JMIR Rehabilitation and Assistive Technologies

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

Contents

Original Papers

XSL•FO
RenderX

XSL•FO
RenderX

Review

XSL•FO
RenderX

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

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

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

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

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

Abstract

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

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

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

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

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

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

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

KEYWORDS

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

Introduction

Technical support in the field of health care is an emerging trend fostering the development of telecommunication between specialists and other medical professionals with patients [1,2]. Owing to advancements in telemonitoring and telerehabilitation, various settings are being explored for diverse clinical conditions. Mortality rate due to noncommunicable diseases is represented at 72.3% globally, requiring closer attention toward the enhancement of prevention among the vulnerable population [3]. An important measure to be taken is to educate the patient on self-care and self-management strategies to prevent serious illness and its secondary complications. Satisfaction with doing physical activity among the patient population during the COVID-19 pandemic lockdown was high, and the telerehabilitation enhanced the patient-reported outcomes but was not limited to patient adherence [4,5].

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

Vital monitoring and supervised exercise program at remote settings would be a needed feature for prescribing the physical activity for an adult with cardiovascular disease, especially during postsurgical revascularization. Components with multiple domain-based software have appeared to be highly effective in engagement and the level of participation [12]. We aim to develop a mobile software that enables effective communication and feasible applicability, with safety measures incorporated for postsurgical revascularization patients, with provision for e-media–based exercise prescriptions. Patient adherence rate and rehospitalization, functional capacity, quality of life, and patient satisfaction would be analyzed after the effective implementation of the software application during hybrid cardiac rehabilitation.

Methods

This was a qualitative feasibility trial, which was a part of a pilot trial before the main trial. An interface was developed for engaging cardiovascular disease patients and health care professionals with an effective mobile software–based technology, which was analyzed for its usefulness and applicability. Purposive sampling was performed among participants to reflect the exposure to cardiac rehabilitation enabled with technology, including physicians (n=5), nurses, cardiac intensive care nurses (n=10), and physiotherapists (n=15) of a total of 25 (Table 1). The systematic face-to-face interview was conducted, and the study was checked for its quality using Consolidated Criteria for reporting qualitative studies (COREQ) [13,14]. The responses received from the health care professionals are stated using the LIKERT scale. Software application would include medication adherence, wound care education, nutritional education, exercise counseling and training, 24/7 communication with the health care team, and associated telemonitoring with postexercise feedback.

The procedure of the experiment trial includes that the intervention group and control group would be trained for physical activity as per routine hospital protocol. Following traditional care, the interventional group would be introduced to a software interface for training, and the control group would receive training based on the routine phase II center-based rehabilitation (Table 1). The intervention group would be trained specifically for regular practice of "physical activity at home" with a safer and effective mode using prescribed dosage. Training would be provided to the participants for recording their vitals, including heart rate, oxygen saturation, blood pressure, and rating of perceived exertion, and they would be trained for documenting these in the software interface.

Table . Baseline characteristics of interview participants

Characteristic	Doctors	Nurses	Physiotherapists
Sex			
Male, n (%)	4 (80)	2 (20)	9 (60)
Female, n (%)	1 (20)	8 (80)	6 (40)
Age (years), mean (SD)	42 (5)	37 (12)	37 (8)
Experience, mean (SD)	16.8 (10)	14.9 (13)	15 (8)
Occupation	• Cardiologist (n=2) • Physicians (n=3)	• Assistant nurse (n=6) • Clinical in-charge (n=2) • Senior nurse (n=2)	• Clinical therapist (n=8) • Senior therapist (n=7)

Education

The management of the postoperative recovery period after a cardiovascular incident would be challenging, requiring an advanced approach for management due to various postoperative complications (Table 2). SRCardioCare enables patients to be effectively educated, where information about conditions, risk factors, and secondary prevention as well as exercise counseling and training is provided through video presentations. Telemonitoring data were collected from the patients, along with postexercise feedback and regular reminders of medications, including communication with the patient.

XSL•FO
RenderX

Table . An outline implementation of software to the participants with a follow-up for 12 weeks.

Module	Week 1 - 2	Week 3 - 4	Week 5 - 6	Week 7 - 8	Week 9 - 10	Week 11 - 12
Introduction and orientation	✓					
Training to self-evaluate vitals	✓					
Evaluation for risk stratification	✓					
Counseling - Medication		✓				
Counseling - Nutrition		✓				
Counseling - Functional capacity		✓				
Counseling - Setting a goal		✓				
Counseling - Physical activity			✓			
Re-evaluation - Functional capacity			✓			
Re-evaluation – Recovery vitals				✓		
Counseling - Energy conservation				✓		
Counseling - Relaxation and coping strategies				✓		
Counseling - Secondary prevention strategies					✓	
Counseling - Behavior changing techniques					✓	✓
Counseling - Leading a quality and comfortable life						✓

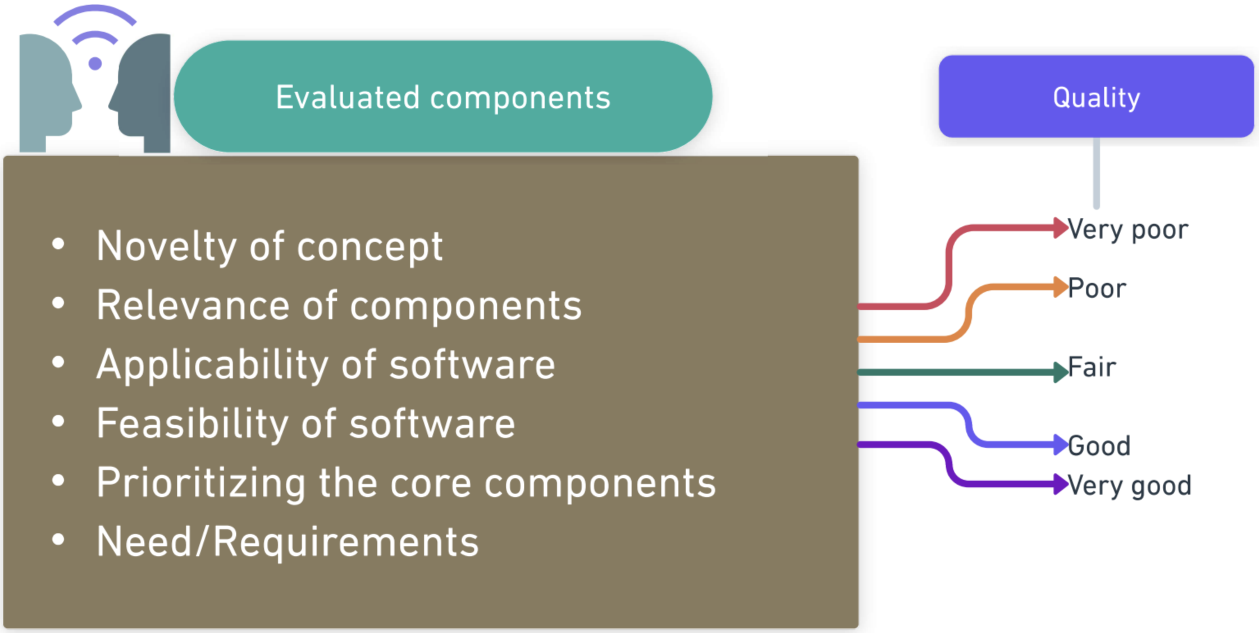
Narrative View of Feasibility Checking Among Health Care Professionals

A systematic interview was conducted among various health care professionals using the following domain (Figure 1), which was stratified from the interview guide. Data were collected from the participants using descriptive field notes and interviews, which were repeated once after the required corrections were made based on feedback provided by the participants in the software. Literature supports that there have been deprived scenarios for implementing the physical activity following the cardiac events, including postoperative cardiac revascularization surgeries and other cardiac diseases. Since understanding the trends falling toward the implementation of newer interventions in health care, most of the health care workers agreed to develop the software. The following feedback was added to the prototype design as suggested by participants:

1. Components of psycho-social counseling
2. Nutritional education components
3. Behavioral change techniques
4. Ensuring safety for implementing physical activity in remote settings

XSL•FO
RenderX

Figure 1. Survey among health care professionals for assessing the quality of software.



System Implementation

To build this mobile app for the rehabilitation of patients, Flutter was used to build the frontend and user interfaces, Node and Express for the backend, and MySQL as the primary database.

Technology Stack

Front-End Development With Flutter

The frontend of the SRCardioCare app has been entirely made using Flutter—a UI framework developed by Google to create cross-platform apps. Some of the advantages of Flutter include the following:

1. Cross-platform support: Flutter can compile to both iOS and Android, our primary targets for this app, from a single code base, maintaining and updating new features with a small development team.
2. Developer experience: Flutter is paired with Dart, a language convenience that was high with building and debugging code, especially on physical devices. The hot reload allows us to observe changes live and assist in preventing the reinstallation of the app frequently.
3. Ecosystem: the ecosystem of packages and libraries available in Flutter. Some of those libraries include: ‘fl_chart’ for visualizations and ‘chewie’ for video playback.

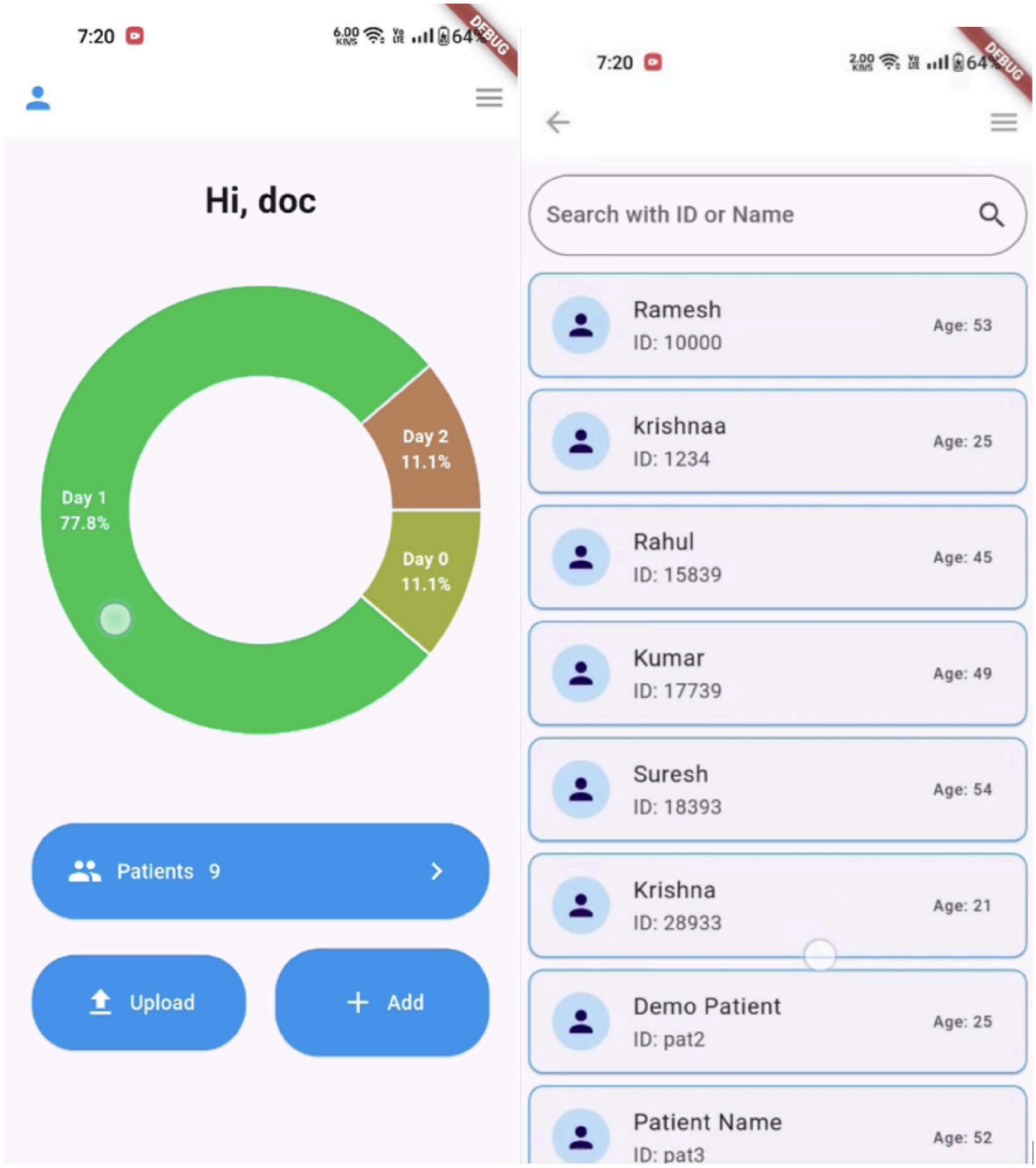
Interface Design

The user interface (UI) of SRCardioCare focuses on simplicity and accessibility, ensuring that patients with varying levels of tech literacy can navigate the platform with ease. The UI comprises several key features:

Dashboard

A personalized dashboard is given for every user who logs into the app, for both patients and physicians (Figure 2).

Figure 2. Developed software design home page for doctors with uploading videos options and patient progress.



Ramesh
ID: 10000
Age: 53

krishnaa
ID: 1234
Age: 25

Rahul
ID: 15839
Age: 45

Kumar
ID: 17739
Age: 49

Suresh
ID: 18393
Age: 54

Krishna
ID: 28933
Age: 21

Demo Patient
ID: pat2
Age: 25

Patient Name
ID: pat3
Age: 52

- Patients: The dashboard shows the average of their vital parameters during their course of therapy. Links to all important actions such as viewing the current day's exercises, help, and contact information of their dedicated physician are featured.
- Physicians: The dashboard shows a visualization representing the different days different patients are in; it provides them with a broad overview of their patients' progress and offers them the ability to add a new patient.

Video Playback

The Flutter `chewie` package on top of a video controller was used to present the video playback of the exercise videos uploaded by the Physician (Figure 3). Each patient can only view the video prescribed to them, in order, and only on the specific day selected by the Physician. This ensures patients cannot share sensitive information, and their physician can always update the exercise based on the feedback from the patient.

XSL•FO
RenderX

Figure 3. Uploaded videography of the exercises prescribed for the patient.

Feedback

After the patients complete the day’s exercises, they are provided with a form to input their vital parameters, such as heart rate, rating of perceived exertion, oxygen saturation, total body weight, and any remarks or doubts (Figure 4). This allows us to create a valuable record of their parameters throughout their journey and helps the physician tailor the exercise program further.

XSL•FO
RenderX

Figure 4. Documentation of post-physical activity vitals for patients.

Technical Challenges

The primary challenge faced when developing this system was integrating both the frontend and backend together with APIs. We used a simple REST (Representational State Transfer) setup, but for safety, we implemented the interfaces of the models and data both in TypeScript and Flutter independently. The difficulty level is high due to Flutter's rigid typing system and inability to convert JSON directly to appropriate types in Flutter. The manually written parsers for each model were used.

Backend

The backend was developed with Express; it runs using a bleeding-edge runtime called Bun. We preferred Bun instead of Node because of the following advantages: The typescript support directly runs a typescript file without having to transpile it to JavaScript. In HTTP/ Web APIs, the performance of Bun appears hastened compared to Node. Bun offers a great standard library, providing useful functions like hashing, which was used to verify our passwords on authentication.

Example- Bun hashing:

let isValid=await Bun.password.verify (password, user[0].password)

Data Storage, Privacy, and Management

Currently, SRCardioCare uses MinIO, a S3-compatible block storage service that is self-hosted. This was used to store the videos uploaded by the physicians. It is a drop-in replacement for AWS S3; therefore, no code changes are required when moving the data. According to the Healthcare Data Protection Regulations by HIPAA (Health Insurance Portability and Accountability Act) and GDPR (General Data Protection Regulation), a clear consent would be obtained from the participants, data usage would be strictly limited and restricted, and participants would be given full rights to remove and access their data as required.

Advantages are that the files are isolated in different buckets, and the access control is strictly granted based on the rules from the database; access is restricted to patients and physicians for maintaining the integrity of data.

The access is limited based on the mechanism of presigning URLs, ensuring no patient ever has permanent access to the videos and preventing them from sharing the exercises without a physician's prescription. Timely backups of the data for redundancy using external scripts such as *cron* and volume backups provided by the Cloud Provider could be scheduled.

System Integration and Module Development

The emerging growth of coronary artery disease and Acute Coronary Syndrome needs an efficient rehabilitation technique that satisfies the desire of affected patients. Traditional methods have more limitations, including monitoring the patient's risk when they are away and the need for continuous personalized care. This design addresses the above-mentioned limitation by supporting a personalized rehabilitation module and assisting them in a fast recovery from the disease. The proposed design methodology is a real-time cloud-based design to continuously monitor the patient and collect all the vital data of the patient.

The main goal of this proposed mobile app is to analyze and manage risk factors to reduce disability, improve functional capacity, lessen activity-related adverse effects, educate patients, and improve the quality of life.

To support the above-mentioned goal, the following design principles are adhered to: designed specifically for individual patients, collect vitals, processed offline, and prepared into a report specific to patients; development of a mobile application integrating technological advancements and cloud computing into health care, and continuous tracking of the patient based on the exercise.

Ethical Considerations

Ethical clearance was obtained from the Sri Ramachandra Ethical Committee IEC/24/AUG/187/26. The registration number for this trial is CTRI/2025/03/082747. This study is a part of a randomized controlled trial, in which the developed SRCardioCare app would be implemented to compare the effects of home-based cardiac rehabilitation with center-based cardiac rehabilitation for adult surgical revascularization patients (CTRI/2025/03/082747). Written informed consent was obtained from all participants prior to enrollment. Participants were informed about the purpose of the study, the voluntary nature of participation, and their rights to decline to answer any questions or withdraw from the study at any time without any consequences. All interview data were deidentified prior to analysis. Personal identities were removed and confidentiality was maintained through secure, password-protected data storage accessible only to the research team. Participants did not receive any financial or material compensation for their participation in the study.

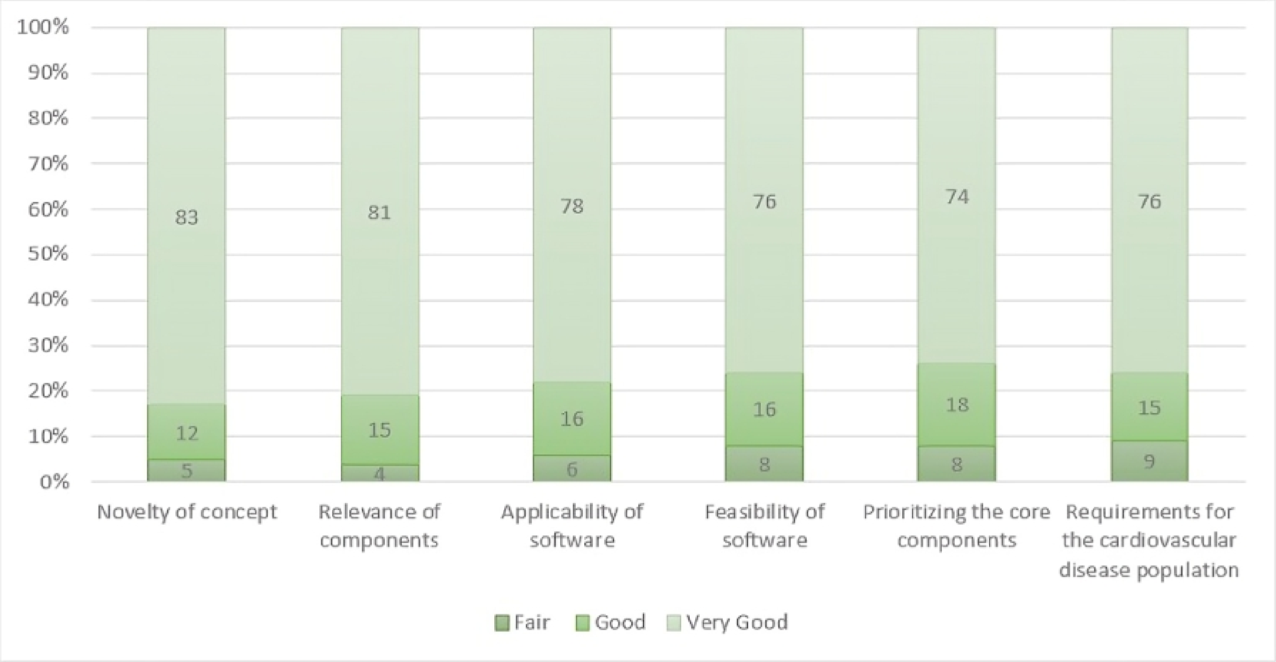
Results

Principal Results

After the presentation of the developed mobile software (app) SRCardioCare, with various expert health care professionals, a total of 25 participants were interviewed for their opinion about the developed software. The acceptability of developed software from the professionals was high due to its application of multidisciplinary approaches, along with high relevance and clinical need (Figure 5). UI-enabled ease toward navigation to their dashboards and exercise videos without difficulties. Respondents provided positive feedback about the app and highlighted their satisfaction with features such as comments after exercises and vitals documentation. But vital documentation emerged as a poor efficiency due to varying levels of digital literacy among participants. Workflow issues addressed by the clinician would be time constraints in managing and replying to the vital details and comments provided by the participant's end.

XSL•FO
RenderX

Figure 5. Responses from health care professionals: the built software was introduced to health care professionals in Sri Ramachandra Hospital and Medical Center, including 5 physicians, 10 nurse practitioners, and 15 physiotherapists. Responses were obtained using survey questions.



- **Acceptability:** Health care professionals described the app-based platform as highly acceptable due to its integration with technology advancements and applicability in cardiac rehabilitation. The novelty of such innovation with tele-rehabilitation contributed to its acceptance.

"The app is found useful for cardiac rehabilitation, especially for the heart failure population…"

- **Usability:** Participants perceived the app as easy to navigate with its simplified dashboards and order of sequence for training without hurdles. Order of week-wise module with training videos supported hassle-free provision for participants.

"The *instruction provided in the app makes easy navigation in the app, providing good follow-ups…"*

- **Satisfaction:** Health care professionals expressed a completeness in app features that include a multidisciplinary counseling forum and improved patient engagement. The essential feature considered mandatory would be the acquisition of the post-vital monitoring and communication module by health care professionals.

"I personally felt safe suggesting the patient get enrolled in cardiac rehabilitation using this software…"

- **Barriers:** The manual entry of vital signs by the participants could affect the accuracy and compliance, as per the current trial protocol. Unless a monitoring device is prepared, the defects might be considered noncompliant. Technical challenges in linking the frontend and backend components were described.

"Preloaded video components were slow occasionally and could reduce the medical terminologies used…, The app could also be used for heart failure, where more precautionary measures need to be taken..."

- **Facilitators:** The components, which included educational and personalized training, motivated health care professionals to prescribe the application. The visual dashboards and consistent communication between patients and providers engage patients.

"It is good that the app displays the number of days the participants have to work on… and the exercises provided were kept changing every week, making participants more involved…"

- **Workflow Issues:** Data flow and coordination between modules were concerns of workflow-related issues, and manual data entry by the participants would be considered a difficulty, indeed an issue. The current software is limited to such a feature; the second phase of the current trial would be an innovation of a real-time monitoring device that would gather data automatically from the participants' continuing exercise training. The integration of real-time monitoring devices was recommended to enhance workflow efficiency.

"The alerting features and regular reminders would enhance the participation of patients effectively if done properly…"

"Rather than doing exercises independently, this app helps participants to perform without missing the sessions…"

Experts supported the SRCardioCare mobile application, emphasizing the necessity of the innovation, stating that telehealth adoption is rapidly increasing. Content analysis of health care professionals' suggestions and feedback revealed that they valued safety as a major requirement, stressing the necessity of continuous monitoring and postexercise feedback by the enrolled participants toward cardiac rehabilitation.

XSL•FO
RenderX

App Outputs

Personalized Treatment

After the presentation of the developed mobile software (app), SRCardioCare, with various expert health care professionals, a total of 25 participants were interviewed for their opinion about the developed software (Figure 1). Based on the vital measurements, the physician will prescribe personalized exercises suited for each patient (Figure 3). The exercises are uploaded to their login and assigned to that patient based on the patient ID. The physician will upload a sequence of six nondownloadable videos that comprise the exercises to be done and will change after four days. The patient can log into the app but can only view the first exercise, which is assigned to their ID. These exercises may vary after four days. The same progress can also be viewed from the portal.

Baseline Assessment of Patient

This module is basic for the Remote Rehabilitation system. Before using this remote system (Mobile App), orientation for the patient is given on this app and the need of this mobile app and the postoperative self-evaluation, importance of their continuous monitoring of their vital parameters of the patient such as heart rate, oxygen saturation, rating of perceived exertion, total body weight, and provided ways of managing the risk in postoperative periods are explained (Figures 4 and 6).

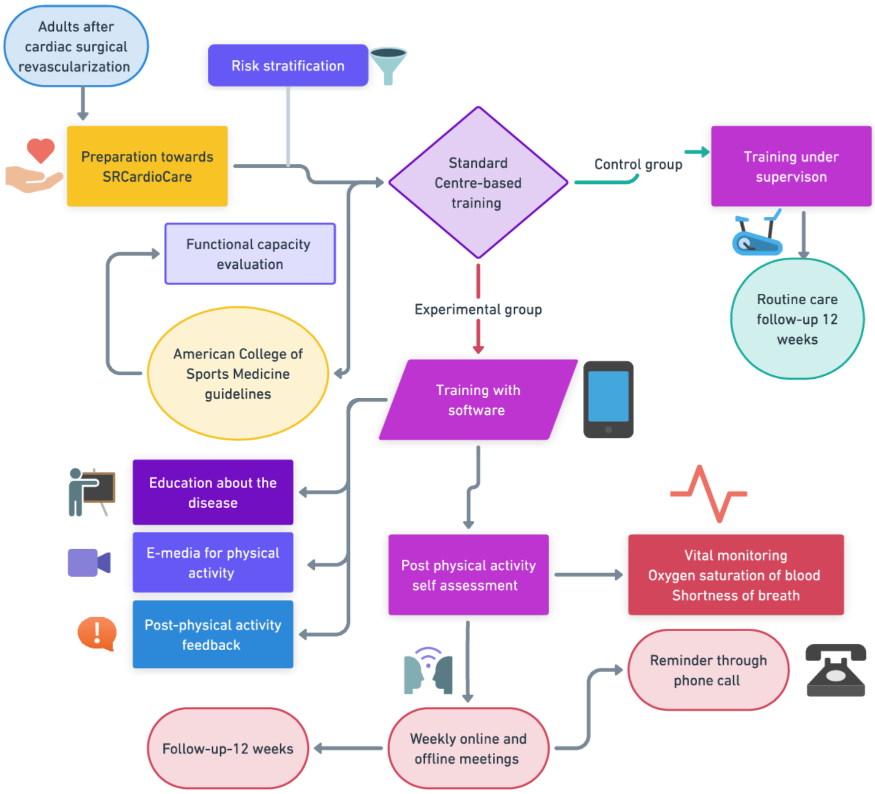
XSL•FO
RenderX

Figure 6. The screenshot shows the history of patients in the physician's login, along with the vital parameters measured.

XSL•FO
RenderX

To support the patients for the above-mentioned program, a Mobile App was designed. The mobile app is facilitated with role-based access, such as physician and patient. A physician will be assigned to many patients. During the first two weeks of this program physician will record the vitals of the patient before they are discharged from the hospital, and based on the vital signs, a personalized rehabilitation program will be designed for the patient. When the patient returned home, they needed constant treatment with remote monitoring (Figure 7). The patient can log into the system and view the personalized exercises loaded to them by their physicians (Figure 3).

Figure 7. SRCardioCare Trial protocol for participants after introducing the software. The chart was developed using the Whimsical tool.



Risk Factor Analysis

The main objective of this module is to assess the patient and provide continuous monitoring of their vital parameters. If needed, changes in the exercises can also be made. In the current app, the patient has to manually enter the vital parameter in the mobile app. As a future enhancement, it is planned to incorporate telemonitoring tools to record vital parameters. The vital parameters of the patients will be tracked and monitored by the physician. If any abnormality arises in the vital parameter, alert the caretaker or speed up the recovery process. This tracking prevents or reduces the adverse effects on cardiac activity.

This data is used to identify patterns and correlations between physical activity and vital signs, helping us assess the effectiveness of the rehabilitation plan and make informed adjustments. If a risk factor is detected (eg, prolonged elevated heart rate), the system sends alerts to both the patient and their health care provider, prompting immediate review and intervention. It also helps to predict the patient's health before the risk arises.

Data Analysis and Feedback Module

This mobile app establishes a communication between the physician and the patient at a remote setup. Through the app, the patient can readily present queries to the physician. They can record their vitals and post them to the physician. Physicians will resolve the queries raised by the patient on a priority basis. Physicians can analyze the patient's vital data stored in the cloud and get detailed insights about the patient's health, and they can predict the potential risk and avoid or reduce the adverse effects.

Discussion

Principal Findings

Cardiovascular disease remains the first among the major causes for years of life lost globally, ischemic heart disease being one of the causes if grouped globally [3,15]. The global risk of noncommunicable diseases is rising and requires immediate intervention to prevent complications [16]. Face-to-face implementation of exercises for chronic disabilities such as low back pain did not prove to be more effective than an exercise-based mobile app for managing the patients' issues, according to a recent randomized controlled trial [17]. Decisions

XSL•FO
RenderX

toward the development of a newer approach to deliver effective telemedicine and telerehabilitation are highly warranted [18]. A trial shows that intervention using a mobile app resulted in promising outcomes, stating that enhanced vitals, exercise capacity, and satisfaction toward exercise participation [19]. Device-dependent cardiac rehabilitation program for patients enhances the effective implementation of the protocol, as well as participants' interest toward participation is also promisingly enabled [20].

Behavioral change techniques would be an emerging intervention that would initiate the hopeful scope toward physical fitness in society. The awareness should be obtained at the guardian level and the care provider level to ensure the implementation of physical activity among the vulnerable population [21]. A systematic review reveals that the mobile app–based intervention would enhance diet, physical activity, and sedentary behavior when applied modestly [12]. Another systematic review suggests that despite strong evidence, people fail to follow the suggested physical activity and guidance from clinicians to lead a healthy life [22].

Cardiac rehabilitation has been prescribed to patients after cardiac events, but effective implementations are becoming poor due to poor participation. Which might require an alternative method of implementation, including technology-enabled implementation of cardiac rehabilitation at remote setups as well as in-center setups [23-25]. Effective implementation of a telehealth program depends on the various domains that are to be incorporated to gain effective clinical outcomes for the participants under safer practice. That might include telemonitoring, medication adherence, education, self-awareness about their functional capacity, understanding, and mastery of the telehealth model for self-implementation by the patients [26,27].

Patients would be thoroughly evaluated by a health care professional before introducing the software app, and the assessed comorbid conditions would be considered for exercise prescriptions. Prior training would be provided based on the suggestions provided by physicians. The concepts providing the e-supported rehabilitation criteria would include implementing cardiac rehabilitation through an economically cheaper method. Henceforth, the app usage would be based on subscription, which depends on the requirements of the duration of rehabilitation sessions suggested by physicians.

Conclusion

Globally, the literature supports that remote rehabilitation for cardiac diseases is effective. A mobile app offers trustworthy standards for the long-term adherence of patients with noncommunicable diseases. The developed software may help us to build a protocol for implementing remote rehabilitation following cardiac surgeries, heart failure, and conservatively managed acute coronary syndromes. Further upgradation may include telemonitoring support using the hardware gadget that would be designed to provide precise data of vitals during and post-exercise sessions. This leads the researchers to focus more on areas of telemonitoring devices that are more feasible and economical; deficient procurement of telemonitoring would be a limitation of this study.

Acknowledgments

The module was developed with the help of our co-author from Sri Ramachandra Engineering Technology; hence, no developmental cost was included. Maintenance of the software would be taken up by the principal investigator and the co-author team until the software is used for the complete implementation. Whimsical AI tool was used to draft Figure 6 (for design purposes only) [28].

