology in the online CBE intervention. Together, intrinsic and extrinsic contextual factors influenced how participants approached technology, the challenges they anticipated, and the strategies they used to overcome barriers.

Intrinsic contextual factors included prior exposure to technology that influenced a participant's anticipated comfort or challenges with the forthcoming technology in the online CBE intervention. Six (54.5%) of the 11 participants previously or currently owned WPAMs, such as an Apple Watch, Samsung Watch, or a Fitbit. Many participants had previously used Zoom prior to the intervention. Participants with prior exposure to technology often reported increased confidence of use at baseline and anticipated greater comfort with technology use, as this participant stated:

> *That's okay. I have a degree in computer sciences so it's not an issue.* [P7, T1]

When another participant was asked about perceived challenges regarding updating software with technology associated with the CBE intervention, they responded:

> *No. As far as updating the equipment and stuff I'm fairly... that goes fairly easy for me.* [P10, T1]

Extrinsic contextual factors included the COVID-19 pandemic and available technical and social support throughout the intervention. Many participants reported that the COVID-19 pandemic normalized the use of technologies, such as Zoom. For some participants, the COVID-19 pandemic led them to their first exposure to technology as it forced them to engage with technology in new ways. This participant reflected on both the safety and convenience of technology as well as the convenience of exercising from the comfort of their own home:

> *One is the times that we're in with COVID, I mean I think there's a safety component to this. …. for a lot of people too because a person's immunity is a vital part of the constitution. I think the convenience of being able to just be at home and in the comfort of your own home is helpful.* [P10, T1]

Another participant acknowledged the motivational pull that comes from an in-person gym setting yet still preferring online sessions for safety concerns that began during the COVID-19 pandemic:

> *To be honest while in the pandemic I prefer to going online. That way I'm safe, I'm good. But also going into the gym and seeing others working out kind of pushes you like listen that person can do it, why can't I. So, it kind of gives you that push compared to online because online is just me and my trainer. I know they will probably push you but it's not the same as actually seeing somebody else, maybe somebody else bigger or somebody older actually doing it easier than you. So, it kind of gives you that push. But I'm more than happy with online right now.* [P7, T1]

Availability of technical and social support played a key role in how some participants planned to address technical challenges in the study. Some participants mentioned the technology support services having insufficient troubleshooting resources and resulting in alternative methods to troubleshoot. This participant reported:

> *I wouldn't say they [technical issues] become an obstacle. Normally when I can't figure out something I just Google or go on YouTube and figure out what it is, what's going on and how to fix it.* [P6, T1]

Other participants depended on their social support network to assist with device setup or to resolve technical issues. When

XSL•FO
RenderX

asked about navigating technical difficulties at home, 1 participant shared how they relied on their children for assistance:

> *If I can't function something I just call them, 'hey show me how this works.* [P1, T1]

Changes in Experiences With Technology Within the CBE Intervention Over Time

Interview data revealed participants had inconsistent experiences with uptake and usage of technology over the duration of the intervention. Some participants experienced increasing levels of technology fatigue and burden over the course of the intervention. Some participants found technology a challenge at both T1 and T2. For 1 participant, by T2, their concerns evolved into frustrations with the technology expressing technology fatigue and burden, particularly with the requirement to download and sync their physical activity each week for the study:

> *This synching up for the Fitbit and the recording on what's gone on over the week, it gets to be a bit of a pain to be honest.* [P5, T2]

Although participants attempted to problem-solve their challenges, frustrations continued with efforts to use the Fitbit. The requirement to sync their Fitbit data weekly became a burden and at times resulted in disengagement and frustration. In contrast, another participant was optimistic with using the Fitbit at T1:

> *Having the coach itself is a very good benefit right...it will motivate me -- oh I'm wearing a Fitbit, so I have to do, let's just do it' and also do the exercise.* [P8, T1]

By T2, this participant's expectations appeared to have been met, and they expressed their intent to continue using the technology beyond the study*:*

> *even giving this Fitbit, I will use it unless it will not damage. If it damage, then definitely I'm going to buy another one because Fitbit gives all that report.* [P8, T2]

Another participant expressed uncertainty about technology usage in T1 and had an indifferent attitude toward the Fitbit and its role in the intervention, stating:

> *I mean I think they are all the same. They don't have any difference. So I don't know. Like what information do you want from these watches?* [P3, T1]

However, by T2 their perspectives regarding the Fitbit had changed:

> *It's good. Fitbit is… after the first or second time that you enter the information online it becomes very easy. Yeah, it's good.* [P3, T2]

This participant experienced a learning curve but was able to reduce barriers and increase usability of the Fitbit with continued use. Overall, initial attitudes toward the technology did not always predict participants' engagement with participants' engagement or satisfaction during the study. Some participants encountered technology fatigue around repetitive tasks such as Fitbit syncing, while others reported increased comfort and motivation.

Telehealth Indicator Questionnaire Responses

Eight (73%) participants completed the Telehealth Indicator Questionnaire at months 2 and 6 of the intervention. Across the 8 participants, the majority (n=6, 75%) demonstrated no change in most responses to the 22 items at months 2 and 6. Two (25%) participants had a change in a positive direction at month 6 in 10 (46%) and 12 (55%) of the 22 items. Supplemental File 3 in Multimedia Appendix 3 provides each participant's direction of responses to the 22 items of the Telehealth Indicator Questionnaire, related to usability, helpfulness, satisfaction, and reliability of the technology in the CBE intervention captured at months 2 and 6.

Across the 22 items, the majority of participants (50% - 88%) reported similar responses (representing no change) for most items (14/22, 63%) at months 2 and 6 of the intervention. An exception was responses to the usability and helpfulness of self-directed learning modules, whereby 5 participants (62%) demonstrated a positive direction of change in their response from months 2 to 6. Conversely, 4 (50%) participants showed a negative change in direction in responses to the usability of the Fitbit and helpfulness of the YMCA Virtuagym website from months 2 to 6 (Supplemental File 3 in Multimedia Appendix 3).

Discussion

Overview of Findings

Experiences with technology uptake and usage among adults aging with HIV were characterized by 4 components: (1) preparations for technology, (2) interactions with technology, (3) facilitators and satisfaction with technology, and (4) challenges and frustrations with technology. Experiences with technology interacted with intrinsic contextual factors (prior exposure linked to anticipated comfort or challenges with technology) and extrinsic contextual factors (COVID-19 pandemic; and technological and social support). Some participants experienced increasing technology burden and fatigue over time, others maintained neutral perspectives throughout, while others developed intentions for continued use postintervention.

Experiences With Technology Uptake and Usage

Technology in the online CBE intervention included a combination of digital tools and platforms, including WPAMs (Fitbit Inspire 2), a videoconferencing platform (Zoom), webcams, the Sweat for Good YMCA app, and YMCA Virtuagym website. Participants demonstrated variation in experiences with technology uptake and usage across different technologies, with preferences emerging for specific types of technology and modes of delivery. Synchronous coaching sessions via Zoom achieved relatively high amounts of uptake following initial setup, likely facilitated by prior exposure gained during the COVID-19 pandemic [53]. The widespread adoption of Zoom during the COVID-19 pandemic likely created a foundational level of comfort that translated well to the synchronous coaching sessions, with participants expressing

XSL•FO
RenderX

appreciation for the flexibility and immediate feedback available through this platform [26,53]. In contrast, participants in this study had variable experiences with the Sweat for Good YMCA app and WPAMs. While WPAMs were primarily used for real-time monitoring of physical activity metrics such as step counts and sleep patterns, the Sweat for Good YMCA app was used to administer asynchronous exercise content. Participants had varied preferences for synchronous or asynchronous tools that had implications for engagement with technology within the intervention. This has direct implications for considerations of technology uptake and usage in digital intervention design, suggesting the need to align online CBE interventions with user preferences for both the type of technology and the mode of delivery. Hence, it is important to recognize that adults aging with HIV may respond differently to synchronous tools that offer real-time human connection versus asynchronous tools that require greater independence [17,18,26].

All participants recruited for the telecoaching CBE study were required to have access to and comfort with the usage of technology. Despite all participants reporting comfort with different forms of technology at baseline, participants expressed challenges with uptake and usage of technology in the intervention. Of the 11 participants, 6 (66%) had prior exposure to a WPAM; however, this was insufficient for sustained uptake and usage of the Fitbit Inspire 2. Additionally, participants voiced inconsistent usage of more complex platforms, such as the Sweat for Good YMCA app. The disconnect between high baseline comfort and subsequent struggles with technology may reflect the requirement for participants to simultaneously navigate multiple technology platforms in the intervention that exceeded individual comfort with each component of the technology [18,26]. These findings suggest that technological complexity, rather than basic digital technology skills, posed a greater barrier to sustained use among adults living with HIV [39]. Future interventions should move beyond one-time onboarding to include tailored, continuous support that addresses and simplifies multiplatform navigation and adapts to participants' evolving needs among adults aging with HIV [17,39,54]. Additionally, future CBE interventions may consider offering a broader range of WPAMs, such as commercially available smartwatches and fitness trackers. Tailoring digital health wearables based on individual preferences, digital health literacy, and accessibility may provide customizable feedback and enhance usability and sustained use over time.

Participants highlighted several facilitators and points of satisfaction with the technology used in the online CBE intervention, including flexible access to workouts, real-time interaction with personal trainers, and digital tools for self-monitoring physical activity. Zoom was consistently highlighted for its convenience and flexibility, supporting real-time interaction with personal trainers, accountability, and routine building, factors shown to improve exercise adherence among adults aging with HIV [17,54,55]. The WPAM served as a motivator for some, enabling self-monitoring of physical activity, steps, and sleep, which some participants found satisfying and empowering. Similar to prior studies, gamification features of WPAMs and real-time feedback supported users' sense of progress and enhanced intrinsic motivation among adults aging with HIV [30,32,40]. The Sweat for Good YMCA app provided asynchronous, on-demand workouts that aligned with participants' preferences for flexibility and autonomy, reinforcing the importance of choice between synchronous and asynchronous formats among adults aging with HIV [55]. These findings suggest that digital tools and processes can support engagement in exercise among adults aging with HIV when designed to be adaptable, user-friendly, and reinforced through personalized feedback and social accountability [18,41,55]. Future interventions should leverage these features by integrating reliable self-monitoring technologies and offering multiple modes of delivery to enhance engagement in online CBE interventions among adults aging with HIV.

Challenges and frustrations with technology use in this study included device syncing failures and navigation difficulties with the Sweat for Good YMCA app and connectivity problems with Zoom videoconferencing. WPAMs were valued for their feedback features but also generated frustration when step counts were perceived as inaccurate or syncing required repeated effort. This aligns with systematic measurement errors in consumer WPAM devices and the importance of balancing the motivational benefits of self-monitoring technology against potential user frustration from technical limitations for adults aging with HIV [40,56]. Participants described difficulty finding exercises to match their activity levels due to navigation difficulties on the Sweat for Good YMCA app. While Zoom was generally reliable, several participants reported audio lag and video freezing, issues commonly experienced in online digital health interventions with synchronous videoconferencing among adults aging with HIV [26,39,55]. Difficulties with device accuracy, navigation of fitness apps, and technical troubleshooting are widely reported in studies involving adults living with HIV using WPAMs or engaging in online exercise interventions [17,44,55]. While these technical experiences are not unique to this sample population, they occur within a context shaped by the challenges of aging with HIV, including health unpredictability, stigma, and the risk of loneliness [57-60]. Embedding lived experiences into intervention design and ongoing technical support are essential strategies to overcome digital barriers and promote uptake and usage with online CBE interventions among adults aging with HIV.

Changes in Technology Experiences Over Time

Changes in experiences with technology over time demonstrated initial attitudes toward technology that did not always predict sustained use with the technology. Some participants experienced increasing technology burden and fatigue, others maintained neutral perspectives throughout, while others developed optimistic intentions for continued use postintervention. Participants used Zoom consistently throughout, as evidenced by consistent usability responses on the Telehealth Indicator Questionnaire, likely due to its relative simplicity, prior exposure, and use as a communication platform with their personal trainer. While the use of WPAMs and the Sweat for Good YMCA app may have declined, this may be attributed to persistent technical issues and limited adaptability. This mirrored patterns of engagement with digital health technology seen in adults aging with HIV [32,55,61]. Some participants reported feeling overwhelmed by syncing

XSL•FO
RenderX

requirements, app navigation, and inconsistent device feedback in the online CBE intervention, factors that diminished motivation and created friction in maintaining regular technology use. This aligns with research that prolonged exposure to poorly optimized tools can lead to emotional and cognitive fatigue, especially among users managing multimorbidity [19,41]. Positive changes in technology use were most apparent in self-directed components, suggesting that asynchronous, exploratory technology use may become more rewarding over time as users gain confidence. For adults aging with HIV, these findings reinforce the need for sustained technical support and simplified interface design, as digital interventions must be responsive, flexible, and grounded in lived experience to facilitate long-term engagement in online CBE interventions [17,18,55].

Strengths, Limitations, and Implications for Future Research

A strength of this study was the longitudinal observational design involving both interview and questionnaire data, allowing for an in-depth understanding of participants' experiences with technology over time. Interview and questionnaire data were analyzed separately, with the questionnaires used descriptively to provide context to the interview findings. We conducted a collaborative content analysis among the team, allowing for diverse perspectives to be integrated throughout coding and interpretation. This iterative, team-based process supported reflexivity and rigor in the analysis. Additionally, input from community experts living with HIV helped ground the analysis in the lived experiences of the HIV community.

Our study sample size was fixed based on the existing data in the telecoaching CBE study. Reaching thematic saturation was not a goal of our study. However, in the results, we indicate the participants via participant number and time point alongside interview quotes to indicate the contribution of participant perspectives across the breadth of categories. Other strategies to enhance rigor included our inclusion of all available interview and questionnaire data for participants across the study. Collectively, we believe we were able to capture an ample description of participants' perspectives related to experiences with uptake and usage, enabling us to achieve our study objectives.

Regarding study limitations, the interviews were not solely focused on technology; hence, the amount of data and depth of insights specific to technology use was limited. Not all participants completed interviews at T1 and T2 or the Telehealth Indicator Questionnaire at T2, which limited the ability to explore participants' changes in technology experiences over time. Our sample was drawn from community-dwelling adults living with HIV in Toronto, a large urban center with strong digital infrastructure, limiting transferability to rural or remote settings with greater barriers to connectivity and access to the internet. Additionally, future work should examine sex and gender differences in technology experiences and engagement within online CBE interventions among adults aging with HIV to promote equitable access, engagement, and effectiveness [62].

Conclusions

Experiences with technology uptake and usage among adults aging with HIV were characterized by 4 components, including preparations for technology use, interactions with the technology, facilitators and satisfaction with technology, and challenges and frustrations with technology. Experiences with technology interacted with intrinsic contextual factors (prior exposure linked to anticipated comfort or challenges with technology) and extrinsic contextual factors (COVID-19 pandemic and technological and social support). Among adults aging with HIV, changes in technology experiences over time highlighted the unpredictability of initial impressions, with some participants expressing technology fatigue and burden, whereas others maintained consistent attitudes of usage over time with intentions to continue using digital tools and processes beyond the intervention. Findings with experiences of technology uptake and usage in this study underscore the importance of incorporating personalized and tailored onboarding and sustained technological support with an online CBE intervention to promote engagement in physical activity among adults aging with HIV.

Acknowledgments

This research was completed in partial fulfillment of the requirements for an MScPT degree in the Department of Physical Therapy, Temerty Faculty of Medicine at the University of Toronto. We thank Maple Leaf Medical Clinic for their support with recruitment via the Ontario HIV Treatment Network Cohort Study for this study. We acknowledge members of the original and ongoing telecoaching community-based exercise study team who were instrumental to the acquisition of funding and implementation of the study: Francisco Ibáñez-Carrasco, Ahmed M. Bayoumi, Ada Tang, Kiera McDuff, Tizneem Jiancaro, George Da Silva, Brittany Torres, Mona R. Loutfy, Shaz Islam, Joanne Lindsay, Colleen Price, Mehdi Zobeiry, Zoran Pandovski, Ivan Illic, Puja Ahluwalia, Darren A. Brown, Lisa Avery, and Patricia Solomon. We specifically thank Kiera McDuff and Julia Nathanson for their support with the telecoaching community-based exercise study. We acknowledge the individuals who volunteered their time to participate in the research as well as the staff, volunteers, and community-based settings who supported the implementation of the study. Specifically, we thank community experts, John, George, and Richard, who generously shared their time, perspectives, and lived experiences throughout the course of this study.

XSL•FO
RenderX

Funding

This research was supported by the Ontario HIV Treatment Network HIV Endgame Funding Program – Breaking New Ground (grant EFP-1121-BNG). This study is a substudy of the Ontario HIV Treatment Network Cohort Study. KKO is supported by a Canada Research Chair in Episodic Disability and Rehabilitation from the Canada Research Chairs Program (CRC-2022-00510).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide (technology questions highlighted).
[PDF File, 260 KB - rehab_v13i1e86785_app1.pdf]

Multimedia Appendix 2
Telehealth indicator questionnaire.
[PDF File, 229 KB - rehab_v13i1e86785_app2.pdf]

Multimedia Appendix 3
Direction of change of participant responses to telehealth indicator questionnaire item responses at months 2 and 6.
[DOCX File, 26 KB - rehab_v13i1e86785_app3.docx]

Checklist 1
Standards for Reporting Qualitative Research checklist.
[PDF File, 554 KB - rehab_v13i1e86785_app4.pdf]

References

1. Canada's progress towards ending the HIV epidemic, 2022. Public Health Agency of Canada. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/canada-progress-towards-ending-hiv-epidemic-2022.html [accessed 2026-05-15]
2. AIDS, crisis and the power to transform: UNAIDS global AIDS update 2025. : Joint United Nations Programme on HIV/AIDS (UNAIDS); 2025 URL: https://www.unaids.org/sites/default/files/2025-07/2025-global-aids-update-JC3153_en.pdf [accessed 2026-05-15]
3. Hasse B, Ledergerber B, Furrer H, et al. Morbidity and aging in HIV-infected persons: the Swiss HIV cohort study. Clin Infect Dis 2011 Dec;53(11):1130-1139. [doi: 10.1093/cid/cir626] [Medline: 21998280]
4. McArthur JC, Johnson TP. Chronic inflammation mediates brain injury in HIV infection: relevance for cure strategies. Curr Opin Neurol 2020 Jun;33(3):397-404. [doi: 10.1097/WCO.0000000000000807] [Medline: 32209807]
5. O'Brien KK, Bayoumi AM, Strike C, Young NL, Davis AM. Exploring disability from the perspective of adults living with HIV/AIDS: development of a conceptual framework. Health Qual Life Outcomes 2008 Oct 4;6:76. [doi: 10.1186/1477-7525-6-76] [Medline: 18834538]
6. O'Brien KK, Hanna S, Gardner S, et al. Validation of the Episodic Disability Framework with adults living with HIV. Disabil Rehabil 2014;36(4):319-329. [doi: 10.3109/09638288.2013.793408] [Medline: 23688295]
7. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131. [Medline: 3920711]
8. Leach LL, Bassett SH, Smithdorf G, Andrews BS, Travill AL. A systematic review of the effects of exercise interventions on body composition in HIV+ adults. Open AIDS J 2015;9:66-79. [doi: 10.2174/1874613601509010066] [Medline: 26587075]
9. Aminde JAA, Burton NW, Thng C, Clanchy K. A systematic review and meta-analysis evaluating the effectiveness of minimally supervised home and community exercise interventions in improving physical activity, body adiposity and quality of life in adults living with HIV. Prev Med 2024 Dec;189:108144. [doi: 10.1016/j.ypmed.2024.108144]
10. O'Brien KK, Tynan AM, Nixon SA, Glazier RH. Effectiveness of Progressive Resistive Exercise (PRE) in the context of HIV: systematic review and meta-analysis using the Cochrane Collaboration protocol. BMC Infect Dis 2017 Apr 12;17(1):268. [doi: 10.1186/s12879-017-2342-8] [Medline: 28403830]
11. Gomes-Neto M, Saquetto MB, Alves IG, Martinez BP, Vieira JPB, Brites C. Effects of exercise interventions on aerobic capacity and health-related quality of life in people living with HIV/AIDS: systematic review and network meta-analysis. Phys Ther 2021 Jul 1;101(7):pzab092. [doi: 10.1093/ptj/pzab092] [Medline: 33704496]
12. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]