Data Availability

Datasets generated during or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Gallagher R, Roach K, Sadler L, et al. Mobile technology use across age groups in patients eligible for cardiac rehabilitation: survey study. JMIR Mhealth Uhealth 2017 Oct 24;5(10):e161. [doi: 10.2196/mhealth.8352] [Medline: 29066425]
2. Oclaman JM, Murray ML, Grandis DJ, Beatty AL. The association between mobile app use and change in functional capacity among cardiac rehabilitation participants: cohort study. JMIR Cardio 2023 May 15;7:e44433. [doi: 10.2196/44433] [Medline: 37184917]
3. Naghavi M, Abajobir AA, Abbafati C, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet 2017 Sep;390(10100):1151-1210. [doi: 10.1016/S0140-6736(17)32152-9]
4. Krzowski B, Boszko M, Peller M, et al. Mobile application and digital system for patients after myocardial infarction: early results from a randomized trial. Pol Arch Intern Med 2023 Sep 29;133(9):16452. [doi: 10.20452/pamw.16452] [Medline: 36876854]

XSL•FO
RenderX

5. Bordas-Martinez J, Matéu Gómez L, Cámara Menoyo D, et al. Patient-reported outcomes measures (PROMs) and patient-reported experience measures (PREMs) of COVID-19 telerehabilitation: prospective pilot program. Medicine (Baltimore) 2022 Aug 5;101(31):e29639. [doi: 10.1097/MD.0000000000029639] [Medline: 35945781]
6. Salarvand S, Farzanpour F, Gharaei HA. The effect of personalized mobile health (mHealth) in cardiac rehabilitation for discharged elderly patients after acute myocardial infarction on their inner strength and resilience. BMC Cardiovasc Disord 2024 Feb 19;24(1):116. [doi: 10.1186/s12872-024-03791-5] [Medline: 38373888]
7. Li X, Zhao L, Xu T, et al. Cardiac telerehabilitation under 5G internet of things monitoring: a randomized pilot study. Sci Rep 2023;13(1). [doi: 10.1038/s41598-023-46175-z]
8. Senanayake S, Halahakone U, Abell B, et al. Hybrid cardiac telerehabilitation for coronary artery disease in Australia: a cost-effectiveness analysis. BMC Health Serv Res 2023 May 20;23(1):512. [doi: 10.1186/s12913-023-09546-w] [Medline: 37208666]
9. Najafi F, Nalini M. Hospital-based versus hybrid cardiac rehabilitation program in coronary bypass surgery patients in western Iran: effects on exercise capacity, risk factors, psychological factors, and quality of life. J Cardiopulm Rehabil Prev 2015;35(1):29-36. [doi: 10.1097/HCR.0000000000000087] [Medline: 25402170]
10. Batalik L, Dosbaba F, Hartman M, Batalikova K, Spinar J, Palazón-Bru A. Benefits and effectiveness of using a wrist heart rate monitor as a telerehabilitation device in cardiac patients: a randomized controlled trial. Medicine (Baltimore) 2020 Mar;99(11):e19556. [doi: 10.1097/MD.0000000000019556] [Medline: 32176113]
11. Dalli Peydró E, Sanz Sevilla N, Tuzón Segarra MT, Miró Palau V, Sánchez Torrijos J, Cosín Sales J. A randomized controlled clinical trial of cardiac telerehabilitation with a prolonged mobile care monitoring strategy after an acute coronary syndrome. Clin Cardiol 2022 Jan;45(1):31-41. [doi: 10.1002/clc.23757] [Medline: 34952989]
12. Schoeppe S, Alley S, Van Lippevelde W, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 7;13(1):127. [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
13. Buus N, Perron A. The quality of quality criteria: Replicating the development of the Consolidated Criteria for Reporting Qualitative Research (COREQ). Int J Nurs Stud 2020 Feb;102:103452. [doi: 10.1016/j.ijnurstu.2019.103452]
14. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care ;19(6):349-357. [doi: 10.1093/intqhc/mzm042]
15. Dorje T, Zhao G, Tso K, et al. Smartphone and social media-based cardiac rehabilitation and secondary prevention in China (SMART-CR/SP): a parallel-group, single-blind, randomised controlled trial. Lancet Digit Health 2019 Nov;1(7):e363-e374. [doi: 10.1016/S2589-7500(19)30151-7] [Medline: 33323210]
16. Fullman N, Yearwood J, Abay SM, et al. Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. The Lancet 2018 Jun;391(10136):2236-2271. [doi: 10.1016/S0140-6736(18)30994-2]
17. López-Marcos JJ, Díaz-Arribas MJ, Valera-Calero JA, et al. The added value of face-to-face supervision to a therapeutic exercise-based app in the management of patients with chronic low back pain: a randomized clinical trial. Sensors (Basel) 2024 Jan 16;24(2):567. [doi: 10.3390/s24020567] [Medline: 38257659]
18. Beatty AL, Beckie TM, Dodson J, et al. A new era in cardiac rehabilitation delivery: research gaps, questions, strategies, and priorities. Circulation 2023 Jan 17;147(3):254-266. [doi: 10.1161/CIRCULATIONAHA.122.061046] [Medline: 36649394]
19. Snoek JA, Prescott EI, van der Velde AE, et al. Effectiveness of home-based mobile guided cardiac rehabilitation as alternative strategy for nonparticipation in clinic-based cardiac rehabilitation among elderly patients in europe: a randomized clinical trial. JAMA Cardiol 2021 Apr 1;6(4):463-468. [doi: 10.1001/jamacardio.2020.5218] [Medline: 33112363]
20. Ruano-Ravina A, Pena-Gil C, Abu-Assi E, et al. Participation and adherence to cardiac rehabilitation programs. A systematic review. Int J Cardiol 2016 Nov 15;223:436-443. [doi: 10.1016/j.ijcard.2016.08.120] [Medline: 27557484]
21. Babu AS, Turk-Adawi K, Supervia M, Jimenez FL, Contractor A, Grace SL. Cardiac rehabilitation in India: results from the International Council of Cardiovascular Prevention and Rehabilitation’s Global Audit of Cardiac Rehabilitation. Glob Heart 2020 Apr 3;15(1):28. [doi: 10.5334/gh.783] [Medline: 32489801]
22. Satyamurthy A, Prabhu N, Padmakumar R, Babu AS. Feasibility of an exercise-based cardiac rehabilitation algorithm in patients following percutaneous coronary intervention for acute coronary syndrome. Indian Heart J Teach Ser 2020 Jul;72(4):289-292. [doi: 10.1016/j.ihj.2020.07.011]
23. McGillion M, Yost J, Turner A, et al. Technology-Enabled remote monitoring and self-management - vision for patient empowerment following cardiac and vascular surgery: user testing and randomized controlled trial protocol. JMIR Res Protoc 2016 Aug 1;5(3):e149. [doi: 10.2196/resprot.5763] [Medline: 27480247]
24. Patterson K, Davey R, Keegan R, et al. Testing the Effect of a smartphone app on hospital admissions and sedentary behavior in cardiac rehabilitation participants: ToDo-CR randomized controlled trial. JMIR Mhealth Uhealth 2023 Oct 3;11:e48229. [doi: 10.2196/48229] [Medline: 37788043]
25. Zou H, Chair SY, Feng B, et al. A social media-based mindfulness psycho-behavioral intervention (MCARE) for patients with acute coronary syndrome: randomized controlled trial. J Med Internet Res 2024 Feb 20;26(26):e48557. [doi: 10.2196/48557] [Medline: 38376899]

XSL•FO
RenderX

26. Chang HY, Wu HW, Hung CS, et al. Costs and cardiovascular benefits of a fourth-generation synchronous telehealth program on mortality and cardiovascular outcomes for patients with atrial fibrillation: retrospective cohort study. J Med Internet Res 2024 Jan 8;26(1):e48748. [doi: 10.2196/48748] [Medline: 38190237]
27. Patterson K, Davey R, Keegan R, et al. A smartphone app for sedentary behaviour change in cardiac rehabilitation and the effect on hospital admissions: the ToDo-CR randomised controlled trial study protocol. BMJ Open 2020 Dec 15;10(12):e040479. [doi: 10.1136/bmjopen-2020-040479] [Medline: 33323435]
28. Whimsical AI. URL: https://whimsical.com/protocol-NeQDxTXwXKwR1sTbcPUzc [accessed 2026-02-03]

Abbreviations

ACS: acute coronary syndrome
CAD: coronary artery disease
NCD: noncommunicable diseases
UI: User Interface

Edited by J Sarvestan; submitted 26.Nov.2024; peer-reviewed by A Suresh, D Singhal, ST Kulnik; revised version received 28.Jul.2025; accepted 15.Sep.2025; published 23.Feb.2026.

Please cite as:
Kumar Pichai A, Sathya A, SenthilKumar T, Shanmugasundaram S, Karthik R
Development of Mobile Software "SRCardioCare" Prototype for Implementing Home-Based Exercise Program Among Patients After Adult Cardiac Surgical Revascularization: Qualitative Feasibility Study
JMIR Rehabil Assist Technol 2026;13:e69197
URL: https://rehab.jmir.org/2026/1/e69197
doi:10.2196/69197

© Ajith Kumar Pichai, A Sathya, Thillaigovindarajan SenthilKumar, Sadhanandham Shanmugasundaram, R Karthik. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging Large Language Models for Early Detection of Anomaly Work Injury Cases: Data-Driven Approach to Rehabilitation Efficiency

Peter Q Chen[1], MSc; Hayley Y W Gu[1], MSc; Heidi K Y Lo[2], MBChB, MRCPsych; Wing Chung Chang[2], MBChB, MD; Cameron J M Lai[3], MBA; Sun H S Lai[3], BSc, PhD; Andy S K Cheng[4], BSc, PhD; Peter H F Ng[1], MSc, PhD

[1]Department of Computing, Hong Kong Polytechnic University, PQ710, Mong Man Wai Building, Hong Kong, China (Hong Kong)
[2]Department of Psychiatry, University of Hong Kong, Hong Kong, China (Hong Kong)
[3]Total Rehabilitation Management (HK) Limited, Hong Kong, China (Hong Kong)
[4]Department of Health and Physical Education, The Education University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Peter H F Ng, MSc, PhD
Department of Computing, Hong Kong Polytechnic University, PQ710, Mong Man Wai Building, Hong Kong, China (Hong Kong)

Abstract

Background: Large language models (LLMs) have demonstrated potential in automating the analysis of unstructured clinical data, yet their application in rehabilitation therapy for work injury cases remains underexplored.

Objective: We aimed to evaluate the performance of an LLM-assisted approach for the rapid identification of anomalous rehabilitation cases related to work injuries to enhance scalability and precision in case management.

Methods: We retrospectively analyzed 110,346 deidentified work injury cases between 2001 and 2024 from a leading rehabilitation coordination company in Hong Kong, representing approximately 20% of all work injury incidents in the region. LLMs were used to estimate the expected duration of recovery based on free-text injury descriptions. The cases in which the actual number of medically certified sick leave days exceeded the LLM-predicted maximum were classified as anomalies.

Results: The LLM-assisted method achieved high accuracy, with GPT-4o achieving over 73% accuracy in nonanomalous classification and 79% accuracy in all dataset detection, outperforming comparator models. The model maintained high accuracy across subgroups and demonstrated the reliable extraction of information from free-text notes.

Conclusions: The proposed method demonstrated robustness when evaluated on a large-scale dataset with a bimodal age distribution. This study highlights the potential of LLMs to transform rehabilitation workflows by automating anomaly detection at scale. The method also shows promise in tailoring rehabilitation strategies to age-specific needs and leveraging LLM tools for efficient case management. However, a key limitation is that the dataset includes only injury cases from a single geographic region, potentially limiting the generalizability of the findings to other populations or health care systems.

(JMIR Rehabil Assist Technol 2026;13:e80607) doi:10.2196/80607

KEYWORDS

anomaly detection; large language model; work injury; rehabilitation prediction; machine learning

Introduction

Efficient and targeted rehabilitation management is essential to ensure that individuals with severe conditions receive timely and appropriate care [1]. However, current work injury management often lacks the precision needed to allocate resources optimally, resulting in delays and inefficiencies in addressing high-priority cases [2]. A core challenge lies in the misallocation of attention and services, where relatively minor injuries with predictable recovery trajectories are sometimes treated with the same urgency as more complex cases. This misdirection not only burdens the health care system but also diverts valuable clinical time and expertise away from anomalies, cases involving severe injuries or irregular patterns of recovery that require specialized evaluation or intervention [3]. Without a robust mechanism to distinguish these cases early in the workflow, rehabilitation systems risk compromised patient outcomes, inefficient use of limited resources, and waste critical resources. In Hong Kong, the term “anomalies” refers to cases where the injured worker’s records indicate a severe injury requiring special or intensive care [4]. These cases may also suggest potential discrepancies, such as claims that have been overstated for additional compensation or legal benefits, indicating that the incident may not follow typical recovery

XSL•FO
RenderX

patterns according to industrial practice. In the Asia-Pacific context, including Singapore’s employment practices guidance and Australian public sector leave management, cases that exceed expected or allowable sick leave provisions are treated as requiring attention, which aligns with our operational use of sick leave exceedance to flag potential anomalies [5-7]. Nonanomalous data refer to cases in which the injured worker experiences a light injury expected to heal in the usual course, with a standard recovery process leading to a timely return to work. These records represent the standard outcomes, without complications or indications of potential fraud. Some anomalies may also signal potential inconsistencies, such as exaggerated claims made for extended compensation or legal advantage. In industrial practice, comprehensive annotations about why a case is “special” are typically unavailable. The only consistent, objective post hoc indicator of atypical recovery trajectory is the realized count of medically certified sick leave days. Consequently, this study operationalizes cases requiring attention via a fast filter that flags cases whose realized sick leave exceeds a large language model (LLM)−estimated expected range. We emphasize that this is a pragmatic triage proxy, not a clinical determination of pathophysiology or fraud.

Addressing this challenge requires innovative methods for quickly and accurately identifying severe cases to optimize the distribution of resources. With the advancement of artificial intelligence (AI) techniques, clinical decision support systems have been increasingly employed across various domains to assist therapists in decision-making [8-11]. In the context of workplace injury, recent research has integrated machine learning methodologies, such as the variational autoencoder, for predicting sick leave outcomes and establishing a high alertness cliff [12]. Nevertheless, the prediction process still partially depends on the initial judgment of work injury case managers, who serve as the primary decision-makers in these cases. Senior work injury case managers consistently achieve higher accuracy compared to AI-based predictions [13]. Even with the assistance of neural networks, AI cannot rapidly achieve an acceptable level of performance without proper data preprocessing and customization. The research gap lies in the lack of efficient, data-driven methods to proactively identify anomalies in work injury management workflows. Current practices predominantly rely on random case assignment and retrospective corrections, resulting in wasted time and resources.

Recently, LLMs have demonstrated exceptional capabilities in processing language-related data, even passing the United States Medical Licensing Examination [14,15]. It has also been studied in several medical fields [16-18]. Numerous studies and surveys have investigated LLMs’ ability to assume specific roles based on provided profiles, with results indicating that LLMs can effectively simulate profiled characters [19-21]. Simply prompting LLMs with a data description can generate responses in less than 1 minute without requiring additional model training. However, a critical research gap remains in determining how to constrain the outputs of LLMs and how to effectively design methods that leverage LLMs to detect anomalies in incoming injury cases efficiently.

Unlike traditional rehabilitation workflows, where senior work injury case managers must spend considerable time manually identifying anomalous cases, we developed an LLM-assisted method to streamline and accelerate this process. By leveraging prompt engineering techniques, we structured the input and constrained the output format to support accurate and efficient initial screening. The method is grounded in clinical reasoning: each injured worker is expected to follow a typical recovery trajectory, reaching a work-ready state within a medically anticipated time frame based on the nature and severity of the injury. LLMs, trained on extensive digital corpora that include medical and occupational content, are well positioned to infer such expectations and assist in detecting deviations from normative recovery patterns [22].

The scenario mirrors current practices in work injury management, where cases are often assigned randomly to junior or senior work injury case managers, only to discover later that certain cases would have benefited from senior-level expertise from the outset. This misallocation frequently results in delays and inefficient use of resources.

To this end, we proposed a novel approach leveraging LLMs to detect anomalies in occupational rehabilitation in the context of work injury management. Our method offered a potentially fast, scalable, and highly accurate solution for identifying severe cases based on data from work injury cases. Furthermore, this research collected over 110,000 work injury cases from a local company in Hong Kong, which handles nearly 20% of the total work injury cases yearly [23]. The data were used to validate our method, and several pilot studies were conducted for feasibility assessment, including model selection. Meanwhile, this research aims to uncover the key factors that characterize anomalies in this dataset, providing deeper insights into the decision-making process and facilitating a more informed allocation of resources. Our objectives are 2-fold: (1) developing a robust LLM-based method for anomaly detection in work injury management and (2) utilizing exploratory data analysis to uncover potential age and work-injury patterns in these anomaly cases.

Methods

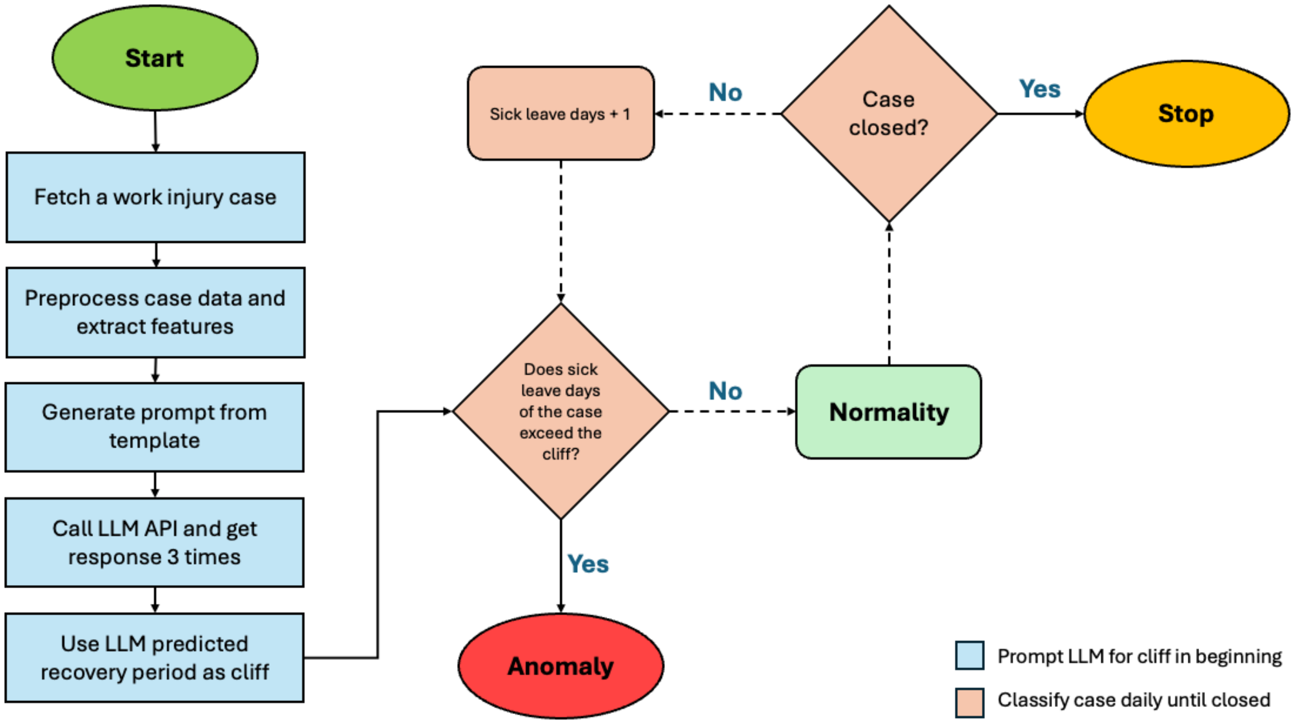
LLM-Assisted Anomaly Fast Detection Method

This method uses a fast and reliable alertness cliff to classify cases more effectively. If the total number of medically certified sick leave days for an injured worker exceeds this cliff, the case is flagged as potentially anomalous and prioritized for review by senior-level personnel or a detailed evaluation. An LLM predicts the expected duration of sick leave for each case from injury and accident descriptions. By comparing realized sick leave with the model-predicted recovery days, cases exceeding the cliff are classified as anomalies, and those within the expected range are considered normal. To enhance precision, 3 aggregation rules are used to define the decision cliff, referred to as the cliff: the maximum, the average, and the median of 3 independent LLM−generated recovery estimates. A case whose realized sick leave exceeds this cliff is classified as an anomaly. To mitigate variability and improve reliability, the LLM is queried 3 times per case using the same prompt structure, and the final decision is derived from the aggregated predictions to produce a robust, data-driven anomaly detection process [24].

XSL•FO
RenderX

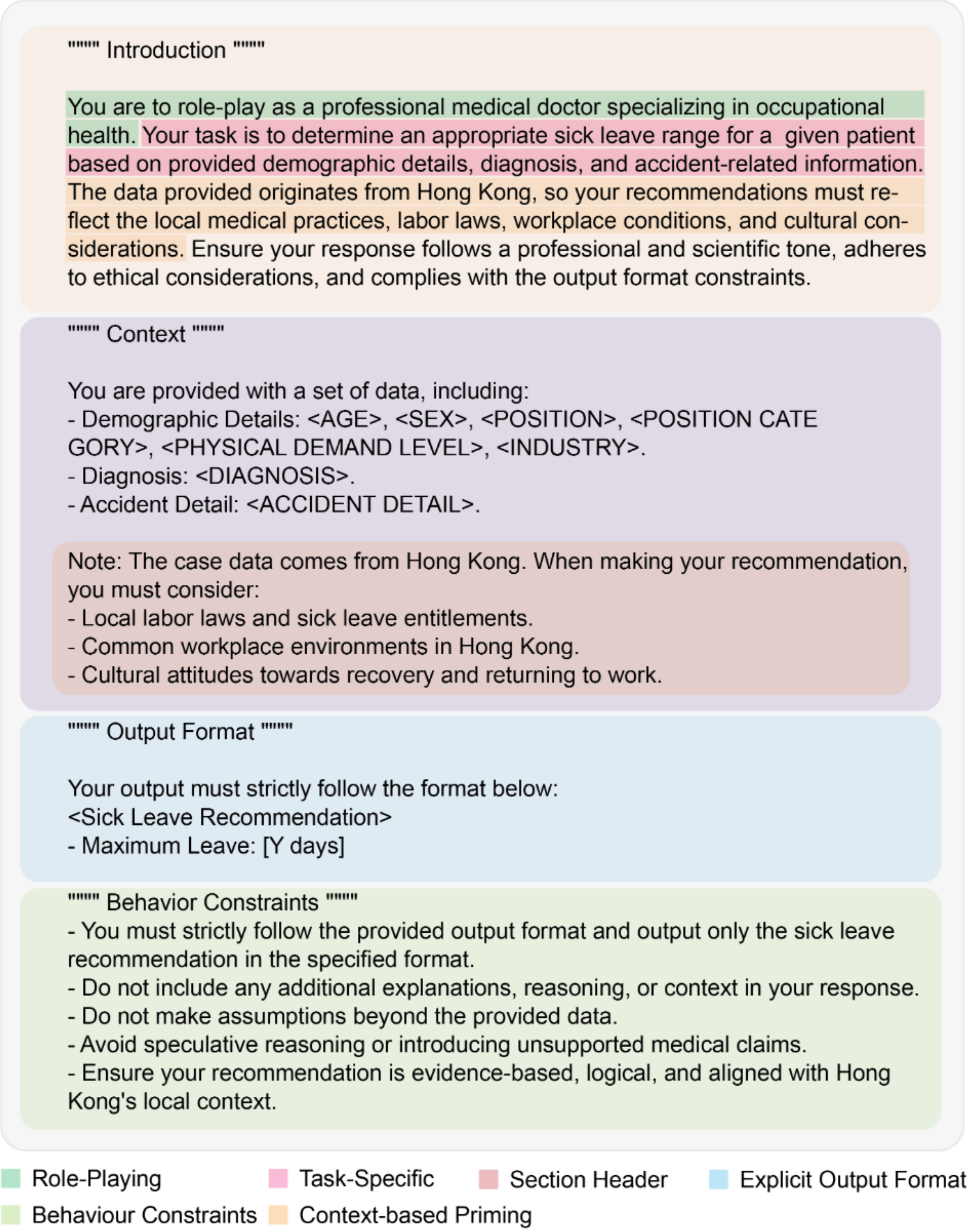
As illustrated in Figure 1, the workflow begins when a new work injury case is received. The “Query LLM 3 times and aggregate results” step rapidly determines the cliff using an LLM. The procedure begins by preparing case information, followed by preprocessing to retrieve demographics and extract key details, such as accident and diagnosis information. This content is embedded into a prompt based on the template shown in Figure 2. For each case, the LLM application programming interface (API) is queried 3 times to obtain predicted recovery periods, which serve as the cliff indicator for anomaly classification. A case is classified as nonanomalous if its sick leave has not yet exceeded the predicted cliff. For ongoing cases, sick leave days are incremented and reassessed against the cliff until the case is closed. Once the sick leave surpasses the cliff, the case is classified as an anomaly and referred to a senior work injury case manager for intervention.

Figure 1. Large language model (LLM)−assisted anomaly fast detection method. API: application programming interface.



XSL•FO
RenderX

Figure 2. Concrete prompt used to acquire information from a large language model.



"""""" Introduction """"""

You are to role-play as a professional medical doctor specializing in occupational health. Your task is to determine an appropriate sick leave range for a given patient based on provided demographic details, diagnosis, and accident-related information. The data provided originates from Hong Kong, so your recommendations must reflect the local medical practices, labor laws, workplace conditions, and cultural considerations. Ensure your response follows a professional and scientific tone, adheres to ethical considerations, and complies with the output format constraints.

"""""" Context """"""

You are provided with a set of data, including:
- Demographic Details: <AGE>, <SEX>, <POSITION>, <POSITION CATEGORY>, <PHYSICAL DEMAND LEVEL>, <INDUSTRY>.
- Diagnosis: <DIAGNOSIS>.
- Accident Detail: <ACCIDENT DETAIL>.

Note: The case data comes from Hong Kong. When making your recommendation, you must consider:
- Local labor laws and sick leave entitlements.
- Common workplace environments in Hong Kong.
- Cultural attitudes towards recovery and returning to work.

"""""" Output Format """"""

Your output must strictly follow the format below:
<Sick Leave Recommendation>
- Maximum Leave: [Y days]

"""""" Behavior Constraints """"""
- You must strictly follow the provided output format and output only the sick leave recommendation in the specified format.
- Do not include any additional explanations, reasoning, or context in your response.
- Do not make assumptions beyond the provided data.
- Avoid speculative reasoning or introducing unsupported medical claims.
- Ensure your recommendation is evidence-based, logical, and aligned with Hong Kong's local context.

The primary outcome is the accuracy of anomaly and nonanomalous classification in work injury cases, assessed on a dataset from a leading local work injury management company with expert-verified labels. The systematic use of LLM predictions enhances detection sensitivity while reducing manual workload, enabling case managers to focus on genuinely complex cases that require expert attention.

In routine operations, rich clinical detail is often unavailable at intake, so triage relies on minimal text and basic demographics. To enable low-latency and low-cost prioritization, a case-specific cliff is defined as the LLM-estimated expected duration of sick leave, serving as a data-driven prior on typical recovery given available notes. As the case progresses, if the running tally of medically certified sick leave exceeds this prior, particularly early in the timeline, the case is automatically queued for senior review. This mechanism functions as a workload triage heuristic rather than a diagnostic judgment, triggering timely escalation in cases of information scarcity,

XSL•FO
RenderX

improving allocation precision, and deferring definitive clinical determinations to expert assessment.

The method was validated using a dataset of 110,346 real-world work injury cases provided by a leading work injury management company in Hong Kong.

Prompt Template

Figure 2 demonstrates the detailed prompt engineering template, which utilizes multiple prompting techniques to enhance LLM performance. For clarity, the section header technique, such as "Introduction," "Context," "Output Format," and "Behavior Constraints," is used in the prompt template [25]. In the first part of the prompt template, the role-playing technique enhances contextual understanding, adaptability, and response accuracy by simulating specific personas, perspectives, or expertise in a given scenario [26]. To improve response relevance and coherence, the template specifies that all cases occurring in Hong Kong should utilize the context-based priming technique [27]. In the "Context" section of the prompt template, the input data for the injury case, including demographic details (eg, age, occupation), diagnosis details (textual description), and accident details (textual description), were included. In the "Output Format" part, the Explicit Output Format avoids undesired reasoning steps (eg, Chain-of-Thought) or other deviations, ensuring the LLM generates responses strictly in the intended structure without adding irrelevant content [28]. The "Behavior Constraints" part also serves a similar purpose, ensuring that the model's responses remain factual, precise, and contextually appropriate. Prompts are explicitly contextualized to Hong Kong's legal and clinical environment to enhance ecological validity, with strict output schemas for parsability and consistency [29]. Explicit instructional constraints are embedded to reduce hallucination and enforce adherence to the analytical task.

Pilot Study

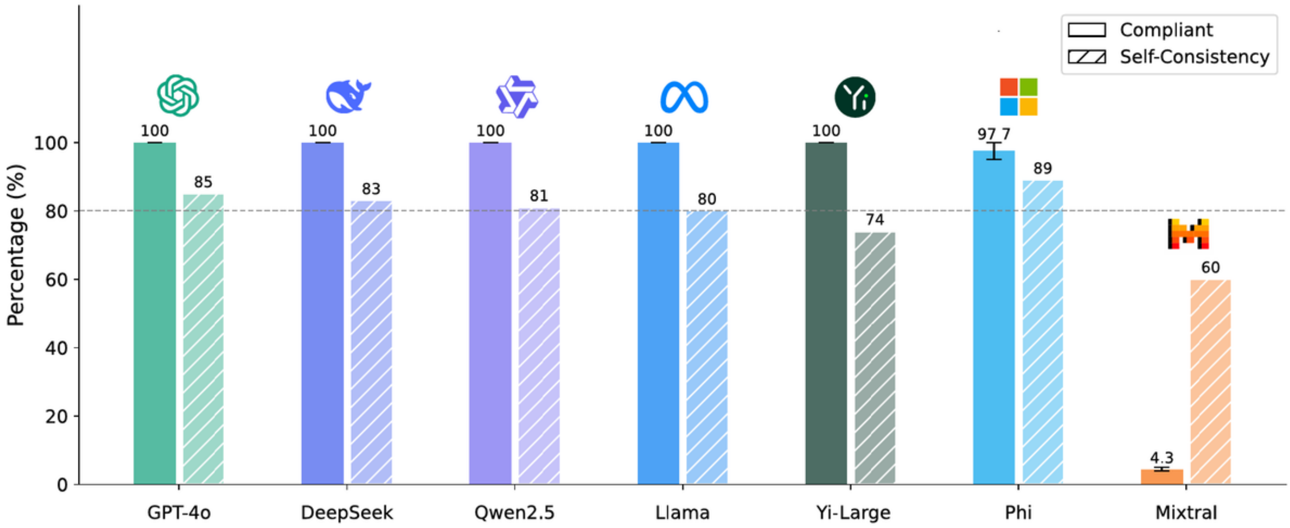
In our pilot study, we examined the varying strengths of different LLMs (eg, mathematical reasoning) and recognized that model size and architecture significantly influence performance [25,30-32]. To ensure a robust evaluation, we selected the largest and most widely recognized models from diverse LLM families, encompassing a broad range of architectures and capabilities.

Our primary objective was to determine whether these models could interpret a predefined prompt template and generate outputs that conformed to the required structure. Rather than examining their reasoning or predictive abilities, we focused on consistency, introducing 2 metrics: Compliance, which measures adherence to guidelines for producing the desired content, and Self-Consistency, which assesses whether the same response is generated across 3 repeated trials. We randomly sampled 100 cases from the dataset, prompted each model 3 times per case, and recorded the number of outputs that followed the required format. We focus on per-case anomaly triage using an upper bound on expected sick leave, so the prior work's center-focused reproducibility framework and large-run benchmarking do not fit our objective, evaluation needs, or deployable low-stochasticity protocol [32].

Figure 3 presents the results of this pilot study. Compliance represents the proportion of responses that satisfied our prescribed guidelines, while Self-Consistency quantifies each model's consistency across repeated prompts. Among 7 leading commercial LLMs, ChatGPT-4o, DeepSeek-V3, Qwen2.5-72B Instruct, Llama-3.1-405B Instruct, and Yi-Large achieved perfect Compliance; other models failed to avoid generating undesired content. We also examined internal consistency by comparing outputs across 3 prompts, categorizing them as identical across all trials, identical in at least 2, or unique each time. ChatGPT-4o exhibited the highest Self-Consistency, and several other models met our chosen 80% cliff for the subsequent experiments.

However, certain models with high Self-Consistency struggled with Compliance. This discrepancy is especially concerning, given our objective of providing a fast and reliable anomaly detection system for work injury management companies. Strict adherence to instructions is critical: any deviation can introduce erroneous or fabricated data, ultimately undermining the detection process. Ensuring compliance is thus essential to maintain the integrity of anomaly detection.

Figure 3. Compliance and self-consistency of different large language models.



XSL•FO
RenderX

Work Pipeline

Based on the results of the pilot study, we selected DeepSeek-V3, ChatGPT-4o, LLaMA-3.1-405B-Instruct, and Qwen2.5-72B to serve as LLM agents in our framework. All selected models achieved 100% compliance with the prompt template and demonstrated over 80% self-consistency across repeated responses.

A request-response framework was implemented using the FireWorks API in Python. Decoding parameters were configured for low-variance outputs using a temperature of 0.2 and a top-p of 1.0, which empirically reduced variability while maintaining robust adherence to the required output schema. A temperature of 0.0 was considered for full determinism; however, some models exhibited occasional output truncation or schema noncompliance at strictly deterministic settings during preliminary checks.

The demographic information, incident accident details, and clinical diagnoses were embedded into a standardized prompt template, as shown in Figure 2. Each prompt was submitted to the LLM via an API. A Python script extracted the responses to generate a predicted maximum duration of medically certified sick leave for each case. The data were then visualized to assess the accuracy of the LLM-based method.

Data Sources

Our dataset originates from a leading work injury management company in Hong Kong, which manages approximately 20% of work injury cases annually, covering records from 2001 to 2024. The dataset comprises 110,346 cases, with a gender distribution of 41.3% female and 53.6% male. The study population is predominantly Chinese, with individuals ranging in age from 18 to 83 years. This broad demographic coverage provides a robust basis for analyzing patterns across different age groups and genders within a relatively homogeneous ethnic context. Input leakage was not possible in this study. All records are confidential medical data, fully classified, and never publicly released or exposed to the models beyond the controlled evaluation pipeline, thereby preventing any external contamination that could bias LLM outputs or compromise validity.

Within this dataset, 15,575 cases were recorded as having zero sick leave days, which were treated as potential data entry errors. We exclude zero-day legitimate cases as outliers because they are immediately escalated to senior rehabilitation coordinators on day zero and handled outside our fast detection system, which targets anomalies only after predicted sick leave durations are exceeded. An additional 9230 cases had nonzero sick leave durations but contained missing values. For the primary predictive analysis, we used 85,541 cases that reported a nonzero number of medically certified sick leave days and had no missing data. These were inputs for the LLMs to predict the expected duration of normal recovery. Although excluded from the prediction task, the remaining data groups were also analyzed to extract relevant insights, given their substantial size.

Data Preprocessing

The dataset underwent staged preprocessing to ensure consistency and analytical suitability. Records outside the target time window were removed, implausible values were constrained within reasonable bounds, and entries with nulls in critical analytical fields were excluded. Noncritical descriptive fields with missing values were imputed using a neutral placeholder to preserve coverage while signaling incompleteness.

Categorical features were standardized through controlled vocabulary mapping, consolidation of multivalued entries into explicit multicategory indicators, and aggregation of low-frequency categories to mitigate sparsity. Text fields were sanitized by removing noninformative placeholders, and duplicates were eliminated based on content equivalence. A focused set of salient variables was retained for downstream analysis.

Data Analysis

In this study, we utilize a comprehensive set of metrics to rigorously evaluate the performance of the LLM-assisted anomaly detection method, encompassing classification accuracy, error magnitude, and model reliability. Classification accuracy, a core metric, is calculated based on 3 cliff-based methods: method 1 (maximum of 3 LLM predictions), method 2 (average of 3 predictions), and method 3 (median of 3 predictions). To further assess prediction deviations between realized sick leave days and LLM-predicted cliffs, we compute mean absolute error (MAE), mean squared error, root mean squared error (RMSE), mean absolute percentage error (MAPE), and mean percentage error (MPE). Additionally, Compliance (adherence to structured output formats) and Self-Consistency (reproducibility of outputs across repeated prompts) are quantified to ensure model reliability. We evaluate misclassification deviation, summarized through percentiles (eg, 50th and 75th percentiles), to analyze error distribution. Furthermore, exploratory data analysis is conducted to provide insights into the dataset, including descriptive statistics such as injury frequency, demographic distributions (eg, age, gender, occupation), anomaly prevalence, and misclassification patterns across key variables like body part and industry type. Together, these metrics and calculations form a robust framework for assessing the precision, reliability, and operational effectiveness of the proposed LLM-based anomaly detection system.

We evaluated triage performance using cliff-based classification derived from LLM-predicted “cliffs” of expected sick leave duration. For each case, the LLM was queried 3 times with the same prompt. The per-case decision cliff was then defined by 1 of the 3 aggregation rules: method 1, method 2, and method 3. A case was classified as an anomaly if it realized medically certified sick leave days exceeded the chosen cliff; otherwise, it was classified as nonanomalous. The primary performance metric was accuracy, computed as the proportion of correctly classified cases over the evaluation set, and reported overall and stratified by anomaly and nonanomalous subsets to characterize trade-offs across decision rules and models.

XSL•FO
RenderX

To assess reliability and error magnitude, we further quantified misclassification deviation, defined for errors as the absolute difference between realized sick leave days and the decision cliff, summarized via percentiles (eg, 50th and 75th).

To gain a deeper understanding of the phenomenon in the dataset, salary is treated as a composite proxy that captures differences across job types, seniority, contract structures, and work experience. Given the absence of granular role-level pay scales, salary should not be interpreted as a pure measure of experience but as an indicator shaped by occupational category and tenure.

Ethical Considerations

This study introduced an innovative anomaly detection method for work injury rehabilitation, validated using real-world cases from Hong Kong. The project has been approved by the PolyU Institutional Review Board (reference HSEARS20250406002). A pilot study was first conducted to evaluate the performance of several well-known commercial LLMs, which informed the selection of the most effective models for the subsequent experiments.

The dataset was provided by one of the largest work injury management companies in Hong Kong, which manages nearly 20% of the government-reported work injury rehabilitation cases. The data were shared exclusively for research purposes under a strict confidentiality agreement. Prior to transfer, all records were anonymized by the provider to ensure the protection of personal information.

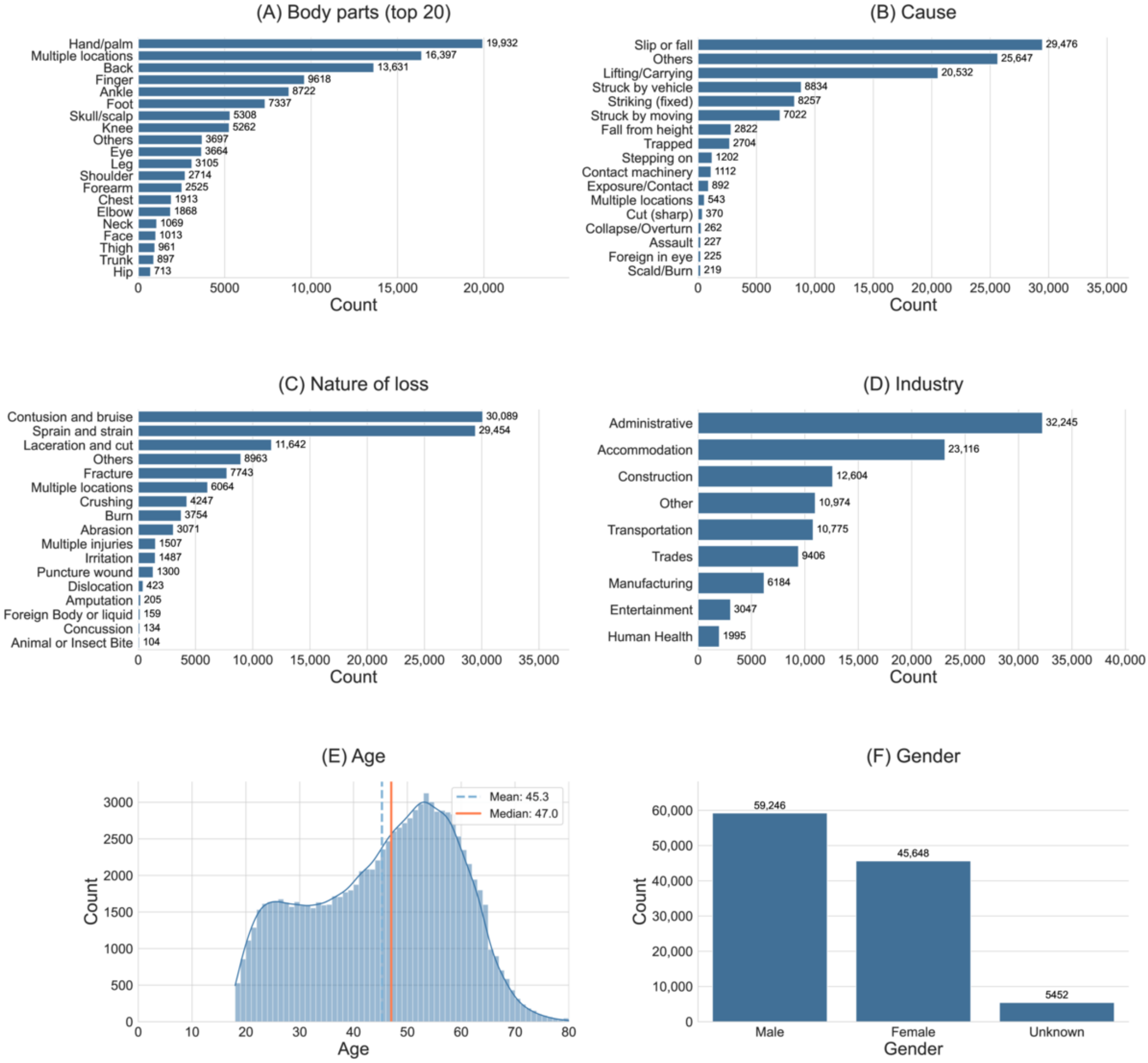
Results

Work Injury Dataset

Figure 4 presents a comprehensive analysis of 110,346 deidentified occupational injury cases managed by a leading rehabilitation coordination company in Hong Kong between 2001 and 2024. Injuries predominantly involved peripheral anatomical regions, with fingers (n=16,397) and backs (n=13,631) collectively accounting for nearly one-quarter of all incidents, followed by hand or palm injuries (n=9618) and ankle injuries (n=8722). Consistent with these anatomical findings, the most common types of loss were contusions or bruises (n=30,089) and sprains or strains (n=29,454), whereas open wounds, such as lacerations and cuts (n=11,642), and fractures (n=7743), occurred less frequently. Industry-specific data indicate a substantial burden arising from labor-intensive service sectors, notably “Administration and support services” (n=32,245) and “Accommodation and food service activities” (n=23,116), collectively accounting for more than half of all cases and surpassing the construction sector (n=12,604) in this dataset. Precipitating events were predominantly same-level slips, trips, and falls (n=29,476) and manual lifting or carrying tasks (n=20,532). Demographically, male workers represented a modest majority (53% of claims); nevertheless, female workers comprised a substantial proportion (42%). The age distribution was right-skewed, with a mean age of 45.3 (SD 13.3) years and a median of 47.0 (IQR 21.0) years.

XSL•FO
RenderX

Figure 4. Demographic data statistics in the dataset. (A) The top 20 most common categories of injured body parts. (B) Most common categories in the industry. (C) Most common categories in the nature of loss. (D) Most common categories in the cause of injury. (E) Gender distribution in the dataset. (F) Age distribution in the dataset.



Performance Assessment of LLMs

Table 1 shows the LLM classification accuracy across the selected models for different data categories. All LLMs mentioned in Figure 3 have been tested. For the entire dataset, all selected LLMs achieved more than 70% accuracy in the maximax, expected maximum, and median cliff criteria methods. Among the selected LLMs, GPT-4o achieved the best performance in all cliff criteria, while DeepSeek-V3 had the lowest performance. In the separate anomaly dataset containing only anomalies, the prediction accuracy of all selected models exceeds 95%. The best performance for the nonanomalous dataset comes from GPT-4o as well, achieving more than 76% under maximax cliff criteria, over 73% accuracy under expected maximum cliff criteria, and a median cliff criterion.

XSL•FO
RenderX

Table . Case classification accuracy across models[a].

	Method 1 (%)	Method 2 (%)	Method 3 (%)
Qwen	78.71	78.26	78.19
DeepSeek	76.31	75.75	75.66
Llama	78.15	76.93	76.76
GPT-4o	81.77	79.95	79.51
Anomaly			
Qwen	97.72	97.79	97.79
DeepSeek	97.59	97.81	97.80
Llama	97.69	98.02	97.97
GPT-4o	96.32	97.16	97.21
Nonanomalous			
Qwen	71.71	71.07	70.98
DeepSeek	68.48	67.62	67.51
Llama	70.96	69.16	68.95
GPT-4o	76.40	73.61	72.99

[a] Method 1 uses the maximum value among the 3 large language models (LLMs) predictions as a cliff to classify anomalies and nonanomalous. Method 2 uses the average of the 3 LLM predictions as a cliff for classification. Method 3 uses the median of the 3 LLM predictions as a cliff for classification.

Table 1 further highlights the trade-offs in classification accuracy when different criteria are applied. When using the expected maximum as the classification criterion, the model achieves higher accuracy in anomaly detection but at the cost of reduced accuracy in nonanomalous classification. However, misclassifying nonanomalous cases is relatively less consequential, as such cases typically resolve quickly, with injured workers returning to work in a short period. In contrast, misclassifying anomalies can have significant financial and operational implications. Suppose an anomaly is incorrectly classified as a normal case. In that case, the company may need to allocate additional resources to reassign a senior work injury case manager later in the process, leading to prolonged recovery times and potentially missed rehabilitation windows. From an anomaly detection perspective, Llama demonstrates the highest detection rate. However, when considering overall performance across both anomaly and nonanomalous classification, GPT-4o outperforms other models, making it the most balanced and practical choice for real-world deployment. These findings highlight the importance of selecting an LLM that optimally balances accuracy, adherence to instruction, and overall classification performance.

Table 2 shows that across the 3 aggregation strategies, absolute and squared errors remain high: MAE is approximately 72 days, and RMSE is around 158 days for all methods, indicating substantial pointwise deviations and volatility in predicting the maximum sick leave duration. Relative errors are also large: MAPE ranges from 169.86% to 195.48%, and MPE exceeds 100% for all methods, evidencing pronounced systematic overestimation. Among the alternatives, the median-based cliff (method 3) yields the lowest relative error (MAPE=169.86%, MPE=106.77%) and slightly lower dispersion, whereas using the maximum prediction as the cliff (method 1) amplifies both bias and variance; the mean (method 2) lies in between. Despite these differences, the small gaps in MAE or RMSE across methods suggest that aggregation choice alone does not resolve the core error magnitude, and bias calibration or robustness enhancements are warranted. Directly using LLMs to predict sick leave duration from demographics yields large absolute and relative errors (≈72-day MAE, ≈158-day RMSE, MAPE >169% with systematic overestimation), revealing unstable and biased point forecasts, which motivates our shift to a fast exceedance−based detection method rather than relying on raw predictions.

XSL•FO
RenderX

Table . Standard metrics for 3 methods[a].

	MAE[b]	MSE[c]	RMSE[d]	MAPE[e] (%)	MPE[f] (%)
Method 1	72.37	24,741.36	157.29	195.48	136.39
Method 2	72.42	25,173.20	158.66	172.39	110.32
Method 3	72.54	25,266.41	158.95	169.86	106.77

[a] Method 1 uses the maximum value among the 3 large language models (LLMs) predictions as a cliff to classify anomalies and nonanomalous. Method 2 uses the average of the 3 LLM predictions as a cliff for classification. Method 3 uses the median of the 3 LLM predictions as a cliff for classification.

[b] MAE: mean absolute error.

[c] MSE: mean squared error.

[d] RMSE: root mean squared error.

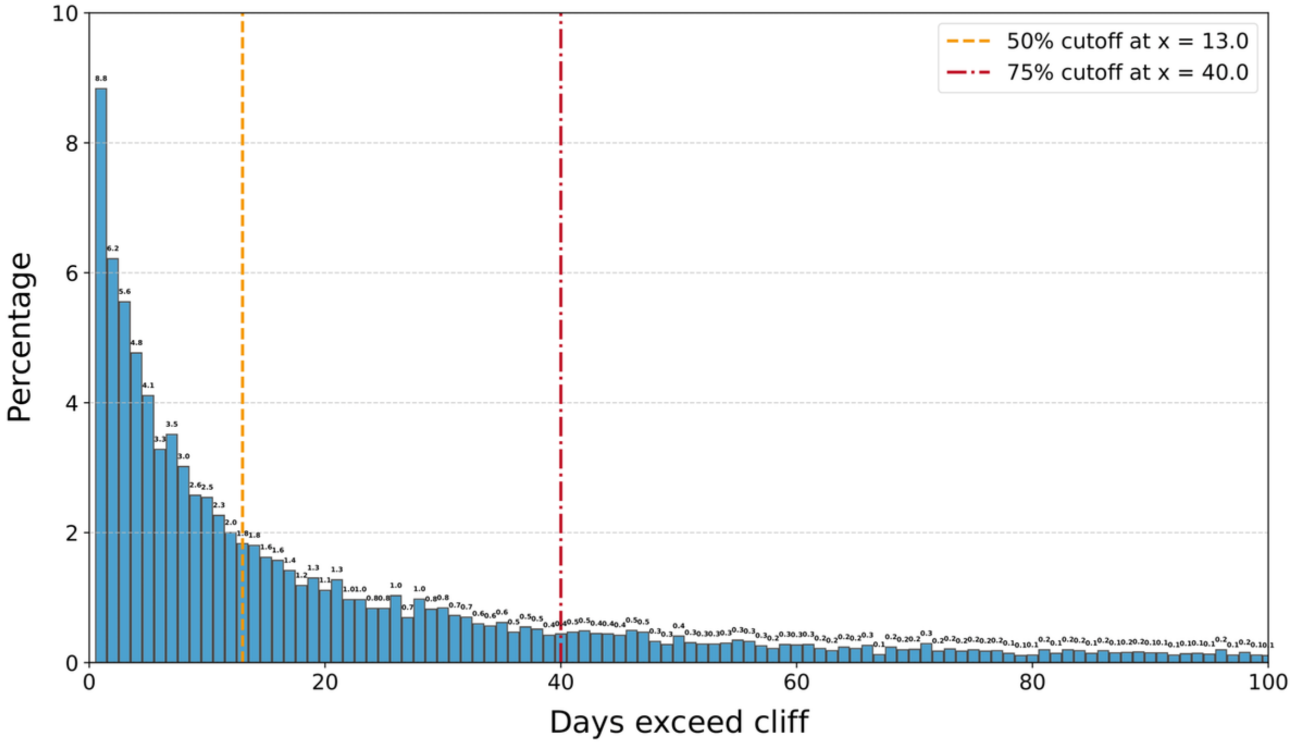
[e] MAPE: mean absolute percentage error.

[f] MPE: mean percentage error.

LLM Misclassifications Case Study

Since method 1 achieved the highest overall accuracy, this section primarily focuses on analyzing the predictions of the LLM based on method 1, as well as those generated by ChatGPT-4o, which demonstrated the highest average accuracy among all the selected models. Figure 5 illustrates the distribution of erroneous predictions that exceed the specified cliff when method 1 is applied. The yellow dashed line represents the 50th percentile of the cumulative distribution, corresponding to a deviation of 13 days. In contrast, the 75th percentile of the cumulative distribution corresponds to a deviation of 40 days. These results indicate that half of the erroneous predictions deviate from the ground truth by no more than 13 days, further underscoring the robustness of the proposed method. The previous work on similar datasets shows that the traditional variational autoencoder could only achieve an average error of 107.447 days, and they failed to predict a cliff for classifying the anomaly and nonanomalous cases [12,13]. Table 2 shows the absolute errors of direct LLM predictions, demonstrating substantial bias, while Figure 5 highlights erroneous cases flagged by our fast detection method, underscoring that our approach is intended to assist rather than replace rehabilitation coordinators.

Figure 5. Distribution of erroneous predictions beyond the cliff.



To gain a deeper insight into the LLM's prediction performance, we examined the distribution of GPT-4o outputs using method 1 (as described in the previous section). Specifically, 14,751 nonanomalous cases were misclassified as anomalies, and 847 anomalies were misclassified as normal. Figure 6 illustrates the distribution of the key variables within these misclassified nonanomalous cases. Finger has the highest proportion of misclassifications, at approximately 23%, whereas other body parts each account for around 10%. In the "Nature of Loss" variable, "Sprain & Strain" accounts for over 28% of the misclassifications, followed by "Contusion & Bruise" and "Laceration & Cut," both of which exceed 15%; the remaining

XSL•FO
RenderX

categories each fall below 10%. Regarding the Industry, the Administrative and Support Service sector and the Accommodation and Food Service Activities sector exhibit notably higher proportions of misclassifications (over 20%) relative to others. Finally, in the “Position Category” variable, most misclassifications occur under the “unknown” category.

Figure 6. Distribution of large language model misclassifications nonanomalous across the key variables (top 10).

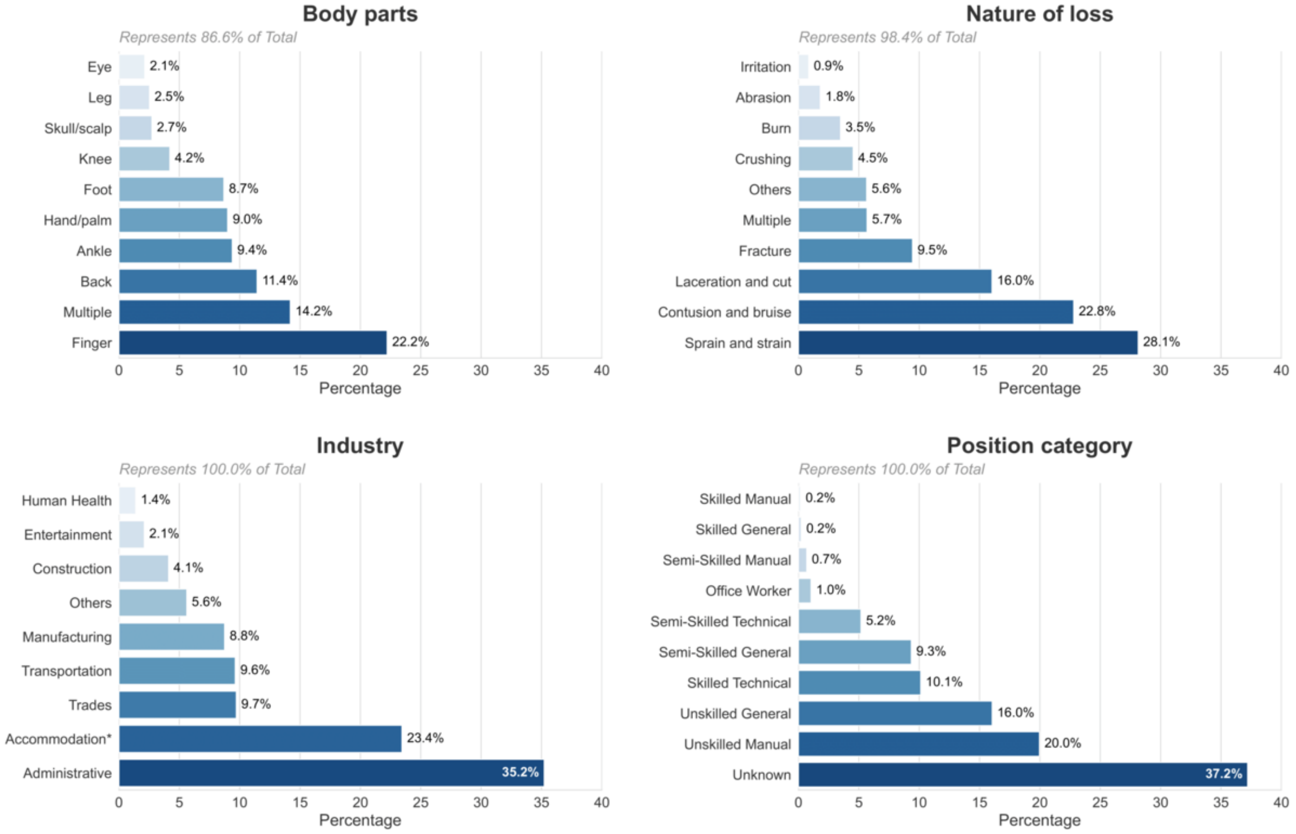


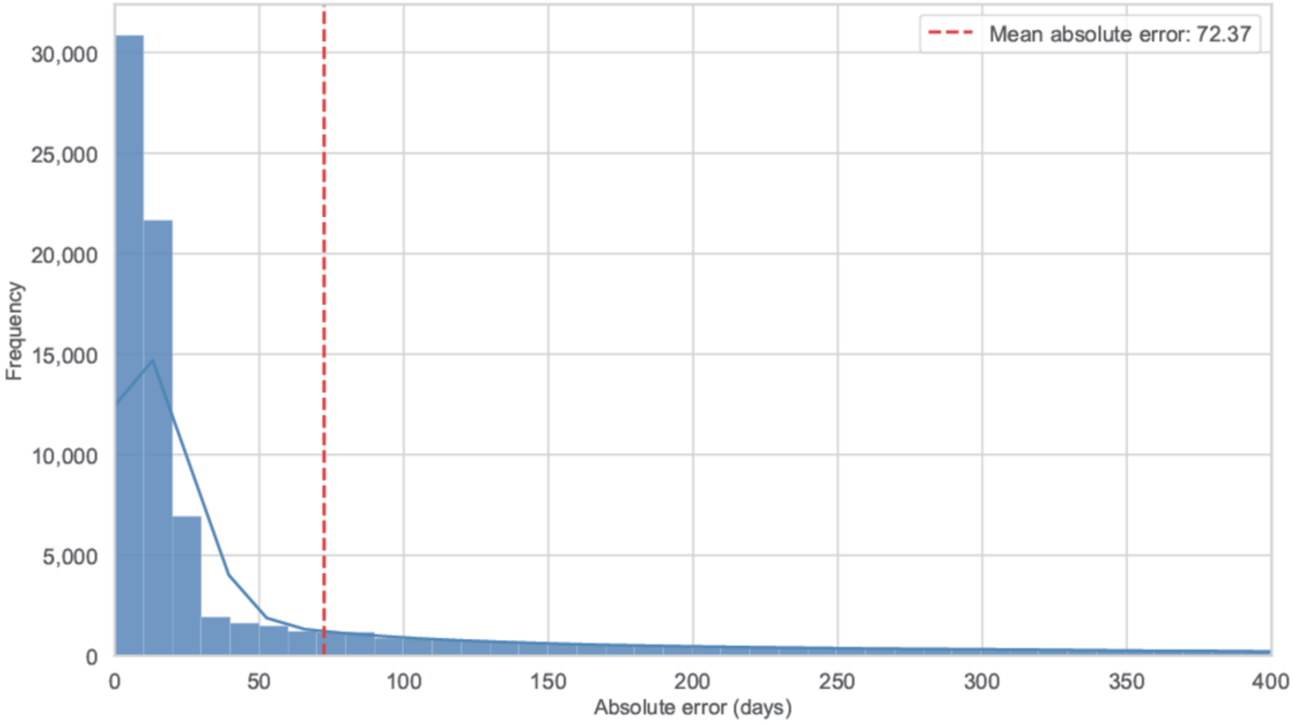
Figure 6 also suggests that certain misclassifications may arise from incomplete data, such as the absence of a position category, which impairs the LLM’s ability to predict the cliff accurately. Without this critical information, the model must rely on its intrinsic knowledge, leading to increased variability and uncertainty. Additionally, for variables such as the nature of loss, injured body parts, and industry type, human judgment is also prone to significant bias, particularly in those variables with high percentages of misclassification [13].

LLM Direct Prediction Error Distribution

Figure 7 shows the distribution of absolute errors for raw LLM day predictions; the mean error is 72 days, which indicates that direct prediction yields deviations too large for practical use. Therefore, we use the LLM as a fast anomaly filter that flags cases whose realized sick leave exceeds an estimated upper bound, rather than predicting sick leave days directly.

XSL•FO
RenderX

Figure 7. Distribution of the large language model raw prediction absolute error.



Case Study of Our Approach

In this case study, we analyze a 51-year-old female worker who sustained a back injury while lifting a heavy basket of packaged bread. The injury was classified under Body Parts: Back, Nature of Loss: Sprain & Strain, and Cause: Injured whilst lifting or carrying, with the worker employed as a packer in the Accommodation and Food Service Activities industry. The actual sick leave duration was 6 days, while the LLM predictions for recovery duration were consistent across all methods: result 1=14 days, result 2=14 days, and result 3=14 days, leading to cliff values of max_max=14 days.

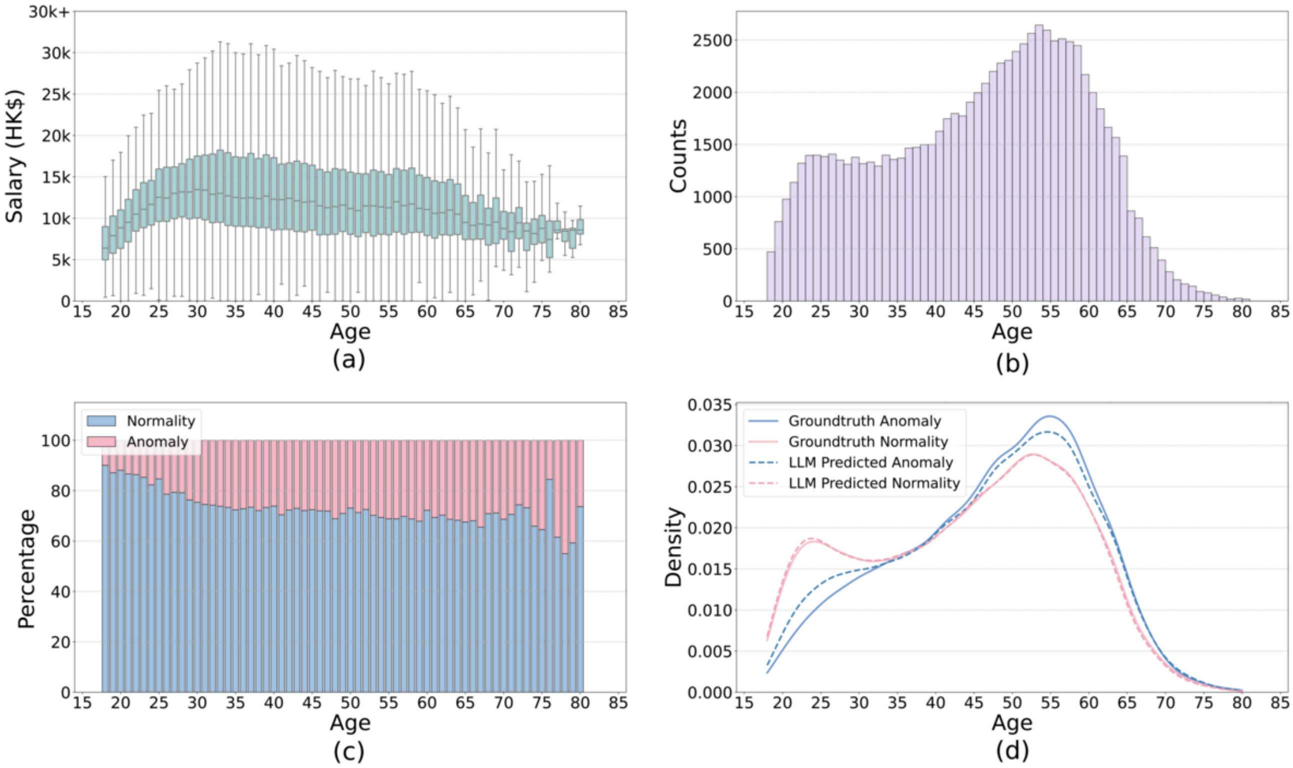
The difference between the actual sick leave and the maximum predicted cliff sick leave duration max_max was –8 days, indicating that the actual recovery period was well within the predicted range and did not exceed the anomaly cliff. This demonstrates the effectiveness of the LLM-assisted method in identifying cases that conform to expected recovery patterns, ensuring that resources are not misallocated to cases that align closely with typical recovery trajectories.

Exploratory Insights From the Dataset

To examine anomalies, we present a salary box plot against age in Figure 8. The top-left panel of the figure presents a box plot of salary against age, where the median salary for each age group is extracted to reveal a clearer trend. The median salary initially increases, peaking between ages 35 and 40, before gradually declining. The top-right panel illustrates the distribution of age counts from 18 to 80 years, showing an initial peak in the early 20s, followed by a decline until the 40s, after which the number of cases rises again, reaching its highest peak around the age of 55 years. The bottom-left panel displays the percentage of anomalies across different age groups, revealing a steady increase in anomaly prevalence with age. Finally, the bottom-right panel illustrates the normalized age distributions of both anomalies and nonanomalous, alongside the LLM's predicted anomalies and nonanomalous. The red line (representing nonanomalous) closely mirrors the overall age distribution, while the blue line (representing anomalies) steadily increases and peaks near the age of 60 years. Notably, the LLM's predicted distributions align closely with the ground-truth trends.

XSL•FO
RenderX

Figure 8. Overview of salary and age distributions with anomaly analysis. (A) Salary boxplot across ages. (B) Age distribution of all data. (C) Percentage of anomaly across ages. (D) Age distribution of anomaly or normality data.



Discussion

Principal Findings

The performance of the proposed LLM-assisted anomaly fast detection method demonstrates promising results, 110,346 deidentified work-injury cases between 2001 and 2024 from a leading work injury management company in Hong Kong, representing approximately 20% of all work injury incidents in the region. Compared to previous research, which primarily assessed AI prediction accuracy by comparing it to work injury case managers' judgments, our study provides a more comprehensive evaluation [12]. Prior studies demonstrated that AI predictions could surpass human work injury case managers in both nonanomalous and anomalous cases. However, while they also attempted to predict anomaly cliffs, their findings did not explicitly report the accuracy of such predictions. In contrast, our approach ensures that the accuracy of anomaly cliff prediction is systematically analyzed, contributing to a more reliable anomaly detection framework.

In Figure 7, we illustrate the trend of salary changes, considering salary as a key indicator of work experience. It is well known that physical ability peaks in the 20s, remains stable or slightly declines until around 30 - 35, and then drops significantly thereafter [33,34]. Comparing this physical ability trend with the salary trend, we observe that workers in their early 20s possess peak physical strength but lack experience. As a result, they may be more prone to injuries; however, their quick recovery often prevents these cases from becoming anomalies. This explains the initial peak in injury cases around the age of 20 years. By the time workers reach their 30s, they have gained significant experience (as indicated by higher salaries), while their physical ability has not yet declined substantially. Consequently, the number of work injury cases decreases between the ages of 30 and 40 years. However, after 40, salaries may begin to decline slightly as workers are unable to maintain the same working hours as before, while their physical ability drops sharply. This results in a second peak in work-related injury cases. At this stage, injuries are more severe, and recovery is less likely, contributing to an increase in both the number and proportion of anomalies.

Limitations and Future Directions

This study tests LLMs for spotting unusual rehabilitation cases. The models read injury diagnoses and accident notes, estimate a normal recovery time, and flag cases that fall outside that range. We judge the approach by how well it separates nonanomalous from anomalies, not by exact prediction accuracy. However, the limitations are that no comparison is made with human rehab coordinators, only general-purpose LLMs without medical models, and data from a single region; therefore, the findings may not be generalizable. As LLMs improve, they could streamline the rehabilitation triangle and resource planning. This research can be readily transferred to other rehabilitation systems within Hong Kong, but adaptation to other regions would be domain specific due to significant differences in work injury frameworks, legal and policy requirements, and cultural practices.

Deterministic decoding with temperature 0.0 could further enhance reproducibility. As a robustness check, future or supplementary analyses can compare performance and schema compliance at 0.0 versus 0.2 to quantify any trade-off between determinism and adherence to output constraints. If comparable, deployment would favor 0.0 to maximize reproducibility; if not, a small nonzero temperature remains justified to preserve formatting reliability under operational constraints.

Conclusions

We present an LLM-based approach that estimates expected recovery time from injury records and flags deviations as anomalies, streamlining rehabilitation triage. Tested on more than 110,000 Hong Kong work-injury cases, the method improved classification efficiency; GPT-4o delivered the most balanced accuracy, with DeepSeek-V3 and Qwen2·5-72B Instruct close behind. Demographic analysis reveals that injuries are more frequent yet milder in younger workers, whereas those aged 40 and above experience more anomalies, reflecting reduced resilience. The approach advances data-driven rehabilitation coordination and optimizes resource allocation.

Funding

We acknowledge the support of the Innovation and Technology Fund (ITP/039/24LP) of the Government of the Hong Kong Special Administrative Region for providing the support.

Data Availability

Medical notes and the full list of responses from large language models will not be shared to protect the privacy of the patients. The responses of the large language models or derived data are available from the corresponding author.

Authors' Contributions

PQC and PHFN conceived the study, oversaw its administration, and provided supervision. PQC, HYWG, CJML, and SHSL accessed, verified, and curated the data; performed the formal analysis and methodology; developed the software; validated the results; and produced the visualizations. PQC prepared the original draft, while HYWG, HKYL, WCC, ASKC, and PHFN reviewed and edited the manuscript. PHFN additionally secured funding and supplied resources. All authors had full access to all study data and approved the final version for publication.

Conflicts of Interest

None declared.