XSL•FO
RenderX

13. Vancampfort D, Mugisha J, De Hert M, et al. Global physical activity levels among people living with HIV: a systematic review and meta-analysis. Disabil Rehabil 2018 Feb;40(4):388-397. [doi: 10.1080/09638288.2016.1260645] [Medline: 27929355]
14. Vancampfort D, Mugisha J, De Hert M, Probst M, Stubbs B. Sedentary behavior in people living with HIV: a systematic review and meta-analysis. J Phys Act Health 2017 Jul;14(7):571-577. [doi: 10.1123/jpah.2016-0507] [Medline: 28290767]
15. Simonik A, Vader K, Ellis D, et al. Are you ready? Exploring readiness to engage in exercise among people living with HIV and multimorbidity in Toronto, Canada: a qualitative study. BMJ Open 2016 Mar 8;6(3):e010029. [doi: 10.1136/bmjopen-2015-010029] [Medline: 26956163]
16. Johs NA, Kellar-Guenther Y, Jankowski CM, Neff H, Erlandson KM. A qualitative focus group study of perceived barriers and benefits to exercise by self-described exercise status among older adults living with HIV. BMJ Open 2019 Mar 7;9(3):e026294. [doi: 10.1136/bmjopen-2018-026294] [Medline: 30850416]
17. Solomon P, Carusone SC, Davis AM, Aubry R, O'Brien KK. Experiences of people living with HIV in community based exercise: a qualitative longitudinal study. J Int Assoc Provid AIDS Care 2021;20:2325958221995344. [doi: 10.1177/2325958221995344] [Medline: 33611978]
18. Lau B, Sharma I, Manku S, et al. Considerations for developing and implementing an online community-based exercise intervention with adults living with HIV: a qualitative study. BMJ Open 2022 Apr 15;12(4):e059294. [doi: 10.1136/bmjopen-2021-059294] [Medline: 35428647]
19. Montgomery CA, Henning KJ, Kantarzhi SR, Kideckel TB, Yang CFM, O'Brien KK. Experiences participating in a community-based exercise programme from the perspective of people living with HIV: a qualitative study. BMJ Open 2017 Apr 4;7(4):e015861. [doi: 10.1136/bmjopen-2017-015861] [Medline: 28377397]
20. Li A, McCabe T, Silverstein E, et al. Community-based exercise in the context of HIV: factors to consider when developing and implementing community-based exercise programs for people living with HIV. J Int Assoc Provid AIDS Care 2017;16(3):267-275. [doi: 10.1177/2325957416686836] [Medline: 28074681]
21. O'Brien KK, Davis AM, Chan Carusone S, et al. Examining the impact of a community-based exercise intervention on cardiorespiratory fitness, cardiovascular health, strength, flexibility and physical activity among adults living with HIV: a three-phased intervention study. PLoS ONE 2021;16(9):e0257639. [doi: 10.1371/journal.pone.0257639] [Medline: 34559851]
22. Galea MD. Telemedicine in rehabilitation. Phys Med Rehabil Clin N Am 2019 May;30(2):473-483. [doi: 10.1016/j.pmr.2018.12.002] [Medline: 30954160]
23. Russell TG. Physical rehabilitation using telemedicine. J Telemed Telecare 2007;13(5):217-220. [doi: 10.1258/135763307781458886] [Medline: 17697506]
24. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS. Telerehabilitation: review of the state-of-the-art and areas of application. JMIR Rehabil Assist Technol 2017 Jul 21;4(2):e7. [doi: 10.2196/rehab.7511] [Medline: 28733271]
25. Jones J, Knox J, Meanley S, et al. Explorations of the role of digital technology in HIV-related implementation research: case comparisons of five ending the HIV epidemic supplement awards. J Acquir Immune Defic Syndr 2022 Jul 1;90(S1):S226-S234. [doi: 10.1097/QAI.0000000000002983] [Medline: 35703775]
26. Piraux E, Reychler G, Forget P, Yombi JC, Caty G. Feasibility and preliminary effects of a telerehabilitation program for people living with HIV: a pilot randomized study. J Assoc Nurses AIDS Care 2019;30(2):176-185. [doi: 10.1097/JNC.0000000000000005] [Medline: 30822290]
27. Loeckx M, Rabinovich RA, Demeyer H, et al. Smartphone-based physical activity telecoaching in chronic obstructive pulmonary disease: mixed-methods study on patient experiences and lessons for implementation. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e200. [doi: 10.2196/mhealth.9774] [Medline: 30578215]
28. Ramage ER, Fini NA, Lynch EA, Patterson A, Said CM, English C. Supervised exercise delivered via telehealth in real time to manage chronic conditions in adults: a protocol for a scoping review to inform future research in stroke survivors. BMJ Open 2019 Mar 20;9(3):e027416. [doi: 10.1136/bmjopen-2018-027416] [Medline: 30898833]
29. Lai B, Bond K, Kim Y, Barstow B, Jovanov E, Bickel CS. Exploring the uptake and implementation of tele-monitored home-exercise programmes in adults with Parkinson's disease: a mixed-methods pilot study. J Telemed Telecare 2020;26(1-2):53-63. [doi: 10.1177/1357633X18794315] [Medline: 30134777]
30. Albergoni A, Hettinga FJ, La Torre A, Bonato M, Sartor F. The role of technology in adherence to physical activity programs in patients with chronic diseases experiencing fatigue: a systematic review. Sports Med Open 2019 Sep 12;5(1):41. [doi: 10.1186/s40798-019-0214-z] [Medline: 31512075]
31. Kapadia V, Ariani A, Li J, Ray PK. Emerging ICT implementation issues in aged care. Int J Med Inform 2015 Nov;84(11):892-900. [doi: 10.1016/j.ijmedinf.2015.07.002] [Medline: 26250987]
32. Turner JR, Chow J, Cheng J, et al. Wireless physical activity monitor use among adults living with HIV in a community-based exercise intervention study: a quantitative, longitudinal, observational study. BMJ Open 2023 Apr;13(4):e068754. [doi: 10.1136/bmjopen-2022-068754]
33. Wright PJ, Raynor PA, Bowers D, et al. Leveraging digital technology for social connectedness among adults with chronic conditions: a systematic review. Digit Health 2023;9:20552076231204746. [doi: 10.1177/20552076231204746] [Medline: 37799504]

XSL•FO
RenderX

34. Berry ECJ, Sculthorpe NF, Warner A, Mather JD, Sanal-Hayes NEM, Hayes LD. A scoping review of the feasibility, usability, and efficacy of digital interventions in older adults concerning physical activity and/or exercise. Front Aging 2025;6:1516481. [doi: 10.3389/fragi.2025.1516481] [Medline: 40290578]
35. Ramage ER, Fini N, Lynch EA, et al. Look before you leap: interventions supervised via telehealth involving activities in weight-bearing or standing positions for people after stroke-a scoping review. Phys Ther 2021 Jun 1;101(6):pzab073. [doi: 10.1093/ptj/pzab073] [Medline: 33611602]
36. Nittas V, Zecca C, Kamm CP, Kuhle J, Chan A, von Wyl V. Digital health for chronic disease management: an exploratory method to investigating technology adoption potential. PLoS ONE 2023;18(4):e0284477. [doi: 10.1371/journal.pone.0284477] [Medline: 37053272]
37. Nittas V, Daniore P, Chavez SJ, Wray TB. Challenges in implementing cultural adaptations of digital health interventions. Commun Med (Lond) 2024 Jan 5;4(1):7. [doi: 10.1038/s43856-023-00426-2] [Medline: 38182750]
38. Apolinário-Hagen J, Menzel M, Hennemann S, Salewski C. Acceptance of mobile health apps for disease management among people with multiple sclerosis: web-based survey study. JMIR Form Res 2018 Dec 12;2(2):e11977. [doi: 10.2196/11977] [Medline: 30684408]
39. Campbell BR, Ingersoll KS, Flickinger TE, Dillingham R. Bridging the digital health divide: toward equitable global access to mobile health interventions for people living with HIV. Expert Rev Anti Infect Ther 2019 Mar;17(3):141-144. [doi: 10.1080/14787210.2019.1578649] [Medline: 30721103]
40. Dagenais M, Cheng D, Salbach NM, Brooks D, O'Brien KK. Wireless physical activity monitor use among adults living with HIV: a scoping review. Rehabil Oncol 2019;37(1):17-28. [doi: 10.1097/01.REO.0000000000000153]
41. Oursler KK, Marconi VC, Briggs BC, Sorkin JD, Ryan AS, FIT VET Project Team. Telehealth exercise intervention in older adults with HIV: protocol of a multisite randomized trial. J Assoc Nurses AIDS Care 2022;33(2):168-177. [doi: 10.1097/JNC.0000000000000235] [Medline: 33481463]
42. O'Brien KK, Ibáñez-Carrasco F, Carusone SC, et al. Piloting an online telecoaching community-based exercise intervention with adults living with HIV: protocol for a mixed-methods implementation science study. BMJ Open 2023 Mar 30;13(3):e067703. [doi: 10.1136/bmjopen-2022-067703] [Medline: 36997255]
43. Rourke SB, Gardner S, Burchell AN, et al. Cohort profile: the Ontario HIV Treatment Network Cohort Study (OCS). Int J Epidemiol 2013 Apr;42(2):402-411. [doi: 10.1093/ije/dyr230] [Medline: 22345312]
44. Ibáñez-Carrasco F, McDuff K, Da Silva G, et al. Qualitative insights from an online community-based exercise intervention for persons living with HIV. Front Rehabil Sci 2025;6. [doi: 10.3389/fresc.2025.1602007]
45. Huggins AC, Ritzhaupt AD, Dawson KM. Measuring information and communication technology literacy using a performance assessment: validation of the student tool for technology literacy (ST2L). Comput Educ 2014 Aug;77:1-12. [doi: 10.1016/j.compedu.2014.04.005]
46. Manual for measuring ICT access and use by housholds and individuals. International Telecommunication Union. 2014. URL: https://www.itu.int/en/ITU-D/Statistics/Pages/publications/manual.aspx [accessed 2026-05-15]
47. Molnar A, Weerakkody V. Defining key performance indicators for evaluating the use of high definition video-to-video-services in ehealth. In: IFIP International Federation for Information Processing 2013, Vol. 412:452-461. [doi: 10.1007/978-3-642-41142-7_46]
48. Flicker S, Nixon SA. The DEPICT model for participatory qualitative health promotion research analysis piloted in Canada, Zambia and South Africa. Health Promot Int 2015 Sep;30(3):616-624. [doi: 10.1093/heapro/dat093]
49. NVivo qualitative data analysis software. Version 12. QSR International. 2018. URL: https://lumivero.com/products/nvivo/ [accessed 2026-06-21]
50. Microsoft Excel. Microsoft Corporation. 2016. URL: https://www.microsoft.com/en-ca/microsoft-365/excel [accessed 2026-06-21]
51. Flynn A, Preston E, Dennis S, Canning CG, Allen NE. Utilising telehealth to support exercise and physical activity in people with Parkinson disease: a program evaluation using mixed methods. BMC Health Serv Res 2023 Mar 7;23(1):224. [doi: 10.1186/s12913-023-09194-0] [Medline: 36882780]
52. Dunston ER, Malouf A, Podlog LW, et al. Experiences participating in a telehealth exercise program among older adults with cancer: a qualitative study. J Cancer Surviv 2025 Oct;19(5):1462-1471. [doi: 10.1007/s11764-024-01564-9] [Medline: 38492170]
53. Budak JZ, Scott JD, Dhanireddy S, Wood BR. The impact of COVID-19 on HIV care provided via telemedicine-past, present, and future. Curr HIV/AIDS Rep 2021 Apr;18(2):98-104. [doi: 10.1007/s11904-021-00543-4] [Medline: 33616811]
54. Solomon P, Chan Carusone S, Davis AM, Aubry R, O'Brien KK. A qualitative study of fitness coaches' experiences in community based exercise with people living with HIV. J Int Assoc Provid AIDS Care 2021;20:23259582211046762. [doi: 10.1177/23259582211046762] [Medline: 34668422]
55. Jiancaro T, Bayoumi AM, Ibáñez-Carrasco F, et al. Factors influencing initial implementation of an online community-based exercise intervention with adults living with HIV: a systems approach. Front Rehabil Sci 2023;4:1176960. [doi: 10.3389/fresc.2023.1176960]
56. Dagenais M, Salbach NM, Brooks D, O'Brien KK. Assessing the measurement properties of the Fitbit Zip® among adults living with HIV. J Phys Act Health 2020 Mar 1;17(3):293-305. [doi: 10.1123/jpah.2019-0242] [Medline: 31958771]

XSL•FO
RenderX

57. Solomon P, O'Brien K, Wilkins S, Gervais N. Aging with HIV and disability: the role of uncertainty. AIDS Care 2014 Feb;26(2):240-245. [doi: 10.1080/09540121.2013.811209] [Medline: 23799874]
58. Yoo-Jeong M, Hepburn K, Holstad M, Haardörfer R, Waldrop-Valverde D. Correlates of loneliness in older persons living with HIV. AIDS Care 2020 Jul;32(7):869-876. [doi: 10.1080/09540121.2019.1659919] [Medline: 31462066]
59. Grov C, Golub SA, Parsons JT, Brennan M, Karpiak SE. Loneliness and HIV-related stigma explain depression among older HIV-positive adults. AIDS Care 2010 May;22(5):630-639. [doi: 10.1080/09540120903280901] [Medline: 20401765]
60. Pollak C, Cotton K, Winter J, Blumen H. Health outcomes associated with loneliness and social isolation in older adults living with HIV: a systematic review. AIDS Behav 2025 Jan;29(1):166-186. [doi: 10.1007/s10461-024-04471-3] [Medline: 39231919]
61. Ibáñez-Carrasco F, Jiancaro T, Torres B, et al. HIV in MOTION: a community of practice on physical rehabilitation for and by people living with HIV and their allies. Front Rehabil Sci 2023;4:1154692. [doi: 10.3389/fresc.2023.1154692] [Medline: 37869573]
62. Sahel-Gozin N, Loutfy M, O'Brien KK. Exploring experiences engaging in exercise from the perspectives of women living with HIV: a qualitative study. PLoS One 2023;18(6):e0286542. [doi: 10.1371/journal.pone.0286542] [Medline: 37267270]

Abbreviations

CBE: community-based exercise
WHO: World Health Organization
WPAM: wireless physical activity monitor

Edited by S Munce; submitted 30.Oct.2025; peer-reviewed by A Samuel, V Santos; revised version received 15.May.2026; accepted 21.May.2026; published 02.Jul.2026.

Please cite as:
Mucha J, Hamdy R, Marini M, Aasem R, Duong C, Su TT, Chan Carusone S, O'Brien KK
Experiences With Technology Among Adults Aging With HIV Engaged in an Online Community–Based Exercise Intervention Study: Longitudinal Qualitative Descriptive Study and Secondary Data Analysis
JMIR Rehabil Assist Technol 2026;13:e86785
URL: https://rehab.jmir.org/2026/1/e86785
doi:10.2196/86785

© Julia Mucha, Rana Hamdy, Melina Marini, Reda Aasem, Chung Duong, Tai-Te Su, Soo Chan Carusone, Kelly K O'Brien. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 2.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Fidelity of an Online Telecoaching Community-Based Exercise Intervention Among Adults Living With HIV in Toronto, Ontario, Canada: Observational Longitudinal Study

Marfy Abousifein[1*]; Julia Nathanson[2*]; Tizneem Jiancaro[2]; Ada Tang[3]; Puja Ahluwalia[4]; Ahmed M Bayoumi[5,6,7]; Lisa Avery[8,9]; George Da Silva[2,4]; Francisco Ibáñez-Carrasco[9]; Ivan Ilic[10]; Kiera McDuff[2]; Mona Loutfy[5,7,11]; Annamaria Furlan[10]; Zoran Pandovski[10]; Helen Trent[10]; Soo Chan Carusone[1,2*]; Kelly K O'Brien[2,5,12*]

[1]McMaster Collaborative for Health and Aging, McMaster University, Hamilton, ON, Canada

[2]Department of Physical Therapy, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[3]School of Rehabilitation Science, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[4]Realize, Toronto, ON, Canada

[5]Institute of Health Policy, Management and Evaluation (IHPME), Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[6]MAP Centre for Urban Health Solutions, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

[7]Department of Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[8]Biostatistics Department, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

[9]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[10]YMCA of Greater Toronto, Toronto, ON, Canada

[11]Women's College Hospital Research Institute, Women's College Hospital, Toronto, ON, Canada

[12]Rehabilitation Sciences Institute (RSI), Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Kelly K O'Brien
Department of Physical Therapy
Temerty Faculty of Medicine
University of Toronto
160-500 University Avenue
Toronto, ON, M5G 1V7
Canada
Phone: 1 416 978 0565
Email: kelly.obrien@utoronto.ca

Abstract

Background: Community-based exercise (CBE) can prevent and mitigate disability and improve health outcomes in people aging with HIV; however, engagement in exercise can vary.

Objective: This study aims to describe the fidelity of implementation (FOI) of an online telecoaching CBE intervention with adults living with HIV.

Methods: We conducted an observational longitudinal study involving a 6-month online CBE intervention involving thrice-weekly 60-minute exercise sessions; online biweekly supervised one-on-one exercise sessions with a personal trainer (13 sessions); and online monthly group educational sessions. We assessed fidelity using (1) online structured interviews with participants living with HIV at months 2 and 6, comprising 20 fidelity items on trainer performance, participant engagement, and exercise experiences; and (2) coaching logs completed by personal trainers after biweekly training sessions (frequency, intensity, time, and type of physical activity during supervised sessions and self-reported physical activity in the prior week). We considered FOI criteria to be met if ≥80% of the participants living with HIV: (1) attended ≥80% of training sessions (11 of 13 sessions); (2) engaged in thrice-weekly exercise the week of the personal training session ≥80% of the time over the 6-month intervention, as reported on the coaching logs; (3) reported "complete criteria met" for ≥80% of the fidelity items (≥16 of 20 items) at month 2 and month 6, as reported from the interviews; and (4) engaged in a combination of aerobic, strength, balance, and flexibility exercise ≥80% of the time over the 6-month intervention, as reported in the coaching logs.

XSL•FO
RenderX

Results: A total of 29 of 32 participants completed at least 1 fidelity interview; 69% (22/32) were male. FOI was met for 1 of our 4 criteria among the 18 participants who completed the intervention and FOI assessment at month 6, specifically criterion 1, whereby 83% (15/18; 95% CI 66%-100%) of participants attended ≥80% of the 13 coaching sessions. For criterion 2, 50% (9/18; 95% CI 27%-73%) of participants engaged in thrice-weekly exercise in the week of the personal training session ≥80% of the time over the 6-month intervention; criterion 3: ≥80% of participants reported criteria as "completely met" for 70% (14/20 items; 95% CI 50%-89%) of items at month 2 (n=27 participants) and 30% (6/20 items; 95% CI 10%-50%) of items at month 6 (n=18 participants); and criterion 4: 22% (4/18; 95% CI 3%-41%) of participants engaged in a combination of all 4 exercise types ≥80% of the time. Barriers to engagement included scheduling issues, lack of interest, technology problems, and episodic disability.

Conclusions: FOI was achieved for supervised components of the CBE intervention and less so with independent exercise. Health and rehabilitation providers should tailor exercise interventions to personal preferences and abilities to improve engagement among adults living with HIV.

(JMIR Rehabil Assist Technol 2026;13:e84902) doi:10.2196/84902

KEYWORDS

exercise; fidelity; HIV/AIDS; implementation science; telerehabilitation

Introduction

Overview

In Canada, over 65,000 people were estimated to be living with HIV at the end of 2022 [1]. More individuals are living longer and aging with HIV; globally, persons living with HIV aged 50 years or older increased from 5.4 million in 2015 to 8.1 million in 2020 [2]. As people living with HIV age, they can experience health-related consequences of HIV, including an increased risk of cardiovascular disease, diabetes, bone and joint disorders, neurocognitive disorders, and certain cancers [3-5]. Rehabilitation, such as physical therapy and occupational therapy, can mitigate disability associated with HIV and multimorbidity, such as fatigue, pain, cognitive impairments, body composition changes, and employment challenges [6-9]. Exercise is a rehabilitation intervention and self-management strategy that can address disability and improve or sustain the health of individuals with HIV. Systematic review evidence indicates that exercise is safe and can lead to benefits in cardiopulmonary fitness, strength, weight and body composition, and psychological outcomes for people with HIV [10,11]. However, this evidence primarily involved supervised interventions by physical therapists and exercise physiologists that can be costly and unsustainable [10,11]. With up to an estimated 73% of people living with HIV classified as sedentary or physically inactive, it is essential to consider exercise interventions that are accessible and practical for people with HIV to sustain over the long term [12].