References

1. Szeto GPY, Cheng ASK, Lee EWC, Schonstein E, Gross DP. Implementing the work disability prevention paradigm among therapists in Hong Kong: facilitators and barriers. J Occup Rehabil 2011 Mar;21(1):76-83. [doi: 10.1007/s10926-010-9256-2] [Medline: 20652377]
2. Remus L, Grope M, Lemke S, Bethge M. An innovative case management intervention for people at high risk of permanent work disability to improve rehabilitation coverage and coordination of health services: a randomized controlled trial (AktiFAME, DRKS00024648). BMC Health Serv Res 2022 Mar 15;22(1):342. [doi: 10.1186/s12913-022-07482-9] [Medline: 35292005]
3. Brownlee S, Chalkidou K, Doust J, et al. Evidence for overuse of medical services around the world. Lancet 2017 Jul 8;390(10090):156-168. [doi: 10.1016/S0140-6736(16)32585-5] [Medline: 28077234]
4. Cooper DA, Slaughter JE, Gilliland SW. Reducing injuries, malingering, and workers' compensation costs by implementing overt integrity testing. J Bus Psychol 2021 Jun;36(3):495-512. [doi: 10.1007/s10869-020-09681-9]
5. Paid sick and carer's leave. Fair Work Ombudsman. URL: https://www.fairwork.gov.au/leave/sick-and-carers-leave/paid-sick-and-carers-leave [accessed 2025-12-11]
6. Management of staff leave in the Australian public service. : Australian National Audit Office (ANAO); 2022 URL: https://www.anao.gov.au/work/performance-audit/management-staff-leave-the-australian-public-service [accessed 2025-12-11]
7. Tripartite standards: frequently asked questions (FAQs). TAFEP. 2024. URL: https://www.tal.sg/tafep/-/media/tal/tafep/getting-started/files/tripartite-standards-faqs [accessed 2025-11-12]
8. Lutz W, Deisenhofer AK, Rubel J, et al. Prospective evaluation of a clinical decision support system in psychological therapy. J Consult Clin Psychol 2022 Jan;90(1):90-106. [doi: 10.1037/ccp0000642] [Medline: 34166000]
9. Antoniadi AM, Du Y, Guendouz Y, et al. Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review. Appl Sci 2021;11(11):5088. [doi: 10.3390/app11115088]
10. Musen MA, Middleton B, Greenes RA. Clinical decision-support systems. In: Biomedical Informatics: Computer Applications in Health Care and Biomedicine: Springer; 2021:795-840. [doi: 10.1007/978-3-030-58721-5_24]
11. Borges do Nascimento IJ, Abdulazeem HM, Vasanthan LT, et al. The global effect of digital health technologies on health workers' competencies and health workplace: an umbrella review of systematic reviews and lexical-based and sentence-based meta-analysis. Lancet Digit Health 2023 Aug;5(8):e534-e544. [doi: 10.1016/S2589-7500(23)00092-4] [Medline: 37507197]

XSL•FO
RenderX

12. Ng PHF, Chen PQ, Sin ZPT, Lai SHS, Cheng ASK. Smart Work Injury Management (SWIM) system: a machine learning approach for the prediction of sick leave and rehabilitation plan. Bioengineering (Basel) 2023 Jan 28;10(2):172. [doi: 10.3390/bioengineering10020172] [Medline: 36829666]
13. Yeung YYK, Chen PQ, Ng PHF, Cheng ASK. Evaluation of the accuracy of the Smart Work Injury Management (SWIM) system to assist case managers in predicting the work disability of injured workers. J Occup Rehabil 2025 Jun;35(2):320-332. [doi: 10.1007/s10926-024-10199-7] [Medline: 38874680]
14. Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023 Feb 8;9(1):e45312. [doi: 10.2196/45312] [Medline: 36753318]
15. Malgaroli M, Schultebraucks K, Myrick KJ, et al. Large language models for the mental health community: framework for translating code to care. Lancet Digit Health 2025 Apr;7(4):e282-e285. [doi: 10.1016/S2589-7500(24)00255-3] [Medline: 39779452]
16. Deiner MS, Deiner RY, Fathy C, et al. Use of large language models to classify epidemiological characteristics in synthetic and real-world social media posts about conjunctivitis outbreaks: infodemiology study. J Med Internet Res 2025 Jul 2;27:e65226. [doi: 10.2196/65226] [Medline: 40601927]
17. Huang W, Wei W, He X, et al. ChatGPT-assisted deep learning models for influenza-like illness prediction in mainland China: time series analysis. J Med Internet Res 2025 Jun 27;27:e74423. [doi: 10.2196/74423] [Medline: 40577658]
18. Huang J, Lai H, Zhao W, et al. Large language model–assisted risk-of-bias assessment in randomized controlled trials using the revised risk-of-bias tool: evaluation study. J Med Internet Res 2025 Jun 24;27:e70450. [doi: 10.2196/70450] [Medline: 40554779]
19. Li N, Gao C, Li M, Li Y, Liao Q. EconAgent: large language model-empowered agents for simulating macroeconomic activities. 2024 Presented at: Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); Aug 11-16, 2024; Bangkok, Thailand p. 15523-15536 URL: https://aclanthology.org/2024.acl-long [accessed 2025-12-12]
20. Zhang Z, Zhang-Li D, Yu J, et al. Simulating classroom education with LLM-empowered agents. Presented at: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics; Apr 29 to May 4, 2025; Albuquerque, NM p. 10364-10379 URL: https://aclanthology.org/2025.naacl-long [accessed 2025-12-12] [doi: 10.18653/v1/2025.naacl-long.520]
21. Wang L, Ma C, Feng X, et al. A survey on large language model based autonomous agents. Front Comput Sci 2024 Dec;18(6):186345. [doi: 10.1007/s11704-024-40231-1]
22. Longwell JB, Hirsch I, Binder F, et al. Performance of large language models on medical oncology examination questions. JAMA Netw Open 2024 Jun 3;7(6):e2417641. [doi: 10.1001/jamanetworkopen.2024.17641] [Medline: 38888919]
23. Occupational safety and health statistics. Labour Department. URL: https://www.labour.gov.hk/tc/osh/content10.htm [accessed 2025-12-11]
24. Wang X, Wei J, Schuurmans D, et al. Self-consistency improves chain of thought reasoning in language models. arXiv. Preprint posted online on Mar 21, 2022. [doi: 10.48550/arXiv.2203.11171]
25. Brown TB, Mann B, Ryder N, et al. Language models are few-shot learners. 2020 Presented at: 34th Conference on Neural Information Processing Systems (NeurIPS 2020); Dec 6-12, 2020; Vancouver, Canada p. 1877-1901 URL: https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf [accessed 2025-12-11]
26. Shanahan M, McDonell K, Reynolds L. Role play with large language models. Nature 2023 Nov;623(7987):493-498. [doi: 10.1038/s41586-023-06647-8] [Medline: 37938776]
27. Dong Q, Li L, Dai D, et al. A survey on in-context learning. arXiv. Preprint posted online on Dec 31, 2022. [doi: 10.48550/arXiv.2301.00234]
28. Kojima T, Gu SS, Reid M, Matsuo Y, Iwasawa Y. Large language models are zero-shot reasoners. 2022 Presented at: 36th Conference on Neural Information Processing Systems (NeurIPS 2022); Nov 28 to Dec 9, 2022; New Orleans, LA p. 22199-22213 URL: https://openreview.net/pdf?id=e2TBb5y0yFf [accessed 2025-12-12]
29. Yu F, Quartey L, Schilder F. Exploring the effectiveness of prompt engineering for legal reasoning tasks. Presented at: Findings of the Association for Computational Linguistics; Jul 9-14, 2023; Toronto, Canada p. 13582-13596 URL: https://pdfs.semanticscholar.org/9661/dda8024192342097647c37423eb7fccee298.pdf [accessed 2025-12-12]
30. Achiam J, Adler S, Agarwal S, et al. GPT-4 technical report. arXiv. Preprint posted online on Mar 15, 2023. [doi: 10.48550/arXiv.2303.08774]
31. Hoffmann J, Borgeaud S, Mensch A, et al. Training compute-optimal large language models. arXiv. Preprint posted online on Mar 29, 2022. [doi: 10.48550/arXiv.2203.15556]
32. Wang JJ, Wang VX. Assessing consistency and reproducibility in the outputs of large language models: evidence across diverse finance and accounting tasks. SSRN. Preprint posted online on Mar 21, 2025. [doi: 10.2139/ssrn.5189069]
33. Burr DB. Muscle strength, bone mass, and age-related bone loss. J Bone Miner Res 1997 Oct;12(10):1547-1551. [doi: 10.1359/jbmr.1997.12.10.1547] [Medline: 9333114]
34. Westerterp KR. Changes in physical activity over the lifespan: impact on body composition and sarcopenic obesity. Obes Rev 2018 Dec;19 Suppl 1:8-13. [doi: 10.1111/obr.12781] [Medline: 30511504]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
API: application programming interface
LLM: large language model
MAE: mean absolute error
MAPE: mean absolute percentage error
MPE: mean percentage error
RMSE: root mean squared error

Edited by S Munce; submitted 14.Jul.2025; peer-reviewed by AA Barr, E Fu, VX Wang; accepted 19.Nov.2025; published 30.Jan.2026.

Please cite as:
Chen PQ, Gu HYW, Lo HKY, Chang WC, Lai CJM, Lai SHS, Cheng ASK, Ng PHF
Leveraging Large Language Models for Early Detection of Anomaly Work Injury Cases: Data-Driven Approach to Rehabilitation Efficiency
JMIR Rehabil Assist Technol 2026;13:e80607
URL: https://rehab.jmir.org/2026/1/e80607
doi:10.2196/80607

© Peter Q Chen, Hayley Y W Gu, Heidi K Y Lo, Wing Chung Chang, Cameron J M Lai, Sun H S Lai, Andy S K Cheng, Peter H F Ng. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predicting Geriatric Rehabilitation Stays of ≤4 Weeks After Hip Fracture Surgery: Machine Learning Approach Using Physical Activity and Patient Data

Sanne M Krakers[1,2], MSc; Frank J Wouda[1], PhD; Dieuwke van Dartel[1,3], MSc; Miriam MR Vollenbroek-Hutten[1,4], Prof Dr; Johannes H Hegeman[1,2], MD, Prof Dr; Up&Go After a Hip Fracture Group[5]

[1]Department of Biomedical Signals and Systems, University of Twente, Enschede, The Netherlands
[2]Department of Trauma Surgery, Ziekenhuisgroep Twente, Almelo, The Netherlands
[3]Reggeborgh Research Fellow, ZGT Academy, Ziekenhuisgroep Twente, Almelo, The Netherlands
[4]Board of Directors, Medisch Spectrum Twente, Enschede, The Netherlands
[5]See Acknowledgments

Corresponding Author:
Sanne M Krakers, MSc
Department of Biomedical Signals and Systems
University of Twente
Drienerlolaan 5
Enschede, 7522 NB
The Netherlands
Phone: 31 (0)53 489 9111
Email: s.m.krakers@utwente.nl

Abstract

Background: In 2022, over 18,000 patients aged ≥70 years were hospitalized in the Netherlands for a hip fracture, with 50% requiring geriatric rehabilitation after surgery. Increasing geriatric rehabilitation patient numbers, staff shortages, and rising pressure on health care budgets make adequate care challenging. To make geriatric rehabilitation more future-proof, a stronger focus on home-based rehabilitation is needed. Early identification of patients likely to be discharged soon enables timely discharge planning and coordination of support at home. Early geriatric rehabilitation discharge planning may help organize home-based rehabilitation more effectively by arranging home care services in advance. This can facilitate smoother transitions toward home and prevent discharge delays, which is important to ensure optimal bed occupancy.

Objective: This study aims to develop machine learning (ML) models to predict a geriatric rehabilitation stay of 4 weeks or less in a skilled nursing home for older patients after hip fracture surgery, using continuously monitored physical activity data from the first week of geriatric rehabilitation and patient characteristics.

Methods: This prospective cohort study (January 2019-August 2024) included 100 patients. Patient characteristics and physical activity data from the MOX1 accelerometer (Maastricht Instruments BV) were collected during the first rehabilitation week. Principal component analysis was used to reduce the physical activity features. ML models were developed using Bayesian hyperparameter optimization and refined if necessary. The performance of the single best-performing configuration per remaining ML model type was evaluated, and the most important features for predicting the length of geriatric rehabilitation stay were identified.

Results: Of the 3 ML models evaluated (support vector machine [SVM], ensemble of decision trees, and neural network), the SVM achieved the highest performance, with 19 out of 20 correct predictions (accuracy=0.95, 95% CI 0.85-1.00; precision=0.91, 95% CI 0.71-1.00; recall=1.00, 95% CI 1.00-1.00; F_1-score=0.95238, 95% CI 0.83-1.00; area under the curve [AUC]=0.97, 95% CI 0.83-1.00). The most important features for predicting the length of geriatric rehabilitation stay across the best-performing ML models included the continuously monitored physical activity data, time in the emergency room, functional ambulation category (FAC) at hospital discharge, age, Katz Index of Independence in Activities of Daily Living–6 (Katz-ADL6) at hospital discharge, Montreal Cognitive Assessment (MoCA), availability of nonprofessional help, surgery type, Charlson Comorbidity Index (CCI), gender, and hemoglobin level at hospital admission.

XSL•FO
RenderX

Conclusions: This study developed several ML models to predict a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home for older patients after hip fracture surgery. Among these models, the SVM proved to be highly accurate in its predictions with an accuracy of 0.95 (95% CI 0.85-1.00), precision of 0.91 (95% CI 0.71-1.00), recall of 1.00 (95% CI 1.00-1.00), F_1-score of 0.95 (95% CI 0.83-1.00), and AUC of 0.97 (95% CI 0.88-1.00).

(JMIR Rehabil Assist Technol 2026;13:e79331) doi:10.2196/79331

KEYWORDS

accelerometer; continuous physical activity monitoring; geriatric rehabilitation; hip fracture; length of stay; machine learning; prediction

Introduction

In the Netherlands, more than 18,000 patients aged 70 years and older were hospitalized with a hip fracture in 2022 [1]. After hip fracture surgery, more than 50% of the patients are admitted to a geriatric rehabilitation department at a skilled nursing home [2,3]. Geriatric rehabilitation is defined by the European Consensus group as “a multidimensional approach involving diagnostic and therapeutic interventions. Its purpose is to optimize functional capacity, promote activity, and preserve functional reserve and social participation in older people with disabling impairments” [4]. Based on incidence trends and increasing life expectancy, the number of hip fracture patients is expected to double by 2050 [5]. The expectation is that the group of geriatric rehabilitation patients will also increase. However, the number of professionals able to provide care for these patients will not grow at the same pace. In addition, health care budgets are under increasing pressure [6]. Together, these developments make the provision of geriatric rehabilitation care increasingly challenging.

To make geriatric rehabilitation future-proof, a stronger focus on home-based rehabilitation is essential, especially given the shifting vision of Dutch health care that emphasizes treatment closer to home [7]. Achieving this requires the early identification of patients who are likely to be discharged within 4 weeks of geriatric rehabilitation admission. Such early identification enables timely discharge planning and the coordination of necessary support at home. The 4-week cutoff marks the transition between shorter and longer rehabilitation trajectories, as defined by the Dutch Diagnosis Treatment Combinations. Early geriatric rehabilitation discharge planning may help organize home-based rehabilitation more effectively by arranging home care services in advance. This can facilitate smoother transitions toward home and prevent unnecessary discharge delays, which is important for maintaining optimal bed occupancy.

Early geriatric rehabilitation discharge planning requires prediction models to make objective predictions for the length of geriatric rehabilitation stay. By relying on data rather than assumptions, these models enable more accurate discharge planning. Machine learning (ML) is a subset of artificial intelligence that enables computers to learn from data without being explicitly programmed [8-10]. The use of ML for developing length of stay prediction models in various medical fields has shown promising results [8,10-12]. However, to the best of our knowledge, no ML models exist that predict the length of geriatric rehabilitation stay at a skilled nursing home for older patients following hip fracture surgery. In a recent study, we identified factors that influence the length of geriatric rehabilitation stay [13]. However, this study did not aim to predict the exact length of geriatric rehabilitation stay, and its multivariate model explained only 32% of the variance. This suggests that additional factors influencing the length of stay exist that were not included in the study. Since early mobilization and physical activity enhance functional recovery, continuously monitored physical activity using accelerometers during the first week of geriatric rehabilitation might be a relevant factor to include [14-16]. Another important factor could be cognitive functioning, assessed with a cognitive screening test, as patients with cognitive impairments may have difficulty performing physiotherapy exercises [17].

The aim of this study is to develop ML models to predict a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home for older patients after hip fracture surgery, using continuously monitored physical activity data from the first week of geriatric rehabilitation and patient characteristics.

Methods

Study Design

A prospective cohort study was conducted from the first of January 2019 until the first of August 2024. Older patients aged ≥70 years who underwent hip fracture surgery at the Centre for Geriatric Traumatology at Ziekenhuisgroep Twente (ZGT) and rehabilitated at one of the 3 participating skilled nursing homes (TriviumMeulenbeltZorg, ZorgAccent, and Carintreggeland) were included. Exclusion criteria were pathological or periprosthetic fractures, severe cognitive impairment, total hip replacement, plaster allergy, or contact isolation. Patients were enrolled one day after hip fracture surgery or one day before discharge to one of the collaborating skilled nursing homes.

Data Collection

The primary outcome was the length of geriatric rehabilitation stay at the skilled nursing homes, stratified into 2 lengths of stay groups, including ≤4 and >4 weeks. This division was based on the structure of the Dutch Diagnosis Treatment Combinations, which fund geriatric rehabilitation and classify rehabilitation trajectories into one of 5 length of stay groups (1-2, 2-4, 4-8, 8-13, and 13-17 weeks) [18,19]. The 4-week cutoff corresponds to the transition point between shorter and longer rehabilitation pathways.

For this study, the physical activity of all enrolled patients was continuously monitored during the first week of geriatric

XSL•FO
RenderX

rehabilitation. Patient characteristics up to the first week of rehabilitation were collected from a transmural care pathway database. This care pathway was implemented to synchronize the care processes between the hospital and geriatric rehabilitation departments and provide more insight into the rehabilitation process of older patients after hip fracture surgery. Patients with an incomplete data collection (<5 days of physical activity data of the first week of geriatric rehabilitation and/or missing patient characteristics) were excluded from the data analysis.

The physical activity of all enrolled patients was continuously monitored using an MOX1 accelerometer (Maastricht Instruments BV) attached with a custom-made patch 10 cm proximal of the patella on the ipsilateral side of the operated hip. This small, lightweight, waterproof device (35x35x10 mm, 11 g, IPX8) contains a tri-axial accelerometer sensor, which measures raw acceleration data for the X-, Y-, and Z-axes at a sampling frequency of 25 Hz, storing the data directly in its internal memory (1.5 GB) for up to 7 days. Data analysis was performed offline after the raw acceleration data were uploaded to a computer via a USB connection.

The collected patient characteristics (baseline, in-hospital, hospital discharge, and geriatric rehabilitation variables; Textbox 1) were selected based on previous studies demonstrating their value in predicting rehabilitation outcomes or length of stay [20-31].

Textbox 1. Patient characteristics included in this study.

Baseline variables

- Age
- Gender
- Premorbid living situation
- Living alone versus living together
- Availability of nonprofessional help
- Need for climbing stairs
- Prefracture Mobility Score (PFMS)
- Katz Index of Independence in Activities of Daily Living-6 (Katz-ADL6)
- Charlson Comorbidity Index (CCI)
- American Society of Anesthesiologists physical status classification (ASA) score
- Short Nutritional Assessment Questionnaire (SNAQ)

In-hospital variables

- Hemoglobin level at hospital admission
- Hip fracture type
- Time spent in the emergency room
- Surgical treatment
- Postoperative weight-bearing protocol
- In-hospital complications

Hospital discharge variables

- Functional Ambulation Categories (FAC) at hospital discharge
- Fracture Mobility Score (FMS) at hospital discharge
- Katz-ADL6 at hospital discharge
- Discharge destination (skilled nursing home)
- Length of hospital stay

Geriatric rehabilitation variables

- Montreal Cognitive Assessment (MoCA)

Data Analysis for Accelerometer

The raw acceleration data were analyzed using a MATLAB algorithm (R2024b; The MathWorks, Inc) developed and validated for older hospitalized patients by Van Dijk-Huisman et al [32]. This resulted in the total intensity of physical activity per day by determining the signal magnitude area of the acceleration data; the total time spent sitting or lying, standing,

XSL•FO
RenderX

and walking per day; the position classification (sitting or lying, standing, walking) every 4 seconds; as well as the total measurement time per day. The intensity of physical activity reflects the raw accelerometer output and is expressed as counts per day, not as categorical levels of activity. If the position classification changed between 2 consecutive 4-second time steps, it was defined as a transition. For each patient, the following metrics were calculated for each day:

- The intensity of physical activity (counts per hour)
- Time spent sitting or lying (minutes per hour)
- Time spent standing (minutes per hour)
- Time spent walking (minutes per hour)
- Transitions between sitting or lying and standing (number of transitions per hour)
- Transitions between sitting or lying and walking (number of transitions per hour)
- Transitions between standing and walking (number of transitions per hour)

Different features were extracted from each metric, including the overall mean, overall SD, weekend mean, weekend SD, weekday mean, weekday SD, median, IQR, root mean square, variance, minimum value, maximum value, minimum-maximum range, coefficient of variance, skewness, kurtosis, and characteristics of the third-degree polynomial curve describing the shape of the pattern. The characteristics of the third-degree polynomial curve were defined as coefficients a, b, c, and d, which were derived from the third-order polynomial equation: $y = ax^3 + bx^2 + cx + d$. More detailed information about the different continuously monitored physical activity features can be found in Multimedia Appendix 1.

After data analysis, data within each of the 2 length-of-stay groups were randomly divided into a training (80%) and testing (20%) set, which were then combined to ensure class balance.

Feature Reduction

Principal component (PC) analysis, a common method for feature dimensionality reduction, was applied exclusively to the features extracted from the physical activity data of the training set. As a preprocessing step, the data were standardized to have zero mean and unit variance. Each PC, obtained after applying PC analysis, consists of a weighted combination of the features extracted from the physical activity data. The PCs were sorted in decreasing order of their eigenvalues, indicating how much of the data's variance is captured by its corresponding PC. The PCs contributing to a cumulative explained variance of 95% were selected as input for the ML models.

Feature selection was also applied to the patient characteristics of the training set. Features that did not occur or exhibited an identical distribution between the 2 lengths of stay groups were excluded, as they provided no discriminative value for the ML models. Feature selection was further guided by prior research [13]. When variables were conceptually linked or measured similar aspects of a patient's condition, priority was given to the variable that was more commonly used in clinical practice or provided a more direct assessment of functional status. The remaining patient characteristics were incorporated into the ML models along with the selected PCs.

ML Models

First, the performance of 5 commonly used supervised ML models for classification was considered: support vector machine (SVM), neural network (NN), decision tree, naïve Bayes (NB), and an ensemble of decision trees (EoDT). These models were selected because they represent a diverse set of widely used supervised learning models, allowing for a comprehensive evaluation of which modeling strategy is most suitable for this clinical prediction task. For each model, hyperparameters were optimized using Bayesian optimization with an expected-improvement-plus acquisition function and tenfold cross-validation on the training set. Bayesian optimization was chosen because it enables efficient hyperparameter tuning and allows for the development and comparison of multiple ML models simultaneously.

For the SVM, 4 kernel functions (linear, radial basis function, polynomial, and Gaussian) were evaluated separately, allowing the kernel-related model performance to be systematically compared. For each kernel, Bayesian optimization tuned the remaining hyperparameters. For the NN, decision tree, NB, and EoDT, all model-specific hyperparameters were included in the Bayesian optimization procedure. The optimal hyperparameter values for each ML model were determined by the minimum observed value of the objective function.

Running Bayesian optimization for all models over many iterations can be computationally expensive. Therefore, model performance was evaluated after the first 30 iterations to identify a subset of well-performing models. In this screening phase, multiple configurations of the same model type, for example, the SVM models with different kernels, could be retained if they showed similar performance. The area under the receiver operating characteristic curve was used for initial model comparison, after which the models with the highest area under the curve (AUC) underwent continuous Bayesian optimization for an additional 30 iterations. Subsequently, the final hyperparameters of these ML models were evaluated. If notable deviations were identified, a new ML model was trained with adjusted hyperparameters.

Finally, only the single best-performing configuration per remaining ML model type was evaluated on the test set. Model performance was assessed using the confusion matrix, actual length of stay of the misclassified patients, accuracy score, precision score, recall score, F_1-score, and AUC. CIs were obtained using bootstrap resampling; the test set was resampled with replacements 1000 times while keeping the training set and ML models fixed. The 95% CIs were defined by the 2.5th and 97.5th percentiles. A Shapley Additive Interpretation (SHAP) summary plot based on the test set was used to determine the relationship between the 2 length of stay groups and its main predictors for each of the best-performing ML models [33]. The overall best-performing ML model was retrained and retested using 2 additional dataset configurations, (1) patient characteristics only and (2) physical activity data only, to conduct an ablation study evaluating the contribution of each dataset compared with the combined datasets.

XSL•FO
RenderX

Ethical Considerations

The Medical Ethical Committee of Twente stated that this study did not require an assessment by the Medical Ethical Committee according to Dutch law. The study was approved by the Institutional Review Board of ZGT (ZGT17-40). All patients gave written informed consent to participate. Study data obtained from the enrolled patients were deidentified and managed in accordance with relevant data protection regulations. Patients did not receive any financial or material compensation for their participation.

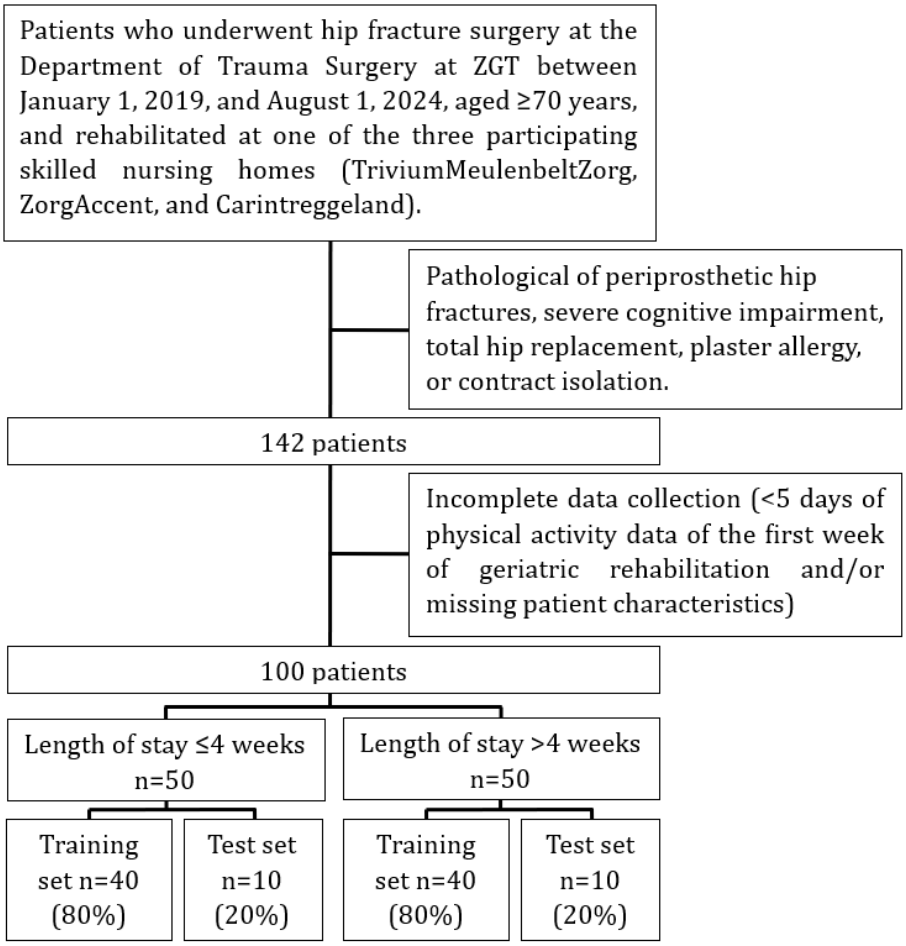
Results

Patient Characteristics and Physical Activity Data

A total of 142 older patients were included in this study (Figure 1). Of these, 42 were excluded due to incomplete data (<5 days of physical activity data of the first week of geriatric rehabilitation and/or missing patient characteristics). The missing physical activity data were mainly caused by sensor-related issues, such as battery depletion and removal of the device for charging and data upload. This resulted in a final sample of 100 patients. Patients were divided into ≤4 weeks (n=50) and >4 weeks (n=50) length of geriatric rehabilitation stay and subsequently split into a training and test set. The distribution of the length of geriatric rehabilitation stay of both sets can be seen in Figure 2. Patient characteristics of the training and test set are shown in Table 1. The mean age in the ≤4 weeks group was 84.5 (SD 5.9) years, while the mean age in the >4 weeks group was 82.8 (SD 6.9) years. The median Montreal Cognitive Assessment (MoCA) score was 21.5 (IQR 18-24) and 19.5 (IQR 17-22.5), respectively. The patient characteristics of the test set showed similar results.

Individual patient trajectories of intensity of physical activity during the first 7 days of geriatric rehabilitation are displayed, along with the mean intensity for patients with a length of stay of 4 weeks or less and for those with more than 4 weeks, including 95% CI (Figure 3).

Figure 1. Flowchart patient recruitment.



XSL•FO
RenderX

Figure 2. Length of geriatric rehabilitation stay (days): training set vs test set.

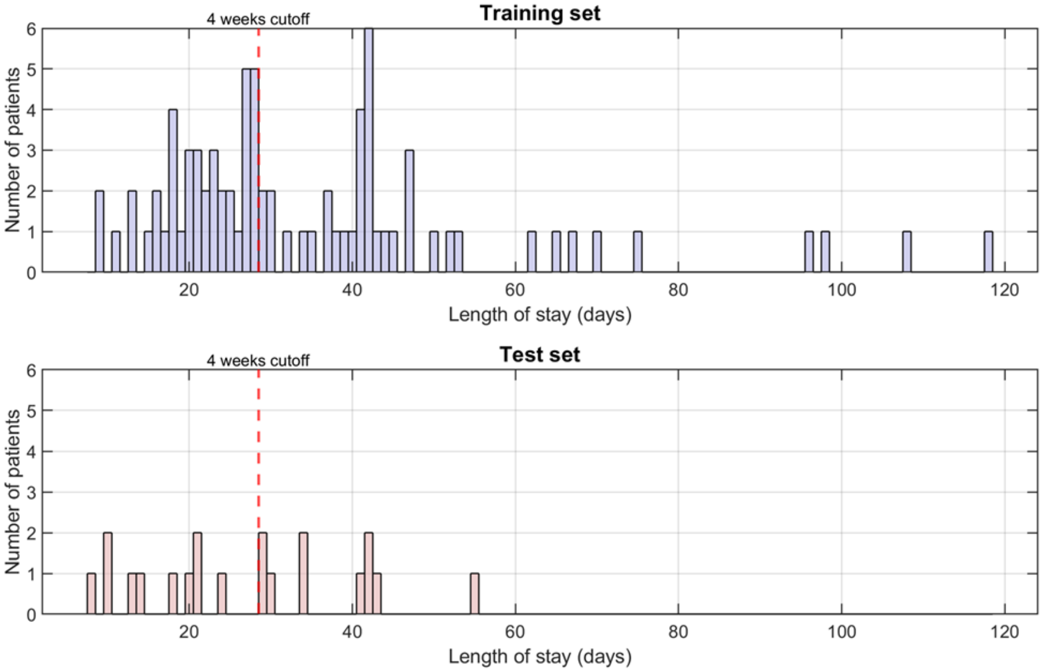


Table 1. Patient characteristics training and test set.

Patient characteristics training and test set	Training set		Test set	
	Length of geriatric rehabilitation stay		Length of geriatric rehabilitation stay	
	≤4 weeks (n=40)	>4 weeks (n=40)	≤4 weeks (n=10)	>4 weeks (n=10)
Baseline variables				
Age (years)[a], mean (SD)	84.5 (5.9)	82.8 (6.9)	82.8 (6.0)	82.2 (6.9)
Female (gender)[a], n (%)	27 (68)	33 (83)	7 (70)	8 (80)
Premorbid living situation[a], n (%)				
Independent without home care services	30 (75)	26 (65)	8 (80)	9 (90)
Independent with home care services	9 (23)	14 (35)	1 (10)	1 (10)
Residential home	1 (3)	0 (0)	1 (10)	0 (0)
Living together[a], n (%)	12 (30)	14 (35)	2 (20)	7 (70)
Availability nonprofessional help[a], n (%)	31 (78)	38 (95)	6 (60)	7 (70)
Need for climbing stairs[a], n (%)	11 (28)	7 (18)	4 (40)	3 (30)
PFMS[a,b], n (%)				
1-Freely mobile without aids	20 (50)	11 (28)	5 (50)	4 (40)
2-Mobile outdoors with one aid	4 (10)	1 (3)	1 (10)	2 (20)
3-Mobile outdoors with 2 aids or frame	16 (40)	27 (68)	4 (40)	4 (40)
4-Some indoor mobility but never goes outside without help	0 (0)	1 (3)	0 (0)	0 (0)
5-No functional mobility	0 (0)	0 (0)	0 (0)	0 (0)
Premorbid Katz-ADL6[a,c], median (IQR)	0 (0-0)	0 (0-1)	0 (0-0)	0 (0-1)
CCI[a,d], median (IQR)	1 (0-2)	1 (0-2)	0 (0-1)	1 (0-3)
ASA[e]score[a], n (%)				
ASA 1	0 (0)	2 (5)	1 (10)	1 (10)
ASA 2	15 (38)	10 (25)	4 (40)	1 (10)
ASA 3	23 (58)	22 (55)	3 (30)	6 (60)
ASA 4	2 (5)	6 (15)	2 (20)	2 (20)
ASA 5	0 (0)	0 (0)	0 (0)	0 (0)
SNAQ[a,f], median (IQR)	0 (0-0)	0 (0-0)	0 (0-0)	0 (0-0)
In-hospital variables				
Hemoglobin level at hospital admission[a], median (IQR)	7.9 (7.6-8.7)	7.8 (7.5-8.6)	7.9 (7.5-8.1)	7.9 (6.0-8.3)
Hip fracture type, n (%)				
Femoral neck	24 (60)	17 (43)	5 (50)	3 (30)
Subtrochanteric femur	16 (40)	23 (58)	5 (50)	7 (70)
Time in the emergency room in minutes[a], median (IQR)	133.5 (103.5-168.0)	133.0 (106.5-161.0)	104.5 (82.0-155.0)	158.5 (124.0-184.0)
Surgical treatment[a], n (%)				
Dynamic hip screw	3 (8)	3 (8)	2 (20)	1 (10)
Intramedullary nail	15 (38)	25 (63)	5 (50)	6 (60)
Hemiarthroplasty	22 (55)	12 (30)	3 (30)	3 (30)

XSL•FO
RenderX

Patient characteristics training and test set	Training set Length of geriatric rehabilitation stay		Test set Length of geriatric rehabilitation stay	
	≤4 weeks (n=40)	>4 weeks (n=40)	≤4 weeks (n=10)	>4 weeks (n=10)
Postoperative weight-bearing protocol[a], n (%)				
Full weight-bearing	40 (100)	36 (90)	10 (100)	9 (90)
Partial or non–weight-bearing	0 (0)	4 (10)	0 (0)	1 (10)
In-hospital complications, n (%)				
Anemia[a]	6 (15)	7 (18)	1 (10)	3 (30)
Heart failure[a]	1 (3)	2 (5)	0 (0)	1 (10)
Pressure ulcers[a]	3 (8)	2 (5)	0 (0)	0 (0)
Delirium[a]	1 (3)	5 (13)	0 (0)	0 (0)
Pulmonary embolism[a]	0 (0)	1 (3)	0 (0)	0 (0)
Kidney failure[a]	0 (0)	1 (3)	0 (0)	0 (0)
Pneumonia[a]	3 (8)	4 (10)	0 (0)	0 (0)
Urinary tract infection[a]	2 (5)	1 (3)	1 (10)	0 (0)
Fall incident	0 (0)	0 (0)	0 (0)	0 (0)
Wound infection	0 (0)	0 (0)	0 (0)	0 (0)
Reoperation	0 (0)	0 (0)	0 (0)	0 (0)
Others	8 (20)	8 (20)	3 (30)	0 (0)
Hospital discharge variables				
FAC[g] at hospital discharge[a], n (%)				
0-No functional mobility	0 (0)	5 (13)	0 (0)	1 (10)
1-Dependent in mobility level 2	1 (3)	4 (10)	0 (0)	4 (40)
2-Dependent in mobility level 1	9 (23)	17 (43)	0 (0)	2 (20)
3-Independent mobility under supervision	22 (55)	9 (23)	8 (80)	1 (10)
4-Independent mobility on a flat surface	8 (20)	5 (13)	2 (20)	2 (20)
FMS[h] at hospital discharge, n (%)				
1-Freely mobile without aids	0 (0)	0 (0)	0 (0)	0 (0)
2-Mobile outdoors with one aid	0 (0)	0 (0)	0 (0)	0 (0)
3-Mobile outdoors with 2 aids or frame	3 (8)	2 (5)	1 (10)	0 (0)
4-Some indoor mobility but never goes outside without help	37 (93)	33 (83)	9 (90)	9 (90)
5-No functional mobility	0 (0)	5 (13)	0 (0)	1 (10)
Katz-ADL6 at hospital discharge[a], median (IQR)	4 (4-4)	4 (4-5)	4.5 (4-5)	4 (4-5)
Rehabilitation department skilled nursing home, n (%)				
Skilled nursing home A	22 (55)	23 (58)	9 (90)	5 (50)
Skilled nursing home B	12 (30)	13 (33)	1 (10)	3 (30)
Skilled nursing home C	6 (15)	4 (10)	0 (0)	2 (20)
Length of hospital stay in days[a], median (IQR)	7 (6-9)	8 (6.5-9)	6.5 (6-8)	7 (6-8)
Geriatric rehabilitation variables				

XSL•FO
RenderX

Patient characteristics training and test set	Training set Length of geriatric rehabilitation stay		Test set Length of geriatric rehabilitation stay	
	≤4 weeks (n=40)	>4 weeks (n=40)	≤4 weeks (n=10)	>4 weeks (n=10)
MoCA[a,i], median (IQR)	21.5 (18-24)	19.5 (17-22.5)	21 (20-25)	23 (18-26)

[a] Patient characteristics included in the ML models.

[b] PFMS: Prefracture Mobility Score.

[c] Katz-ADL6: Katz Index of Independence in Activities of Daily Living-6.

[d] CCI: Charlson Comorbidity Index.

[e] ASA: American Society of Anesthesiologists physical status classification.

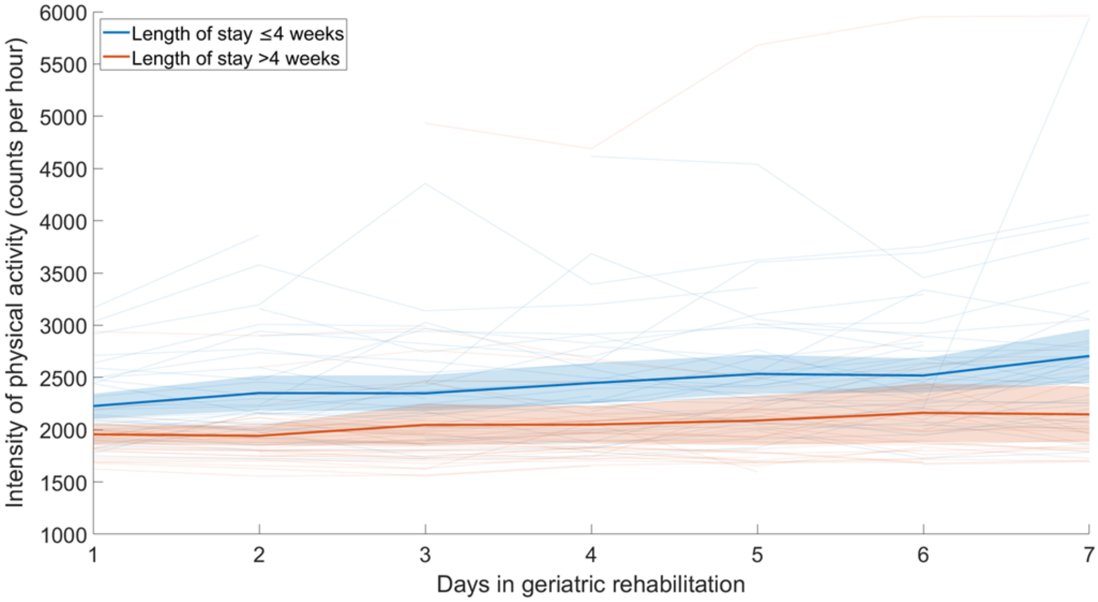
[f] SNAQ: Short Nutritional Assessment Questionnaire.

[g] FAC: Functional Ambulation Categories.

[h] FMS: Fracture Mobility Score.

[i] MoCA: Montreal Cognitive Assessment.

Figure 3. Intensity of physical activity with individual trajectories of the training set and mean physical activity intensity with 95% CI.



Feature Reduction

After applying PC analysis for the feature reduction of the features extracted from the physical activity data, 79 PCs were obtained. The graph of cumulative explained variance versus the number of PCs shows that 19 PCs explained 95% of the variance (Multimedia Appendix 2; Figure S1). These 19 PCs were included in the ML models.

In-hospital fall incidents, in-hospital wound infections, in-hospital reoperations, and other in-hospital complications were excluded because none of these complications occurred or the distributions were identical between the 2 lengths of stay groups. Hip fracture type was excluded from the ML models because it is strongly associated with surgical treatment [13]. Surgical treatment was regarded to be more patient-specific, as it depends on both fracture type and patient condition. FMS at hospital discharge was also excluded due to its strong association with functional ambulation category (FAC) at hospital discharge [13]. The FAC was considered more relevant because it is more widely used in geriatric rehabilitation to assess patient progress. Similarly, admission to which skilled nursing home was excluded due to its association with FAC at hospital discharge [13]. Notably, skilled nursing home A had more patients with lower FAC scores at hospital discharge compared to B and C. This was a coincidence, as the skilled nursing home placement procedure does not depend on the FAC. The FAC at hospital discharge was prioritized for its greater clinical relevance.

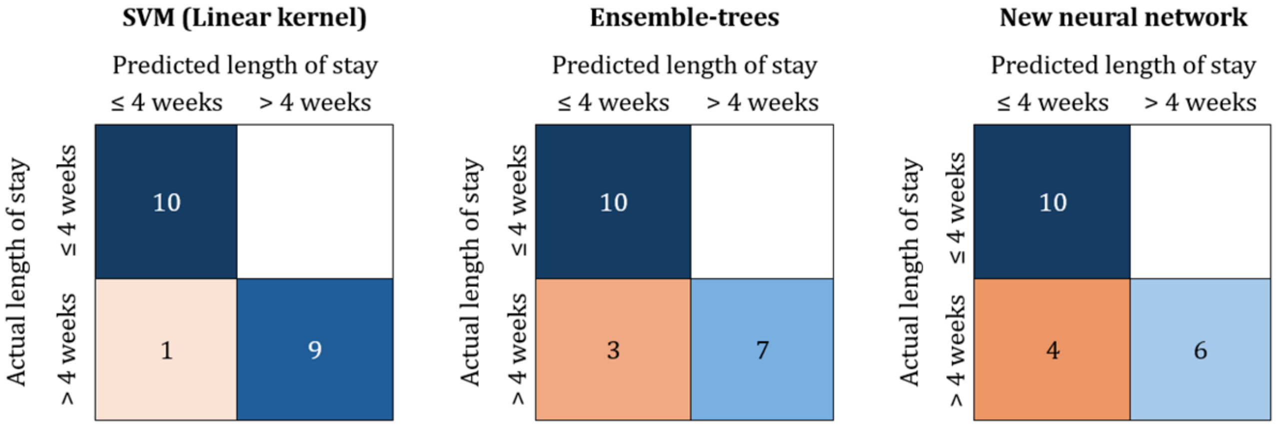
ML Models

Based on the AUC of each ML model after the first 30 iterations (Multimedia Appendix 2; Figure S2), the SVM, NN, and EoDT ML models were selected for continuous Bayesian optimization.

The evaluation of the final hyperparameters after Bayesian optimization resulted in an NN consisting of only one layer with 2 neurons. Since this architecture was considered too simple and likely insufficient for capturing the complexity of the data, a new NN with 2 layers containing 102 and 51 neurons, respectively, was trained for further evaluation.

Figure 4 presents the confusion matrices of the final best-performing configuration of each remaining ML model type (SVM, EoDT, and NN) for the test set. The confusion matrix evaluates the performance of the ML models by comparing the actual length of stay with the predicted length of stay generated by the models. This provides an overview of the number of correctly classified and misclassified patients.

Figure 4. Confusion matrices of the final best-performing machine learning (ML) models (support vector machine [SVM], ensemble of decision trees [EoDT], and neural network [NN]) tested on the test set (n=20).



The performance of the final best-performing ML models, along with the actual length of stay for patients misclassified as ≤4 weeks, is presented in Table 2. The overall best results were achieved using the SVM model with a linear kernel (accuracy=0.95, 95% CI 0.85-1.00; precision=0.91, 95% CI 0.71-1.00; recall=1.00, 95% CI 1.00-1.00; F_1-score=0.95238, 95% CI 0.83-1.00; AUC=0.97, 95% CI 0.83-1.00). The misclassified ≤4 weeks group included one patient with an actual length of stay of 42 days. The EoDT misclassified 3 patients with lengths of stay of 29, 29, and 42 days, while the new NN misclassified 4 patients with lengths of stay of 29, 30, 41, and 42 days.

Table 2. Evaluation of final best-performing machine learning (ML) models (SVM[a], EoDT[b], and neural network).

Machine learning models	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)	AUC[c] (95% CI)	Actual LOS[d] misclassified ≤4 weeks patients
SVM (linear kernel)	0.95 (0.85-1.00)	0.91 (0.71-1.00)	1.00 (1.00-1.00)	0.95 (0.83-1.00)	0.97 (0.88-1.00)	42 days
Ensemble-trees	0.85 (0.70-1.00)	0.77 (0.50-1.00)	1.00 (1.00-1.00)	0.87 (0.67-1.00)	0.97 (0.88-1.00)	29, 29, and 42 days
New neural network	0.80 (0.60-0.95)	0.71 (0.47-0.93)	1.00 (1.00-1.00)	0.83 (0.64-0.97)	0.92 (0.78-1.00)	29, 30, 41, and 42 days

[a] SVM: support vector machine.

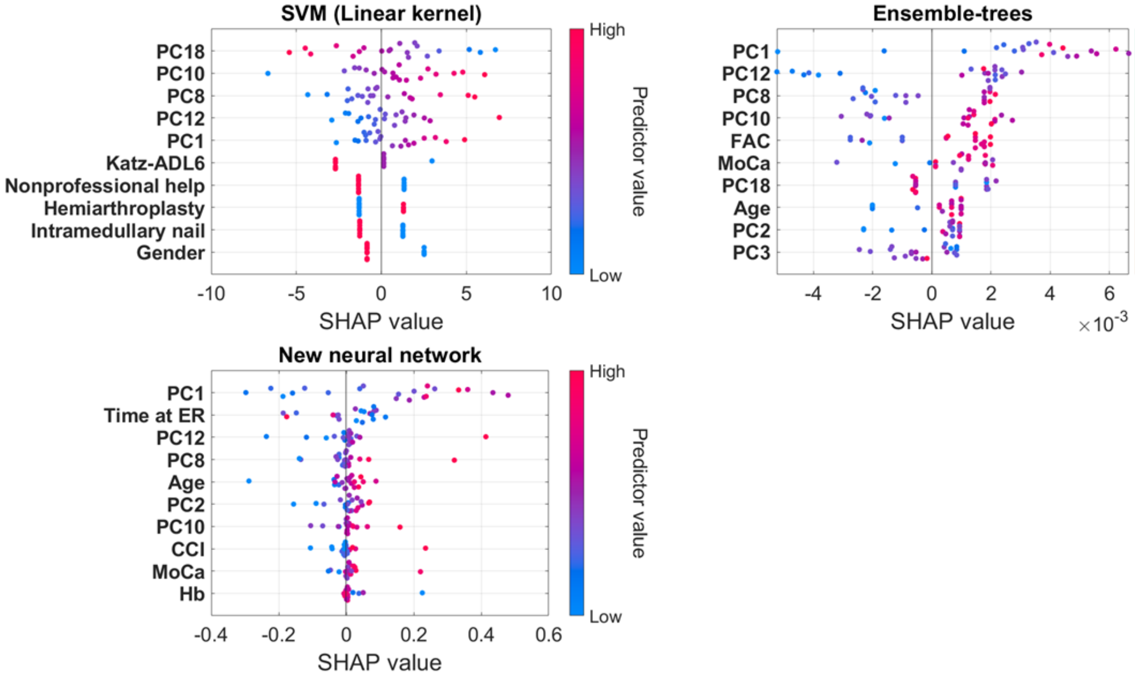
[b] EoDT: ensemble of decision trees.

[c] AUC: area under the curve.

[d] LOS: length of stay.

The 10 most important features of the final best-performing ML models, ranked from most to least impactful, and their influence on predictions can be seen in the results of the SHAP analysis (Figure 5). Each dot represents a patient from the test set, with its color indicating the actual feature value (predictor value: red for high values and blue for low values). The SHAP value of each dot represents the contribution of that feature to the prediction for an individual patient. It indicates how much the feature pushes the prediction toward one length of stay group or the other. Negative SHAP values push the prediction toward a length of stay of ≤4 weeks, while positive SHAP values push it toward a length of stay of >4 weeks.

Figure 5. Shapley Additive Interpretation (SHAP) summary plots: the 10 most important features ranked from most to least impactful and their influence on predictions.



For example, in the Shapley summary plot of the SVM (linear kernel) model, the feature availability of nonprofessional help appears blue on the right. This indicates that patients with no availability of nonprofessional help (0=no availability, 1=availability) have positive SHAP values, pushing predictions toward a length of stay of >4 weeks. Based on the SHAP analysis, it can be seen that the most important features for predicting the length of geriatric rehabilitation stay, across the 3 best-performing ML models, were the PCs (weighted combinations of features extracted from the physical activity data); time in the emergency room (ER); FAC at hospital discharge; age; Katz Index of Independence in Activities of Daily Living-6 (Katz-ADL6) at hospital discharge; MoCA; availability of nonprofessional help; surgery type; Charlson Comorbidity Index (CCI); gender; and hemoglobin level at hospital admission.

In the ablation study, the best-performing ML model trained on physical activity data only outperformed the model trained on patient characteristics only. The combined dataset yielded the highest predictive performance. The confusion matrices and detailed performance metrics for the ablation study are provided in Multimedia Appendix 3.

Discussion

Principal Findings

This study developed ML models for the early prediction of a geriatric rehabilitation stay of ≤4 weeks in older patients after hip fracture surgery, using continuously monitored physical activity data and patient characteristics. This resulted in several ML models with varying overall predictive performance. The SVM was the best-performing model, demonstrating excellent predictive capability with an accuracy of 0.95 (95% CI 0.85-1.00), precision of 0.91 (95% CI 0.71-1.00), recall of 1.00 (95% CI 1.00-1.00), F_1-score of 0.95 (95% CI 0.83-1.00), and AUC of 0.97 (95% CI 0.88-1.00). These initial results suggest that early prediction of a geriatric rehabilitation stay of ≤4 weeks is promising.

In our previous study, the FAC at hospital discharge, premorbid living situation, postoperative weight-bearing protocol, surgery type, in-hospital delirium, and in-hospital heart failure were identified as key factors affecting the length of geriatric rehabilitation stay [13]. Among all these previously identified predictors, the current study confirmed that FAC at hospital discharge and surgery type are key features for predicting a geriatric rehabilitation stay of ≤4 weeks. Postoperative weight-bearing protocol, in-hospital delirium, and in-hospital heart failure did not emerge as key predictors in the best-performing models. A possible explanation is that these factors occurred relatively infrequently in our study population, limiting their predictive power in this context. Furthermore, whereas the previous study focused solely on identifying predictors for the length of geriatric rehabilitation stay, the current study developed and evaluated actual prediction models, offering potential for early discharge planning.

A substantial amount of other research also exists on the factors influencing the length of stay after hip fracture surgery. However, many of these studies focus on factors influencing the length of hospital stay [28,30], or the length of stay in other rehabilitation settings (eg, private care rehabilitation, in-hospital rehabilitation, or both the acute hospital and rehabilitation phases) [21-24], making them difficult to compare with this study. These previously identified predictors for the length of geriatric rehabilitation stay include age [22-24]; fracture type [22]; American Society of Anesthesiologists physical status classification score [20,21]; complications (wound infection,

XSL•FO
RenderX

delirium, urinary tract infection, and pneumonia) [22,23]; comorbidities (Parkinson disease, diabetes, and dementia) [22-24]; gait status at hospital discharge [20]; living situation prior to the hip fracture (alone or together) [24,25]; use of mobility aids prior to the hip fracture [21,25]; Barthel score at geriatric rehabilitation admission [25]; Functional Independence Measure score at rehabilitation admission [23]; Abbreviated Mental Test Score (AMTS) [21]; and pain score at rehabilitation admission [24]. Among these previously identified predictors, this study further confirmed that age, comorbidities (CCI), gait status at hospital discharge (FAC), an assessment tool to evaluate a patient's functional ability to perform activities of daily living (Barthel score or Katz-ADL6), and a cognitive score (AMTS or MoCA) are key features for predicting a geriatric rehabilitation stay of ≤4 weeks.

Unlike previous studies, which did not incorporate continuously measured physical activity, this study included this aspect, providing a more detailed insight into the effect of physical activity on the length of geriatric rehabilitation stay. Consequently, the PCs (weighted combinations of features extracted from the physical activity data) of this study emerged as one of the most important features for predicting a geriatric rehabilitation stay of ≤4 weeks. The findings of this study underscore the critical role of physical activity and align with previous studies emphasizing the positive impact of early mobilization and physical activity on the functional recovery of hip fracture patients [14-16,34]. Recent research further supports this by demonstrating that higher levels of physical activity at the time of rehabilitation admission are associated with greater independence in activities of daily living at discharge [35]. Early mobilization may help maintain postural muscle function, leading to improved functional recovery [35]. In contrast, prolonged bed rest, associated with low levels of physical activity, has been linked to a decline in physical function and reduced independence in activities of daily living due to decreased oxygen consumption, metabolic changes, and muscle atrophy [36,37]. Patients with lower physical activity levels at rehabilitation admission may experience slower functional recovery, prolonged dependence in activities of daily living, and consequently a longer geriatric rehabilitation stay.

In addition to physical activity, this study identified other important predictors of the length of geriatric rehabilitation stay that have not been examined before in previous research. Specifically, the availability of nonprofessional help, time in the ER, and hemoglobin level at hospital admission emerged as important predictors. The availability of nonprofessional help was included as it represents a key aspect of functional social support (social relationships that fulfill particular functions in times of need), which has been shown to be associated with rehabilitation outcomes in older adults with hip fractures [27]. Time spent in the ER was included, as recent research has found an association with hospital length of stay [30], although the reason for this association remains unclear. Hemoglobin level at hospital admission was also included, as recent research has found an association with hospital length of stay due to delayed recovery and increased complication risk in patients with low hemoglobin levels [28].

The remaining most important feature across all ML models, gender, also does not align with previously identified predictors of the length of geriatric rehabilitation stay. However, gender has been included as a variable in earlier studies, although it was not found to be a significant predictor. A possible explanation for its role as a key feature in this study, but not in previous ones, may lie in the use of different analytical techniques. In this study, we used ML techniques, whereas previous research relied on traditional statistical analyses.

Overall, while many of the key features for predicting a geriatric rehabilitation stay of ≤4 weeks identified in this study align with previous research, this study also uncovered other predictors, including weighted combinations of features extracted from continuously measured physical activity data, availability of nonprofessional help, and time spent in the ER. These findings highlight the potential influence of factors often overlooked in traditional statistical analyses.

When examining the misclassified ≤4 weeks patients of the 3 best-performing ML models in more detail, the EoDT has 2 patients with a length of stay of 29 days, and the new NN has one patient with a length of stay of 29 days and another with a length of stay of 30 days. However, misclassifying a 29- and 30-day stay as ≤4 weeks may not be a clinically significant issue, as it is close to 28 days. Furthermore, the new NN misclassified a patient with a length of stay of 41 days. However, this patient developed a complication during rehabilitation, which may have contributed to the incorrect prediction. Since complications occurring after the first seven days of geriatric rehabilitation are not taken into account in the models, this may explain the misclassification. In clinical practice, however, treatment plans are adjusted when complications arise, meaning that such a case would likely have been identified as requiring a longer stay regardless. The other 3 misclassified patients across the 3 best-performing ML models had a length of stay of 42 days. No clear explanation was found based on the available data.

Strengths and Limitations of This Study

To the best of our knowledge, this is the first study to develop and compare ML models for predicting a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home at an early rehabilitation stage, using a novel experimental approach with continuously monitored physical activity data and patient characteristics in older patients after hip fracture surgery. Another strength of our study is the inclusion of a diverse range of features from both the physical and social domains, with particular emphasis on the integration of continuously monitored physical activity data. In addition, the MOX1 accelerometer was well tolerated by patients and was easily integrated into the clinical workflow, requiring minimal effort from clinical staff and demonstrating the feasibility of continuous activity monitoring in this population.

Regarding the limitations, a complete dataset was required for ML model development, leading to the exclusion of many patients due to missing a single feature. No data imputation was applied in this study, as the aim was to avoid introducing potential bias from imputed values. This requirement for complete data reduced the sample size. A practical limitation

XSL•FO
RenderX

contributing to missing physical activity data was the inability to check the battery level of the MOX1 accelerometer during use. Although the lights of the accelerometer show whether it is actively measuring, clinical staff did not routinely check this, as the wearable required no daily attention. As a result, wearables occasionally ran out of battery unnoticed, leading to the loss of multiple days of activity data. While model performance was already high, having complete data for all patients could have potentially increased the reliability and robustness of the analyses. Second, given the focus on ML model development, the sample size may be relatively small, making some models prone to overfitting. To mitigate this risk, techniques such as cross-validation and feature reduction were applied. Third, although length of geriatric rehabilitation stay is a continuous variable, this study formulated the problem as a classification task (≤4 weeks vs >4 weeks). This cutoff reflects a clinically meaningful decision point with substantial impact on early discharge planning. For clinicians, knowing after one week whether a patient is likely to be discharged within 4 weeks is far more actionable than attempting to predict the exact number of days. Predicting the exact length of stay in days would introduce unnecessary complexity, require substantially more data, and is not essential for guiding early clinical decisions. Therefore, we consider the current classification framework to be the most suitable and clinically relevant approach for guiding early discharge planning. Future research is recommended to explore appropriate data imputation strategies in a cautious manner, to reduce data loss while ensuring that bias is not introduced. For categorical patient characteristics such as the availability of nonprofessional help, imputation may be challenging and not always be desirable, highlighting the need for careful consideration of which variables to impute. In addition, future work is recommended to validate the models with a larger patient population, refine them if necessary, and integrate the best-performing model into clinical practice.

Clinical Interpretation

The results of this study indicate that early prediction of a geriatric rehabilitation stay of ≤4 weeks or less in a skilled nursing home for older patients after hip fracture surgery is feasible. This marks the first step toward a more future-proof geriatric rehabilitation system. Identifying patients with an expected stay of ≤4 weeks or less after 7 days of rehabilitation enables timely discharge planning and coordination of support at home. Such planning can smooth the transition toward home and prevent discharge delays, which is important to ensure optimal bed occupancy.

Conclusion

Several ML models were developed to predict a geriatric rehabilitation stay of ≤4 weeks in a skilled nursing home for older patients after hip fracture surgery, using continuously monitored physical activity data from the first week of geriatric rehabilitation and patient characteristics. Among these models, the SVM proved to be highly accurate in its predictions with an accuracy of 0.95 (95% CI 0.85-1.00), precision of 0.91 (95% CI 0.71-1.00), recall of 1.00 (95% CI 1.00-1.00), F_1-score of 0.95 (95% CI 0.83-1.00), and AUC of 0.97 (95% CI 0.88-1.00).

Acknowledgments

The authors would like to thank the Up&Go after a Hip Fracture Group, including the health care professionals of the rehabilitation departments of the nursing home institutions TriviumMeulenbeltZorg, ZorgAccent, and Carintreggeland for their contribution to the transmural Up&Go after a hip fracture project.

Funding

This research received no external funding.

Data Availability

The datasets analyzed during this study are not publicly available due to the fact that our patients did not give permission to share the data publicly but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: SK, FJW, DvD, MMRV-H, JHH
Data curation: SK
Formal analysis: SK, FJW
Investigation: SK
Methodology: SK, FJW
Supervision: FJW, DvD, MMRV-H, JHH
Validation: SK, FJW
Visualization: SK
Writing – original draft: SK
Writing – review & editing: SK, FJW, DvD, MMRV-H, JHH

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed explanation of the extracted continuously monitored physical activity features.
[DOCX File , 19 KB - rehab_v13i1e79331_app1.docx]

Multimedia Appendix 2
Feature reduction and machine learning model development.
[DOCX File , 312 KB - rehab_v13i1e79331_app2.docx]

Multimedia Appendix 3
Confusion matrices and detailed performance metrics for the ablation study.
[DOCX File , 40 KB - rehab_v13i1e79331_app3.docx]

References

1. CBS. Hospital admissions and patients; diagnostic classification VTV. Statistics Netherlands (CBS). 2024. URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84067NED/table?ts=1570535289637 [accessed 2025-01-28]
2. Mattiazzo GF, Drewes YM, van Eijk M, Achterberg WP. Geriatric rehabilitation care after hip fracture. Eur Geriatr Med 2023 Apr;14(2):295-305 [FREE Full text] [doi: 10.1007/s41999-023-00755-4] [Medline: 36788193]
3. van Dartel D, Vermeer M, Folbert EC, Arends AJ, Vollenbroek-Hutten MMR, Hegeman JH, Dutch Hip Fracture Audit (DHFA) Group. Early Predictors for Discharge to Geriatric Rehabilitation after Hip Fracture Treatment of Older Patients. J Am Med Dir Assoc 2021 Dec;22(12):2454-2460 [FREE Full text] [doi: 10.1016/j.jamda.2021.03.026] [Medline: 33933417]
4. Grund S, Gordon AL, van Balen R, Bachmann S, Cherubini A, Landi F, et al. European consensus on core principles and future priorities for geriatric rehabilitation: consensus statement. Eur Geriatr Med 2020 Apr;11(2):233-238 [FREE Full text] [doi: 10.1007/s41999-019-00274-1] [Medline: 32297191]
5. Sing CW, Lin TC, Bartholomew S, Bell JS, Bennett C, Beyene K, et al. Global Epidemiology of Hip Fractures: Secular Trends in Incidence Rate, Post-Fracture Treatment, and All-Cause Mortality. J Bone Miner Res 2023 Aug;38(8):1064-1075 [FREE Full text] [doi: 10.1002/jbmr.4821] [Medline: 37118993]
6. Minderhout RN, van Ede AF, Voragen L, Verheijen C, Vos HM, Numans ME, et al. Reforming healthcare in the Netherlands: practical population health management and the Plot model: A questionnaire survey and focus group study to assess the willingness and readiness of six regions in the Netherlands. SAGE Open Med 2023;11:20503121231160830 [FREE Full text] [doi: 10.1177/20503121231160830] [Medline: 36949828]
7. Integraal Zorgakkoord: 'samenwerken aan gezonde zorg'. The National Government. For the Netherlands. 2022. URL: https://www.rijksoverheid.nl/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg [accessed 2025-01-28]
8. Lin WT, Wu TY, Chen YJ, Chang YS, Lin CH, Lin YJ. Predicting in-hospital length of stay for very-low-birth-weight preterm infants using machine learning techniques. J Formos Med Assoc 2022 Jun;121(6):1141-1148 [FREE Full text] [doi: 10.1016/j.jfma.2021.09.018] [Medline: 34629242]
9. Deo RC. Machine Learning in Medicine. Circulation 2015 Nov 17;132(20):1920-1930 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.001593] [Medline: 26572668]
10. Tian CW, Chen XX, Shi L, Zhu HY, Dai GC, Chen H, et al. Machine learning applications for the prediction of extended length of stay in geriatric hip fracture patients. World J Orthop 2023 Oct 18;14(10):741-754 [FREE Full text] [doi: 10.5312/wjo.v14.i10.741] [Medline: 37970626]
11. Daghistani TA, Elshawi R, Sakr S, Ahmed AM, Al-Thwayee A, Al-Mallah MH. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. Int J Cardiol 2019 Aug 01;288:140-147. [doi: 10.1016/j.ijcard.2019.01.046] [Medline: 30685103]
12. Muhlestein WE, Akagi DS, Davies JM, Chambless LB. Predicting Inpatient Length of Stay After Brain Tumor Surgery: Developing Machine Learning Ensembles to Improve Predictive Performance. Neurosurgery 2019 Sep 01;85(3):384-393 [FREE Full text] [doi: 10.1093/neuros/nyy343] [Medline: 30113665]
13. Krakers SM, Woudsma S, van Dartel D, Vermeer M, Vollenbroek-Hutten MMR, Hegeman JH. Rehabilitation of Frail Older Adults after Hip Fracture Surgery: Predictors for the Length of Geriatric Rehabilitation Stay at a Skilled Nursing Home. J Clin Med 2024 Aug 03;13(15):4547 [FREE Full text] [doi: 10.3390/jcm13154547] [Medline: 39124813]
14. Riemen AH, Hutchison JD. The multidisciplinary management of hip fractures in older patients. Orthop Trauma 2016 Apr;30(2):117-122 [FREE Full text] [doi: 10.1016/j.mporth.2016.03.006] [Medline: 27418950]
15. Morri M, Forni C, Marchioni M, Bonetti E, Marseglia F, Cotti A. Which factors are independent predictors of early recovery of mobility in the older adults' population after hip fracture? A cohort prognostic study. Arch Orthop Trauma Surg 2018 Jan;138(1):35-41. [doi: 10.1007/s00402-017-2803-y] [Medline: 28956152]

XSL•FO
RenderX

16. van Dartel D, Wang Y, Hegeman JH, Vermeer M, Vollenbroek-Hutten MMR. Patterns of physical activity over time in older patients rehabilitating after hip fracture surgery: a preliminary observational study. BMC Geriatr 2023 Jun 16;23(1):373 [FREE Full text] [doi: 10.1186/s12877-023-04054-2] [Medline: 37328743]
17. Makizako H, Shimada H, Doi T, Tsutsumimoto K, Lee S, Hotta R, et al. Cognitive functioning and walking speed in older adults as predictors of limitations in self-reported instrumental activity of daily living: prospective findings from the Obu Study of Health Promotion for the Elderly. Int J Environ Res Public Health 2015 Mar 11;12(3):3002-3013 [FREE Full text] [doi: 10.3390/ijerph120303002] [Medline: 25768239]
18. van de Haar HJ, Alme MN. Geriatric rehabilitation in the Netherlands and their preparedness for the increase in older adults: a scoping review. Discov Public Health 2025 Jan 13;22(1):13. [doi: 10.1186/s12982-025-00399-8]
19. Frowijn T, Vermeer M, Koop R, Schreuder R. USER draagt bij aan voorspelling revalidatieduur GRZ. Tijdschrift voor Ouderengeneeskunde 2016:4 [FREE Full text]
20. Adrados M, Wang K, Deng Y, Bozzo J, Messina T, Stevens A, et al. A Simple Physical Therapy Algorithm Is Successful in Decreasing Skilled Nursing Facility Length of Stay and Increasing Cost Savings After Hip Fracture With No Increase in Adverse Events. Geriatr Orthop Surg Rehabil 2021;12:2151459321998615 [FREE Full text] [doi: 10.1177/2151459321998615] [Medline: 33815865]
21. Richards T, Glendenning A, Benson D, Alexander S, Thati S. The independent patient factors that affect length of stay following hip fractures. Ann R Coll Surg Engl 2018 Sep;100(7):556-562 [FREE Full text] [doi: 10.1308/rcsann.2018.0068] [Medline: 29692191]
22. Ireland AW, Kelly PJ, Cumming RG. Total hospital stay for hip fracture: measuring the variations due to pre-fracture residence, rehabilitation, complications and comorbidities. BMC Health Serv Res 2015 Jan 22;15(1):17 [FREE Full text] [doi: 10.1186/s12913-015-0697-3] [Medline: 25609030]
23. Thornburgh Z, Samuel D. Factors Influencing Length of Stay and Discharge Destination of Patients with Hip Fracture Rehabilitating in a Private Care Setting. Geriatrics (Basel) 2022 Mar 31;7(2):44 [FREE Full text] [doi: 10.3390/geriatrics7020044] [Medline: 35447847]
24. Daly N, Fortin C, Jaglal S, MacDonald SL. Predictors of Exceeding Target Inpatient Rehabilitation Length of Stay After Hip Fracture. Am J Phys Med Rehabil 2020;99(7):630-635. [doi: 10.1097/phm.0000000000001386]
25. Faut L, van Schieveen C, Schalk BWM. Met een heupfractuur op de Geriatrische Revalidatie: Factoren gerelateerd aan een kortere opnameduur. Tijdschrift voor Ouderengeneeskunde 2021:5 [FREE Full text]
26. Orwig D, Hochberg MC, Gruber-Baldini AL, Resnick B, Miller RR, Hicks GE, et al. Examining Differences in Recovery Outcomes between Male and Female Hip Fracture Patients: Design and Baseline Results of a Prospective Cohort Study from the Baltimore Hip Studies. J Frailty Aging 2018;7(3):162-169 [FREE Full text] [doi: 10.14283/jfa.2018.15] [Medline: 30095146]
27. Zhu Y, Xu BYX, Low SG, Low LL. Association of Social Support with Rehabilitation Outcome among Older Adults with Hip Fracture Surgery: A Prospective Cohort Study at Post-Acute Care Facility in Asia. J Am Med Dir Assoc 2023 Oct;24(10):1490-1496 [FREE Full text] [doi: 10.1016/j.jamda.2023.03.034] [Medline: 37156471]
28. Zhang N, Zhang D, Ren S, Gao Y, Sun W, Yang S. Relationship between preoperative hemoglobin levels and length of stay in elderly patients with hip fractures: A retrospective cohort study. Medicine (Baltimore) 2024 Jun 21;103(25):e38518 [FREE Full text] [doi: 10.1097/MD.0000000000038518] [Medline: 38905374]
29. Millrose M, Schmidt W, Krickl J, Ittermann T, Ruether J, Bail HJ, et al. Influence of Malnutrition on Outcome after Hip Fractures in Older Patients. J Pers Med 2023 Jan 03;13(1):109 [FREE Full text] [doi: 10.3390/jpm13010109] [Medline: 36675770]
30. Clement ND, Farrow L, Chen B, Duffy A, Murthy K, Duckworth AD. Delayed admission of patients with hip fracture from the emergency department is associated with an increased mortality risk and increased length of hospital stay. Emerg Med J 2024 Oct 23;41(11):654-659. [doi: 10.1136/emermed-2023-213085] [Medline: 39379165]
31. Ek S, Meyer AC, Hedström M, Modig K. Hospital Length of Stay After Hip Fracture and It's Association With 4-Month Mortality-Exploring the Role of Patient Characteristics. J Gerontol A Biol Sci Med Sci 2022 Jul 05;77(7):1472-1477 [FREE Full text] [doi: 10.1093/gerona/glab302] [Medline: 34622920]
32. van Dijk-Huisman HC, Bijnens W, Senden R, Essers JMN, Meijer K, Aarts J, et al. Optimization and Validation of a Classification Algorithm for Assessment of Physical Activity in Hospitalized Patients. Sensors (Basel) 2021 Feb 27;21(5):1652 [FREE Full text] [doi: 10.3390/s21051652] [Medline: 33673447]
33. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. Clin Transl Sci 2024 Nov;17(11):e70056. [doi: 10.1111/cts.70056] [Medline: 39463176]
34. Mashimo S, Kubota J, Sato H, Saito A, Gilmour S, Kitamura N. The impact of early mobility on functional recovery after hip fracture surgery. Disabil Rehabil 2023 Dec;45(26):4388-4393. [doi: 10.1080/09638288.2022.2151652] [Medline: 36448297]
35. Shimizu T, Kanai C, Asakawa Y. Relationship between independence in activities of daily living at discharge and physical activity at admission of older postoperative hip fracture rehabilitation inpatients: A retrospective case-control study. Physiother Res Int 2024 Jan;29(1):e2070. [doi: 10.1002/pri.2070] [Medline: 38284469]

XSL•FO
RenderX

36. Graham ZA, Lavin KM, O'Bryan SM, Thalacker-Mercer AE, Buford TW, Ford KM, et al. Mechanisms of exercise as a preventative measure to muscle wasting. Am J Physiol Cell Physiol 2021 Jul 01;321(1):C40-C57 [FREE Full text] [doi: 10.1152/ajpcell.00056.2021] [Medline: 33950699]
37. Creditor MC. Hazards of hospitalization of the elderly. Ann Intern Med 1993 Feb 01;118(3):219-223. [doi: 10.7326/0003-4819-118-3-199302010-00011] [Medline: 8417639]

Abbreviations

AMTS: Abbreviated Mental Test Score
ASA: American Society of Anesthesiologists physical status classification
AUC: area under the curve
CCI: Charlson Comorbidity Index
EoDT: Ensemble of decision trees
ER: emergency room
FAC: functional ambulation categories
Katz-ADL6: Katz Index of Independence in Activities of Daily Living-6
ML: machine learning
MoCA: Montreal Cognitive Assessment
NB: naïve Bayes
NN: neural network
PFMS: Prefracture Mobility Score
PC: principal component
SHAP: Shapley Additive Interpretation
SVM: support vector machine
ZGT: Ziekenhuisgroep Twente

Edited by S Munce; submitted 19.Jun.2025; peer-reviewed by M Santos, H Namba; comments to author 19.Nov.2025; accepted 13.Jan.2026; published 23.Feb.2026.