Community-Based Exercise

Community-based exercise (CBE) is a promising approach that may help people living with HIV increase their engagement in exercise and physical activity [13-16]. Exercise is purposeful, planned, and repetitive bodily movement that has the objective of improving or maintaining physical fitness, while physical activity is any bodily movement that results in energy expenditure [17]. Persons living with HIV perceive benefits from exercise and physical activity as a component of health promotion, with the potential to improve physical function and mental health and reduce isolation [18]. CBE can be provided under the guidance of a health or fitness instructor, with the aim of encouraging consistent exercise and physical activity within the community [13,14,16]. CBE enhances the benefits of traditional exercise as it can improve social interaction, provide support and motivation, and promote emotional, cognitive, and behavioral self-management strategies to help individuals manage health challenges [14,19,20]. CBE has been associated with engagement in physical activity and improvements in health outcomes for adults living with HIV [19,21]; however, the factors influencing the implementation and uptake of CBE are less clear.

Implementation of CBE

Using the Consolidated Framework for Implementation Research (CFIR), Cooper et al [22] conducted a qualitative systematic review of CBE interventions to explore barriers and facilitators of implementing CBE interventions in the general population [22]. Barriers included lack of participant motivation, insufficient resources allocated for implementation, the complexity of the CBE intervention, and lack of participant knowledge about the intervention [22]. Facilitators included the flexibility to adapt the intervention to the context, intervention compatibility with participants, and participants having a positive perception of those implementing the intervention [22]. The authors concluded that making minor changes to tailor an intervention while maintaining core components, such as clear implementation strategies and coordination of delivery, can facilitate the implementation of a CBE intervention.

Members of this authorship team assessed the impact of an in-person CBE intervention with adults living with HIV at the YMCA in Toronto, Ontario, Canada [21]. While the team reported improvements in cardiovascular health, strength, flexibility, and self-reported physical activity, barriers to taking part in CBE were documented among women, persons who are gender diverse, and persons who did not live geographically close to the YMCA [21]. Persons with HIV may experience obstacles to exercising in traditional gym environments, including interpersonal, financial, and geographical barriers, self-image issues and stigma, and difficulty initiating exercise following periods of inactivity [23,24].

XSL•FO
RenderX

Telerehabilitation in the Context of HIV

Telerehabilitation, and more specifically telecoaching, is a potential strategy for implementing CBE with persons living with HIV. Telerehabilitation is the delivery of rehabilitation programs or services via technology, such as phone and video calls, web-based platforms, or mobile apps, used initially to increase efficiency, quality of care, and access to health care for populations with geographical [25,26], financial, or time-limiting barriers for persons living with HIV [27]. Telecoaching is a component of telerehabilitation that uses technology for remote supervision, guidance, and communication of an exercise program to improve physical activity levels and manage symptoms for people living with chronic diseases [28-33].

Telecoaching CBE Study

We established an online telecoaching CBE intervention with the aim of enhancing engagement in physical activity among adults living with HIV in Toronto, Canada [34]. The intervention was developed based on foundational considerations for developing and implementing online CBE with adults living with HIV that included: (1) person-specific considerations (episodic nature of HIV, stigma, and HIV disclosure), (2) accessibility of the program, (3) program delivery and technology, (4) attributes of program personnel, (5) program content and design, and (6) building community and peer support [35].

Fidelity of implementation (FOI) is important to consider in online CBE interventions to develop practical interventions that lead to sustained physical activity. FOI, or integrity, is the degree to which programs are implemented as intended by the program developers [36-38]. The National Institutes of Health Behavior Change Consortium (BCC) treatment fidelity framework specifically refers to treatment fidelity and the link between training the providers, delivery, receipt, and enactment of the intervention for research involving health behavior interventions [39]. The fidelity with which an intervention is implemented can influence its level of success and uptake, and the impact on the outcome, with high fidelity promoting improved outcomes. However, it is important to consider planned adaptations to the intervention to optimize treatment fidelity while acknowledging the need to accommodate practical needs, study demands, and provider and participant burden, requiring pragmatic approaches in real-world settings [39]. Lack of fidelity can lead to failure in intervention implementation in a real-life context [40]. Hence, reporting fidelity is essential to determine the credibility, validity, and replicability of intervention study findings [41]. Despite the importance of understanding intervention fidelity, lack of knowledge or understanding about assessing fidelity may be barriers to assessment and reporting [42]. While quantitative and qualitative methods are recommended to comprehensively assess fidelity, this guidance is not consistently followed [39,42-44]. To our knowledge, there is limited evidence assessing the FOI of exercise interventions for persons living with HIV. Assessing the FOI of exercise interventions in the context of HIV is important because (1) sustained exercise can mitigate HIV-related disability, and (2) people living with HIV can face unique barriers and facilitators to exercise, such as episodic disability, which can influence the FOI of exercise interventions [45].

Our aim was to describe the FOI of a 6-month online telecoaching CBE intervention with adults living with HIV from the perspectives of participants (adults living with HIV) and personal trainers.

Findings from this study may help inform policy and programming for broader scale-up and optimal CBE implementation that can be adopted by fitness centers and community-based organizations to proactively prevent multimorbidity and improve health outcomes among persons living with HIV.

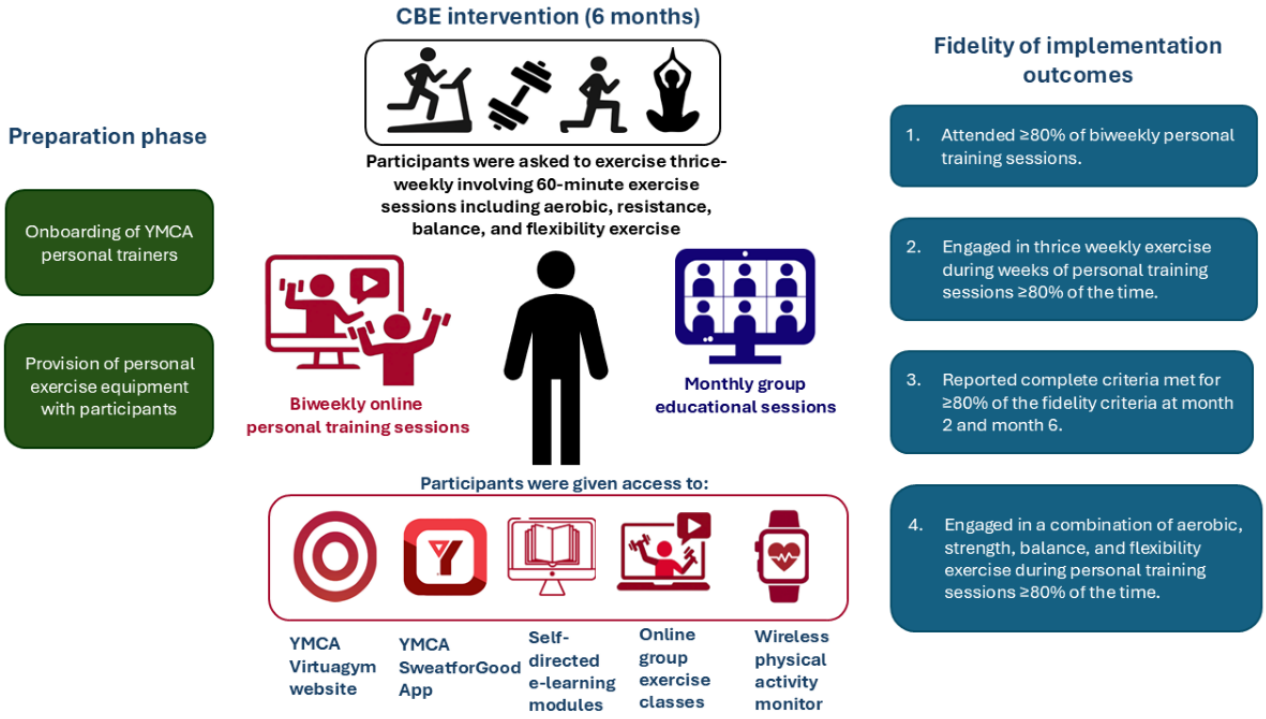
Methods

Study Design

We conducted an observational longitudinal study involving (1) online structured interviews with adults living with HIV who participated in a 6-month online CBE intervention and (2) coaching logs of training sessions completed by personal trainers. Details of the study protocol have been previously published [34]. Figure 1 provides an overview of the CBE intervention and fidelity outcomes.

XSL•FO
RenderX

Figure 1. Overview of online community-based exercise (CBE) intervention and fidelity of implementation outcomes. YMCA: Young Men's Christian Association.



Participants

We included adults living with HIV who participated in the telecoaching study and personal trainers from the YMCA who were involved in supervising personalized online exercise sessions with the study participants [34]. Eligibility criteria for the telecoaching study were adults (aged 18 years or older) who considered themselves safe to engage in exercise as determined by the self-administered Physical Activity Readiness Questionnaire (PAR-Q) and had access to an electronic device with an internet connection and a webcam, as well as space at home for exercise [34]. We recruited potential participants living with HIV through community-based organizations, an HIV clinic in Toronto via the Ontario HIV Treatment Network Cohort Study (OCS), and participants from a prior CBE study who agreed to be contacted about future research [21,46].

Telecoaching Online CBE Intervention

Prior to the CBE implementation, fitness instructors (certified personal trainers and staff at the YMCA) took part in workshops on HIV, rehabilitation and exercise for adults living with HIV, goal setting, and research procedures [47]. We piloted the setup of the technology, including (but not limited to) Zoom software [48], a webcam, YMCA membership, web-based questionnaire software, and the Fitbit app [49] with the participants living with HIV and fitness instructors.

The telecoaching online CBE intervention included a 6-month structured intervention phase followed by an independent 6-month exercise phase [34]. During the 6-month intervention phase, participants were asked to engage in the following three components of the intervention: (1) an exercise regimen of aerobic, resistance, balance, and flexibility exercise occurring thrice weekly at moderate to vigorous intensity for ≥60 minutes per session; (2) attend online biweekly individually tailored one-on-one exercise sessions with a certified personal trainer from the YMCA, involving supervision and progression of aerobic, resistance, balance, and flexibility exercise; and (3) engage in online monthly group educational sessions on sleep, nutrition, goal setting, cognitive health, chronic pain, and other topics related to self-management, health, and physical activity while living with HIV, delivered by members of the research team. Each participant met with their personal trainer prior to the intervention to tailor the intervention to their goals, abilities, and interests [50]. Participants received a YMCA membership allowing them to access online YMCA classes through the YMCA SweatforGood app. Personal trainers uploaded the individualized exercise plans to the SweatforGood app, allowing participants to access, review, and track their exercises [34].

After the 6-month intervention, participants entered a 6-month follow-up independent exercise phase, during which they were asked to continue with thrice-weekly exercise. They continued to have access to the YMCA online membership, providing access to group online classes, and the SweatforGood app, but the study no longer provided one-on-one personal training sessions or monthly educational sessions.

Participants were given home exercise equipment (Therabands and a wooden step) and a wireless physical activity monitor (Fitbit Inspire 2) to monitor and encourage exercise during both the 6-month CBE intervention and the 6-month follow-up period. Participants documented their physical activity through a weekly online exercise log. Participants were given the Fitbit Inspire 2 and exercise equipment to keep at the end of the study as a token of appreciation for their participation. For more details, see the published protocol [34].

We aimed to recruit 40 adults living with HIV and 30 to complete the intervention, enabling us to achieve our primary

XSL•FO
RenderX

study aim of assessing the impact of the CBE intervention [23,51].

Data Collection

Overview

We used a combination of structured interviews with participants living with HIV and information collected in coaching logs completed by personal trainers to assess FOI.

Structured Fidelity Interviews: Perspectives of Persons Living With HIV

We conducted structured interviews with participants living with HIV to obtain perspectives on experiences with the CBE intervention and assess the extent to which the prescribed exercises were delivered according to the established program (FOI). Two members of the team conducted the interviews on Zoom. Interviews were approximately 10 minutes in duration and occurred partway through (month 2) and at the completion of the intervention (month 6). Participants were asked to share their experiences based on their most recent online personal training session and provided insights on any deviations from the prescribed regimen in an open-ended comment section. Responses were recorded electronically on an FOI form developed in Microsoft Word (Multimedia Appendix 1). Each question was scored by research staff as 0 (no criteria for fidelity were met), 1 (partial criteria met), or 2 (complete criteria met). For example, for the item regarding whether the "fitness provider provided direct and clear instructions during...[the] exercise session," the item would be rated 0 if the participant stated that the instructions provided by the fitness provider were never clear or direct; 1 if the participant stated that the instructions provided by the fitness provider were sometimes clear and direct; or 2 if the participant stated that the fitness provider's instructions were always clear and direct. TJ also documented open-ended responses from the participants to further describe barriers and facilitators to fidelity and provide additional context to the numeric ratings.

We evaluated FOI based on 20 items (defined a priori) captured in the interviews. The items were divided into four sections: (1) the completion of their exercise session within the allotted time (1 item), (2) their personal trainer's performance during the session (7 items), (3) participant engagement and experience within the session (6 items), and (4) the broader online CBE study (6 items; Multimedia Appendix 1). FOI for a participant was defined as met if a participant met ≥80% (16/20) of the indicators at months 2 and 6 (from the perspective of the participant living with HIV). Overall, we considered FOI as met if ≥80% of the participants in the study met (achieved) ≥80% of FOI indicators (16/20 items; criterion 3).

Coaching Logs of Personal Training Sessions: Perspectives of Personal Trainers

After each supervised online training session with participants living with HIV, the personal trainers completed an online web-based questionnaire, referred to as the "coaching log." Personal trainers were asked to document (1) whether the participants attended their biweekly personal training session; (2) whether they engaged in aerobic, strength, balance, and flexibility exercise during their biweekly personal training session; and (3) the number of days each participant exercised in the week prior to their session (Multimedia Appendix 2).

FOI for a participant was defined as met, from the perspectives of the personal trainers, if a participant attended ≥ 80% (11 out of 13) of their biweekly coaching sessions (criterion 1), if they exercised 3 times over the past week for ≥1 hour per session (criterion 2), and if they completed a combination of all 4 exercise components (aerobic, strength, balance, and flexibility) ≥80% of the time across their personal training sessions (criterion 4). Collectively, we considered FOI as met if ≥80% of the participants in the study achieved these 4 criteria (Table 1).

Table 1. FOI[a] interpretation criteria.

FOI criteria for each participant[b]	Target population	Data source
Criterion 1: attended ≥80% of biweekly training sessions (at least 11 of 13) over the 6-month intervention	Personal trainers	Coaching logs
Criterion 2: engaged in thrice-weekly exercise the week of the personal training session ≥80% of the time over the 6-month intervention	Personal trainers	Coaching logs
Criterion 3: reported "complete criteria met" for ≥80% of the fidelity items (at least 16 of 20 items) at months 2 and 6 of the intervention	Participants living with HIV	Structured interview at months 2 and 6
Criterion 4: engaged in a combination of aerobic, strength, balance, and flexibility exercise during supervised personal training sessions ≥80% of the time over the 6-month intervention	Personal trainers	Coaching logs

[a]FOI: fidelity of implementation.

[b]Overall FOI was considered "met" for the community-based exercise intervention if the 4 criteria were achieved by ≥80% of participants.

Participant Demographic Questionnaire

We administered a web-based demographic questionnaire to participants living with HIV to collect baseline characteristics such as age, gender, time since HIV diagnosis, and race and ethnicity.

Analysis

Overview

We entered numeric responses from the FOI form collected during the interviews, coaching logs, and demographic questionnaires into Microsoft Excel and analyzed the data using

XSL•FO
RenderX

descriptive statistics [52]. We calculated frequencies and percentages for categorical variables and the median and IQR for continuous variables.

Structured Fidelity Interview

We reported the number and percentage of participants living with HIV who reported "no criteria met" (0), "partial criteria met" (1), and "complete criteria met" (2) for each of the 20 items assessed at months 2 and 6, based on the participants' most recent training session (criterion 3). Open-ended responses from the structured interviews were analyzed using content analysis [53].

Coaching Logs

We calculated the median number of personal training sessions attended over the 6-month intervention (out of 13 sessions total). We calculated the number and proportion of participants who engaged in a combination of all 4 exercise types (aerobic, strength, balance, and flexibility) during the supervised coaching session. We calculated the number and proportion of participants who reported engaging in thrice-weekly exercise in the week of the personal training session, as recorded by the personal trainers in the coaching logs. We summed the number and proportion of participants who attended ≥80% (11/13) of biweekly coaching sessions (criterion 1), summed the number and proportion of participants who engaged in thrice-weekly exercise ≥80% of the time over the 6-month intervention (criterion 2), and the number and proportion of participants who completed the combination of all 4 exercise types in ≥80% of the sessions they attended (criterion 4) [54-56].

To account for attrition across the 6-month intervention, we reported baseline characteristics of participants at baseline and at months 2 and 6. We reported fidelity outcomes among participants who remained in the study at month 2 and who completed the intervention at month 6, as well as the original sample of participants who initiated the intervention.

Fidelity Interpretation

We defined overall FOI of the CBE intervention as achieved if the following 4 criteria were met by ≥80% of the participants in the study: criterion 1, attended ≥80% (11/13) of their 13 training sessions; criterion 2, engaged in (at least) thrice-weekly exercise in the week of the personal training session ≥80% of the time over the 6-month intervention, as recorded by the personal trainers on the coaching logs; criterion 3, reported "complete criteria met" for ≥80% of the fidelity elements (at least 16 of 20 items) at months 2 and 6, as reported in the interviews with participants living with HIV; and criterion 4, engaged in a combination of aerobic, strength, balance, and flexibility exercises during the supervised coaching sessions ≥80% of the time over the 6-month intervention, as reported by the personal trainers in the coaching logs. Our criteria are based on the literature, whereby 80% to 100% is considered high fidelity, and ≤50% defines low FOI [54-56].

Ethical Considerations

This study was approved by the University of Toronto Health Sciences Research Ethics Board (#40410). Participants provided informed verbal consent to participate in the study, which was documented on the consent form. Participant data were anonymized and stored on a secure server at the University of Toronto. Participants did not receive compensation for their participation in the exercise intervention. They were able to retain their exercise equipment and wireless physical activity monitor at study completion.

Results

Overview

Adults living with HIV who initiated the intervention and had at least 1 completed coaching log were included in the study. Of the 32 participants, 18 (56%) completed the 6-month intervention and had an FOI interview assessment. A total of 29/32 (91%) participants who completed at least 1 fidelity interview were included in the FOI analysis. Further, 16 (50%) participants completed the FOI assessment (interviews) at both months 2 and 6; 11 (34%) completed the assessment at month 2 only, and 2 (6%) completed the assessment at month 6 only. Of a potential 64 fidelity interviews, 45 (70%) were completed. At month 2, a total of 84% (27/32) of assessments were completed, and at month 6, a total of 56% (18/32) of assessments were completed.

Of the 32 participants at baseline, the majority were male and living with 2 or more concurrent health conditions in addition to HIV. Table 2 provides characteristics of participants enrolled in the study at baseline, month 2, and month 6.

XSL•FO
RenderX

Table 2. Characteristics of participants living with HIV at baseline and among those who completed an FOI[a] interview at months 2 and 6.

Characteristic	Baseline (n=32)	Month 2 (n=27)	Month 6 (n=18)
Age (years), median (IQR)	53 (44-60)	50 (42-59)	57 (47-59)
Sex, n (%)			
Male	22 (69)	18 (67)	13 (72)
Female	10 (31)	9 (33)	5 (28)
Gender, n (%)			
Man, “cis-man”	22 (69)	18 (66)	4 (22)
Woman, “cis-woman”	9 (28)	8 (30)	13 (72)
Nonbinary	1 (3)	1 (4)	1 (6)
Sexual orientation, n (%)			
Gay/homosexual	18 (56)	15 (56)	11 (61)
Heterosexual	7 (22)	7 (26)	3 (17)
Queer	4 (13)	3 (11)	2 (11)
Bisexual	1 (3)	1 (4)	1 (6)
Race/ethnicity[b], n (%)			
White	14 (44)	11 (41)	6 (30)
South Asian	8 (25)	8 (30)	4 (20)
Black	8 (25)	6 (22)	5 (28)
Hispanic	2 (6)	2 (7)	2 (11)
Métis	2 (6)	1 (4)	0 (0)
Southeast Asian	1 (3)	1 (4)	1 (6)
Gross annual income (CAD $[c]), n (%)			
<20,000	7 (22)	0 (0)	5 (28)
20,000-49,999	17 (53)	20 (74)	7 (39)
>50,000	8 (25)	7 (26)	6 (33)
Married or in a committed relationship, n (%)	14 (44)	13 (48)	7 (39)
Live alone, n (%)	12 (38)	7 (26)	7 (39)
Living with 2 or more concurrent health conditions in addition to HIV, n (%)	23 (72)	18 (67)	11 (61)

[a]FOI: fidelity of implementation.