Please cite as:
Krakers SM, Wouda FJ, van Dartel D, Vollenbroek-Hutten MMR, Hegeman JH, Up&Go After a Hip Fracture Group
Predicting Geriatric Rehabilitation Stays of ≤4 Weeks After Hip Fracture Surgery: Machine Learning Approach Using Physical Activity and Patient Data
JMIR Rehabil Assist Technol 2026;13:e79331
URL: https://rehab.jmir.org/2026/1/e79331
doi:10.2196/79331
PMID:

©Sanne M Krakers, Frank J Wouda, Dieuwke van Dartel, Miriam MR Vollenbroek-Hutten, Johannes H Hegeman, Up&Go After a Hip Fracture Group. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Enhancing Knee Joint Proprioception in Healthy Adults Through Exergame Training With Augmented Feedback: Randomized Controlled Pilot Trial

Yiling Zhang[1], MSc; Luis Felipe García Arias[1], MSc; Hans Timmerman[2], PhD; Ming Cao[1], Prof Dr; Elisabeth Wilhelm[1], PhD

[1]Engineering and Technology Institute, Faculty of Science and Engineering, Universtiy of Groningen, Groningen, The Netherlands
[2]Department of Anesthesiology, Pain Center, University Medical Center Groningen, Groningen, The Netherlands

Corresponding Author:
Elisabeth Wilhelm, PhD
Engineering and Technology Institute
Faculty of Science and Engineering
Universtiy of Groningen
Nijenborgh 4
Groningen, 9747AG
The Netherlands
Phone: 31 0631921169
Email: e.wilhelm@rug.nl

Abstract

Background: Proprioception training is essential for restoring knee function in several medical conditions. Open kinetic chain (OKC) and closed kinetic chain (CKC) exercises are used in active movement interventions to enhance proprioception. Exergames, supported by wearable sensors, offer a solution by providing real-time feedback. Auditory feedback (AF) embedded in the serious game training has shown benefits in upper limb rehabilitation compared to visual feedback (VF) alone. However, the potential of AF in exergames that provide knee training is not known.

Objective: This study presents an exergame platform aimed at enhancing knee joint proprioception through stretching and squatting exercises. The platform allows to provide feedback in 2 modes, namely, VF only and a combination of AF and VF. The AF indicates the joint position by adjusting the loudness of the sound. The VF maps the motion of the lower limb into the game space, where this information is used to control a game object. A randomized controlled trial with 14 participants compared AF to VF only. The hypothesis was that AF would improve knee joint position accuracy, enhancing neuromuscular coordination and lower limb stability.

Methods: A randomized controlled trial was conducted within 14 healthy volunteers to test the exergames for knee joint motor learning using augmented feedback. All participants were required to do a pretest consisting of half and full squats and two tasks, in which participants were asked to reproduce a 45-degree knee bend and to stretch their knee fully. After that, participants played 4 rounds of each of the 2 exergames. Then the tasks of the pretest were repeated. A 1-sided Mann-Whitney *U* test was conducted for answering whether AF has a positive effect on the ability of participants to accurately control the knee joint angle. In addition, we calculated the muscle synergies participants used to complete the exercises. Subjective gaming experience was assessed using the Intrinsic Motivation Inventory and the User Experience Questionnaire.

Results: A total of 14 participants were recruited, including 7 in the experimental group with AF, and 7 in the control group only with VF. The result of the Mann-Whitney *U* test demonstrated that augmented feedback improved knee joint accuracy compared to VF in both the CKC (statistics=41.0; *P*=.04) and OKC (statistics=42.0; *P*=.03) tasks. Additionally, muscle synergy analysis revealed high consistency in different muscle synergy patterns between groups across both game types.

Conclusions: Augmented feedback significantly enhanced knee joint motor learning performance (reflected in the knee joint angle positioning ability) in both CKC and OKC exergame training. Consistent muscle synergy patterns across participants show that the developed exergames are suitable for knee training. Studies in patient populations are needed to establish whether the games could be used in lower limb rehabilitation.

Trial Registration: ClinicalTrials.gov NCT07141290; https://clinicaltrials.gov/study/NCT07141290

XSL•FO
RenderX

(JMIR Rehabil Assist Technol 2026;13:e78525) doi:10.2196/78525

KEYWORDS

home-based exergame; augmented feedback; auditory feedback; knee proprioception; Inertial Measurement Unit

Introduction

Background and Rationale

Proprioceptive training has been reported to be beneficial in restoring knee joint function among clinical populations and athletes. For example, rehabilitation by means of physical therapy can restore this proprioception in patients with knee osteoarthritis [1].

According to Aman et al [2], proprioceptive training can be further divided into 5 types. These are active movement and balance training, passive movement training, somatosensory stimulation training, somatosensory discrimination training, and combined and multiple system training. The meta-analysis does not come to a clear conclusion on which of these training types is superior. It also highlights that one major limitation of the field is that there is no clear definition of the term proprioceptive training. To overcome this, the authors suggest that this term should be reserved for studies that train only based on somatosensory signals while excluding other modalities such as vision [2]. Winter et al [3] suggest a broader definition that includes any intervention aiming to improve proprioception with the ultimate goal to enhance motor function. This definition is the one that applies to most serious games, as they usually involve visual feedback (VF). Winter et al [3] also concluded that active movement interventions are an important aspect of proprioceptive training. Therefore, serious games that contain active movement interventions could be beneficial to improve proprioception.

In physiotherapy, active movements such as open kinetic chain (OKC) and closed kinetic chain (CKC) exercises are used to strengthen the muscles of the knee joint [4]. OKC exercises consist of free movement of the distal joint in space without any loading, such as seated leg extensions, terminal knee extension exercises, hamstring curls, and calf pumps [5]. OKC exercises isolate specific muscle groups, enabling precise, targeted strengthening. Furthermore, OKC exercises can increase the range of motion in patients with anterior cruciate ligament injury [6]. In contrast to OKC exercises, CKC exercises fix or load a distal segment, thereby prohibiting free movement of that segment [5]. Prominent examples of CKC exercises are squats and lunges. CKC exercises promote the cocontraction of muscles to stabilize and control joint movements [6].

In the rehabilitation context, patients ideally would train 3 to 7 days per week. However, due to the lack of physiotherapists, several sessions need to be carried out unsupervised. Exergames are an option for providing feedback, while patients with impairments of the upper and lower limbs, gait disorders, balance disorders, visual-spatial disorder, or cognitive disorder train independently [7]. Especially, video exergames become a promising option for engaging patients in rehabilitative exercises. As reviewed by Gelineau et al [8], for the physical rehabilitation exergames, nonspecific video game systems (eg, Nintendo Wi and Xbox Kinect) need to be combined with specific rehabilitation systems (eg, Rehabilitation Gaming System and virtual gloves) or specific rehabilitation devices (eg, Hand Mentor Pro and Polhemus 3) to achieve satisfying training results in patients with stroke [8].

The rehabilitation systems that can be combined with exergames usually contain sensors to track the participant's motion. Inertial Measurement Units (IMUs), which measure linear acceleration and angular acceleration, can be used to monitor kinetic motion parameters [9]. In addition, some devices use electromyography (EMG) to extract information on the activation of different muscle groups [10,11]. Metrics derived from these different sensor modalities can be used to calculate, analyze, and predict whether a user accomplishes the prescribed exercises. The use of sensorized systems makes it possible to inform exergame users about the knowledge of results (KOR), which can be used as extrinsic feedback. The KOR is defined as information about the measured error between the terminal limb position and the given target position [12]. Traditionally, most video games use a visual representation to inform users about the KOR.

Augmented feedback has been suggested as a strategy to enhance the performance of rehabilitation training with exergames. Among the different feedback modalities that have been suggested for motor learning, auditory feedback (AF), which converts movement-related data into sound, has been successfully used in tasks with low functional task complexity. Sounds can either be used to warn participants (alarms), to make them more aware of their movements (motion sonification), or to amplify errors (error amplification) [13]. The decrease in performance when feedback is removed is smaller in AF as compared to VF, indicating that participants who receive AF do not become dependent on the presence of the feedback [14].

For proprioceptive training in older adults, a combination of a commercially available stepping game (Stomp it) and a game in which participants have to kick a virtual ball (Sport Kinect target kick) or prevent a ball from entering the goal (Sport Kinect goal keeper game) has been investigated. These games mainly focus on OKC tasks. Furthermore, the games that are related to soccer playing require a mix of knee movement and balancing. The feedback in the games is mostly VF. To assess the effect of an 8-week training program with these games, participants were asked to reproduce knee angles when seated. The instructions for the knee angle reproduction assessment were given with a tactile interface. Participants were blind folded during this procedure. Participants of the intervention group reduced the knee proprioception error from 5.2 to 3.5 for a target angle of 30. Similar significant improvements were also reported for other target angles [15]. For patients with knee osteoarthritis, a game that provides VF on CKC exercises was proposed. The exercises in this game were squatting and leg presses against an elastic Theraband. The proprioception was assessed by asking participants to reproduce a knee angle that has previously visually been demonstrated to them. For a target position of 30,

XSL•FO
RenderX

the knee proprioception error improved from 6.25 to 3.5 in a training period of 6 weeks [16]. Another study in patients with osteoarthritis focused on balance games with the Nintendo Wii Fit. The games are, in general, challenging the balance of the user. As the solution the player attempts is not limited, they can contain a mix of open and closed kinematic chain motions. In accordance with its main goal, this study focused on balance-related outcome measures [17]. No study so far has systematically investigated whether the efficacy of AF is related to the type of exercise (OCK vs CKC) in lower limb training.

Objectives

In this study, we developed a video exergame system for enhancing position accuracy through targeted training with incorporated motion tracking and augmented feedback. The system contains 2 different games, one for CKC and one for OKC exercises. The CKC game is called Bunny Game and encourages users to do half squats and deep squats. The OKC game, Fish Game, asks users to perform 4 different knee joint stretching tasks. To provide feedback, the knee joint angle was measured by IMU sensors. This information was used to adjust the position of a game object on the screen (VF) and provide motion sonification (AF).

To verify whether augmented feedback training yields better results than VF, a pilot study was conducted with 14 healthy participants, 7 in the augmented feedback group (experimental group [EG]) and 7 in the VF group (control group [CG]). All participants played both games, one with an OKC task and one with a CKC task in a randomized order. We hypothesized that participants in the EG would significantly improve knee joint position accuracy compared to the CG. In total, 7 EMG signals from the nondominant leg of the participants were collected to analyze the muscle synergies used during training.

Methods

This paper adheres to the CONSORT (Consolidated Standards of Reporting Trials). The CONSORT checklist for this paper can be found in Multimedia Appendix 1.

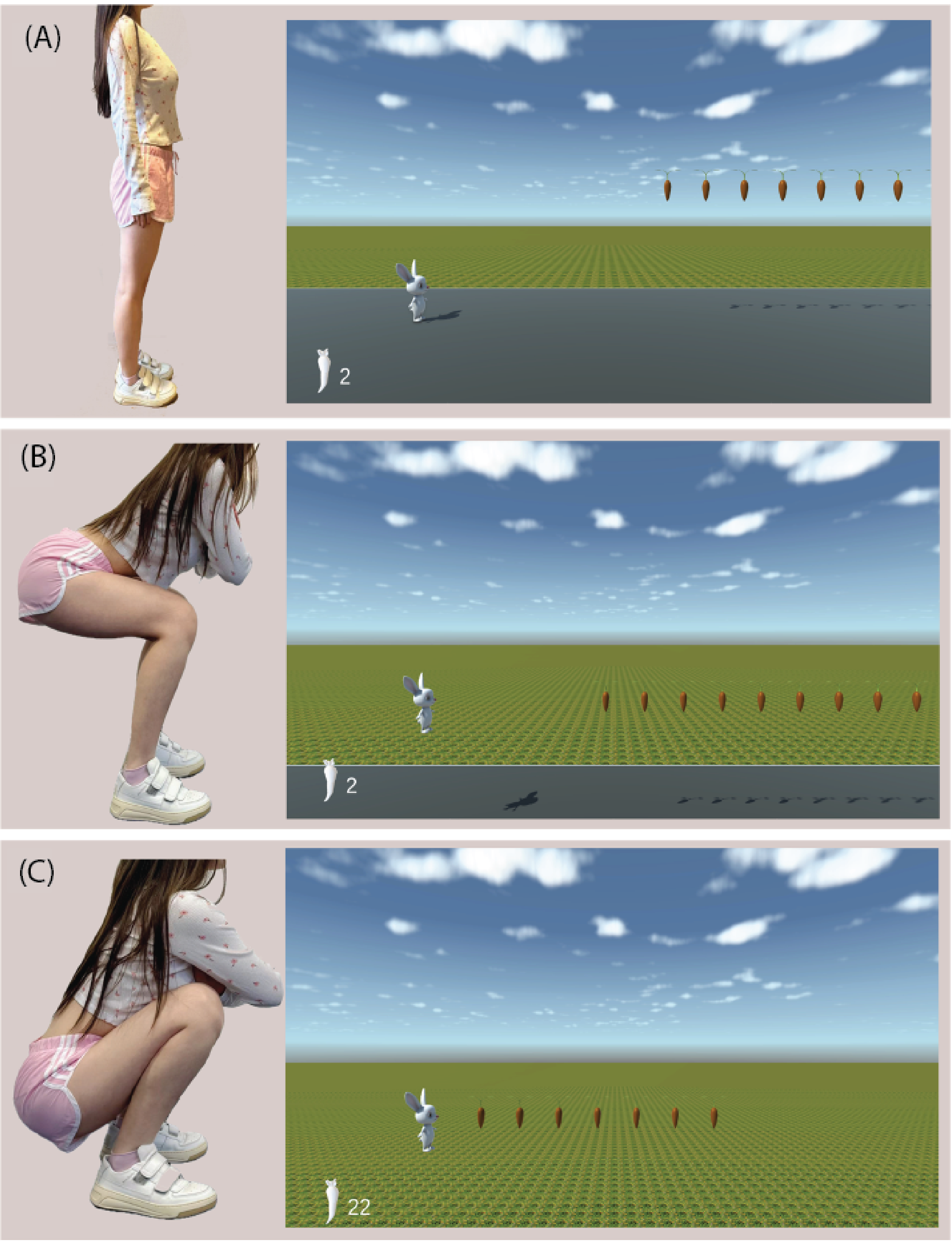
Game Development

CKC Exergame

The user interface (UI) of the CKC exergame, called “Bunny Game,” is shown in Figure 1. In this game, participants performed 2 different types of squats to control the vertical position of the bunny on the screen: the deeper the squat, the higher the bunny jumps. The sonification of the position of the bunny is based on a sequence of brief, percussive tones with varying pitches. The pitch increases as the bunny’s vertical position rises. Carrots were set at 2 heights: one reachable by a half squat, as shown in Figure 1B, and the other by a deep squat, as shown in Figure 1C. When the bunny reaches a carrot, it is collected, and the score increases by 1. Each group consists of 9 carrots, requiring the player to hold the squat position for 2.5 seconds to collect all of them. Each group of carrots is followed by a 20-second gap to allow the player to stand up and rest. One round of the game contains 12 groups of carrots, thereby prompting the player to perform 12 squats per round.

XSL•FO
RenderX

Figure 1. Illustration of the game with closed kinetic chain movements. (A) Preparation in stance gesture. (B) Half squat demonstration and the carrots' height in the user interface. (C) Deep squat demonstration and the height of the carrots that should trigger this motion.

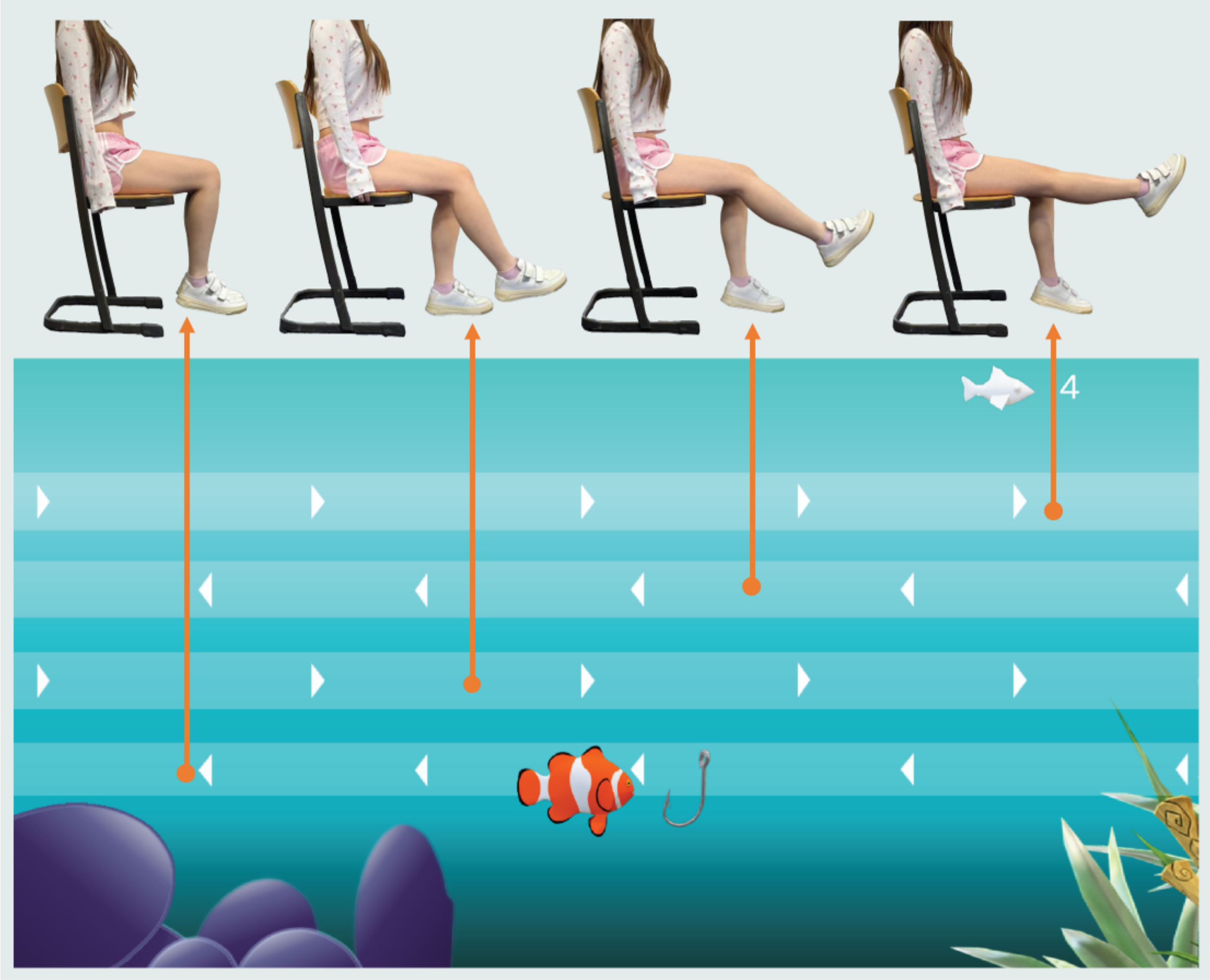


OKC Exergame

The UI of the stretch exergame "Fish Game" is shown in Figure 2. In this game, 12 fish appear randomly along 4 different trails, each corresponding to a specific knee stretch movement. The lowest trail matched the movement of the 90° knee joint, the second lowest trail matched the 120° knee joint bending, the second highest trail matched the 150° knee joint bending, and the highest trail matched the total stretching. The participant rotates the knee joint of the tested leg to move the hook from one trail to the other in a vertical direction. The pitch of the sonification prompt increases as the vertical position of the hook rises. If the hook stays for more than 2 seconds on the trail where a fish appears, the fish is caught, and the score increases by 1. If the fish is not caught within 10 seconds, it reaches the other side of the screen, where it disappears. The word "Miss" is shown on the screen to communicate to the player that the fish escaped.

XSL•FO
RenderX

Figure 2. Illustration of the game with open kinetic chain motion. From left to right are the demonstrations of knee joint angles in 90°, 120°, 150°, and 180°, which correspond to 4 trials.



Auditory Feedback

The sound used in AF mode consisted of a short, percussive tone extracted from a rain sound recording, which gave a crisp, bell-like quality with natural texture. The pitch of the sonification was modulated according to the vertical position of the task object, such that the higher the bunny or hook, the louder the sound. The audio was delivered through the Unity engine, with amplitude scaled linearly from silence to the maximum nonclipping volume. The duration of 1 auditory stimulus was 0.1-0.2 seconds. After this period, the next fragment with a different pitch was played.

Game Integration With Sensors

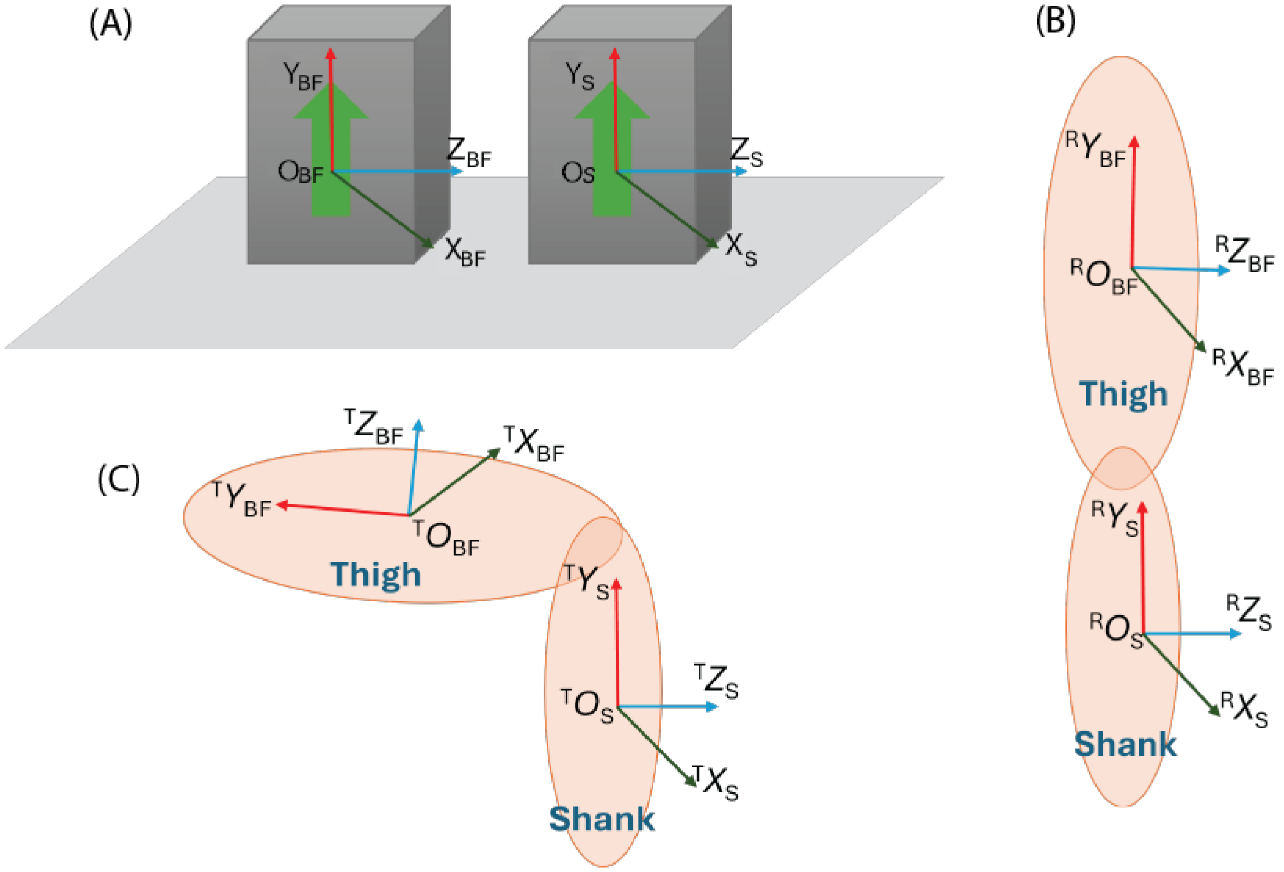
This knee joint training game system was developed using Unity (version 2021.3.0.f1; Unity Technologies). Both kinematic data and EMG data were collected with 9 Avanti sensors (Delsys Europe Ltd). Kinematic data were obtained from 3 Avanti sensors placed on midshank and biceps femoris (BF) muscle on the participant's nondominant leg and the one on the collarbone region of the trunk. Muscular activity of tibialis anterior, gastrocnemius medialis, gluteus medius, gastrocnemius lateralis, peroneus longus, BF, and rectus femoris in the nondominant leg were measured with 6 Avanti sensors. The application programming interface of Delsys was used to stream the data from the IMU sensors in real time into the game environment.

Sensor Calibration

Prior to sensor attachment, the sensors were lying next to each other in a parallel orientation on a flat desk. In this state, the sensors were connected to the game system; when the UI showed that all of the sensors were connected, the initial IMU orientation was recorded to determine the offset of individual sensors. Once all sensors started streaming data, they were attached to the legs of the participant, and the electrode location followed the SENIAM (Surface Electromyography for the Non-Invasive Assessment of Muscle) guidelines [18]. Before testing, the participants were asked to stand in an upright posture to establish a baseline orientation of the IMUs on the user's body. This position offset was used in the kinematic analysis to prevent that difference in body shape affects the game performance. The whole calibration process is shown in Figure 3.

XSL•FO
RenderX

Figure 3. Sensor calibration process illustration: (A) sensors aligned in parallel, (B) sensors attached to the leg in an upright standing position, and (C) sensors in a movement. BF: biceps femoris.



The data obtained during the calibration process were used for the following in-game calculations. To obtain the initial reference quaternion for knee joint angle measurements, the IMUs attached to the BF and the midshank on the nondominant leg were placed parallel to each other on a flat surface. Sensor data were then recorded for 14 seconds at a sampling frequency of 74 Hz, resulting in a buffer of 1000 unit quaternion samples, which was used to compute the reference orientation. Each quaternion in this buffer is represented as $q_i=(w_i, x_i, y_i, z_i)$. The original reference quaternion (q_{avg}) is obtained by averaging all 1000 quaternions component-wise.

To ensure that this average remains a unit quaternion, q_{avg} is normalized by dividing it by its magnitude:

This q_0 serves as the baseline orientation of the sensors.

After this calculation, all of the sensors were attached to the participant by the experimenter, with the midshank sensor placed in the middle of the shank and aligned with the nondominant leg BF sensor. The player was asked to stand upright for 14 seconds, avoiding moving. During this time, the computer calculated the offset (q_{ref}). This information was entered in equation 2 to calculate the rotation q_T between the original quaternion $q_{original}$ and the offset $q_{reference}$ to account for any orientation bias due to the leg shape.

During game training, the real-time sensor orientation is computed by multiplying the measured quaternion (q_R) and the offset (q_T):

The knee angle (q_{knee}) is defined as the angle between the calibrated coordinate system of the midshank and that of the BF of the nondominant leg. This angle can be obtained from equation 4. The Euler angle of the knee in the Y direction is used as the game controller in games.

Power Calculation

The effect size of AF was estimated to be 0.78, based on the results of Ghai et al [19] and Fujii et al [20]. The main results of this study were evaluated using a repeated-measures, between-factors ANOVA. The power was set to 80%, and an α of .05. G*Power (version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf) was used to calculate the sample size a priori. This calculation indicated that a minimum number of 10 participants, including 5 participants in the EG and 5 participants in the CG, were needed to draw a conclusion. Considering a dropout rate of 30%, the recruitment target was set to 14 participants.

XSL•FO
RenderX

Recruitment

All the participants were recruited among the employees and students of the University of Groningen. To be eligible to participate in this study, participants had to be at least 18 years of age, be healthy based on self-report, and be fluent in either English or Dutch to ensure that they understand the instructions given by the experimenter. Exclusion criteria were pregnancy, orthopedic problems, musculoskeletal disorders, prior surgery on the lower extremities, or abuse of drugs, growth hormones, anabolic steroids, or performance-enhancing substances based on self-report. Furthermore, participants with a known adhesive allergy were excluded to reduce the risk of skin irritation due to sensor attachment. To be able to attach the EMG electrodes, participants were required to wear shorts during the test. Therefore, we also had to exclude participants who were not willing to expose the skin of the legs for any reason.

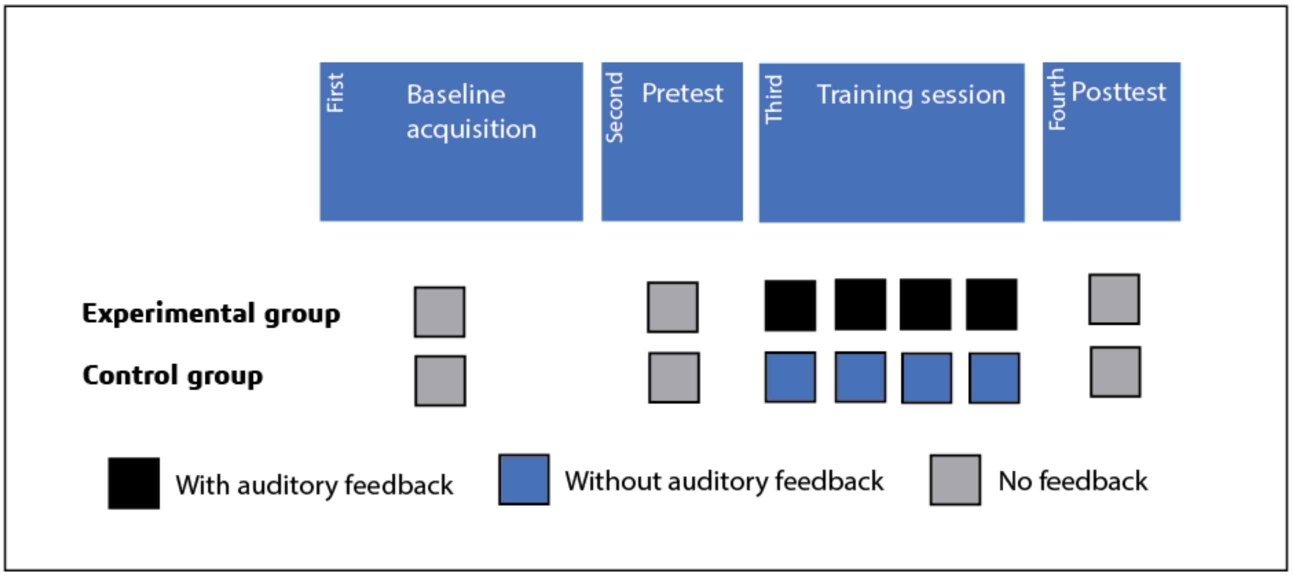
Experiment Design

Overview

The study design of this experiment was a randomized between-group comparison. Participants were randomly assigned to groups using a Microsoft Excel–generated sequence to minimize bias. For comparing the game effectiveness within and without AF, all participants tested both game levels in randomized order. The group the participant belongs to was predetermined via a vector generated with the random function in Excel. This vector was used to assign each participant to 1 group based on the order in which they were recruited. AF was assigned to participants in the EG, but not to the participants in the CG. To rule out confounders, the participants were assigned to the EG and CG at random, apart from the AF; both groups followed the same experimental procedure.

There were 4 blocks in the experiment, as shown in Figure 4.

Figure 4. Flow chart of the study design overview and the main procedures that subjects will undergo in the course of research.



After the sensor calibration, the pretest was performed. During the pretest, each participant performed 10 squats and 10 knee stretches. Participants during the pretest were instructed by written information on the screen on which movement they should perform. Each movement was held for approximately 5 seconds. During the squatting sequence of the pretest, the commands “half squat” or “deep squat” appeared in randomized order. During the stretching sequence of the pretest, the prompts “half stretch (45°)” or “totally stretch (0°)” appeared in random order by the script used for prompt text generation. Participants were seated for this test, and they were previously instructed to obtain a position that is in the middle between totally stretching and having their feet in a 90°. As all participants had a mathematical background, no further explanation of the instructions was necessary. Participants were not blindfolded during the test. Psychometric properties of the test are not available. Participants who were randomly allocated to start with the bunny game started the pretest with the squatting session, while participants who started with the fish game were first prompted with the stretching exercises. To initiate the pretest, the experimenter selected the “Pre-Test” scene in the game menu. This game menu consists of 3 buttons: “Bunny Game,” “Fish Game,” and “Pre-Test,” which allow players to select the different games. In the “Pre-Test” scene, the screen displays 2 buttons: “Stretch Test” and “Squat Test.” The experimenter used these buttons to select the tests in the order indicated on the random allocation list. After completing a test, the exit button returns to the main menu scene, with the Pre-Test button now labeled Post-Test. The Fish Game and Bunny Game options remain available.

After the pretest, each participant completed 4 rounds of training for both games. The training platform for the EG included AF, while for the CG, AF was turned off, leaving only VF.

In the final block, a posttest was conducted, consisting of 10 nondominant knee stretches and 10 squats. The test content and format were identical to those of the pretest. During the pre- and the posttest, participants were asked to perform 10 half squats (target knee angle [TKA]=90°) and full squats (TKA=130°). In addition, participants were asked to sit on a chair and stretch their leg 10 times with a TKA of 135° or 180°, which equals a half and full stretch, respectively.

XSL•FO
RenderX

Study Parameters

The main outcome was the error of position between TKA and actual knee angle (AKA) of the participant along the femoral-tibial axis. The AKA was calculated based on the IMU sensors as described in the Game Development section. The error of position was calculated by subtracting TKA from AKA (equation 5).

$$\text{Error}=\text{AKA}–\text{TKA} \quad \textbf{(5)}$$

According to both the IMU signal and the recorded video during the test, the IMU data are segmented, with each segment representing a test movement. The error of these data segments was calculated. In each test, average error (AE) is classified by the type of movements, with 2 groups (squat test group and stretch test group) of AE in the pretest and 2 groups of AE in the posttest for each participant. The absolute mean value of each group represented the average performance of the participant in the corresponding movement during the test (average mean error [AME]), calculated by equation 6, where n denotes the number of times this type of movement appeared in the test, and AE denotes the average error of a movement.

For each participant, the improvement of training was represented by the difference in performance of each group between pretest and posttest. Equation 7 expresses the calculation of the improvement for a participant in 1 movement classification, which is denoted by IE (improvement of error trajectory), AME_{Pre} denotes the absolute mean value of a group of movement errors in the pretest, and AME_{Post} denotes the absolute mean value of the group of movement error in the posttest.

$$\text{IE}=\text{AME}_{\text{Pre}}–\text{AME}_{\text{Post}} \quad \textbf{(7)}$$

Self-Report Instruments

Participants completed the Task Evaluation Questionnaire of Intrinsic Motivation Inventory (IMI) [21], a multidimensional measurement tool used to assess participants' subjective motivation during game training. The following 4 subscales: interest or enjoyment, perceived competence, perceived choice, and pressure or tension, were assessed based on the Task Evaluation Questionnaire. Each subscale consisted of 5-7 items rated on a 7-point Likert scale (1=not at all true, 4=somewhat true, and 7=very true) [22].

The User Experience Questionnaire (UEQ) [23] was also filled after the experiment. Six dimensions were measured: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. Each dimension was assessed through a set of bipolar adjective pairs (eg, "annoying-enjoyable"), rated on a 7-point scale from –3 to +3. Higher scores indicated a more positive user experience [23].

Muscle Synergy

In this study, muscle synergy analysis was performed to examine the differences in muscle weighting patterns across different synergies during game training, following the methodology described in Hayes et al [24].

EMG signals were collected from 7 muscles (tibialis anterior, gastrocnemius medialis, gluteus medius, gastrocnemius lateralis, peroneus longus, BF, and rectus femoris) on the nondominant leg. The EMG signals were segmented based on task-specific event timestamps. For the Fish Game, segments were extracted from 2 seconds before the timestamp at which a fish was caught up to the event itself. For the Bunny Game, a valid task was defined as the collection of a complete group of 9 consecutive carrots; the corresponding segment spanned from the timestamp of the first carrot to that of the ninth in a single group. Each segment was processed with a fifth-order Butterworth low-pass filter (cutoff: 30 Hz) and high-pass filter (cutoff: 4 Hz). Following filtering, the signals were detrended by subtracting the mean and then rectified by taking the absolute value. Finally, each signal was normalized by dividing it first by its maximum value and then by the SD of the resulting signal to achieve unit variance.

Muscle weights were extracted using nonnegative matrix factorization. To determine the appropriate number of synergies required to reconstruct the original EMG signal, the variance accounted for (VAF) was calculated. For each trial, the number of synergies was identified as the minimum number for which the average VAF dropped below thresholds of 90%, 85%, and 80%, respectively, ensuring a balance between model simplicity and signal reconstruction accuracy.

Statistics

Q-Q plots were used to test for the distribution of the data. As the data were not normally distributed, the nonparametric alternative for a repeated measures ANOVA was used. Therefore, a 1-sided Mann-Whitney U test was conducted for answering whether AF has a positive effect on the knee joint training based on the exergame platform. To ensure data quality and reduce the impact of extreme values, outliers were identified and removed using the z score method. Specifically, for each variable, z scores were calculated, and data points with absolute z scores greater than a predefined threshold ($|z|>2.5$) were regarded as extreme data and excluded from the analysis. The threshold was chosen to balance outlier detection while retaining sufficient data for robust statistical analysis.

The null hypothesis (H_0) is that AF does not significantly improve trajectory error compared to without AF during stretching and squatting. The observations are squat test IE in the EG, squat test IE in the CG, stretch test IE in the EG, and stretch test IE in the CG. A significance level of $\alpha=.05$ was used for hypothesis testing.

Ethical Considerations

The study was carried out in April 2025 in the laboratories of the Discrete Technology and Production Automation group of the Engineering and Technology Institute of the University of Groningen in Groningen, The Netherlands. The study was conducted in accordance with the Declaration of Helsinki. The institutional ethics committee (Research Ethics Review Committee [CETO]) of the Faculty of Arts, University of Groningen, reviewed the experiment protocol and had no objection to the proposal (ID 93318312). Before the sessions started, participants had received an information letter describing

XSL•FO RenderX

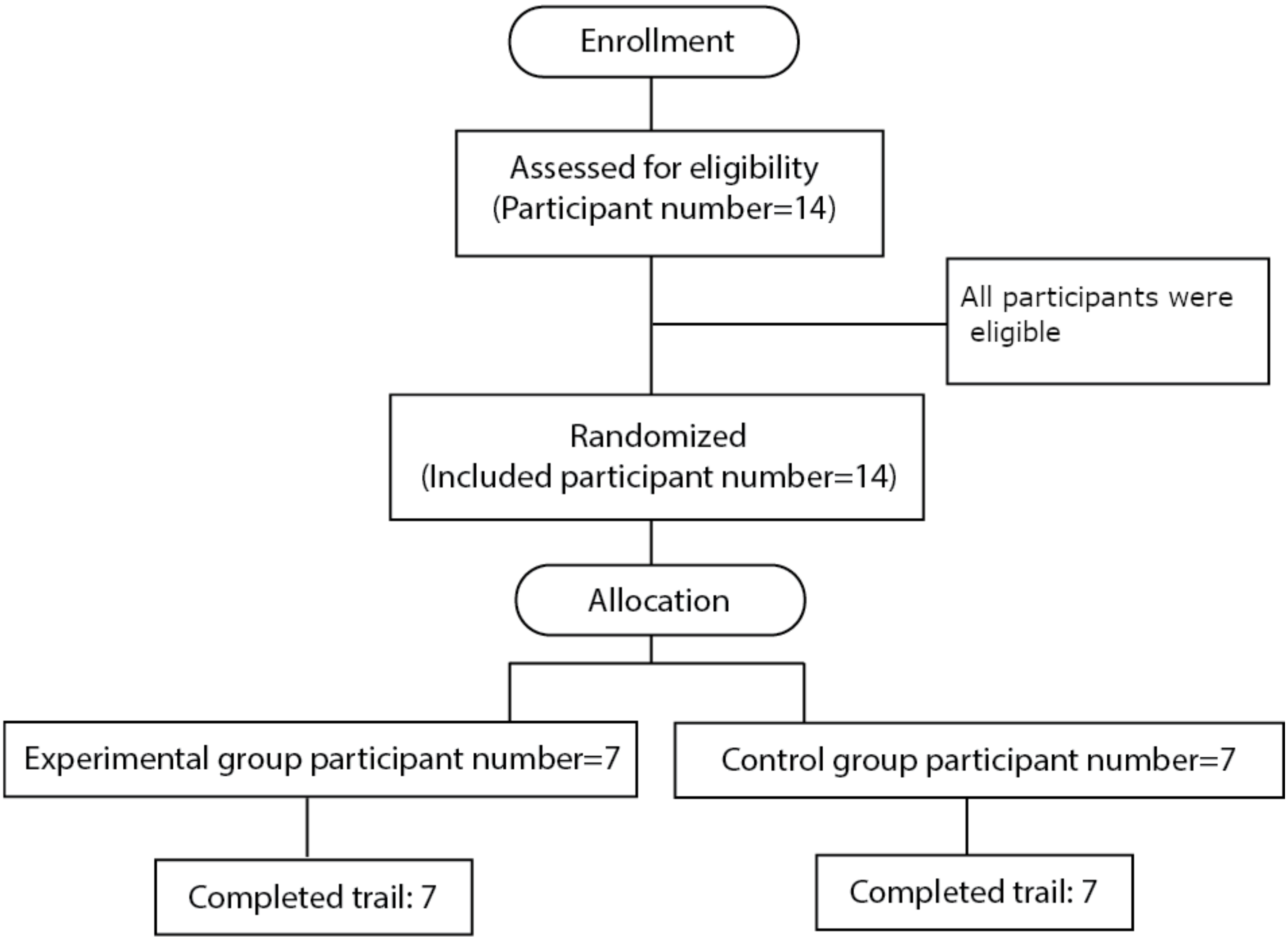
the aim and procedures carried out during the study. They were given the opportunity to ask questions about their participation. Participants, who signed the informed consent form, were required to fill in a short demographic questionnaire and the Waterloo Footedness Questionnaire [25]. Afterward, the experimenter tested their hearing capability by Audiometer 9910 (Amplivox Ltd). Prior to data acquisition, a Data Protection Impact Assessment was carried out to ensure that the data processing is in line with European regulations. As a conclusion of this assessment, the data were stored in a secure virtual research workspace of the University of Groningen. The results are reported as median or average to ensure that individuals cannot be identified. Participants did not receive any compensation for their participation in these experiments.

Results

Participants

A total of 14 participants were recruited for this study: 7 participants were in the EG, and 7 participants were in the CG. All of the participants were right-leg dominant. In the augmented AF group, there were 4 female and 3 male participants; the average age was 27.58 (SD 3.09) years, and the mean BMI was 21.07 (SD 3.04) kg/m^2. The VF group consisted of 3 female and 4 male participants; the average age was 28.69 (SD 4.17) years, and the mean BMI was 21.15 (SD 2.86) kg/m^2. The trial ended when the recruitment target was reached. A flowchart describing the recruitment process is depicted in Figure 5.

Figure 5. Flowchart depicting the recruitment process.



Kinetic Analysis

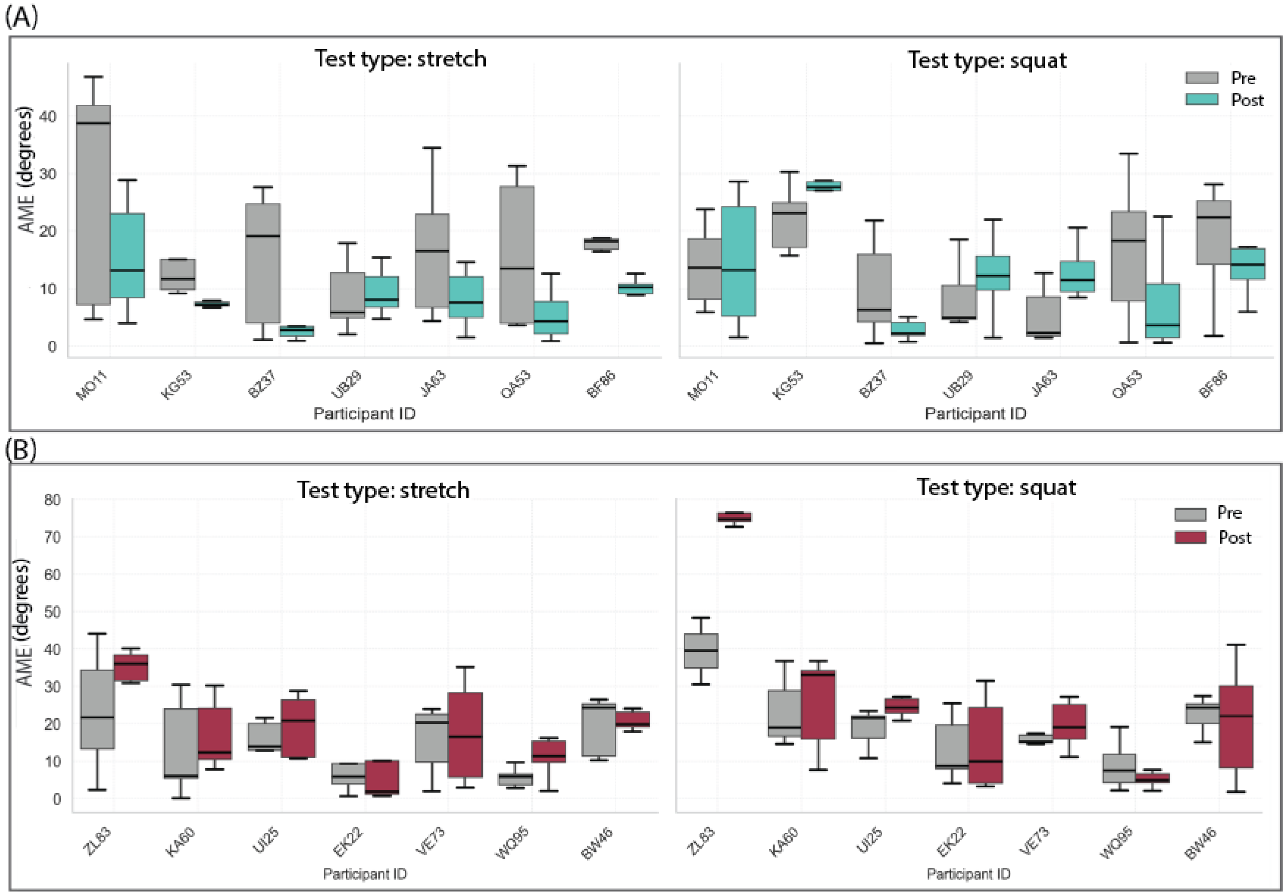
Figure 6 depicts the AME during pre- and posttest per movement category and participant. Figure 6A illustrates the AME of 7 participants in the EG. In the stretch test, all EG participants showed a smaller AME, indicating an improvement in their ability to reproduce the angles. In the squat test, only 3 participants of the EG showed obvious improvement in the posttest. The other 4 participants either had a huge spread in the AME or did not improve in the posttest compared to the pretest.

Figure 6B displays the AME for 7 participants in the CG. During both the stretch and the squat test, the participants of this group showed nearly no improvement in the AME.

The violin plot depicted in Figure 7 illustrates the improvement in knee joint performance for 2 test movements in both the EG and CG. Each violin plot represents the distribution of IE values within 1 category.

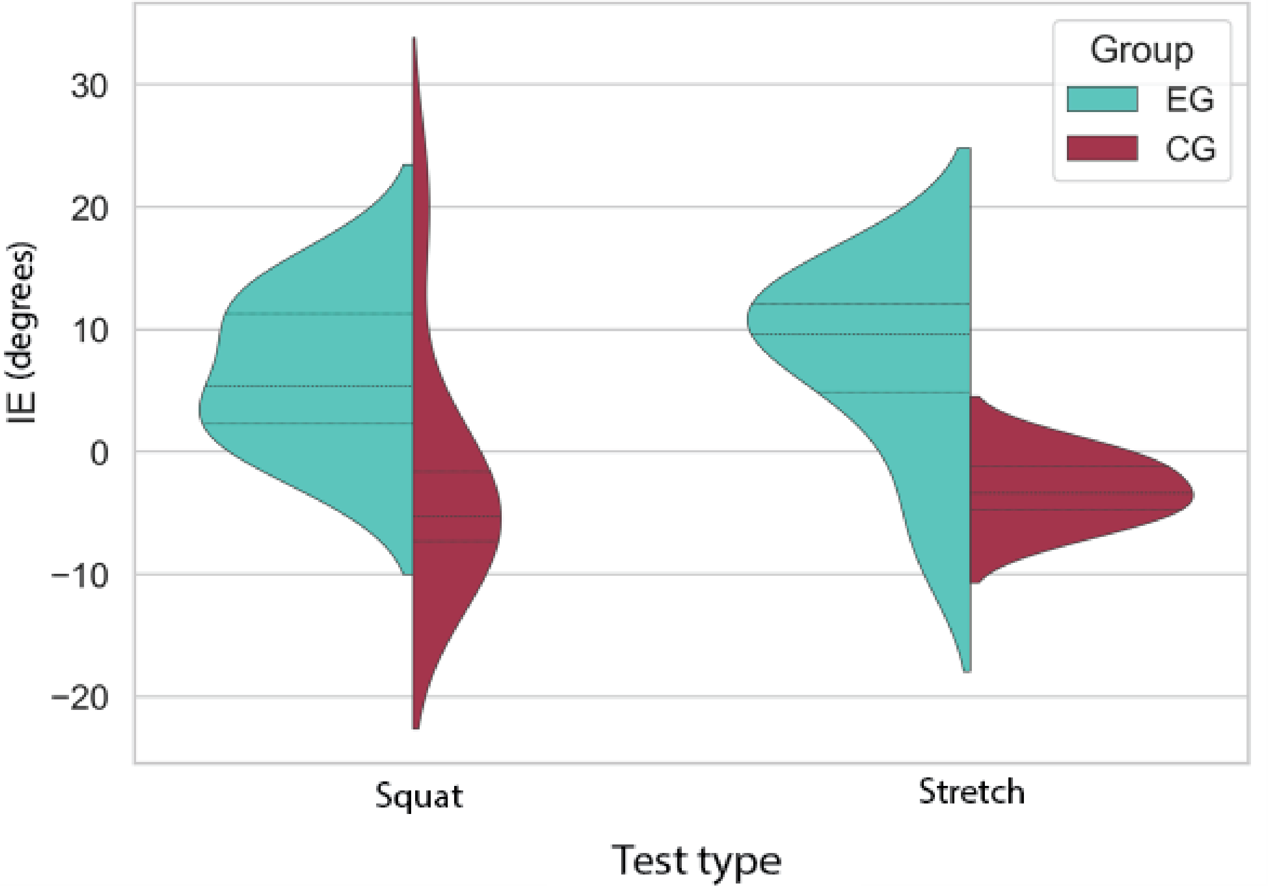
XSL•FO
RenderX

Figure 6. AME in the stretch and squat test in both pre- and posttest for the experimental group (EG) and the control group (CG). (A) AME of 7 participants in the EG during pre- and posttest. (B) AME of 7 participants in the CG during pre- and posttest. AME: average mean error.



XSL•FO
RenderX

Figure 7. Squat and stretch test IE after intervention (with and without augmented feedback) illustration. CG: control group; EG: experimental group; IE: improvement of error trajectory.



For the squat and stretch test groups, the mean value in the EG is higher than in the CG, indicating a notable improvement in squatting performance for the EG compared to the CG. A similar, though slightly smaller, trend is observed between the stretch test in EG and CG. The quartiles (first quartile Q1, median, and third quartile Q3) of each group of tests are depicted in Figure 7. The median improvement of the experimental group was 5.36 (IQR 2.27-11.25) and 10.46 (IQR 5.45-12.07) in the squat and stretch test, respectively. The control group achieved an IE of –5.31 (IQR –7.32 to –1.62) and –3.37 (IQR –5.73 to –1.22) in the squat and stretch test, respectively.

A 1-sided Mann-Whitney *U* test was conducted on the IE datasets, revealing a statistically significant difference between the EG and the CG (statistics=41.0; *P*=.04) for squatting and between the EG and the CG (statistics=42.0; *P*=.03) for stretching.

Participants who received augmented feedback demonstrated greater improvements in both squat and stretch tasks compared to the CG. H_0 was rejected, indicating that integrating augmented AF into the exergame platform positively impacts knee joint motor training.

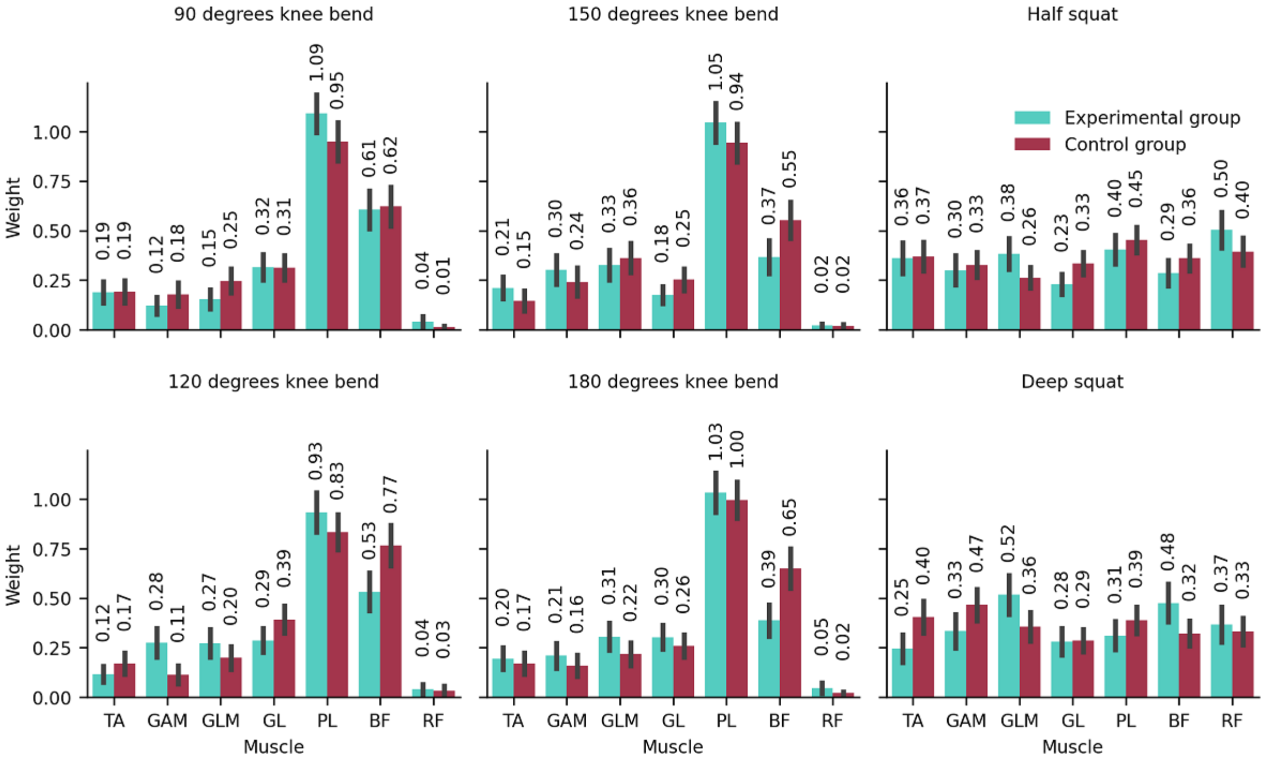
Muscle Synergy

EMG data from all 14 participants (1054 exercise movements in total) were included in the muscle synergy analysis. The number of synergies required to meet the VAF thresholds varied across trials. For the 90% VAF threshold, synergy numbers ranged from 4 to 7; a total of 3 movements had 4 synergies, 98 movements had 5 synergies, 787 movements had 6 synergies, and 166 movements had 7 synergies. For 85% VAF, synergy numbers ranged from 3 to 6; a total of 3 movements had 3 synergies, 29 movements had 4 synergies, 342 movements had 5 synergies, and 682 movements had 6 synergies. For 80% VAF, synergy numbers ranged from 3 to 6; a total of 4 movements had 3 synergies, 115 movements had 4 synergies, 683 movements had 5 synergies, and 252 movements had 6 synergies. The 85% VAF threshold was selected for subsequent analysis to balance signal reconstruction quality and model simplicity, 6 synergies appeared most frequently across participants and movements, suggesting a stable underlying structure in muscle coordination.

One of the extracted muscle synergy weight patterns is presented in Figure 8. The complete results of the synergy analysis can be found in Multimedia Appendix 2. The first and second columns display the 4 tasks from the Fish Game, comparing 2 groups. Overall, the muscle weight distributions between the groups were largely consistent. The third column shows the 2 tasks from the Bunny Game, half squat and deep squat, with similarly consistent muscle weighting patterns observed between the 2 groups.

XSL•FO
RenderX

Figure 8. Muscle synergy analysis including tibialis anterior (TA), gastrocnemius medialis (GAM), gluteus medius (GLM), gastrocnemius lateralis (GL), peroneus longus (PL), biceps femoris (BF), and rectus femoris (RF).



Quantitative Analysis of Self-Report Measures

Table 1 summarizes the IMI subscale scores for EG and CG. The EG reported higher perceived choice, and the CG reported higher interest or enjoyment, perceived competence, and pressure or tension.

The results of UEQ were analyzed from the following 5 subscales: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. In EG, the mean of subscales is 2.738 (SD 0.31), 2.821 (SD 0.14), 2.143 (SD 0.46), 2.500 (SD 0.27), 2.536 (SD 0.43), and 2.000 (SD 1.15), respectively. In CG, the mean (variance) of subscales is 2.357 (SD 0.70), 2.714 (SD 0.03), 2.286 (SD 0.28), 2.286 (SD 0.38), 1.929 (SD 1.31), and 1.679 (SD 2.08), respectively. According to the rating scheme of the questionnaire, these scores indicate that the experience was rated “excellent” [23]. The EG showed higher scores across all dimensions, particularly in attractiveness (mean 2.738, SD 0.31) and novelty (mean 2.000, SD 1.15), indicating a more engaging and novel user experience. Furthermore, to access the overall experience, the subscales were grouped into pragmatic and hedonic quality according to UEQ guidelines [23]. The EG scored higher on both pragmatic quality (mean 2.49) and hedonic quality (mean 2.27) compared to the CG (mean 2.43 and mean 1.80, respectively).

Table 1. Intrinsic Motivation Inventory subscale scores by group.

Subscale	Experimental group, mean (SD)	Control group, mean (SD)
Interest or enjoyment	35 (10)	35 (9)
Perceived competence	29 (7)	29 (7)
Perceived choice	8 (4)	8 (4)
Pressure or tension	–6 (4)	–5 (4)

Discussion

Main Results

The aim of this study was to explore whether augmented AF applied in motor learning during game-based knee joint exercise (involving both CKC and OKC) improves the knee angle proprioception more than VF alone. The results demonstrated that augmented AF exergame training improved the AME of the knee joint angle significantly more than VF alone. After both CKC and OKC game training, participants in the EG demonstrated significantly better knee angle positioning skills. This suggests that the benefit of auditory sonification cues can positively support both CKC and OKC knee joint tasks.

The results of self-report instruments, IMI and UEQ, showed that the experience of the participants was positive in both games. The IMI suggests that the participants felt low pressure while experiencing enjoyment, which is desired in an exergame. EG participants reported a more immersive, attractive, and novel experience compared with CG participants. These findings are consistent with prior research that supports the role of

XSL•FO
RenderX

augmented feedback, especially auditory, in facilitating sensorimotor integration and enhancing motor skill learning [13]. Auditory cues provide immediate, real-time reinforcement, which can guide users' movements, keep users engaged, and increase interactions during motor learning.

As shown in Figure 6, most participants in EG exhibited a reduction in AME during the OKC test following the augmented feedback game training. In addition, the SD of the error decreased from pretest to posttest for each participant, indicating reduced variability and more stable task performance after training. In CG, 6 of 7 participants also had a decreased SD after training. However, the mean AME did not show improvement in the OKC tasks after training. These findings are not in line with previous literature, which reports that improvements can be reached with both CKC [16] and OKC [15] exercises. However, the studies in the literature observe participants over a duration of 6 weeks, while we only investigated a single game session. Furthermore, the participants in the studies reported in the literature were older male individuals and patients with osteoarthritis. Our population was much younger and healthy. Therefore, the baseline might already have been better. Due to differences in the instructions during the assessment, a direct comparison of the absolute values is unfortunately not possible.

For the CKC test (squat test), the performance in EG after training does show smaller improvements than the OKC test result. In total, 4 of 7 participants in the EG decreased the AME after interventions, and 6 participants had a decreased SD after training. Compared to the EG, AME results of the CG show less improvement. Figure 7 demonstrates a more intuitive comparison. In EG, the improvement of knee joint positioning skills is generally greater than the CG in both squat and stretch tests. The OKC game (Fish Game) showed greater improvements in motor learning compared to the CKC game (Bunny Game).

Another observation in this study is the consistency in muscle synergy weight pattern between EG and CG across both game contexts. As shown in Figure 8, the muscle weight distributions are remarkably similar between 2 group participants, suggesting that motor control strategies in the muscle remained stable regardless of group assignment. This trend is also evident in both games. The stability in muscle synergy organization supports the use of game-based interventions as a tool for training repetitive motions.

Limitations

Several limitations should be acknowledged in this study. First, all participants were recruited from the University of Groningen, consisting exclusively of students and researchers. This relatively homogenous sample limits the generalizability of the findings to broader populations. Second, the experiment was conducted within a single day, and the study does not capture long-term retention or progression of motor learning over time. Third, due to the nature of AF, it was not possible to blind the participants with respect to group assignment. Therefore, the subjective measures taken with the questionnaires might be influenced by the fact that participants were aware whether they were in the control or the intervention group.

Fourth, the system has not yet been tested in a clinical population. While the current results are promising, validation in clinical settings is essential to confirm the utility and effectiveness of the developed augmented AF game training system in rehabilitation contexts for the knee joint. Finally, both Fish Game and Bunny Game include 2 distinct types of motor tasks: spatial positioning of the knee joint and knee joint extreme movements. These task types were not separately analyzed in this study.

The most noticeable limitation of our paper is that participants were not blindfolded during the pre- and posttest. Therefore, the extent to which they used vision to compensate for eventual proprioceptive deficits is unknown. Due to this limitation, we can only say that the study had positive effects on motor learning in general. Future studies that test proprioception while occluding vision and apply psychometric testing methods are needed to evaluate the precise contribution of proprioception.

Implication for Research and Future Work

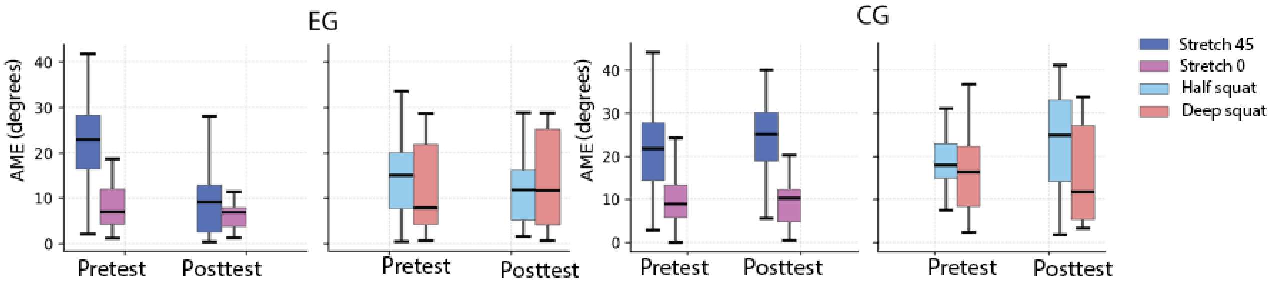
Despite the limitations, this study provides a prototype of an exergame for knee joint motor learning. The result shows the benefit of augmented feedback in interactive, feedback-driven games to support effective motor learning in the lower limbs.