[b]Participants could select more than 1 response.

[c]CAD $1=US $0.71 as of August 1, 2026.

Of the 32 participants, a total of 8 (25%) who initiated the intervention completed all 13 coaching sessions, 7 (24%) of whom were among the 29 participants who completed an FOI assessment interview. We received responses to 72% (298/416) of the coaching logs out of the total possible for 32 participants living with HIV to determine FOI from the perspectives of personal trainers.

Fidelity Results

Overall Fidelity

FOI was met for attendance criterion 1 among participants who completed the intervention, whereby 83% (15/18; 95% CI 66%-100%) of participants attended ≥80% of their personal training sessions (Table 3). We describe results as they pertain to each criterion below.

XSL•FO
RenderX

Table 3. Summary of FOI[a] criterion achievement.

FOI criterion	Data source	Participants who met the criteria among those who initiated the intervention (n=32), n (%); 95% CI	Participants who met the criteria among those who completed the intervention and month 6 FOI assessment (n=18), n (%) 95% CI	Fidelity met: achieved by ≥80% of participants (yes or no)
Criterion 1: attended ≥80% of biweekly personal training sessions (at least 11 of 13 sessions)	Coaching log attendance documented by personal trainers	18 (56); 39-73	15 (83); 66-100	Yes[b]
Criterion 2: engaged in thrice-weekly exercise the week of the personal training session ≥80% of the time over the 6-month intervention[c]	Coaching logs documented by personal trainers	11 (34); 18-50	9 (50); 27-73	No
Criterion 3: reported "complete criteria met" for ≥80% of the fidelity criteria (≥16 of the 20 items) at months 2 and 6[d]	Structured interview at months 2 and 6 with participants living with HIV	Month 2 (n=27): 21 (78) 62-93 and month 6 (n=18): 8 (44); 27-67	Month 2 (n=16): 12 (75); 54-96 and month 6 (n=18): 8 (44); 21-67	No
Criterion 4: engaged in a combination of aerobic, strength, balance, and flexibility exercises during supervised personal training sessions ≥80% of the time over the 6-month intervention	Coaching logs documented by personal trainers	8 (25) 10-40	4 (22) 3-41	No

[a]FOI: fidelity of implementation.

[b]Indicates that criterion 1 was met among participants who completed the intervention (≥80%).

[c]For criterion 2, the week of the first personal training session was excluded; baseline was defined as the start of the intervention.

[d]For criterion 3, the proportion of participants who selected "complete criteria met" for ≥80% of the fidelity criteria also were calculated based on the total number of participants who initiated the intervention (n=32). Participants who did not complete the fidelity assessments were assumed not to have met the criteria: At month 2, n=20/32 (63%; 95% CI 46%-79%); at month 6, n=8/32 (25%; 95% CI 10%-40%).

Criterion 1: Attended ≥80% of Personal Training Sessions by ≥80% of Participants

This FOI criterion was met for the participants who completed the intervention and an FOI assessment at month 6; 83% (15/18; 95% CI 66%-100%) of the participants attended ≥80% of the 13 coaching sessions (Table 3). Of all 32 participants, 18 (56%; 95% CI 39%-73%) attended ≥80% of the 13 coaching sessions (Table 3). Of all 32 (100%) participants who attended at least 1 personal training session, the median number of training sessions attended was 11 of 13 (IQR 6-12). Reasons for not attending or having to reschedule the online coaching sessions included scheduling issues, health reasons, or unknown reasons.

Criterion 2: Engaged in Thrice-Weekly Exercise the Week of the Personal Training Session ≥80% of the Time Over the 6-Month Intervention by ≥80% of Participants

This FOI criterion was not met. A total of 34% (11/32; 95% CI 18%-50%) of all participants and 50% (9/18; 95% CI 27%-73%) of those who completed the month 6 FOI assessment engaged in thrice-weekly exercise in the week of the personal training session ≥80% of the time over the 6-month intervention (Table 3). The median number of days of exercise per week for all participants (n=32) and for those who completed the intervention and month 6 FOI assessment (n=18) was 3 (IQR 3-5) and 4 days (IQR 3-5), respectively.

Criterion 3: Participants Living With HIV Reported ≥80% of Criteria as "Completely Met" by ≥80% of Participants

This FOI criterion, defined as ≥80% of participants reporting ≥80% of FOI elements (items) as "completely met," was not achieved at months 2 or 6 (Table 3). Among the participants who completed an FOI assessment, 78% (21/27; 95% CI 62%-93%) and 44% (8/18; 95% CI 27%-67%) of the participants completely met the FOI elements for ≥80% of the 20 items at months 2 and 6, respectively. When considering the entire sample of 32 participants, fewer items met the criteria, whereby ≥80% of participants reported the FOI element as "completely met" for 4 of 20 items at month 2 and none of the 20 items at month 6 (details are provided in Multimedia Appendix 3).

Table 4 provides a detailed breakdown of the FOI assessment across the 20 interview items. At least 80% of participants reported the FOI element as "completely met" for 70% (14/20; 95% CI 50%-89%) of items at month 2 (n=27) and 30% (6/20; 95% CI 10%-50%) of items at month 6 (n=18). At month 2, the following 6 elements did not achieve fidelity at the threshold of ≥80% of participants: (1) participant completed their session in the allotted time (1 hour with coach online; 21/27, 78%; 95% CI 62%-94% of participants completely met this item), (2) participant responded to and completed the most recent weekly exercise email or SMS text message (20/27, 74%; 95% CI 58%-91%), (3) participant completed their additional individual independent exercise sessions (60 minutes twice per week, for a total of 60 minutes thrice per week) in the past week (13/27,

XSL•FO
RenderX

48%; 95% CI 29%-67%), (4) participant managed the required technology (eg, Zoom and SweatforGood app; 15/27, 55%; 95% CI 36%-74%), (5) participant engaged in monthly self-management sessions online (12/27, 45%; 95% CI 26%-63%), and (6) participant attended weekly group-based classes online at the YMCA (8/27, 30%; 95% CI 13%-47%; Table 4). At month 6, the following eight additional elements were not met: (1) fitness trainer instructed participant to complete resistive exercises according to their set exercise program (13/18, 72%; 95% CI 52%-93% of participants completely met this criterion), (2) fitness trainer instructed participant to complete flexibility exercises according to their set exercise program (13/18, 72%; 95% CI 52%-93%), (3) participant engaged in the prescribed strength exercises during the most recent session (12/18, 67%; 95% CI 45%-89%), (4) fitness trainer instructed participant to complete neuromotor/balance exercises according to their set exercise program (12/18, 67%; 95% CI 45%-89%), (5) participant engaged in the prescribed flexibility exercises (12/18, 67%; 95% CI 45%-89%), (6) fitness trainer instructed participant to complete aerobic exercises according to their set exercise program (11/18, 61%; 95% CI 39%-84%), (7) participant engaged in the prescribed neuromotor/balance exercises (11/18, 61%; 95% CI 39%-84%), and (8) participant engaged in the prescribed aerobic exercises during the most recent session (9/18, 50%; 95% CI 27%-73%; Table 4).

Content analysis of the open-ended responses identified barriers to meeting some of the criteria and coaching practices that were helpful for participants. The most common barriers to attending YMCA group-based classes were scheduling issues, lack of interest, and technology issues. Fitbit Inspire 2 wristband discomfort was cited as a barrier to wearing the physical activity monitor regularly. Illness, injury, pain, discomfort, and lack of interest were common barriers to completing the aerobic, resistance, balance, and flexibility exercises. Statements of appreciation for the personal training sessions included “[coach] demonstrates exercises to ensure participant understands proper form” (P16; month 2) and “[coach] demonstrates all exercises/modifications then participant performs, and trainer provides feedback” (P2; month 2). Overall, exercise demonstration, adaptation to participant circumstances, and providing feedback were positive coaching practices identified by participants.

XSL•FO
RenderX

Table 4. FOI[a] (criterion 3)–structured interviews with adults living with HIV. "Fitness trainer" is analogous to "personal trainer."

FOI item	Month 2 (n=27 interviews), %			Month 6 (n=18 interviews), %		
	No criteria met (0), n (%); 95% CI	Partial criteria met (1), n (%); 95% CI	Complete criteria met (2), n (%); 95% CI	No criteria met (0), n (%); 95% CI	Partial criteria met (1), n (%); 95% CI	Complete criteria met (2), n (%); 95% CI
External factor[b]						
Item 1: participant completed their session within the allotted time (1 hour with a coach online)	0 (0)	6 (22); 7-38	21 (78); 62-94	0 (0)	4 (22); 3-41	14 (78); 27-61
Fitness trainer factors: the fitness trainer's performance during the session[b]						
Item 2: fitness trainer provided direct and clear instructions during exercise session	0 (0)	0 (0)	27 (100)[c]	2 (11); 0-26	0 (0)	16 (89); 74-100[c]
Item 3: fitness trainer interacted with participant in a way that was meaningful or helpful for their session (provided eye contact, feedback, ongoing communication, etc)	0 (0)	2 (7); 0-17	25 (93); 83-100[c]	0 (0)	2 (11); 0-26	16 (89); 74-100[c]
Item 4: fitness trainer instructed participant to complete aerobic exercises according to their prescribed exercise program	2 (7); 0-17	3 (11); 0-23	22 (82); 53-84[c]	3 (17); 0-34	4 (22); 3-41	11 (61); 39-84
Item 5: fitness trainer instructed participant to complete resistive exercises according to their set exercise program	1 (4); 0-11	0 (0)	26 (96); 89-100[c]	2 (11); 0-26	3 (17); 0-34	13 (72); 52-93
Item 6: fitness trainer instructed participant to complete neuromotor/balance exercises according to their prescribed exercise program	1 (4); 0-11	1 (4); 0-11	25 (93); 83-100[c]	4 (22); 3-41	2 (11); 0-26	12 (67); 45-89
Item 7: fitness trainer instructed participant to complete flexibility exercises according to their prescribed exercise program	0 (0)	3 (11); 0-23	24 (89); 77-100[c]	3 (17); 0-34	2 (11); 0-26	13 (72); 52-93
Item 8: fitness trainer changed the program from previous sessions to progress the participant's activity (either increase or decrease in intensity)	0 (0)	1 (4); 0-11	26 (96); 89-100[c]	1 (6); 0-16	0 (0)	17 (94); 84-100[c]
Participant factors: participant engagement and experience[c]						
Item 9: participant understood the instructions provided by their coach	0 (0)	2 (7); 0-17	25 (93); 83-100[c]	2 (11); 0-26	1 (6); 0-16	15 (83); 66-100[c]
Item 10: participant accepted direction/feedback provided by the coach	0 (0)	0 (0)	27 (100)[c]	2 (11); 0-26	0 (0)	16 (89); 74-100[c]
Item 11: participant engaged in the prescribed aerobic exercises during the most recent session	2 (7); 0-17	3 (11); 0-23	22 (82); 67-96[c]	3 (17); 0-34	6 (33); 12-55	9 (50); 27-73
Item 12: participant engaged in the prescribed resistive exercises during the most recent session	1 (4); 0-11	1 (4); 0-11	25 (93); 83-100[c]	2 (11); 0-26	4 (22); 3-41	12 (67); 45-89
Item 13: participant engaged in the prescribed neuromotor/balance exercises	1 (4); 0-11	1 (4); 0-11	25 (93); 83-100[c]	4 (22); 3-41	3 (17); 0-34	11 (61); 39-84

XSL•FO
RenderX

FOI item	Month 2 (n=27 interviews), %			Month 6 (n=18 interviews), %		
	No criteria met (0), n (%); 95% CI	Partial criteria met (1), n (%); 95% CI	Complete criteria met (2), n (%); 95% CI	No criteria met (0), n (%); 95% CI	Partial criteria met (1), n (%); 95% CI	Complete criteria met (2), n (%); 95% CI
Item 14: participant engaged in the prescribed flexibility exercises	1 (4); 0-11	1 (4); 0-11	25 (93); 83-100[c]	3 (17); 0-34	3 (17); 0-34	12 (67); 45-89
Additional factors: the broader telecoaching study intervention						
Item 15: participant completed additional individual sessions (60 minutes twice per week) during the past week	5 (19); 4-33	9 (33); 15-51	13 (48); 29-67	4 (22); 3-41	3 (17); 0-34	11 (61); 39-84
Item 16: participant responded to or completed the most recent weekly exercise email/text message	6 (22); 7-38	1 (4); 0-11	20 (74); 58-91	4 (22); 3-41	2 (11); 0-26	12 (67); 45-89
Item 17: participant attended weekly group-based classes online at the YMCA	15 (56); 37-74	4 (15); 1-28	8 (30); 13-47	8 (44); 21-67	4 (22); 3-41	6 (33); 12-55
Item 18: participant engaged in monthly self-management sessions online	5 (19); 4-33	10 (37); 19-55	12 (45); 26-63	7 (39); 16-61	7 (39); 16-61	4 (22); 3-41
Item 19: participant wore a Fitbit regularly	0 (0)	3 (11); 0-23	24 (89); 77-100[c]	0 (0)	0 (0)	18 (100)[c]
Item 20: participant managed the required technology (Zoom, SweatforGood app, etc)	0 (0)	12 (44); 26-63	15 (55); 36-74	1 (5); 0-16	7 (39); 16-61	10 (56); 33-79

[a]FOI: fidelity of implementation.

[b]Participants were asked about their experiences at their most recent personal training session.

[c]Indicates ≥80% of participants 'completely met' the criteria.

Discussion

Principal Findings

Among participants who completed the intervention and FOI assessment at month 6 (n=18), FOI was achieved for 1 of the 4 criteria, whereby 83% (95% CI 66%-100%) of participants who completed the study engaged in ≥80% of biweekly online personal training sessions. However, fidelity criteria were not met for independent exercise thrice weekly (50%; 95% CI 27%-73%), participant-reported fidelity criteria at month 6 ("complete criteria met" for at least 16/20, 44% of items; 95% CI 21%-67%), or engagement in all 4 exercise types in ≥80% of sessions (4/18, 22%; 95% CI 3%-41%).

Within the training session, the majority of participants engaged in strength and flexibility exercises, followed by aerobic and then balance exercise. This may be attributed to strength training requiring a higher level of supervision as opposed to aerobic activity that was done independently outside the formal personal training session. In the absence of equipment for aerobic activity, such as a treadmill or stationary bike, participants may have engaged in this type of activity independently outdoors.

Fidelity was achieved for some components of the intervention. At least 80% of participants completely met 14 out of 20 FOI elements at month 2 and 6 out of 20 FOI elements at month 6, as reported in the structured interviews. The FOI items least commonly met were independent engagement in exercise (for a total of thrice weekly or more of 60 minutes or longer), attendance at weekly group-based online classes at the YMCA (8/27, 29% of participants met this criterion), and attendance at monthly online group-based educational sessions (12/27, 45% of participants met this criterion).

Supervised vs Independent Forms of Exercise Intervention

Low engagement in the online group-based YMCA classes and independent exercise outside of supervised one-on-one personal training sessions may be influenced by the unsupervised components of the intervention, as participants had higher engagement rates when supervised (ie, personal training sessions). This is consistent with the literature, which shows increased engagement in supervised exercise programs, with adults aged 65 years or older favoring supervised exercise programs [57]. As found within the open-ended FOI responses in this study, receiving feedback, education, and modified exercise tailored to participant preferences and abilities may be important factors driving adherence to supervised exercise [57-60]. The identification of the least-implemented intervention components can help inform future intervention design and may also inform the supports necessary for participants to sustain regular exercise routines postintervention. This may help to identify components of an intervention that require supervision and suggests the importance for future interventions to tailor

XSL•FO
RenderX

CBE based on personalized preferences to promote engagement. Further, there is a need to explore population-specific factors that impact adherence to unsupervised exercise to better develop strategies that improve long-term behavioral change for adults living with HIV.

Group-Based vs Individual Exercise Intervention

FOI criteria were not met for the group-based elements of the intervention. At month 2, a total of 45% (12/27; 95% CI 26%-63%) of participants reported attending the monthly online self-management sessions (≥80% of the time), and 22% (4/18; 95% CI 3%-41%) at month 6 (item 18 in the FOI assessment). At month 2, a total of 29% (8/27; 95% CI 12%-47%) of participants reported attending weekly online group-based exercise classes at the YMCA, and 33% (6/18; 95% CI 12%-55%) at month 6 (item 17 in the FOI assessment; Table 4). With HIV stigma being a barrier to seeking treatment, care, and health- and wellness-related programs, the lower engagement in group-based interventions may have been attributed to fear of HIV disclosure and discrimination [61-64].

The Influence of Environmental and Personal Factors

Personal and environmental factors may further explain the FOI findings in this study. HIV is characterized by episodic disability, where there are unpredictable periods of wellness and illness [65]. Barriers to engaging in the intervention, as reported by participants, included illness, pain, or injury. The fluctuating nature of chronic conditions like HIV may make it challenging for participants to follow highly rigid schedules and training plans if their health and ability to engage in exercise vary day-to-day [23,51]. Perhaps our FOI criteria were too rigid and did not align with or consider the episodic and fluctuating nature of HIV. We recommend providers adapt the frequency and type of exercise to the nature of the illness to maintain the well-being of an individual living with HIV. Our findings indicate the need for flexibility with CBE interventions, enabling persons to cancel or reschedule sessions due to illness more frequently and engage in exercise at home [23,51]. We recommend that exercise interventions in this population be adaptable and considerate of the dynamic nature of the illness experience to ensure that intervention studies are pragmatic and translational to the real world [23,51,60].

While participants were provided with a YMCA online membership at no cost and access to a personal trainer biweekly, their financial situation still may have influenced engagement in the intervention. Income is positively correlated with exercise adherence [66,67]. Qualitative insights from this telecoaching CBE study reported that participants in this study highlighted cost as a limiting factor to engagement in physical activity [60]. A lower socioeconomic status may be correlated with poor exercise adherence due to a lack of leisure time. Individuals of higher socioeconomic status, on average, allocate more leisure time to exercise than those of lower socioeconomic status [66,67]. An association exists between low socioeconomic status and exercise in persons living with HIV, with barriers to exercise including caregiver responsibilities, time dedicated to childcare, unsafe neighborhoods, strict work schedules, and living in temporary dwellings [68,69]. Traditional exercise interventions may not be suitable for those living with HIV who are socioeconomically disadvantaged because of barriers related to employment, financial constraints, stigmatization, and inaccessible or lack of transportation, requiring interventions to be tailored to the needs of this population [68,69]. Collectively, it is important for health and fitness trainers to consider the environmental and personal factors that can influence exercise engagement and intervention adherence to tailor interventions accordingly. This may include tailoring modes of exercise delivery (eg, online and in-person; individual training or group-based classes) based on individual preferences.

Strengths: Assessing Fidelity From 2 Different Perspectives

Strengths of our approach included our multimethod data collection at multiple time points of the CBE intervention and engaging perspectives of both participants living with HIV and personal trainers involved in implementation of the intervention. FOI in our study was assessed more positively from the perspectives of the participants living with HIV (through structured interviews), as opposed to the personal trainers' logs of the coached sessions. Our fidelity assessment from the perspectives of participants living with HIV pertained to their most recent personal training session, whereas fidelity assessment from the perspective of the personal trainers captured engagement in all coached training sessions and the independent exercise done in the preceding week. Using multiple methods to assess intervention fidelity provided a greater understanding of fidelity [70,71] and allowed for different populations to have different perspectives on a shared phenomenon [37,44,70]. Toomey et al [71] recognized that disagreements between results of different fidelity assessment approaches from providers and persons with lived experiences may occur. Some of the differences in FOI findings in our study may be due to participants' overestimation of their performance when completing self-reported measures (2 time points), as compared to the personal training coaching logs completed by providers more frequently across the biweekly supervised sessions (13 time points). Nevertheless, our assessment captured perspectives from both persons living with HIV and the trainers at different stages of the intervention.

Limitations

Our study protocol included methods for capturing the implementation of the intervention components but did not explicitly state the definition of implementation fidelity. Prior to the analysis, we defined the FOI threshold most commonly reported in the literature as achievement of ≥80% of criteria [54-56]. We considered all 4 criteria for FOI to be weighted equally; however, in reality, specific aspects of the CBE intervention may be more important (eg, attendance at biweekly training sessions) to achieve FOI than others (ie, engaging in all 4 types of exercise in the training session). Given that we did not collect fidelity data in the same manner for both target populations (persons living with HIV and personal trainers), we were unable to directly compare the influence of perspectives on fidelity assessment. Furthermore, participants were asked about fidelity based on their most recent coaching session, meaning that a number of sessions were not captured in this analysis, nor do we have the reasons for noncompletion of the

XSL•FO
RenderX

FOI interviews or coaching logs. We did not report physical activity as measured by the wireless physical activity monitor, as this was beyond the scope of this study. However, as the Fitbit was a component of the intervention, our FOI assessment included Fitbit usage (item 19 on the FOI assessment), whereby 89% to 100% of participants at months 2 and 6 wore the wireless physical activity monitor regularly. The uptake and usage of the Fitbit aligns with experiences of technology uptake and usage in the telecoaching CBE study [72].

Our study is limited to participants who remained in the study and completed the FOI assessment interview at month 6. Of 32 participants, only 18 (56%) completed the intervention and an FOI assessment at month 6, resulting in an overestimation of FOI. In our sensitivity analyses, assuming participants who withdrew or were lost to follow-up did not meet the criteria, only 18 (56%; 95% CI 39%-73%) achieved the criteria of engaging in ≥80% of biweekly online personal training sessions (Table 3 and Multimedia Appendix 3), resulting in none of the 4 FOI criteria being met for the study.

Another limitation was that the majority of participants living with HIV in this study identified as males and men; hence, findings may not be transferable to individuals who identify as females or women living with HIV. Also, we did not assess the association of contextual factors such as level of income, stigma, and discrimination. Future research should consider the intersecting components of sex, gender, culture, socioeconomic status, and stigma and their influence on engagement in exercise [73].

Recommendations

There is a need for balancing FOI with personalized and tailored implementation of interventions for adults living with HIV [74-76]. While fidelity is important, solely prioritizing the implementation of the intervention as intended at the cost of tailoring the intervention to the population may not be the most beneficial for the population of adults living with HIV. To ensure the integrity and quality of intervention fidelity assessments while adapting the intervention to the population and its needs, it is important to define, a priori, the intervention components that may be modified to meet participant needs and circumstances. These include optimal frequency of supervised sessions (weekly vs biweekly), mode of exercise, and exercise setting to support independent exercise.

Conclusions

FOI was achieved for the supervised personal training components of the CBE intervention but was not achieved for independent exercise among adults living with HIV who engaged in this 6-month online CBE intervention. Future research should consider tailoring online CBE interventions to personal preferences to foster uptake and sustained independent engagement in exercise over the long term.

Acknowledgments

This study is a substudy of the Ontario HIV Treatment Network Cohort Study (OCS). We thank the OCS (Abigail Kroch and Tsegaye Bekele) for their collaboration in the study. We thank Maple Leaf Medical Clinic, and specifically Raj Maharaj, for their role in facilitating recruitment of participants from the OCS with the Maple Leaf Medical Clinic. We acknowledge and thank Brittany Torres for her coordination role in the study and for conducting a portion of the fidelity interviews with the participants. We thank the YMCA staff, Annamaria Furlan, Helen Trent, Emilia Przybyl, and Dwayne Campbell, for their roles in the fitness assessments and personal training sessions for the intervention. We acknowledge and thank Mehdi Zobeiry at the YMCA of Toronto for their support with this study.

We did not use AI in the development of this manuscript.

Funding

This research was supported by the Ontario HIV Treatment Network HIV Endgame Funding Program—Breaking New Ground (grant number EFP-1121-BNG). KKO was supported by a Canada Research Chair (Tier 2) in Episodic Disability and Rehabilitation (CRC-2022-00510); AMB was supported by the Foundation Alma and Baxter Ricard Chair in Inner City Health at St Michael's Hospital and the University of Toronto.

Data Availability

Results of the study are supported by the data provided in the manuscript and its supplementary files. The complete dataset is available upon reasonable request from the corresponding author.

Authors' Contributions

Conceptualization: KKO (lead), SCC, AT, AMB, LA, FI-C, ML
Formal analysis: MA (lead), JN (supporting)
Funding acquisition: KKO (lead), SCC, AT, AMB, PA, LA, FI-C, ML
Investigation: TJ, KM, AF, GDS, II, ZP, HT
Methodology: KKO, SCC
Project administration: KKO

XSL•FO
RenderX

Resources: KKO
Supervision: SCC (lead), KKO (supporting)
Validation: KKO, SCC (co-lead)
Visualization: MA, JN, SCC, KKO
Writing—original draft: MA, SCC, KKO
Writing—review and editing: MA, SCC, JN, TJ, AT, PA, AMB, LA, AF, GDS, FI-C, II, KM, ML, ZP, HT, KKO

Conflicts of Interest

None declared.

Multimedia Appendix 1
Fidelity of implementation (FOI) assessment.
[PDF File (Adobe PDF File), 279 KB - rehab_v13i1e84902_app1.pdf]

Multimedia Appendix 2
Personal training coaching log for personal trainers.
[PDF File (Adobe PDF File), 290 KB - rehab_v13i1e84902_app2.pdf]

Multimedia Appendix 3
Fidelity of implementation (FOI) interviews for participants who initiated the study and completed month 2 and month 6 FOI interview assessments.
[PDF File (Adobe PDF File), 292 KB - rehab_v13i1e84902_app3.pdf]

References

1. Canada's progress towards ending the HIV epidemic, 2022. Public Health Agency of Canada. 2024. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/canada-progress-towards-ending-hiv-epidemic-2022.html [accessed 2024-08-05]
2. The Lancet Healthy Longevity. Ageing with HIV. Lancet Healthy Longev 2022;3(3):e119 [FREE Full text] [doi: 10.1016/S2666-7568(22)00041-1] [Medline: 36098283]
3. Pourcher V, Gourmelen J, Bureau I, Bouee S. Comorbidities in people living with HIV: an epidemiologic and economic analysis using a claims database in France. PLoS One 2020;15(12):e0243529 [FREE Full text] [doi: 10.1371/journal.pone.0243529] [Medline: 33332394]
4. Maciel RA, Klück HM, Durand M, Sprinz E. Comorbidity is more common and occurs earlier in persons living with HIV than in HIV-uninfected matched controls, aged 50 years and older: a cross-sectional study. Int J Infect Dis 2018;70:30-35 [FREE Full text] [doi: 10.1016/j.ijid.2018.02.009] [Medline: 29476902]
5. Aberg JA. Aging and HIV infection: focus on cardiovascular disease risk. Top Antivir Med 2020;27(4):102-105 [FREE Full text] [Medline: 32224501]
6. Worthington C, Myers T, O'Brien K, Nixon S, Cockerill R. Rehabilitation in HIV/AIDS: development of an expanded conceptual framework. AIDS Patient Care STDS 2005;19(4):258-271. [doi: 10.1089/apc.2005.19.258] [Medline: 15857198]
7. O'Brien KK, Ibáñez-Carrasco F, Solomon P, Harding R, Cattaneo J, Chegwidden W, et al. Advancing research and practice in HIV and rehabilitation: a framework of research priorities in HIV, disability and rehabilitation. BMC Infect Dis 2014;14:724 [FREE Full text] [doi: 10.1186/s12879-014-0724-8] [Medline: 25551619]
8. Vader K, Carusone SC, Aubry R, Ahluwalia P, Murray C, Baxter L, et al. Strengths and challenges of implementing physiotherapy in an HIV community-based care setting: a qualitative study of perspectives of people living with HIV and healthcare providers. J Int Assoc Provid AIDS Care 2021;20:23259582211005628 [FREE Full text] [doi: 10.1177/23259582211005628] [Medline: 33779374]
9. deBoer H, Cudd S, Andrews M, Leung E, Petrie A, Chan Carusone S, et al. Recommendations for integrating physiotherapy into an interprofessional outpatient care setting for people living with HIV: a qualitative study. BMJ Open 2019;9(5):e026827 [FREE Full text] [doi: 10.1136/bmjopen-2018-026827] [Medline: 31129584]
10. O'Brien KK, Tynan A, Nixon SA, Glazier RH. Effectiveness of progressive resistive exercise (PRE) in the context of HIV: systematic review and meta-analysis using the cochrane collaboration protocol. BMC Infect Dis 2017;17(1):268 [FREE Full text] [doi: 10.1186/s12879-017-2342-8] [Medline: 28403830]
11. O'Brien KK, Tynan A, Nixon SA, Glazier RH. Effectiveness of aerobic exercise for adults living with HIV: systematic review and meta-analysis using the cochrane collaboration protocol. BMC Infect Dis 2016;16:182 [FREE Full text] [doi: 10.1186/s12879-016-1478-2] [Medline: 27112335]
12. Schuelter-Trevisol F, Wolff FH, Alencastro PR, Grigoletti S, Ikeda ML, Brandão ABM, et al. Physical activity: do patients infected with HIV practice? How much? A systematic review. Curr HIV Res 2012;10(6):487-497. [doi: 10.2174/157016212802429794] [Medline: 22762420]

XSL•FO
RenderX

13. Stuart M, Benvenuti F, Macko R, Taviani A, Segenni L, Mayer F, et al. Community-based adaptive physical activity program for chronic stroke: feasibility, safety, and efficacy of the Empoli model. Neurorehabil Neural Repair 2009;23(7):726-734 [FREE Full text] [doi: 10.1177/1545968309332734] [Medline: 19318465]
14. Salbach NM, Howe JA, Brunton K, Salisbury K, Bodiam L. Partnering to increase access to community exercise programs for people with stroke, acquired brain injury, and multiple sclerosis. J Phys Act Health 2014;11(4):838-845. [doi: 10.1123/jpah.2012-0183] [Medline: 23676952]
15. King AC, Haskell WL, Taylor CB, Kraemer HC, DeBusk RF. Group- vs home-based exercise training in healthy older men and women. A community-based clinical trial. JAMA 1991;266(11):1535-1542. [Medline: 1880885]
16. Winward C, LIFE group. Supporting community-based exercise in long-term neurological conditions: experience from the long-term individual fitness enablement (LIFE) project. Clin Rehabil 2011;25(7):579-587. [doi: 10.1177/0269215510392075] [Medline: 21697208]
17. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]
18. Vader K, Simonik A, Ellis D, Kesbian D, Leung P, Jachyra P, et al. Perceptions of ‘physical activity’ and ‘exercise’ among people living with HIV: a qualitative study. Int J Ther Rehabil 2017;24(11):473-482. [doi: 10.12968/ijtr.2017.24.11.473]
19. Sousa CV, Lewis JE, Simoes HG, Campbell CSG, Zaldivar S, Rodriguez A, et al. The effectiveness of a community-based exercise program on depression symptoms among people living with HIV. AIDS Care 2021;33(3):368-374. [doi: 10.1080/09540121.2020.1719278] [Medline: 31996018]
20. Lee L, Arthur A, Avis M. Using self-efficacy theory to develop interventions that help older people overcome psychological barriers to physical activity: a discussion paper. Int J Nurs Stud 2008;45(11):1690-1699. [doi: 10.1016/j.ijnurstu.2008.02.012] [Medline: 18501359]
21. O'Brien KK, Davis AM, Chan Carusone S, Avery L, Tang A, Solomon P, et al. Examining the impact of a community-based exercise intervention on cardiorespiratory fitness, cardiovascular health, strength, flexibility and physical activity among adults living with HIV: a three-phased intervention study. PLoS One 2021;16(9):e0257639 [FREE Full text] [doi: 10.1371/journal.pone.0257639] [Medline: 34559851]
22. Cooper J, Murphy J, Woods C, Van Nassau F, McGrath A, Callaghan D, Irish Physical Activity Research Collaboration (I-PARC). Barriers and facilitators to implementing community-based physical activity interventions: a qualitative systematic review. Int J Behav Nutr Phys Act 2021;18(1):118 [FREE Full text] [doi: 10.1186/s12966-021-01177-w] [Medline: 34493306]
23. Solomon P, Carusone SC, Davis AM, Aubry R, O'Brien KK. Experiences of people living with HIV in community based exercise: a qualitative longitudinal study. J Int Assoc Provid AIDS Care 2021;20:2325958221995344 [FREE Full text] [doi: 10.1177/2325958221995344] [Medline: 33611978]
24. Montgomery CA, Henning KJ, Kantarzhi SR, Kideckel TB, Yang CFM, O'Brien KK. Experiences participating in a community-based exercise programme from the perspective of people living with HIV: a qualitative study. BMJ Open 2017;7(4):e015861 [FREE Full text] [doi: 10.1136/bmjopen-2017-015861] [Medline: 28377397]
25. Hilty DM, Chan S, Torous J, Luo J, Boland RJ. Mobile health, smartphone/device, and apps for psychiatry and medicine: competencies, training, and faculty development issues. Psychiatr Clin North Am 2019;42(3):513-534. [doi: 10.1016/j.psc.2019.05.007] [Medline: 31358129]
26. Framework for the implementation of a telemedicine service. Pan American Health Organization. 2016. URL: https://iris.paho.org/items/2e3cc008-a0a0-487d-bb62-db1fcbb57939 [accessed 2026-08-05]
27. Piraux E, Reychler G, Forget P, Yombi JC, Caty G. Feasibility and preliminary effects of a telerehabilitation program for people living with HIV: a pilot randomized study. J Assoc Nurses AIDS Care 2019;30(2):176-185. [doi: 10.1097/JNC.0000000000000005] [Medline: 30822290]
28. Loeckx M, Rabinovich RA, Demeyer H, Louvaris Z, Tanner R, Rubio N, et al. Smartphone-based physical activity telecoaching in chronic obstructive pulmonary disease: mixed-methods study on patient experiences and lessons for implementation. JMIR Mhealth Uhealth 2018;6(12):e200 [FREE Full text] [doi: 10.2196/mhealth.9774] [Medline: 30578215]
29. Ramage ER, Fini NA, Lynch EA, Patterson A, Said CM, English C. Supervised exercise delivered via telehealth in real time to manage chronic conditions in adults: a protocol for a scoping review to inform future research in stroke survivors. BMJ Open 2019;9(3):e027416 [FREE Full text] [doi: 10.1136/bmjopen-2018-027416] [Medline: 30898833]
30. Lai B, Bond K, Kim Y, Barstow B, Jovanov E, Bickel CS. Exploring the uptake and implementation of tele-monitored home-exercise programmes in adults with Parkinson's disease: a mixed-methods pilot study. J Telemed Telecare 2020;26(1-2):53-63. [doi: 10.1177/1357633X18794315] [Medline: 30134777]
31. Schnall R, Cho H, Mangone A, Pichon A, Jia H. Mobile health technology for improving symptom management in low income persons living with HIV. AIDS Behav 2018;22(10):3373-3383 [FREE Full text] [doi: 10.1007/s10461-017-2014-0] [Medline: 29299790]
32. Simoni JM, Ronen K, Aunon FM. Health behavior theory to enhance eHealth intervention research in HIV: rationale and review. Curr HIV/AIDS Rep 2018;15(6):423-430 [FREE Full text] [doi: 10.1007/s11904-018-0418-8] [Medline: 30511186]
33. Dandachi D, Dang BN, Lucari B, Teti M, Giordano TP. Exploring the attitude of patients with HIV about using telehealth for HIV care. AIDS Patient Care STDS 2020;34(4):166-172. [doi: 10.1089/apc.2019.0261] [Medline: 32324481]

XSL•FO
RenderX

34. O'Brien KK, Ibáñez-Carrasco F, Carusone SC, Bayoumi AM, Tang A, McDuff K, et al. Piloting an online telecoaching community-based exercise intervention with adults living with HIV: protocol for a mixed-methods implementation science study. BMJ Open 2023;13(3):e067703 [FREE Full text] [doi: 10.1136/bmjopen-2022-067703] [Medline: 36997255]
35. Lau B, Sharma I, Manku S, Kobylianski J, Wong LY, Ibáñez-Carrasco F, et al. Considerations for developing and implementing an online community-based exercise intervention with adults living with HIV: a qualitative study. BMJ Open 2022;12(4):e059294 [FREE Full text] [doi: 10.1136/bmjopen-2021-059294] [Medline: 35428647]
36. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implement Sci 2007;2:40 [FREE Full text] [doi: 10.1186/1748-5908-2-40] [Medline: 18053122]
37. Keith RE, Hopp FP, Subramanian U, Wiitala W, Lowery JC. Fidelity of implementation: development and testing of a measure. Implement Sci 2010;5:99 [FREE Full text] [doi: 10.1186/1748-5908-5-99] [Medline: 21192817]
38. Dusenbury L, Brannigan R, Falco M, Hansen WB. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. Health Educ Res 2003;18(2):237-256. [doi: 10.1093/her/18.2.237] [Medline: 12729182]
39. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 2004;23(5):443-451. [doi: 10.1037/0278-6133.23.5.443] [Medline: 15367063]
40. Breitenstein SM, Gross D, Garvey CA, Hill C, Fogg L, Resnick B. Implementation fidelity in community-based interventions. Res Nurs Health 2010;33(2):164-173 [FREE Full text] [doi: 10.1002/nur.20373] [Medline: 20198637]
41. Sprange K, Beresford-Dent J, Mountain G, Craig C, Mason C, Berry K, et al. Assessing fidelity of a community based psychosocial intervention for people with mild dementia within a large randomised controlled trial. BMC Geriatr 2021;21(1):119 [FREE Full text] [doi: 10.1186/s12877-021-02070-8] [Medline: 33573589]
42. McGee D, Lorencatto F, Matvienko-Sikar K, Toomey E. Surveying knowledge, practice and attitudes towards intervention fidelity within trials of complex healthcare interventions. Trials 2018;19(1):504 [FREE Full text] [doi: 10.1186/s13063-018-2838-6] [Medline: 30231917]
43. Poltawski L, Norris M, Dean S. Intervention fidelity: developing an experience-based model for rehabilitation research. J Rehabil Med 2014;46(7):609-615 [FREE Full text] [doi: 10.2340/16501977-1848] [Medline: 24940792]
44. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity criteria: development, measurement, and validation. Am J Eval 2003;24(3):315-340. [doi: 10.1177/109821400302400303]
45. Brown DA, O'Brien KK, Harding R, Sedgwick PM, Nelson M, Boffito M, et al. Prevalence, severity, and risk factors of disability among adults living with HIV accessing routine outpatient HIV care in London, United Kingdom (UK): a cross-sectional self-report study. PLoS One 2022;17(5):e0267271 [FREE Full text] [doi: 10.1371/journal.pone.0267271] [Medline: 35551320]
46. Rourke SB, Gardner S, Burchell AN, Raboud J, Rueda S, Bayoumi AM, et al. Cohort profile: the Ontario HIV Treatment Network Cohort Study (OCS). Int J Epidemiol 2013;42(2):402-411. [doi: 10.1093/ije/dyr230] [Medline: 22345312]
47. E-module for evidence-informed HIV rehabilitation. Realize. 2026. URL: https://realizecanada.org/wp-content/uploads/CWGHR_E-moduleEvidence-InformedHIVRehabilitationfinal-2015.pdf [accessed 2026-08-20]
48. Zoom Health Care Video Communications. Zoom. 2020. URL: https://www.zoom.com/en/industry/healthcare/ [accessed 2026-08-05]
49. Fitbit official site. Fitbit Inc. URL: http://www.fitbit.com/ca/zip [accessed 2026-08-05]
50. Su T, Chan Carusone S, McDuff K, Ibáñez-Carrasco F, Tang A, Bayoumi AM, et al. Goals in motion: exploring goal setting among adults living with HIV who participated in an online community-based exercise intervention. Front Rehabil Sci 2025;6:1644139 [FREE Full text] [doi: 10.3389/fresc.2025.1644139] [Medline: 40820990]
51. Solomon P, Chan Carusone S, Davis AM, Aubry R, O'Brien KK. A qualitative study of fitness coaches' experiences in community based exercise with people living with HIV. J Int Assoc Provid AIDS Care 2021;20:23259582211046762 [FREE Full text] [doi: 10.1177/23259582211046762] [Medline: 34668422]
52. Microsoft Excel 2016. Microsoft Corporation. 2016. URL: https://learn.microsoft.com/en-us/lifecycle/products/excel-2016 [accessed 2026-08-05]
53. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
54. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. J Public Health Dent 2011;71(s1):S52-S63 [FREE Full text] [doi: 10.1111/j.1752-7325.2011.00233.x] [Medline: 21499543]
55. Holcombe A, Wolery M, Snyder E. Effects of two levels of procedural fidelity with constant time delay onchildren's learning. J Behav Educ 1994;4(1):49-73. [doi: 10.1007/BF01560509]
56. Noell GH, Gresham FM, Gansle KA. Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. J Behav Educ) 2002;11(1):51-67. [doi: 10.1023/a:1014385321849]
57. Lacroix A, Hortobágyi T, Beurskens R, Granacher U. Effects of supervised vs. unsupervised training programs on balance and muscle strength in older adults: a systematic review and meta-analysis. Sports Med 2017;47(11):2341-2361. [doi: 10.1007/s40279-017-0747-6] [Medline: 28573401]