Moving forward, implementing longitudinal study designs would allow for assessment of motor learning retention and the long-term impact of knee joint training based on an augmented feedback exergame. Clinical trials are also essential to explore whether the game can be used in the context of rehabilitation. Finally, further analysis should differentiate between the 2 types of motor tasks, knee joint positioning and extreme joint movement, present in the games. As shown in Figure 9, the distributions of AME between "stretch 45," "stretch 0," "half squat," and "deep squat" are obviously different. For OKC tasks, knee joint positioning shows a bigger variance than the extreme joint movement. For CKC tasks, the knee joint positioning shows a smaller variance than the extreme joint movement. Investigating these categories separately could uncover task-specific motor strategies. Together, these directions can build upon the promising outcomes of this study to inform the design of adaptive, personalized game-based interventions for the knee joint motor skill learning and rehabilitation.

The tasks that we selected in this study were focused on specific exercises that are frequently used in knee rehabilitation. Future studies could also investigate how exergaming can be used for activities of daily living such as walking.

XSL•FO
RenderX

Figure 9. AME distribution comparison among 4 test movements. CG: control group; EG: experimental group.



Conclusions

This randomized controlled trial study explored the effectiveness of CKC and OKC game-based interventions in promoting motor learning with augmented feedback in healthy adults. The results demonstrated that augmented feedback significantly enhanced knee joint motor learning performance in both game contexts, highlighting its value as an intuitive and engaging form of feedback. Additionally, consistent muscle synergy patterns across tasks and groups suggest that the neuromuscular strategies used were stable and repeatable, reinforcing the reliability of augmented feedback game-based training for knee joint motor control. In the future, targeted populations such as clinical groups could be considered to test the effectiveness of rehabilitation.

Acknowledgments

The authors thank Andreia Abud and Lisanne Bakker for their guidance on how to apply the electromyography sensors on the leg muscles.

Funding

This work is supported by China Scholarship Council (202006540022) and the Colombian Ministry of Science, Technology and Innovation (Minciencias).

Authors' Contributions

YZ, EW, MC, and HT designed the study protocol. YZ, LFGA, and EW programmed the game and analyzed the data. YZ drafted the first version of the manuscript. All authors reviewed the manuscript and gave feedback.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1292 KB - rehab_v13i1e78525_app1.pdf]

Multimedia Appendix 2
Complete overview of the results of the muscle synergy analysis.
[DOCX File , 570 KB - rehab_v13i1e78525_app2.docx]

References

1. Wang Y, Wu Z, Chen Z, Ye X, Chen G, Yang J, et al. Proprioceptive training for knee osteoarthritis: a systematic review and meta-analysis of randomized controlled trials. Front Med (Lausanne) 2021;8:699921 [FREE Full text] [doi: 10.3389/fmed.2021.699921] [Medline: 34778281]
2. Aman JE, Elangovan N, Yeh I, Konczak J. The effectiveness of proprioceptive training for improving motor function: a systematic review. Front Hum Neurosci 2015;8:1075 [FREE Full text] [doi: 10.3389/fnhum.2014.01075] [Medline: 25674059]
3. Winter L, Huang Q, Sertic JVL, Konczak J. The effectiveness of proprioceptive training for improving motor performance and motor dysfunction: a systematic review. Front Rehabil Sci 2022;3:830166 [FREE Full text] [doi: 10.3389/fresc.2022.830166] [Medline: 36188962]
4. An J, Son YW, Lee BH. Effect of combined kinematic chain exercise on physical function, balance ability, and gait in patients with total knee arthroplasty: a single-blind randomized controlled trial. Int J Environ Res Public Health 2023;20(4):3524 [FREE Full text] [doi: 10.3390/ijerph20043524] [Medline: 36834218]

XSL•FO
RenderX

5. Ellenbecker TS, Davies GJ. Closed Kinetic Chain Exercise: A Comprehensive Guide to Multiple Joint Exercise. Champaign, IL: Human Kinetics; 2001.
6. Pamboris GM, Pavlou K, Paraskevopoulos E, Mohagheghi AA. Effect of open vs. closed kinetic chain exercises in ACL rehabilitation on knee joint pain, laxity, extensor muscles strength, and function: a systematic review with meta-analysis. Front Sports Act Living 2024;6:1416690 [FREE Full text] [doi: 10.3389/fspor.2024.1416690] [Medline: 38887689]
7. Herren S, Seebacher B, Mildner S, Riederer Y, Pachmann U, Böckler NS, et al. Exergame (ExerG)-based physical-cognitive training for rehabilitation in adults with motor and balance impairments: usability study. JMIR Serious Games 2025;13:e66515 [FREE Full text] [doi: 10.2196/66515] [Medline: 39951650]
8. Gelineau A, Perrochon A, Robin L, Daviet J, Mandigout S. Measured and perceived effects of upper limb home-based exergaming interventions on activity after stroke: a systematic review and meta-analysis. Int J Environ Res Public Health 2022;19(15):9112 [FREE Full text] [doi: 10.3390/ijerph19159112] [Medline: 35897472]
9. Fan B, Li Q, Tan T, Kang P, Shull PB. Effects of IMU sensor-to-segment misalignment and orientation error on 3-D knee joint angle estimation. IEEE Sens J 2022 Feb 1;22(3):2543-2552. [doi: 10.1109/jsen.2021.3137305]
10. Kim JJ, Cho H, Park Y, Jang J, Kim JW, Ryu JS. Biomechanical influences of gait patterns on knee joint: kinematic and EMG analysis. PLoS One 2020;15(5):e0233593 [FREE Full text] [doi: 10.1371/journal.pone.0233593] [Medline: 32470052]
11. Yang J, Lu Z, Chen S, Liu C, Zhao H. Continuous knee joint angle prediction with surface EMG. Biomed Signal Process Control 2024 Sep;95:106354. [doi: 10.1016/j.bspc.2024.106354]
12. Salmoni AW, Schmidt RA, Walter CB. Knowledge of results and motor learning: a review and critical reappraisal. Psychol Bull 1984;95(3):355-386. [doi: 10.1037//0033-2909.95.3.355]
13. Sigrist R, Rauter G, Riener R, Wolf P. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. Psychon Bull Rev 2013 Feb;20(1):21-53 [FREE Full text] [doi: 10.3758/s13423-012-0333-8] [Medline: 23132605]
14. Ronsse R, Puttemans V, Coxon J, Goble DJ, Wagemans J, Wenderoth N, et al. Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. Cereb Cortex 2011 Jun;21(6):1283-1294. [doi: 10.1093/cercor/bhq209] [Medline: 21030486]
15. Sadeghi H, Hakim MN, Hamid TA, Amri SB, Razeghi M, Farazdaghi M, et al. The effect of exergaming on knee proprioception in older men: a randomized controlled trial. Arch Gerontol Geriatr 2017;69:144-150 [FREE Full text] [doi: 10.1016/j.archger.2016.11.009] [Medline: 27923177]
16. Mete E, Sari Z. The efficacy of exergaming in patients with knee osteoarthritis: a randomized controlled clinical trial. Physiother Res Int 2022 Jul;27(3):e1952. [doi: 10.1002/pri.1952] [Medline: 35470534]
17. Manlapaz DG, Sole G, Jayakaran P, Chapple CM. Exergaming to improve balance and decrease the risk of falling in adults with knee osteoarthritis: a mixed-methods feasibility study. Physiother Theory Pract 2022 Nov;38(13):2428-2440. [doi: 10.1080/09593985.2021.1952670] [Medline: 34280069]
18. Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. Development of recommendations for SEMG sensors and sensor placement procedures. J Electromyogr Kinesiol 2000 Oct;10(5):361-374. [doi: 10.1016/s1050-6411(00)00027-4] [Medline: 11018445]
19. Ghai S, Schmitz G, Hwang T, Effenberg AO. Training proprioception with sound: effects of real-time auditory feedback on intermodal learning. Ann N Y Acad Sci 2019 Feb;1438(1):50-61. [doi: 10.1111/nyas.13967] [Medline: 30221775]
20. Fujii S, Lulic T, Chen JL. More Feedback Is Better than Less: Learning a Novel Upper Limb Joint Coordination Pattern with Augmented Auditory Feedback. Front Neurosci 2016;10:251 [FREE Full text] [doi: 10.3389/fnins.2016.00251] [Medline: 27375414]
21. Ryan RM, Mims V, Koestner R. Relation of reward contingency and interpersonal context to intrinsic motivation: a review and test using cognitive evaluation theory. J Pers Soc Psychol 1983 Oct;45(4):736-750. [doi: 10.1037/0022-3514.45.4.736]
22. Alhirsan SM, Capó-Lugo CE, Brown DA. Effects of different types of augmented feedback on intrinsic motivation and walking speed performance in post-stroke: a study protocol. Contemp Clin Trials Commun 2021 Dec;24:100863 [FREE Full text] [doi: 10.1016/j.conctc.2021.100863] [Medline: 34841123]
23. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Holzinger A, editor. HCI and Usability for Education and Work. Berlin, Heidelberg: Springer; 2008:63-76.
24. Hayes HB, Chvatal SA, French MA, Ting LH, Trumbower RD. Neuromuscular constraints on muscle coordination during overground walking in persons with chronic incomplete spinal cord injury. Clin Neurophysiol 2014 Oct;125(10):2024-2035 [FREE Full text] [doi: 10.1016/j.clinph.2014.02.001] [Medline: 24618214]
25. Elias LJ, Bryden M, Bulman-Fleming M. Footedness is a better predictor than is handedness of emotional lateralization. Neuropsychologia 1998 Jan;36(1):37-43. [doi: 10.1016/s0028-3932(97)00107-3] [Medline: 9533385]

Abbreviations

AE: average error
AF: auditory feedback
AKA: actual knee angle
AME: average mean error

XSL•FO
RenderX

BF: biceps femoris
CG: control group
CKC: closed kinetic chain
CONSORT: Consolidated Standards of Reporting Trials
EG: experimental group
EMG: electromyography
IE: improvement of error trajectory
IMI: Intrinsic Motivation Inventory
IMU: Inertial Measurement Unit
KOR: knowledge of result
OKC: open kinetic chain
SENIAM: Surface Electromyography for the Non-Invasive Assessment of Muscle
TKA: target knee angle
UEQ: User Experience Questionnaire
UI: user interface
VAF: variance accounted for
VF: visual feedback

Edited by S Munce; submitted 04.Jun.2025; peer-reviewed by P Marsico, PE Roos; comments to author 15.Oct.2025; revised version received 09.Dec.2025; accepted 23.Dec.2025; published 26.Feb.2026.

Please cite as:
Zhang Y, García Arias LF, Timmerman H, Cao M, Wilhelm E
Enhancing Knee Joint Proprioception in Healthy Adults Through Exergame Training With Augmented Feedback: Randomized Controlled Pilot Trial
JMIR Rehabil Assist Technol 2026;13:e78525
URL: https://rehab.jmir.org/2026/1/e78525
doi:10.2196/78525
PMID:

©Yiling Zhang, Luis Felipe García Arias, Hans Timmerman, Ming Cao, Elisabeth Wilhelm. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of a Stretching Forearm Sleeve for Lateral Epicondylitis: Repeated Measures Study

Adriana M Ríos Rincón[1*], OTR, MSc, PhD; Christine Guptill[2], MS, OTR, PhD; Ann Tran[1], OTR; Rija Kamran[1], PT; Salamah Alshammari[1,3], OTR; Antonio Miguel Cruz[1,4*], MSc, DSC

[1]Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

[2]Department of Occupational Therapy, University of Ottawa, Ottawa, ON, Canada

[3]Occupational Therapy Program, College of Applied Medical Sciences, King Saud bin Abdulaziz University for Health Sciences, Alahsa, Eastern Region, Saudi Arabia

[4]Glenrose Rehabilitation Research, Innovation & Technology, Glenrose Rehabilitation Hospital, Edmonton, AB, Canada

[*]these authors contributed equally

Corresponding Author:
Antonio Miguel Cruz, MSc, DSC
Department of Occupational Therapy
Faculty of Rehabilitation Medicine
University of Alberta
8205 114 St NW
2-64 Corbett Hall
Edmonton, AB, AB T6G 2G4
Canada
Phone: 1 7804926104
Email: miguelcr@ualberta.ca

Abstract

Background: Lateral epicondylitis (LE) is a condition that impairs daily activities due to pain exacerbated by wrist and hand movements. The ArmLock sleeve is a novel, nonsurgical intervention to stretch the wrist extensor muscles by maintaining the elbow in extension, forearm in pronation, and wrist and fingers in flexion.

Objective: This study aimed to assess the effectiveness of sustained tension provided by the ArmLock sleeve on pain and functional outcomes in adults with LE. This novel device supports the forearm by aligning the elbow in extension, the forearm in pronation, and the wrist in flexion, while keeping the metacarpophalangeal and proximal interphalangeal joints of fingers II-V flexed.

Methods: A within-subjects repeated-measures design was used to assess outcomes at baseline, week 6, and week 12. Nineteen participants wore the device at home for 30 minutes daily for 12 weeks. Outcome measures were administered by research assistants and included pain intensity, pain-free grip strength, pressure pain threshold, pain during resisted wrist extension, and composite extensibility of wrist and finger extensors. Repeated-measures 1-way ANOVA and Friedman tests were conducted (α=.05), followed by post hoc comparisons (α=.017, Bonferroni correction).

Results: Significant improvements were observed in 6 of 7 (85.7%) outcome variables, including increased muscle extensibility, enhanced grip strength, and reduced pain intensity. Participants reported decreased pain and functional disability in a self-reported questionnaire.

Conclusions: Wearing the device daily for 12 weeks led to significant improvements in extensibility, grip strength, and pain reduction. Participants also reported decreased pain and disability. These results suggest that the ArmLock sleeve may support symptom relief and functional gains in individuals with LE. Larger, controlled studies are needed to confirm its effectiveness.

Trial Registration: ISRCTN Registry ISRCTN13309889; https://www.isrctn.com/ISRCTN13309889

(JMIR Rehabil Assist Technol 2026;13:e80400) doi:10.2196/80400

KEYWORDS

lateral epicondylitis; musculoskeletal; orthopedic device; rehabilitation; tennis elbow; upper extremity

XSL•FO
RenderX

Introduction

Lateral epicondylitis (LE), known as tennis elbow, is a frequent cause of lateral elbow pain that significantly impacts the performance of work, self-care, and leisure activities, as the pain intensifies with wrist and hand movements [1]. Aggravating movements often include resisted wrist extension, middle finger extension, and forearm supination with a fully extended elbow [2]. Prevalence studies indicate that LE affects approximately 1.3% of the general population [1], with higher rates observed in working populations (2%-14.5%) [3] and among tennis players (14%-41%) [4]. The condition is predominantly seen in the dominant upper extremity of individuals aged 30-64 years, with peak incidence occurring between ages 45 and 54. LE is strongly associated with repetitive and forceful tasks [1]. As an increasing percentage of the aging population remains physically active, the incidence of LE is expected to rise significantly in the coming years [5]. Given the high prevalence of LE, its socioeconomic impact can be substantial. For example, in the United States, nontraumatic epicondylitis accounts for an annual workers' compensation claim incidence rate of 10.6 per 10,000 full-time equivalent employees [6].

The etiology of LE remains a subject of ongoing debate within the medical community [7]. The most widely accepted causative factor is repetitive microtrauma, resulting from overuse of the wrist extensor or supinator muscles. Such overuse commonly arises from prolonged heavy workloads, intensive training in racquet sports, or improper technique during physical activities [7,8]. While LE has historically been described as a tendinosis based on histopathological findings [7], contemporary models conceptualize the condition as a degenerative tendinopathy characterized by maladaptive responses to mechanical loading, altered mechanobiology, and pain mechanisms that extend beyond local tissue pathology [9]. This distinction has important therapeutic implications, supporting interventions that emphasize load modulation, sustained low-load mechanical stimuli, and functional restoration rather than anti-inflammatory approaches alone. As a degenerative tendinopathy, LE is characterized by maladaptive responses to mechanical loading, which may include excessive or uneven tensile forces acting on the tendon and contributing to microtrauma and degenerative changes [9,10]. In degenerative tendons, the typically organized structure of collagen architecture is disrupted, with fibroblasts and atypical vascular tissue creating a tangled, disorganized matrix with reduced tensile capacity [10]. LE involves chronic tendon degeneration at the common origin of the forearm extensor muscles. These muscles include the extensor carpi radialis brevis (ECRB), extensor digitorum, extensor digiti minimi, and extensor carpi ulnaris, with the ECRB tendon most frequently affected at the lateral epicondyle of the humerus [9,10].

While many individuals with LE experience symptom resolution over time, including spontaneous improvement without formal treatment, most also demonstrate a positive response to conservative management. However, a clinically meaningful subset of patients does not achieve symptom resolution. Approximately 5%-10% of individuals develop persistent symptoms, with pain extending beyond several weeks, exceeding what would be expected from spontaneous recovery alone and suggesting progression toward a chronic or recalcitrant presentation. Patients whose symptoms persist despite appropriate rehabilitation interventions, particularly beyond 3-6 months, are commonly classified as having recalcitrant LE. In this subgroup, continued passive observation is unlikely to yield meaningful improvement, underscoring the need for targeted interventions aimed at modifying tendon loading and mechanobiological processes [9].

A wide range of rehabilitation and nonsurgical interventions has been described for LE, including corticosteroid injections, iontophoresis, botulinum toxin A injections, prolotherapy, platelet-rich plasma (PRP) or autologous blood injections, as well as physical and occupational therapy approaches such as bracing and splinting, stretching and strengthening programs, kinesiotaping, extracorporeal shockwave therapy, and laser therapy. Among these, PRP has been identified as a conservative treatment with higher levels of supporting evidence [11,12]. However, current research does not clearly demonstrate that PRP is superior to other nonsurgical approaches [11,12]. Surgical intervention is typically reserved for approximately 3% of patients who fail to respond to nonsurgical treatments after 1 year from symptom onset [13].

Although the evidence remains somewhat ambiguous, rehabilitation interventions are generally recommended as first-line treatment over medications, corticosteroid injections, or surgery [14]. Braces and splints represent a rehabilitation intervention for LE, particularly for individuals with limited access to other rehabilitation services. These devices aim to reduce mechanical load on the extensor tendons by providing external support. Commonly used orthoses include forearm straps, which decrease force transmission at the common extensor origin, and wrist extension splints, which offload the extensor musculature by immobilizing the wrist. However, evidence regarding their effectiveness in alleviating LE symptoms remains inconclusive [1]. Therapeutic exercise, including stretching, is supported for short-term improvements in pain and function; however, there is no clear consensus on an evidence-based multimodal rehabilitation program for LE, nor on optimal exercise dosage parameters, including intensity, duration, frequency, and progression [14].

Stretching exercises targeting the wrist extensor musculature are commonly incorporated as part of multimodal conservative management strategies for LE [14-16]. A frequently described technique involves wrist flexion with ulnar deviation, combined with elbow extension and forearm pronation, with or without concurrent finger flexion [14,17]. Although the direct effects of static stretching on tendon structure remain incompletely understood, mechanical loading is recognized as an important stimulus for tendon cell activity. Experimental and clinical literature suggests that the mechanical load generated by this upper limb position may support tendon healing in LE by encouraging fibroblasts to align new collagen fibers in an orderly manner, improving the tendon's structural integrity [18]. Accordingly, the mechanical load generated during wrist extensor stretching is hypothesized to contribute to these cellular responses, although the magnitude and durability of such effects in degenerative tendinopathy remain variable and require further investigation.

XSL•FO
RenderX

A promising rehabilitation intervention for LE is the ArmLock sleeve, which applies principles used in therapeutic stretching. The device positions the elbow in extension, the forearm in pronation, and the wrist in flexion, while maintaining flexion of the metacarpophalangeal and proximal interphalangeal joints of digits II-V. This configuration reproduces a frequently prescribed stretching posture used in multimodal rehabilitation programs and is intended to apply sustained mechanical loading to the musculotendinous structures implicated in LE. Specifically, this alignment targets the ECRB, extensor digitorum, extensor digiti minimi, extensor carpi ulnaris, and the superficial head of the supinator, which attach near the lateral epicondyle. By providing prolonged, externally supported positioning, the ArmLock sleeve may offer a practical means of delivering consistent mechanical stimulus as part of a rehabilitation approach, particularly when adherence to repeated active stretching is challenging. However, the effectiveness of the ArmLock sleeve in reducing LE symptoms has not yet been determined. This study aimed to assess the effectiveness of sustained tension provided by the ArmLock sleeve on pain and functional outcomes in adults with LE.

Methods

Study Research Design

A within-subjects repeated-measures design was used to address the research objective. Outcomes related to pain and function were assessed at 3 time points: baseline (week 0), midintervention (week 6), and postintervention (week 12). The use of a within-subjects repeated-measures design to evaluate the effects of the ArmLock sleeve on pain and function in adults with LE aligns with contemporary frameworks for generating evidence in rehabilitation technologies and reflects the technology's level of maturity. Specifically, this study was guided by the Framework for Accelerated and Systematic Technology-based Intervention Development and Evaluation Research (FASTER), which cautions against the routine application of pharmaceutical-style randomized controlled trials (RCTs) to early-stage, rapidly evolving technology-based interventions in disability and rehabilitation [19]. Increasingly, such expectations have been questioned due to their high cost and logistical burden, the risk that extended trial timelines render technologies obsolete, and the limited external validity of tightly controlled experimental conditions. FASTER therefore recommends quasi-experimental designs as rigorous and appropriate approaches for phase 2 (progressive usability and feasibility evaluation). Within-subject repeated-measures designs, in particular, are well suited to this phase because they allow participants to serve as their own controls while also enabling group analysis, thereby facilitating the examination of individual change trajectories and effect signals. The effect sizes and response patterns observed in this study provide critical parameters for subsequent sample size estimation, refinement of research protocol, and hypothesis development in future controlled trials. By situating this study within an iterative, technology-specific evidence-generation pathway, consistent with the FASTER framework, our findings contribute to generating foundational evidence, including effect size estimates, necessary to design future RCTs.

Materials

As illustrated in Figure 1, the ArmLock sleeve is a soft upper-limb orthosis designed to provide sustained positioning and postural support of the elbow, forearm, wrist, and hand. The orthosis incorporates an adjustable strapping system within a semistructured textile sleeve to maintain the elbow in extension, the forearm in pronation, and the wrist in flexion, while positioning the metacarpophalangeal and proximal interphalangeal joints of digits II-V in flexion. The orthosis is fabricated from a breathable, elastic textile composite (74% polyester, 20% rayon, and 6% elastic). Its design allows adjustable traction to increase or decrease wrist and finger flexion to the degree required to achieve the intended positioning. The ArmLock sleeve is classified as a class I medical device, is licensed by Health Canada (Medical Device Establishment Licence [MDEL] No. 7429), and is commercially available.

XSL•FO
RenderX

Figure 1. The ArmLock sleeve.

Participants

Participants were eligible for inclusion if they were 18 years or older, able to communicate in English, and had a self-reported diagnosis of LE accompanied by the following symptoms: (1) pain on the lateral side of the elbow elicited by resisted wrist extension, (2) tenderness at the lateral epicondyle, (3) measurable loss in composite muscle flexibility (extensibility), and (4) symptoms persisting for at least 12 weeks.

Exclusion criteria were applied to minimize confounding factors and ensure the safety and appropriateness of participants for this study. Potential participants were excluded if they (1) had a steroid injection for LE within the past 3 months (steroid injections are a treatment for LE and could confound the study results) and (2) had undergone previous surgery for LE (approximately 10% of postsurgical cases remain symptomatic, suggesting severe LE or underlying conditions unrelated to this study) [9].

Exclusion criteria for symptoms likely due to other medical conditions. These conditions could mimic or contribute to LE-like symptoms or reflect different underlying pathologies, making them unsuitable for this study. Such exclusions included those who (1) had been diagnosed with elbow, wrist, or finger arthritis; (2) had sensory and/or motor changes distal to the elbow, such as carpal tunnel syndrome or elbow joint instability; (3) experienced pain associated with radiculopathy, cervical nerve compression, or thoracic outlet syndrome; and (4) practiced high-velocity racquet sports, such as tennis or badminton.

Exclusion criteria for conditions preventing safe or appropriate use of the ArmLock sleeve included those who (1) exhibited a difference of less than 15 degrees in the range of motion (ROM)

XSL•FO
RenderX

of wrist flexion with fingers extended compared to fingers flexed as such cases indicate limited capacity for further mechanical stretching using the ArmLock sleeve; (2) experienced pain during ligament stress tests at the elbow, which may indicate ligament instability, precluding safe use of the device; and (3) had a wound or scarring in the area where the ArmLock sleeve would be applied as these conditions could prevent proper application and pose risks to healing.

Settings

All assessments were performed at Lab 1-45 Corbett Hall, University of Alberta. Participants took the Armlock sleeve home.

Sample Size

We aimed to recruit 31 participants with a power of 0.8, an alpha of .05, and a medium effect size (f=0.30) [20]. The sample size calculation was based on the Patient-Rated Tennis Elbow Evaluation (PRTEE). In previous studies, sample sizes between 28 and 33 participants were enough to detect a change in the PRTEE (*P*=.002) [21].

Recruitment

The sampling strategy used was nonprobabilistic. Study details were disseminated through various channels, including social media platforms, online classified advertisement services, regional newspapers, and flyers. Flyers were placed in multiple locations, such as bulletin boards at postsecondary institutions, orthopedic supply stores, and medical supply retail stores. Additionally, the research team reached out to local private clinics and professionals using email lists available on professional association websites in Alberta. Potential participants who expressed interest in the study contacted the research team for more information.

Once participants initiated contact, a member of the research team provided a detailed explanation of the study over the phone and conducted screening questions using a standardized telephone script to confirm eligibility. If an individual met the inclusion criteria and expressed interest in participating, the research assistant emailed the information letter for further review and invited them to an in-person session for upper extremity screening. During this session, researchers explained that the purpose of the screening was to confirm eligibility through a physical evaluation. Participants provided signed consent before proceeding with the assessment. The upper extremity screening included (1) isometric resisted muscle testing, (2) sensory examination (light touch), and (3) ROM assessment of the wrist and fingers.

Any asymmetry between sides or a ROM limitation that prevented the participant from tolerating the position required to wear the ArmLock sleeve resulted in exclusion from the study.

Additionally, the following clinical tests were performed:

1. Elbow Varus Instability Stress Test (to rule out elbow joint instability) [22]
2. Spurling Test (to assess cervical nerve compression) [23]
3. Roos Test (to assess thoracic outlet compression) [24]
4. Upper Limb Reflex Testing (to rule out radiculopathy) [25]

If a potential participant tested positive for any of these assessments, they were deemed ineligible for the study, and the research team informed them that the ArmLock sleeve would not be beneficial for their condition. The research team thanked the participants for their time and provided parking compensation for their participation.

Variables

Dependent Variables (Outcome Variables)

A combination of objective and subjective (self-reported) outcome variables was used to assess the effectiveness of the ArmLock sleeve intervention. The following outcome variables were measured:

Composite Wrist and Finger Extensor Muscles

The composite extensibility of the wrist and finger extensor muscles is the muscle tightness of the wrist and finger extensors (objective outcome variable). For this measure, we followed an adaptation of the Mill Test for LE. Participants were positioned with a closed hand, the wrist in dorsiflexion, and the elbow in extension. The examiner then applied wrist flexion, instructing the participant to resist the movement. The test was considered positive and discontinued when the participant reported pain at the lateral epicondyle [26]. No overpressure was applied, and wrist flexion was measured using a 12-inch clear plastic goniometer (Baseline 360°). The value was recorded in degrees.

Pain-Free Grip Strength

Pain-free grip strength (PFG) is the maximal grip strength at which pain appears (objective outcome variable). This is a common measure of pain in LE. In this study, we followed the protocol published in previous studies of tennis elbow [4,27]. First, to minimize the risk of exacerbating injuries during PFG assessment, participants were first introduced to the concept of PFG. Then, participants were asked to stand with the elbow in a complete extension and the shoulder and radioulnar joints in neutral rotation. They were then instructed to slowly squeeze the dynamometer to their maximum strength using the unaffected arm, allowing them to become familiar with the device. Following this, participants repeated the process with the affected arm, maintaining the same gripping rate as the unaffected side but stopping immediately upon experiencing any discomfort. We asked the participants to perform 3 trials with 1-minute rest intervals, as done in previous tennis elbow studies [4,27]. PFG was measured using the Sammons Jamar Plus+ digital hand dynamometer (200-lb capacity; Performance Health, Sammons Preston). The value was recorded in pounds (lb).

Pressure Pain Threshold

The pressure pain threshold (PPT; objective outcome variable) was defined as the minimum amount of pressure required to elicit pain. Participants were seated facing a table with their forearms resting on its surface. The researcher identified the muscle belly of the wrist extensor group by asking the participant to repeatedly extend their middle finger. Testing began on the unaffected side and was then repeated on the affected side, alternating between sides for a total of 3

XSL•FO
RenderX

measurements. A 20-second interval was provided between each measurement. The PPT was measured using the Commander Echo digital algometer with a 1-cm² rubber probe tip (JTECH Medical), applied to the most tender point over the lateral epicondyle, following the method described in a study by Cho et al [4]. The value was recorded in pounds (lb).

Pain During Resisted Wrist Extension

Pain during resisted wrist extension (2 lb) test (subjective outcome variable) is a common symptom of LE. For this assessment (pain with lifting a 2-lb weight), participants were instructed to hold a 2-lb weight while standing with their arm relaxed. They were then asked to slowly lift the weight by flexing the elbow from 0 to 120 degrees and subsequently return to the starting position, following the method described by Cho et al [4]. Pain experienced during the movement was rated by the participant using a Numerical Rating Scale (NRS) from 0 to 10, with 0 = "no pain" and 10 = "worst possible pain."

Pain and Functional Disability

Pain in the affected arm and functional disability were assessed using the PRTEE (subjective [self-reported] outcome variable). The PRTEE is a 15-item questionnaire that evaluates 2 key components: pain (5 items) and functional disability (10 items). The functional disability section is further divided into 2 parts: specific daily activities (eg, turning a doorknob or key, pulling up pants; 6 items) and usual activities (eg, personal care, household tasks, work, and recreational activities; 4 items). The PRTEE is a widely used, validated assessment tool with excellent psychometric properties [28,29]. The questionnaire provides subscale scores for pain and function, as well as a total score, all ranging from 0 to 100. The total score is the sum of the pain and functional disability subscales (maximum score 100). Higher scores indicate greater symptom severity. In this study, we used the pain subscale, the functional disability subscale, and the total score as outcome variables.

Independent Variable

The independent variable in this study was the use of the ArmLock sleeve. Participants were instructed to wear the device for 30 minutes daily for a 12-week period as an intervention to treat LE. The prescribed duration of ArmLock sleeve use was informed by clinical guidance for stretching protocols in lateral elbow tendinopathy and by parameters reported in prior RCTs. Clinical guidelines commonly describe stretching protocols involving multiple repetitions of 30-45 second holds, typically performed twice daily [30]. Across randomized trials, stretching parameters vary substantially, with hold durations ranging from 6 to 45 seconds, delivered over multiple sets (eg, 1-3), multiple repetitions (eg, 6-15), and performed 2 to 3 times per day [15].

Acknowledging this variability, the decision to prescribe 30 minutes of daily ArmLock sleeve use represented an informed estimation of these parameters, corresponding to an approximate cumulative stretch exposure of ~10-33 minutes when considering total time under stretch across repetitions and sessions reported in the literature. For this study evaluating the effectiveness of a novel technology, we intentionally selected a duration toward the upper end of reported parameter ranges to ensure sufficient stretch exposure while also examining participant tolerance to a sustained, cumulative stretching dose that approximates or exceeds conventional protocols. This approach allowed us to balance theoretical therapeutic dosing with feasibility, safety, and tolerability considerations in our evaluation.

Baseline Participant Information

Participants completed a demographic questionnaire that collected information on age, gender, occupation, side affected, current treatments (eg, hand therapy), and symptom duration.

Procedures

Participants who met the inclusion criteria and provided written consent were invited to attend an initial meeting. During this first meeting, participants completed the demographic questionnaire. A research assistant administered the measures of each outcome variable. We started with the ones that were not painful at all and finished with those that were potentially the most painful for participants. The order of the tests was as follows: PRTEE, composite extensibility, PFG, the pain with lifting a 2-lb weight, and PPT. After these assessments, the research assistant provided each participant with an ArmLock sleeve and instructed them on how to properly wear it. Participants wore the device for 15 minutes to check for any signs of skin irritation (eg, red marks or pressure points) or discomfort. The device was adjusted as needed. Participants were instructed to wear the ArmLock sleeve for 30 minutes daily for a 12-week period, at a time of their choosing. Device use was documented using a Daily Log Sheet provided by the research team. For each session, participants recorded the time the sleeve was donned and doffed, the total duration of wear, and any relevant comments (eg, pain, discomfort, or events that prevented adherence to the recommended 30-minute wear time). Participants were asked to return the completed Daily Log Sheet at their subsequent study visit. The research team used the logged information to monitor adherence and document participant-reported experiences with device use.

Participants attended 2 additional sessions: one at week 6 and a final session at week 12. During these visits, the researcher readministered all outcome measures. At the week 12 session, a brief exit interview was conducted to gather participants' feedback on the use and acceptance of the ArmLock sleeve. Additionally, follow-up phone calls were made during weeks 3 and 9 to monitor participants' use of the device and to remind participants to track use in the Daily Log Sheet.

Statistical Analysis

We described the demographics by reporting the frequency and proportion for categorical variables and the mean (SD) for continuous variables. To determine the normality of the data, histogram analysis and the Shapiro-Wilk and Kolmogorov-Smirnov tests were conducted. Repeated-measures one-way ANOVA and Friedman tests were conducted to evaluate differences among the measurement time points (week 0, week 6, and week 12) for normally distributed and non-normally distributed data, respectively. Post hoc comparisons between the measurement time points were evaluated using paired *t* tests and Wilcoxon signed-rank for normally distributed and non-normally distributed data,

XSL•FO
RenderX

respectively. Bonferroni post hoc correction was used to account for the increased risk of type I errors associated with multiple comparisons. A *P* value <0.05 was considered to indicate statistical significance for ANOVA and Friedman tests. The Bonferroni correction level of statistical significance was corrected at a *P* value <.02 (ie, 0.05/3 tests per outcome variable). All the statistical analyses and visual presentations were performed using IBM SPSS (version 29.0.2.0).

Ethical Considerations

The University of Alberta Research Ethics Board reviewed and approved this study (Pro00095330). Potential participants contacted the research team to express interest, and written consent was obtained before participation. All personal data were deidentified prior to data analysis.

Participants received parking compensation for the time spent in each in-person session, including the Upper Extremity Screening for eligibility determination and the assessment sessions at weeks 0, 6, and 12. Additionally, participants received the ArmLock sleeve at no cost and were allowed to keep it after the study was completed.