XSL•FO
RenderX

58. Ishee JH, Hughes M. From supervised to unsupervised exercise: factors associated with exercise adherence. J Phys Educ Recreat Dance 2004;75(6):13. [doi: 10.1080/07303084.2004.10607247]
59. Gillett PA. Self-reported factors influencing exercise adherence in overweight women. Nurs Res 1988;37(1):25-29. [doi: 10.1097/00006199-198801000-00006]
60. Ibáñez-Carrasco F, McDuff K, Da Silva G, Bayoumi AM, Chan Carusone S, Loutfy M, et al. Qualitative insights from an online community-based exercise intervention for persons living with HIV. Front Rehabil Sci 2025;6:1602007 [FREE Full text] [doi: 10.3389/fresc.2025.1602007] [Medline: 40927745]
61. Stangl AL, Earnshaw VA, Logie CH, van Brakel W, C Simbayi L, Barré I, et al. The health stigma and discrimination framework: a global, crosscutting framework to inform research, intervention development, and policy on health-related stigmas. BMC Med 2019;17(1):31 [FREE Full text] [doi: 10.1186/s12916-019-1271-3] [Medline: 30764826]
62. Steward WT, Bharat S, Ramakrishna J, Heylen E, Ekstrand ML. Stigma is associated with delays in seeking care among HIV-infected people in India. J Int Assoc Provid AIDS Care 2013;12(2):103-109 [FREE Full text] [doi: 10.1177/1545109711432315] [Medline: 22282878]
63. Hatzenbuehler ML, Phelan JC, Link BG. Stigma as a fundamental cause of population health inequalities. Am J Public Health 2013;103(5):813-821 [FREE Full text] [doi: 10.2105/AJPH.2012.301069] [Medline: 23488505]
64. Tesfay F, Javanparast S, Mwanri L, Ziersch A. Stigma and discrimination: barriers to the utilisation of a nutritional program in HIV care services in the Tigray region, Ethiopia. BMC Public Health 2020;20(1):904 [FREE Full text] [doi: 10.1186/s12889-020-09040-6] [Medline: 32522269]
65. O'Brien KK, Bayoumi AM, Strike C, Young NL, Davis AM. Exploring disability from the perspective of adults living with HIV/AIDS: development of a conceptual framework. Health Qual Life Outcomes 2008;6:76 [FREE Full text] [doi: 10.1186/1477-7525-6-76] [Medline: 18834538]
66. Tuakli-Wosornu YA, Selzer F, Losina E, Katz JN. Predictors of exercise adherence in patients with meniscal tear and osteoarthritis. Arch Phys Med Rehabil 2016;97(11):1945-1952 [FREE Full text] [doi: 10.1016/j.apmr.2016.05.011] [Medline: 27296899]
67. Rhodes RE, Martin AD, Taunton JE, Rhodes EC, Donnelly M, Elliot J. Factors associated with exercise adherence among older adults. An individual perspective. Sports Med 1999;28(6):397-411. [doi: 10.2165/00007256-199928060-00003] [Medline: 10623983]
68. Jaggers JR, Dudgeon W, Blair SN, Sui X, Burgess S, Wilcox S, et al. A home-based exercise intervention to increase physical activity among people living with HIV: study design of a randomized clinical trial. BMC Public Health 2013;13:502 [FREE Full text] [doi: 10.1186/1471-2458-13-502] [Medline: 23706094]
69. Mabweazara SZ, Ley C, Leach LL. Physical activity, social support and socio-economic status amongst persons living with HIV and AIDS: a review. Afr J AIDS Res 2018;17(2):203-212. [doi: 10.2989/16085906.2018.1475400] [Medline: 30003848]
70. Ginsburg LR, Hoben M, Easterbrook A, Anderson RA, Estabrooks CA, Norton PG. Fidelity is not easy! Challenges and guidelines for assessing fidelity in complex interventions. Trials 2021;22(1):372 [FREE Full text] [doi: 10.1186/s13063-021-05322-5] [Medline: 34051830]
71. Toomey E, Matthews J, Hurley DA. Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain. BMJ Open 2017;7(8):e015452 [FREE Full text] [doi: 10.1136/bmjopen-2016-015452] [Medline: 28780544]
72. Mucha J, Hamdy R, Marini M, Aasem R, Duong C, Su T, et al. Experiences with technology among adults aging with HIV engaged in an online community-based exercise intervention study: longitudinal qualitative descriptive study and secondary data analysis. JMIR Rehabil Assist Technol 2026;13:e86785 [FREE Full text] [doi: 10.2196/86785] [Medline: 42391560]
73. Sahel-Gozin N, Loutfy M, O'Brien KK. Exploring experiences engaging in exercise from the perspectives of women living with HIV: a qualitative study. PLoS One 2023;18(6):e0286542 [FREE Full text] [doi: 10.1371/journal.pone.0286542] [Medline: 37267270]
74. Pérez D, Van der Stuyft P, Zabala MDC, Castro M, Lefèvre P. A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. Implement Sci 2016;11(1):91 [FREE Full text] [doi: 10.1186/s13012-016-0457-8] [Medline: 27391959]
75. Hasson H, Hedberg Rundgren E, Strehlenert H, Gärdegård A, Uvhagen H, Klinga C, et al. The adaptation and fidelity tool to support social service practitioners in balancing fidelity and adaptations: longitudinal, mixed-method evaluation study. Implement Res Pract 2023;4:26334895231189198 [FREE Full text] [doi: 10.1177/26334895231189198] [Medline: 37790175]
76. Albers B, Verweij L, Blum K, Oesch S, Schultes M, Clack L, et al. Firm, yet flexible: a fidelity debate paper with two case examples. Implement Sci 2024;19(1):79 [FREE Full text] [doi: 10.1186/s13012-024-01406-3] [Medline: 39639379]

Abbreviations

BCC: National Institutes of Health Behavior Change Consortium
CBE: community-based exercise
CFIR: Consolidated Framework for Implementation Research
FOI: fidelity of implementation

XSL•FO
RenderX

OCS: Ontario HIV Treatment Network Cohort Study
PAR-Q: Physical Activity Readiness Questionnaire

Edited by K Kokorelias; submitted 26.Sep.2025; peer-reviewed by MA Shaikh, L Declerck; comments to author 26.May.2026; revised version received 06.Jul.2026; accepted 03.Aug.2026; published 03.Sep.2026.

Please cite as:
Abousifein M, Nathanson J, Jiancaro T, Tang A, Ahluwalia P, Bayoumi AM, Avery L, Da Silva G, Ibáñez-Carrasco F, Ilic I, McDuff K, Loutfy M, Furlan A, Pandovski Z, Trent H, Chan Carusone S, O'Brien KK
Fidelity of an Online Telecoaching Community-Based Exercise Intervention Among Adults Living With HIV in Toronto, Ontario, Canada: Observational Longitudinal Study
JMIR Rehabil Assist Technol 2026;13:e84902
URL: https://rehab.jmir.org/2026/1/e84902
doi:10.2196/84902
PMID:42691484

©Marfy Abousifein, Julia Nathanson, Tizneem Jiancaro, Ada Tang, Puja Ahluwalia, Ahmed M Bayoumi, Lisa Avery, George Da Silva, Francisco Ibáñez-Carrasco, Ivan Ilic, Kiera McDuff, Mona Loutfy, Annamaria Furlan, Zoran Pandovski, Helen Trent, Soo Chan Carusone, Kelly K O'Brien. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 03.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Real-World Use and Effectiveness of Internet-Based Vestibular Rehabilitation Following Nationwide Implementation: Prospective Observational Cohort Study

Hà TN Ngo[1,2], MD; Otto R Maarsingh[1,2], MD, PhD; Marco H Blanker[3], MD, PhD; Jettie Bont[1,2,4], MD, PhD; Vincent A van Vugt[1,2], MD, PhD

[1]Department of General Practice, Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, Noord-Holland, The Netherlands

[2]Amsterdam Public Health, Amsterdam, Noord-Holland, The Netherlands

[3]Department of Primary and Long-term care, Cure and Care in the Community Context (FOUR-C) research program, University Medical Center Groningen, Groningen, The Netherlands

[4]Department of General Practice, Amsterdam UMC location University of Amsterdam, Amsterdam, Noord-Holland, The Netherlands

Corresponding Author:
Hà TN Ngo, MD
Department of General Practice
Amsterdam UMC location Vrije Universiteit Amsterdam
Boelelaan 1117
Amsterdam, Noord-Holland, 1081 BT
The Netherlands
Phone: 31 637385774
Email: t.n.h.ngo@amsterdamumc.nl

Abstract

Background: Vestibular rehabilitation is a safe and effective exercise-based treatment for various vestibular disorders, yet it is underused by general practitioners. Following a randomized controlled trial demonstrating the effectiveness of Vertigo Training, an internet-based vestibular rehabilitation program, a nationwide implementation was initiated in the Netherlands.

Objective: This study aimed to evaluate the real-world use and effectiveness of Vertigo Training on vestibular symptoms.

Methods: We conducted a prospective observational cohort study by using anonymized data from Vertigo Training users (period 2023-2025). Individuals were eligible if they reported vestibular symptoms for at least 1 month, irrespective of the underlying cause. Main analyses focused on users who completed voluntary questionnaires, capturing demographic and clinical characteristics. We evaluated the effectiveness using the Vertigo Symptom Scale—Short Form (VSS-SF). Factors associated with treatment success (defined as ≥3-point reduction in VSS-SF score at 3 months) were analyzed using multivariable logistic regression. To contextualize these outcomes, we used program-level data and additional voluntary questionnaires to describe engagement and adherence.

Results: A total of 2156 users were included for the main analyses (median age 60, IQR 48-70 years; n=1427, 66% women). The most frequently self-reported diagnoses were benign paroxysmal positional vertigo and vestibular neuritis. Median VSS-SF scores improved from 17 (IQR 12-22) at baseline to 11 (IQR 6-18) at 3 months, with 63% (155/247) achieving treatment success. Higher baseline VSS-SF scores were associated with a greater likelihood of treatment success (odds ratio 1.1, 95% CI 1.0-1.1), as was a diagnosis of benign paroxysmal positional vertigo (risk ratio 1.7, 95% CI 1.1-2.0). These outcomes were observed in a context where engagement declined over time, with 43% (1442/3367) of users progressing to week 2, and adherence to the recommended exercise frequency averaged 13%.

Conclusions: In this real-world cohort, Vertigo Training was associated with meaningful reductions in vestibular symptoms following nationwide implementation. These findings confirm its effectiveness beyond the trial setting and support broader integration of internet-based vestibular rehabilitation into routine care. However, sustained engagement remains an important target for future research and optimization.

International Registered Report Identifier (IRRID): RR2-10.1186/s43058-023-00524-1

(JMIR Rehabil Assist Technol 2026;13:e106458) doi:10.2196/106458

XSL•FO
RenderX

KEYWORDS

implementation; dizziness; vertigo; general practice; primary care; mobile health; mHealth; eHealth

Introduction

Patients with vestibular symptoms, such as dizziness and vertigo, are commonly seen in general practice. Approximately 1 in 4 adults experience these symptoms in their lives, and the prevalence increases with age [1]. A significant number of patients develop chronic symptoms. For example, 30% to 40% of patients with vestibular neuritis persistently experience vestibular symptoms after 6 months [2]. This can be caused by a failure of the vestibular compensation process, which encompasses an adaptation of the nervous system to vestibular damage [3]. It is difficult to estimate the prevalence of chronic vestibular symptoms due to varying definitions, but cross-sectional studies on the epidemiology report a prevalence of up to 4.8% in the adult population [4,5].

Vestibular symptoms are associated with a high physical, psychosocial, and economic burden. Anxiety and depression are common among patients with vestibular symptoms, as well as increased health care use and work absenteeism [4,6-8]. Adequate treatment is therefore essential, especially when these symptoms are chronic. Vestibular rehabilitation (VR), an exercise-based treatment that aims to stimulate vestibular compensation, has proven to be safe and effective for a variety of vestibular disorders, including chronic symptoms [9-11]. The majority of patients with vestibular symptoms are treated in primary care [12-15], but general practitioners (GPs) seem to lack knowledge about this treatment. Surveys among Dutch and British GPs report that only 6% to 7% of GPs apply this treatment [16,17]. This gap between evidence and practice highlights the need for scalable strategies to improve access to VR [18].

To address this, we developed Vertigo Training, an internet-based VR program designed for use in general practice [19]. In a randomized controlled trial (RCT), Vertigo Training was shown to be safe and effective in reducing vestibular symptoms, dizziness-related impairment, and anxiety for up to 36 months of follow-up [20,21]. However, evidence from controlled trials does not necessarily translate to effectiveness in routine care, where patient populations and context differ [22]. To bridge this gap, Vertigo Training was made available nationwide in the Netherlands, offering free and easily accessible VR to patients with chronic vestibular symptoms [23].

The objective of this study is to evaluate the real-world use and effectiveness of Vertigo Training in a nationwide user cohort. These findings will inform further optimization of Vertigo Training and support strategies to enhance its uptake and sustained use in general practice.

Methods

Study Design, Setting, and Participants

We conducted a prospective observational cohort study, following the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement (Multimedia Appendix 1).

Vertigo Training was made available nationwide in the Netherlands in April 2023. For this study, we collected anonymized data for 2 years (April 2023 to April 2025) from Vertigo Training users who gave informed consent for the collection and analysis of their data. Users who did not give informed consent could still follow the treatment. Users were eligible for Vertigo Training if they were 18 years or older, had reported vestibular symptoms for at least 1 month on most days of the week (irrespective of the underlying cause), and experienced symptoms provoked or exacerbated by head movements or visual stimuli. Upon registration, all users were required to complete a questionnaire to assess their eligibility. Users who did not meet these criteria were not granted access to the intervention. Apart from these eligibility criteria, no other exclusion criteria were applied. More information on Vertigo Training can be found in Textbox 1.

Our study protocol contains specifics on the implementation and the multicomponent implementation strategy [23]. We used the RE-AIM (reach, effectiveness, adoption, implementation, and maintenance) framework, a widely used framework for planning and evaluating implementation, to guide the evaluation [24]. More detailed findings and reflections on the nationwide implementation process, including implementation outcomes and lessons learned, will be reported elsewhere.

Textbox 1. Information on Vertigo Training.

- Vertigo Training is a freely accessible online intervention that encompasses 6 vestibular rehabilitation exercises. These exercises include eye, head, and body movements that aim to stimulate vestibular recovery through 3 main principles: adaptation (using head movements to improve vestibulo-ocular reflex gain), habituation (using repeated provocative eye, head, and/or body movements to diminish responsiveness), and substitution (using alternative sensory inputs to substitute missing vestibular functions). Additionally, the intervention incorporates a form of exposure-based behavioral therapy by provoking vestibular symptoms in a controlled setting.
- The intervention prescribes a personal exercise program for a period of 6 weeks, based on the severity of symptoms while performing the exercises. Users are asked to perform the exercises twice a day for 10 minutes (total duration 20 minutes).
- Users may register for Vertigo Training on their own initiative or following a referral by a health care provider. They can perform the exercises independently or under the guidance of a health care provider, such as a physiotherapist. In the Netherlands, physiotherapy is accessible without a referral.

XSL•FO
RenderX

Ethical Considerations

The medical ethics committee of the VU University Medical Center approved the study protocol (2022.0569).

Data Collection and Outcome Measures

We collected baseline and follow-up data through online voluntary questionnaires that were integrated into the intervention. Main analyses focused on users who completed these questionnaires. Baseline data included user demographics; self-reported (vestibular) diagnoses and treatments; the source through which users became aware of Vertigo Training; and severity of vestibular symptoms according to the Vertigo Symptom Scale—Short Form (VSS-SF), a validated self-report instrument [25-27]. Follow-up VSS-SF measurements were obtained at 6 weeks, 3 months, and 6 months to evaluate changes in symptom severity over time. We chose the VSS-SF because this measure was used as the primary outcome in the RCT investigating the safety and effectiveness of Vertigo Training, allowing us to compare trial outcomes with the real-world outcomes of this study.

Outcomes were user characteristics, the effectiveness of Vertigo Training, and factors associated with treatment success. The effectiveness of Vertigo Training was assessed as the proportion of users with treatment success, which was defined as a clinically relevant improvement in vestibular symptoms (≥3-point reduction in VSS-SF score between baseline and 3 months) [28-31].

To contextualize these outcomes, we used data on program use to describe engagement (number of users over time) and adherence (exercise frequency). Exercise frequency was obtained weekly from week 2 through week 6 using a voluntary questionnaire administered to all users. Table S1 in Multimedia Appendix 2 [32-34] shows an overview of variables, including their definitions.

Statistical Analyses

We used descriptive statistics to assess user characteristics, program engagement and adherence, and the effectiveness of Vertigo Training. To further describe adherence, we compared exercise frequency between users who completed the baseline questionnaire and those who did not, as well as between users who completed all questionnaires (ie, baseline and follow-up at 6 weeks, 3 months, and 6 months) and those who did not. Furthermore, we described the exercise frequency of users who completed questionnaires at both baseline and 3 months. A Wilcoxon signed-rank test was performed to evaluate whether there was a significant difference in VSS-SF score between baseline and follow-up measurements. Finally, we performed univariable and multivariable logistic regression analyses to evaluate whether certain characteristics were associated with treatment success. Because the relatively favorable prognosis of benign paroxysmal positional vertigo (BPPV) could have influenced our findings, we performed an additional sensitivity analysis excluding these patients. Odds ratios (ORs) and 95% CIs were obtained from the regression analyses. For categorical variables, ORs were converted to risk ratios (RRs), including their 95% CIs. All variables with a P value <.05 were considered statistically significantly associated with treatment success. We used SPSS (version 28.0; IBM Corp) for the analyses.

Results

User Characteristics

During the 2-year observation period, a total of 2156 users completed the baseline questionnaire and were included in the cohort. The median age of users was 60 (IQR 48-70) years, and 66% (n=1427) were women. Most users (n=1489, 69%) lived in moderately to extremely urbanized areas, 57% (n=1223) had a high educational level, and 9% (n=192) reported a migration background. The most common self-reported diagnoses were BPPV, vestibular neuritis, persistent postural perceptive dizziness, and Ménière disease. Users also mentioned other vestibular and nonvestibular disorders, such as vestibular migraine, Mal de Débarquement syndrome, postconcussion syndrome, cerebrovascular accidents, and vestibular schwannomas. Concurrently, 47% (n=1025) of users reported having no diagnosis.

More than half of the users (n=1209, 56%) received treatment for their symptoms from a health care professional, mainly a physiotherapist and, to a lesser extent, a GP or medical specialist. Antivertigo drugs were used by 19% (n=416) of users. Most users were referred to Vertigo Training by their GP or found Vertigo Training online (eg, through search engines or social media). Table 1 shows all baseline user characteristics.

XSL•FO
RenderX

Table 1. Baseline characteristics of Vertigo Training users (N=2156).

Characteristics	Users
Demographic characteristics	
Age (y), median (IQR)	60 (48-70)
Gender, n (%)	
Women	1427 (66)
Men	718 (33)
Other	11 (1)
Degree of urbanization[a], n (%)	
Extremely urbanized	550 (26)
Strongly urbanized	585 (27)
Moderately urbanized	354 (16)
Hardly urbanized	377 (18)
Not urbanized	108 (5)
Missing	182 (8)
Education level[b], n (%)	
Low	897 (42)
High	1223 (57)
Missing	36 (1)
Migration background, n (%)	
Yes	192 (9)
No	1928 (89)
Missing	36 (2)
Clinical characteristics	
Received diagnosis, n (%)	
Yes	1120 (52)
No	1025 (47)
Missing	11 (1)
Self-reported diagnosis[c], n	
BPPV[d]	371
Vestibular neuritis	210
PPPD[e]	182
Ménière disease	146
Other or unknown	340
Received medical care, n (%)	
Yes	1209 (56)
No	931 (43)
Missing	16 (1)
Which health care provider[c] , n	
Physiotherapist	704
Medical specialist	595
General practitioner	574
Other	83

XSL•FO
RenderX

Characteristics	Users
Antivertigo drug use, n (%)	
Yes	416 (19)
No	1729 (80)
Missing	11 (1)
Awareness of Vertigo Training	
Source[c], n	
The internet	705
General practitioner	530
Media	337
Medical specialist	266
Friends or acquaintances	260
Patient association	132
Other	90

[a]On the basis of the municipality of residence: not or hardly urbanized (<1000 surrounding addresses), moderately urbanized (1000-1500 surrounding addresses), strongly urbanized (1500-2500 surrounding addresses), and extremely urbanized (>2500 surrounding addresses).

[b]Low=primary education, prevocational secondary education, senior general secondary education, preuniversity education, or senior secondary vocational education. High=higher vocational education or university education.

[c]Multiple answers possible.

[d]BPPV: benign paroxysmal positional vertigo.

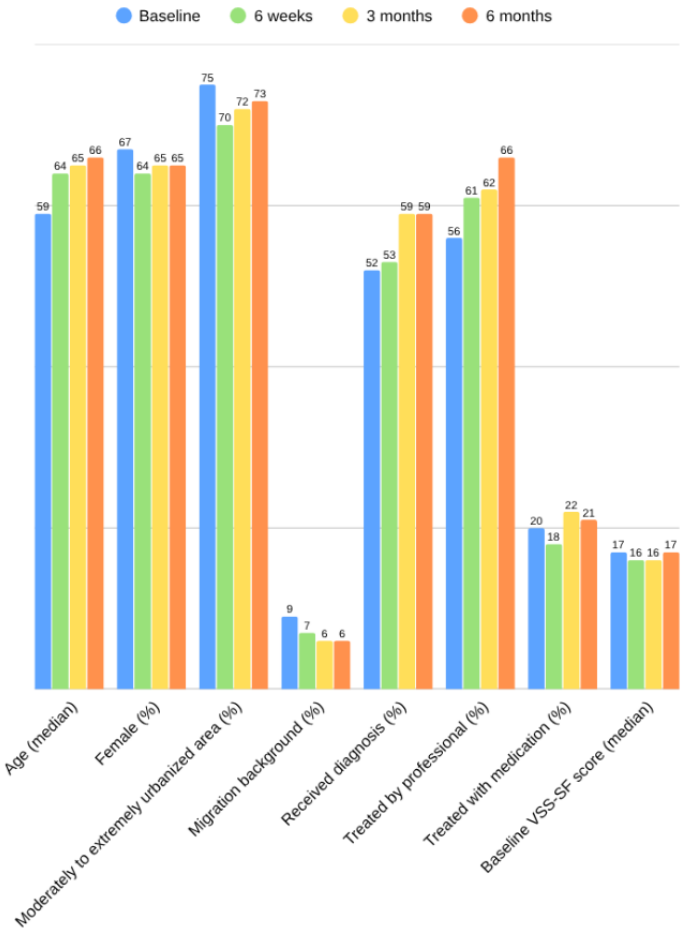
[e]PPPD: persistent postural perceptive dizziness.

Effectiveness

A total of 1905 users completed the questionnaire on the VSS-SF score at baseline. Follow-up data were available for 382 users (20%) at 6 weeks, 275 (14%) at 3 months, and 199 (10%) at 6 months; 88 users (5%) completed all measurements. Figure 1 depicts baseline user characteristics per VSS-SF measurement. The median VSS-SF score was 17 (IQR 12-22) at baseline, 13 (IQR 7-18) at 6 weeks, 11 (IQR 6-18) at 3 months, and 12 (IQR 7-17) at 6 months. Of 247 users with both baseline and 3-month data, 155 (63%) were treated successfully (≥3-point reduction). All follow-up VSS-SF measurements were significantly lower than baseline (P<.001).

XSL•FO
RenderX

Figure 1. Baseline characteristics of users who completed the Vertigo Symptom Scale–Short Form (VSS-SF) at each measurement point. For each characteristic, bars show the values at baseline, 6 weeks, 3 months, and 6 months. Continuous variables are shown as median and categorical variables as proportions.



Factors Associated With Treatment Success

The odds of treatment success increased by 10% with each point higher in VSS-SF score at baseline (OR 1.1, 95% CI 1.0-1.1). A self-reported diagnosis of BPPV also increased the likelihood of treatment success (RR 1.7, 95% CI 1.1-2.0). Table 2 shows the results from the univariable and multivariable regression analyses.

A sensitivity analysis, excluding users with BPPV as a self-reported diagnosis (n=55), yielded a similar proportion of treatment success (119/192, 62%) and similar regression results (Table S2 in Multimedia Appendix 2).

XSL•FO
RenderX

Table 2. Results of univariable and multivariable logistic regression analyses assessing factors associated with treatment success.

Characteristics	Successful treatment (n=155)	No successful treatment (n=92)	Univariable analysis, OR[a] (95% CI)	Univariable analysis, RR[b] (95% CI)	Multivariable analysis, OR[a] (95% CI)	Multivariable analysis, RR[b] (95% CI)
Demographic characteristics						
Age (y), median (IQR)	63 (53-70)	66 (56-72)	1.0 (1.0-1.0)	—[c]	1.0 (1.0-1.0)	—[c]
Gender, %						
Men	34	33	1.1 (0.6-1.8)	1.0 (0.8-1.2)	1.5 (0.8-2.9)	1.3 (0.9-1.5)
Education level [d], %						
High	60	60	1.1 (0.6-1.8)	1.0 (0.8-1.2)	1.0 (0.5-1.9)	1.0 (0.7-1.3)
Migration background, %						
Yes	7		1.0 (0.4-2.9)	1.0 (0.5-1.3)	0.9 (0.2-3.3)	0.9 (0.4-1.8)
	7					
Clinical characteristics						
Baseline VSS-SF[e] score, median (IQR)	17 (13-23)	15 (9-19)	1.1 (1.0-1.1)[f]	—[c]	1.1 (1.0-1.1)[f]	—[c]
Received diagnosis, %						
Yes	57	65	0.8 (0.5-1.3)	0.9 (0.7-1.1)	0.7 (0.3-1.5)	0.9 (0.6-1.1)
Self-reported diagnosis, %						
BPPV[g]	23	21	1.2 (0.6-2.3)	1.1 (0.8-1.3)	2.7 (1.2-6.3)[f]	1.7 (1.1-2.0)[f]
Vestibular neuritis	12	11	1.1 (0.5-2.4)	1.0 (0.7-1.3)	0.9 (0.3-2.5)	1.0 (0.5-1.3)
PPPD[h]	9	9	1.0 (0.4-2.6)	1.0 (0.6-1.3)	0.8 (0.3-2.3)	0.9 (0.5-1.3)
Ménière disease	8	11	0.7 (0.3-1.7)	0.9 (0.5-1.2)	0.7 (0.2-2.2)	0.9 (0.4-1.3)
Received medical care, %						
Yes	59	70	0.7 (0.4-1.1)	0.9 (0.7-1.0)	0.4 (0.2-1.1)	0.7 (0.4-1.0)
Which health care provider, %						
Physiotherapist	34	41	0.7 (0.4-1.3)	0.9 (0.7-1.1)	0.9 (0.4-2.0)	1.0 (0.7-1.2)
General practitioner	28	27	1.1 (0.6-1.9)	1.0 (0.8-1.2)	1.2 (0.6-2.5)	1.1 (0.8-1.3)
Medical specialist hospital	29	33	0.8 (0.5-1.5)	0.9 (0.7-1.1)	0.8 (0.4-1.8)	0.9 (0.6-1.2)
Medical specialist elsewhere	6	7	0.9 (0.3-2.6)	1.0 (0.5-1.3)	0.7 (0.2-2.5)	0.9 (0.4-1.3)
Antivertigo drug use, %						
Yes	23	25	0.9 (0.5-1.6)	1.0 (0.7-1.2)	1.3 (0.5-3.0)	1.1 (0.8-1.3)
Other characteristics						
Exercise frequency, %						
At least once daily	69	58	1.6 (0.9-2.7)	1.2 (1.0-1.4)	1.7 (0.9-3.3)	1.2 (1.0-1.4)

[a]OR: odds ratio.

[b]RR: risk ratio.

[c]Odds ratio not converted to risk ratio, because of continuous variable.

[d]Low=primary education, prevocational secondary education, senior general secondary education, preuniversity education, or senior secondary vocational education. High=higher vocational education or university education.

[e]VSS-SF: Vertigo Symptom Scale—Short Form.

[f]Significant association (P<.05).

XSL•FO
RenderX

[g]BPPV: benign paroxysmal positional vertigo.

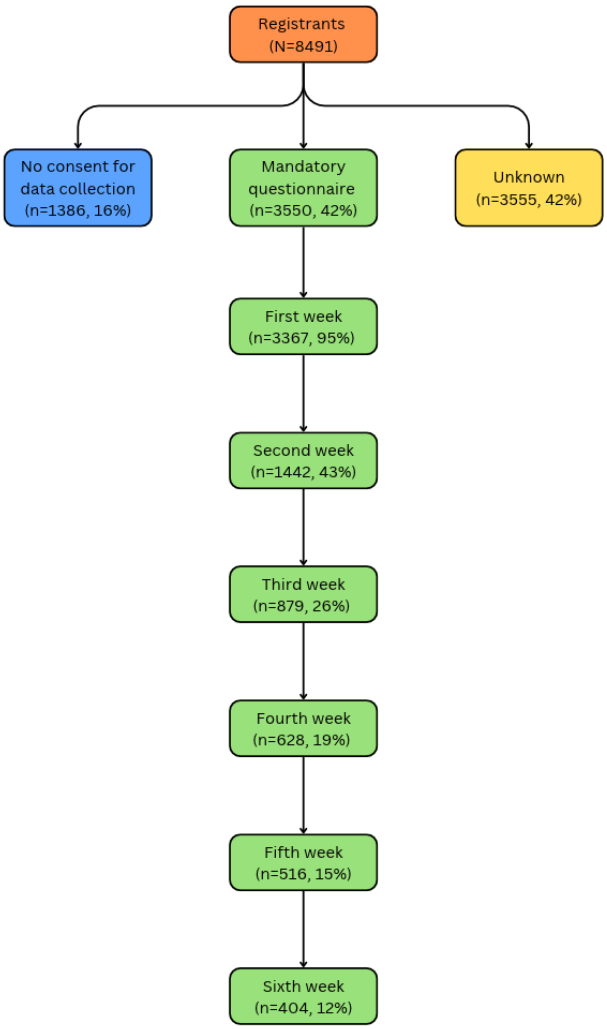
[h]PPPD: persistent postural perceptive dizziness.

Program Engagement and Adherence

During the 2-year observation period (2023-2025), a total of 8491 persons registered for Vertigo Training, of whom 3550 (42%) completed the mandatory questionnaire to assess their eligibility. Of these users, 3367 (95%) entered the first week of the program. Additionally, of 8491 persons who registered for Vertigo Training, 1386 (16%) gave no consent for data collection and analysis. This leaves 3555 (42%) registrants who did not engage with Vertigo Training. Engagement declined over time, with 43% (1442/3367) of users progressing to the second week and 12% (404/3367) of users completing the final week. Reasons for not engaging or discontinuation are unknown. Figure 2 illustrates user engagement from registration to week 6.

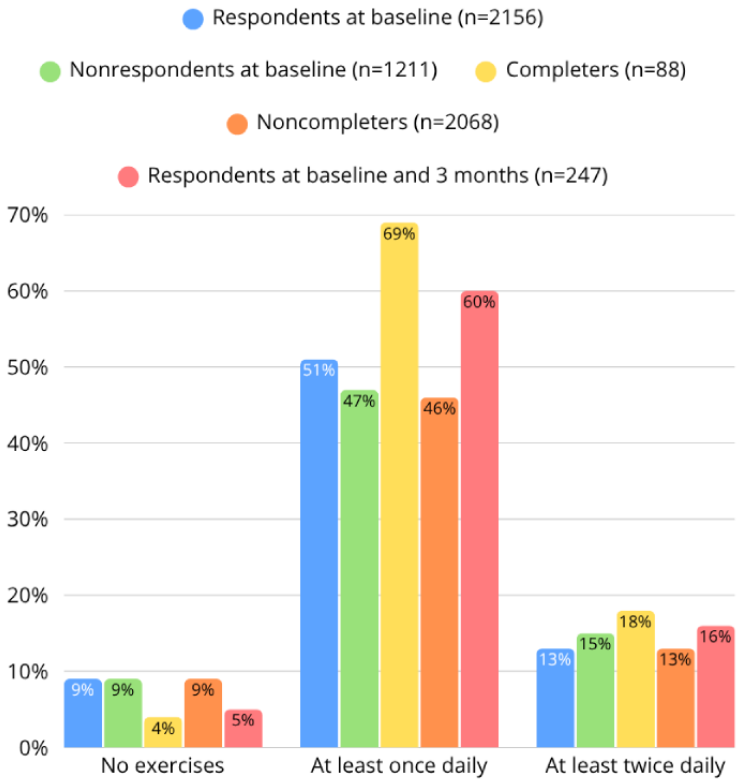
At the beginning of every new week of the program, all users were asked to report their exercise frequency. On average, 51% of users performed exercises at least once daily, 13% adhered to the recommended frequency of ≥14 times per week, and 9% did not perform any exercises. Among 2156 users who completed the baseline questionnaire, an average of 13% adhered to the recommended frequency, compared with 15% of users in the group who did not. Of the 88 users who completed all voluntary questionnaires (ie, baseline and follow-up at 6 weeks, 3 months, and 6 months), an average of 18% performed the exercises according to the recommended frequency, compared with 13% of users who completed the baseline questionnaire but not all follow-up questionnaires. Finally, 16% of users who completed the questionnaire at both baseline and 3 months adhered to the recommended frequency. Figure 3 shows adherence among respondents to the questionnaires compared with nonrespondents.

Figure 2. Flowchart illustrating user engagement from registration to week 6.



XSL•FO
RenderX

Figure 3. Adherence among respondents and nonrespondents to the baseline questionnaire, users who completed all questionnaires (completers), users who completed the baseline questionnaire but not all follow-up questionnaires (noncompleters), and users who completed the questionnaire at both baseline and 3 months.



Discussion

Principal Findings

In this real-world cohort of 2156 Vertigo Training users, we found clinically relevant improvement of vestibular symptoms, with 63% (155/247) of users achieving treatment success at 3 months. Median VSS-SF scores decreased from 17 (IQR 12-22) to 11 (IQR 6-18), indicating substantial symptom relief in routine care. Higher baseline symptom severity and a diagnosis of BPPV were associated with a greater likelihood of treatment success. These outcomes were observed in a context where engagement declined over time, with 43% (1442/3367) of users progressing to week 2, and adherence to the recommended exercise frequency averaged 13%. Our findings demonstrate that the effectiveness of Vertigo Training, previously established in an RCT, is maintained in a real-world setting following nationwide implementation. However, sustained engagement of digital interventions remains a challenge.

Strengths and Limitations

This study had several strengths and limitations. A key strength was the prospective evaluation of a large, real-world user cohort following nationwide implementation. This design enhances the generalizability of our findings and provides insight into how an evidence-based digital intervention performs outside a controlled trial setting. Additionally, the 2-year observation period led to a substantial number of users, which increased the accuracy of our evaluation of the effectiveness and factors associated with treatment success.

Several limitations should be considered. First, data collection relied mostly on voluntary questionnaires, introducing the possibility of selection and nonresponse bias. Respondents may have been more engaged or motivated, potentially leading to an overestimation of the effect of Vertigo Training. We addressed this potential bias in 2 ways. First, we compared user characteristics per VSS-SF measurement. Respondents to the baseline VSS-SF questionnaire were overall comparable to respondents to the follow-up measurements. At 3-month follow-up, respondents were slightly older (median age 65 [IQR 54-72] vs 59 years [IQR 47-69] at baseline), and a larger proportion was diagnosed (161/272, 59% vs 990/1900, 52%) and treated (171/274, 62% vs 1063/1895, 56%). However, research, including this study, has shown that age does not necessarily influence VR outcomes [35]. We expect that having a diagnosis will not significantly impact outcomes either, as VR is known to be effective for chronic vestibular symptoms irrespective of the original cause [11], which is the target population for Vertigo Training. Finally, prior treatment by a health care professional is more likely to be associated with worse VR outcomes, as patients may be less responsive to treatment. This would lead to an underestimation of the effect, rather than an overestimation. Second, we compared adherence between respondents and nonrespondents to the questionnaires. Users who completed all questionnaires performed the exercises more frequently, suggesting that more engaged users are overrepresented. However, adherence did not differ between respondents and nonrespondents to the baseline questionnaire, and the difference in exercise frequency was less pronounced in users who completed questionnaires at both baseline and 3

XSL•FO
RenderX

months. This suggests a limited impact of bias on our effect estimates.

Two other aspects were possibly affected by the voluntary nature of the questionnaires. First, we deliberately kept the questionnaires brief to limit the burden for users, but this restricted the depth of the information obtained. For example, antivertigo drug use was only assessed as "ever used," and it therefore remained unclear whether users were still treated with medication during the program. Second, almost all data, including exercise frequency and VSS-SF score, were self-reported. Exercise frequency may not accurately reflect actual exercise behavior due to overreporting. We collected program-level data as an objective measure, but this reflects progression through the program, rather than exercise performance: users may quickly learn the exercises and continue independently of the program. To our knowledge, there is no gold standard for the measurement of adherence, and future studies are needed on this topic. In our study, self-report was the most feasible method, as users should be able to follow Vertigo Training on their own initiative and independently at home. The VSS-SF is a validated instrument but also relies on self-report. Objective vestibular assessment was not feasible in this study and is not part of routine general practice care. Nevertheless, this limits the interpretation of the results from our analyses on the effectiveness.

Another limitation was the heterogeneity of our cohort: self-reported diagnoses included both vestibular and nonvestibular disorders, and 47% (1025/2156) of users reported no diagnosis at all. This composition reflects routine general practice, in which a substantial proportion receives no diagnosis [36] and, consequently, no appropriate treatment. Vertigo Training aims to address this gap. However, the heterogeneity prevented us from drawing conclusions on the diagnosis-specific effectiveness of Vertigo Training. A future study with a larger number of users with complete follow-up data would allow us to evaluate this reliably.

Our considerable loss to follow-up was also a possible limitation. This is common in digital health studies [37] and observational studies in general but may affect the precision of effect estimates. Despite this limitation, a Cochrane review from 2024 reported that effect estimates of observational studies differ little to not at all from those of RCTs [38].

The final limitation is that our design did not allow evaluation of users who registered but did not consent to data collection or did not engage with the program, limiting insight into barriers to uptake and adherence.

Comparison With Prior Work

The observed effect of Vertigo Training in this study is consistent with the effect reported in the earlier RCT from 2019. In this study, the median VSS-SF score decreased from 17 (IQR 12-22) at baseline to 11 (IQR 6-18) at 3 months. The RCT showed a similar reduction: in the group that only received Vertigo Training as an intervention, the mean baseline VSS-SF score was 14 and decreased to 8 at 3 months [20]. In both studies, the effect remained at 6 months (median VSS-SF score 12 in this study compared with 8 in the RCT). This suggests that treatment effects are preserved when moving from a controlled research setting to everyday practice.

In this study, 63% (155/247) of users who completed both baseline and 3-month VSS-SF measurements achieved treatment success. To our knowledge, only a few other studies have specifically reported the proportion of patients with clinical improvement, as measured by the VSS-SF. A secondary analysis of the RCT assessing Vertigo Training reported that 60% of participants receiving Vertigo Training, either as stand-alone treatment or in combination with physiotherapy support, showed improvement in VSS-SF score at 6 months [39]. A similar British RCT on internet-based VR for chronic vestibular symptoms reported that 62% of patients felt better at 3 months [31]. These proportions are similar to our findings.

Regarding factors for treatment success, we found that a higher baseline VSS-SF score and a self-reported diagnosis of BPPV were associated with a higher likelihood of treatment success. A prior study attempted to develop a prediction model for treatment success of Vertigo Training, but all models demonstrated poor discriminative ability [39]. No other studies have evaluated factors associated with treatment success of internet-based VR, but research has shown that many factors can influence the outcome of VR in general [35]. For example, patients with unilateral vestibular disorders respond well to VR, as reported by a Cochrane review in 2015 [9]. For patients with BPPV specifically, canalith repositioning maneuvers are first-line treatment, as these are more effective than VR in the short term. However, combining these maneuvers with VR is likely to improve longer-term recovery [9]. Vertigo Training users with BPPV represent a specific subgroup, that is, those with persistent symptoms, who may not have responded sufficiently to repositioning maneuvers. This could explain why a BPPV diagnosis was still associated with a higher likelihood of treatment success.

Finally, we observed that 13% of users adhered to the recommended exercise frequency of twice daily and that program engagement declined over time. The updated clinical practice guideline from the Academy of Neurologic Physical Therapy of the American Physical Therapy Association recommends gaze stabilization exercises 3 to 5 times per day, for a total of at least 20 minutes for patients with chronic unilateral and bilateral vestibular hypofunction [40]. Static and dynamic balance exercises are recommended for a minimum of 20 minutes daily. However, these recommendations are graded as weak, and evidence is based on extrapolation from available literature rather than direct evidence. In addition, most underlying studies were performed in the context of supervised VR, whereas Vertigo Training is primarily intended as an unsupervised program. The recommended exercise frequency in the guideline may therefore not be realistic for Vertigo Training users. Most users in our study performed the exercises once daily. Still, 63% (155/247) of users achieved treatment success. This suggests that a lower but sustainable exercise frequency may be sufficient in a real-world setting. It also highlights the need for more research on the minimal effective exercise dosage. The declining engagement in this study—43% (1442/3367) of users progressed to the second week, after which the number of users decreased with each week— is in line with

XSL•FO
RenderX

studies on digital interventions for other chronic conditions. Reasons for not engaging or discontinuation are unknown, but prior research has reported the following reasons that may apply to our study as well: unrelated health issues, finding the exercises too provocative, difficulty of the exercises, family or work conflicts, litigation, lack of time, loss of interest or motivation, or feeling better [40]. It should be noted that discontinuation of the Vertigo Training program does not necessarily imply discontinuation of the exercises, as users may continue doing them independently of the program. The majority of Vertigo Training content is given in the first week; this week provides information on vestibular symptoms, the rationale for VR, and both video and written instructions for the exercises [20,41]. However, our data do not allow us to determine whether users continued exercising after discontinuing the program or how early disengagement affected treatment benefit. Multiple studies report similar dropout or retention rates, including a systematic review and meta-analysis from 2020. This study, evaluating dropout in app-based interventions for chronic diseases, reported a pooled dropout rate of 49% for observational studies, compared with 42% (3555/8491) in our study [37]. A study from 2023 assessed a self-management application for individuals with type 2 diabetes and reported that 50% of users discontinued after registration, but the reasons are unknown [42]. Another study from 2022, which focused on the real-world implementation of a self-management application for bipolar disorder, also found decreasing retention rates over time [43]. They reported the following reasons for discontinuation: repetitiveness, technical issues, not finding it useful, and having an undesired daily reminder of their condition. Furthermore, a study from 2020 analyzed the reach, use, and retention of a mindfulness training application [44]. One week after downloading the application, 43% of users engaged with it, and this proportion decreased over time. The aforementioned studies, including this study, show that sustained engagement remains an important issue that should be addressed in implementation research.

Implications for Research and/or Practice

Our findings support the use of internet-based VR as an effective and scalable treatment option in general practice. The demonstrated effectiveness in a heterogeneous, real-world population strengthens the case for broader integration of Vertigo Training into primary care pathways.

At the same time, improving engagement of users should be a priority. Future research should focus on users who discontinued the app and individuals who are eligible for the app but do not use it. This includes exploring patient preferences, usability, and contextual factors influencing adherence. These findings would give more insight into possible targets for improvement of the implementation. Ongoing work will further evaluate the impact of Vertigo Training on general practice care, including its role in reducing unnecessary medication use and supporting more efficient care pathways in general practice.

Conclusions

In this real-world cohort study, Vertigo Training was associated with meaningful reductions in vestibular symptoms, confirming its effectiveness beyond the trial setting. Effects were most pronounced in users with higher baseline symptom severity and those with BPPV. However, sustained engagement remains an important target for future research and optimization.

Acknowledgments

The authors would like to thank Curavista for their continued effort and aid in making Vertigo Training available to patients with vestibular symptoms. The authors would also like to thank all Vertigo Training users for their contributions to this research.

Funding

This work was supported by ZonMw (grant 08391052110005).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Generative AI Use

The authors declare that generative AI was used in the creation of this manuscript. Generative AI (ChatGPT, Claude) was used for proofreading, editing, or language polishing. The authors verified all outputs of the tool(s) used in the preparation of this manuscript and take full responsibility for the final content of the manuscript" should be changed into "The authors declare that generative AI was used in the creation of this manuscript. Generative AI (ChatGPT GPT-6, Claude Sonnet 5.5) was used for proofreading, editing, or language polishing. The authors verified all outputs of the tool(s) used in the preparation of this manuscript and take full responsibility for the final content of the manuscript.

Authors' Contributions

Conceptualization: VAvV (lead), HTNN (equal), ORM, (supporting), MHB (supporting), JB (supporting)
Data curation: HTNN
Formal analysis: HTNN
Funding acquisition: VAvV

XSL•FO
RenderX

Investigation: HTNN
Methodology: VAvV (lead), HTNN (equal), ORM (supporting), MHB (supporting), JB (supporting)
Project administration: VAvV (lead), HTNN (supporting)
Supervision: VAvV
Validation: VAvV
Visualization: HTNN
Writing—original draft: HTNN
Writing—review & editing: HTNN (lead), ORM, (supporting), MHB (supporting), JB (supporting), VAvV (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE statement.
[DOCX File , 41 KB - rehab_v13i1e106458_app1.docx]

Multimedia Appendix 2
An overview of the included variables and results from the sensitivity analysis.
[DOCX File , 32 KB - rehab_v13i1e106458_app2.docx]

References

1. Neuhauser HK. The epidemiology of dizziness and vertigo. Handb Clin Neurol 2016;137:67-82. [doi: 10.1016/B978-0-444-63437-5.00005-4] [Medline: 27638063]
2. Neuhauser HK, Lempert T. Vertigo: epidemiologic aspects. Semin Neurol 2009 Nov;29(5):473-481. [doi: 10.1055/s-0029-1241043] [Medline: 19834858]
3. Lacour M, Helmchen C, Vidal PP. Vestibular compensation: the neuro-otologist's best friend. J Neurol 2016 Apr;263 Suppl 1:S54-S64 [FREE Full text] [doi: 10.1007/s00415-015-7903-4] [Medline: 27083885]
4. Neuhauser HK, Radtke A, von Brevern M, Lezius F, Feldmann M, Lempert T. Burden of dizziness and vertigo in the community. Arch Intern Med 2008 Oct 27;168(19):2118-2124. [doi: 10.1001/archinte.168.19.2118] [Medline: 18955641]
5. Kim EJ, Song HJ, Lee HI, Kwon E, Jeong SH. One-year prevalence and clinical characteristics in chronic dizziness: the 2019–2020 Korean National Health and Nutrition Examination Survey. Front Neurol 2022 Dec 01;13:1016718 [FREE Full text] [doi: 10.3389/fneur.2022.1016718] [Medline: 36530637]
6. Kovacs E, Wang X, Grill E. Economic burden of vertigo: a systematic review. Health Econ Rev 2019 Dec 27;9(1):37 [FREE Full text] [doi: 10.1186/s13561-019-0258-2] [Medline: 31883042]
7. Ciorba A, Bianchini C, Scanelli G, Pala M, Zurlo A, Aimoni C. The impact of dizziness on quality-of-life in the elderly. Eur Arch Otorhinolaryngol 2017 Mar;274(3):1245-1250. [doi: 10.1007/s00405-016-4222-z] [Medline: 27450383]
8. Kim CH, McCray LR, Nguyen SA, Staab JP, Jafri S, Rizk H. Anxiety and depression in adults with vestibular disorders: a systematic review and meta-analysis. Laryngoscope 2026 Feb;136(2):535-546. [doi: 10.1002/lary.70055] [Medline: 40815545]
9. McDonnell MN, Hillier SL. Vestibular rehabilitation for unilateral peripheral vestibular dysfunction. Cochrane Database Syst Rev 2015 Jan 13;1(1):CD005397. [doi: 10.1002/14651858.CD005397.pub4] [Medline: 25581507]
10. Porciuncula F, Johnson CC, Glickman LB. The effect of vestibular rehabilitation on adults with bilateral vestibular hypofunction: a systematic review. J Vestib Res 2012;22(5-6):283-298. [doi: 10.3233/VES-120464] [Medline: 23302709]
11. Kundakci B, Sultana A, Taylor AJ, Alshehri MA. The effectiveness of exercise-based vestibular rehabilitation in adult patients with chronic dizziness: a systematic review. F1000Res 2018 Mar 05;7:276 [FREE Full text] [doi: 10.12688/f1000research.14089.1] [Medline: 29862019]
12. Grill E, Penger M, Kentala E. Health care utilization, prognosis and outcomes of vestibular disease in primary care settings: systematic review. J Neurol 2016 Apr;263 Suppl 1:S36-S44 [FREE Full text] [doi: 10.1007/s00415-015-7913-2] [Medline: 27083883]
13. Murphy C, Reinhardt C, Linehan D, Katiri R, O'Connor A. A review of primary care referrals for patients with dizziness and vertigo: prevalence and demographics. Ir J Med Sci 2022 Feb;191(1):385-389. [doi: 10.1007/s11845-021-02575-6] [Medline: 33675015]
14. Sloane PD, Dallara J, Roach C, Bailey KE, Mitchell M, McNutt R. Management of dizziness in primary care. J Am Board Fam Pract 1994;7(1):1-8. [Medline: 8135132]
15. Stam H, Harting T, Sluijs MV, van Marum R, Horst HV, Wouden JC, et al. Usual care and management of fall risk increasing drugs in older dizzy patients in Dutch general practice. Scand J Prim Health Care 2016 Jun;34(2):165-171 [FREE Full text] [doi: 10.3109/02813432.2016.1160634] [Medline: 27049170]

XSL•FO
RenderX

16. van Vugt VA, Diaz Nerio PM, van der Wouden JC, van der Horst HE, Maarsingh O. Use of canalith repositioning manoeuvres and vestibular rehabilitation: a GP survey. Scand J Prim Health Care 2017 Mar;35(1):19-26 [FREE Full text] [doi: 10.1080/02813432.2017.1288683] [Medline: 28277043]
17. Jayarajan V, Rajenderkumar D. A survey of dizziness management in general practice. J Laryngol Otol 2003 Aug;117(8):599-604. [doi: 10.1258/002221503768199915] [Medline: 12956912]
18. van Vugt VA, van der Horst HE, Payne RA, Maarsingh OR. Chronic vertigo: treat with exercise, not drugs. BMJ 2017 Aug 23;358:j3727. [doi: 10.1136/bmj.j3727] [Medline: 28835412]
19. van Vugt VA, van der Wouden JC, Bosmans JE, Smalbrugge M, van Diest W, Essery R, et al. Guided and unguided internet-based vestibular rehabilitation versus usual care for dizzy adults of 50 years and older: a protocol for a three-armed randomised trial. BMJ Open 2017 Jan 20;7(1):e015479 [FREE Full text] [doi: 10.1136/bmjopen-2016-015479] [Medline: 28110290]
20. van Vugt VA, van der Wouden JC, Essery R, Yardley L, Twisk JW, van der Horst HE, et al. Internet based vestibular rehabilitation with and without physiotherapy support for adults aged 50 and older with a chronic vestibular syndrome in general practice: three armed randomised controlled trial. BMJ 2019 Nov 05;367:l5922 [FREE Full text] [doi: 10.1136/bmj.l5922] [Medline: 31690561]
21. van Vugt VA, Ngo H, van der Wouden JC, Twisk JW, van der Horst HE, Maarsingh OR. Online vestibular rehabilitation for chronic vestibular syndrome: 36-month follow-up of a randomised controlled trial in general practice. Br J Gen Pract 2023 Aug 31;73(734):e710-e719 [FREE Full text] [doi: 10.3399/BJGP.2022.0468] [Medline: 37487644]
22. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet 2005 Jan;365(9453):82-93. [doi: 10.1016/S0140-6736(04)17670-8] [Medline: 15639683]
23. Ngo HT, Maarsingh OR, van de Berg R, Blanker MH, Bruintjes TD, Castien R, et al. Study protocol for a nationwide implementation of internet-based vestibular rehabilitation for patients with chronic vestibular symptoms (I-RECOVER). Implement Sci Commun 2023 Nov 22;4(1):147 [FREE Full text] [doi: 10.1186/s43058-023-00524-1] [Medline: 37993954]
24. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019 Mar 29;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
25. Yardley L, Masson E, Verschuur C, Haacke N, Luxon L. Symptoms, anxiety and handicap in dizzy patients: development of the Vertigo Symptom Scale. J Psychosom Res 1992 Dec;36(8):731-741. [doi: 10.1016/0022-3999(92)90131-k] [Medline: 1432863]
26. Wilhelmsen K, Strand LI, Nordahl SH, Eide GE, Ljunggren AE. Psychometric properties of the Vertigo Symptom Scale - Short Form. BMC Ear Nose Throat Disord 2008 Mar 27;8:2 [FREE Full text] [doi: 10.1186/1472-6815-8-2] [Medline: 18371190]
27. Noordzij AS, Terluin B, van der Wouden JC, van der Horst HE, Essery R, Yardley L, et al. Cross-cultural validation of the Dutch version of the Vertigo Symptom Scale – Short Form Questionnaire in general practice. Research Square Preprint posted online on February 08, 2022. [doi: 10.21203/rs.3.rs-1325407/v1]
28. Yardley L, Barker F, Muller I, Turner D, Kirby S, Mullee M, et al. Clinical and cost effectiveness of booklet based vestibular rehabilitation for chronic dizziness in primary care: single blind, parallel group, pragmatic, randomised controlled trial. BMJ 2012 Jun 06;344:e2237 [FREE Full text] [doi: 10.1136/bmj.e2237] [Medline: 22674920]
29. Yardley L, Kirby S. Evaluation of booklet-based self-management of symptoms in Ménière disease: a randomized controlled trial. Psychosom Med 2006;68(5):762-769. [doi: 10.1097/01.psy.0000232269.17906.92] [Medline: 17012531]
30. Yardley L, Donovan-Hall M, Smith HE, Walsh BM, Mullee M, Bronstein AM. Effectiveness of primary care-based vestibular rehabilitation for chronic dizziness. Ann Intern Med 2004 Oct 19;141(8):598-605. [doi: 10.7326/0003-4819-141-8-200410190-00007] [Medline: 15492339]
31. Geraghty AW, Essery R, Kirby S, Stuart B, Turner D, Little P, et al. Internet-based vestibular rehabilitation for older adults with chronic dizziness: a randomized controlled trial in primary care. Ann Fam Med 2017 May;15(3):209-216 [FREE Full text] [doi: 10.1370/afm.2070] [Medline: 28483885]
32. Degree of urbanisation. Centraal bureau voor statistiek. URL: https://www.cbs.nl/en-gb/our-services/methods/definitions/degree-of-urbanisation [accessed 2026-11-12]
33. Education level. Centraal bureau voor statistiek. URL: https://www.cbs.nl/en-gb/news/2023/48/satisfaction-high-in-caribbean-nl-in-spite-of-low-material-well-being/education-level [accessed 2026-11-12]
34. Migration background. Centraal bureau voor statistiek. URL: https://www.cbs.nl/en-gb/our-services/methods/definitions/migration-background [accessed 2026-11-12]
35. Whitney SL, Sparto PJ, Furman JM. Vestibular rehabilitation and factors that can affect outcome. Semin Neurol 2020 Feb;40(1):165-172. [doi: 10.1055/s-0039-3402062] [Medline: 31887754]
36. Bösner S, Schwarm S, Grevenrath P, Schmidt L, Hörner K, Beidatsch D, et al. Prevalence, aetiologies and prognosis of the symptom dizziness in primary care - a systematic review. BMC Fam Pract 2018 Feb 20;19(1):33 [FREE Full text] [doi: 10.1186/s12875-017-0695-0] [Medline: 29458336]

XSL•FO
RenderX

37. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
38. Toews I, Anglemyer A, Nyirenda JL, Alsaid D, Balduzzi S, Grummich K, et al. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials: a meta-epidemiological study. Cochrane Database Syst Rev 2024 Jan 04;1(1):MR000034. [doi: 10.1002/14651858.MR000034.pub3] [Medline: 38174786]
39. van Vugt VA, Heymans MW, van der Wouden JC, van der Horst HE, Maarsingh OR. Treatment success of internet-based vestibular rehabilitation in general practice: development and internal validation of a prediction model. BMJ Open 2020 Oct 16;10(10):e038649 [FREE Full text] [doi: 10.1136/bmjopen-2020-038649] [Medline: 33067287]
40. Hall CD, Herdman SJ, Whitney SL, Anson ER, Carender WJ, Hoppes CW, et al. Vestibular rehabilitation for peripheral vestibular hypofunction: an updated clinical practice guideline from the Academy of Neurologic Physical Therapy of the American Physical Therapy Association. J Neurol Phys Ther 2022 Apr 01;46(2):118-177 [FREE Full text] [doi: 10.1097/NPT.0000000000000382] [Medline: 34864777]
41. Geraghty AW, Kirby S, Essery R, Little P, Bronstein A, Turner D, et al. Internet-based vestibular rehabilitation for adults aged 50 years and over: a protocol for a randomised controlled trial. BMJ Open 2014 Jul 22;4(7):e005871 [FREE Full text] [doi: 10.1136/bmjopen-2014-005871] [Medline: 25052178]
42. Blythin AM, Elkes J, van Lindholm T, Bhogal A, Wilkinson TM, Saville C, et al. Can digital health apps provide patients with support to promote structured diabetes education and ongoing self-management? A real-world evaluation of myDiabetes usage. Digit Health 2023 Mar 11;9:20552076221147109 [FREE Full text] [doi: 10.1177/20552076221147109] [Medline: 36923369]
43. García-Estela A, Cantillo J, Angarita-Osorio N, Mur-Milà E, Anmella G, Pérez V, et al. Real-world implementation of a smartphone-based psychoeducation program for bipolar disorder: observational ecological study. J Med Internet Res 2022 Feb 02;24(2):e31565 [FREE Full text] [doi: 10.2196/31565] [Medline: 35107440]
44. Kozlov E, Bantum E, Pagano I, Walser R, Ramsey K, Taylor K, et al. The reach, use, and impact of a free mHealth mindfulness app in the general population: mobile data analysis. JMIR Ment Health 2020 Nov 27;7(11):e23377 [FREE Full text] [doi: 10.2196/23377] [Medline: 33245289]

Abbreviations

BPPV: benign paroxysmal positional vertigo
GP: general practitioner
OR: odds ratio
RE-AIM: reach, effectiveness, adoption, implementation, and maintenance
RR: risk ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
VR: vestibular rehabilitation
VSS-SF: Vertigo Symptom Scale–Short Form

Edited by S Munce; submitted 10.Jul.2026; peer-reviewed by A Wolfovitz, T Ledin, D Meldrumd; comments to author 23.Sep.2026; revised version received 16.Oct.2026; accepted 29.Oct.2026; published 20.Nov.2026.

Please cite as:
Ngo HTN, Maarsingh OR, Blanker MH, Bont J, van Vugt VA
Real-World Use and Effectiveness of Internet-Based Vestibular Rehabilitation Following Nationwide Implementation: Prospective Observational Cohort Study
JMIR Rehabil Assist Technol 2026;13:e106458
URL: https://rehab.jmir.org/2026/1/e106458
doi:10.2196/106458
PMID:

©Hà TN Ngo, Otto R Maarsingh, Marco H Blanker, Jettie Bont, Vincent A van Vugt. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX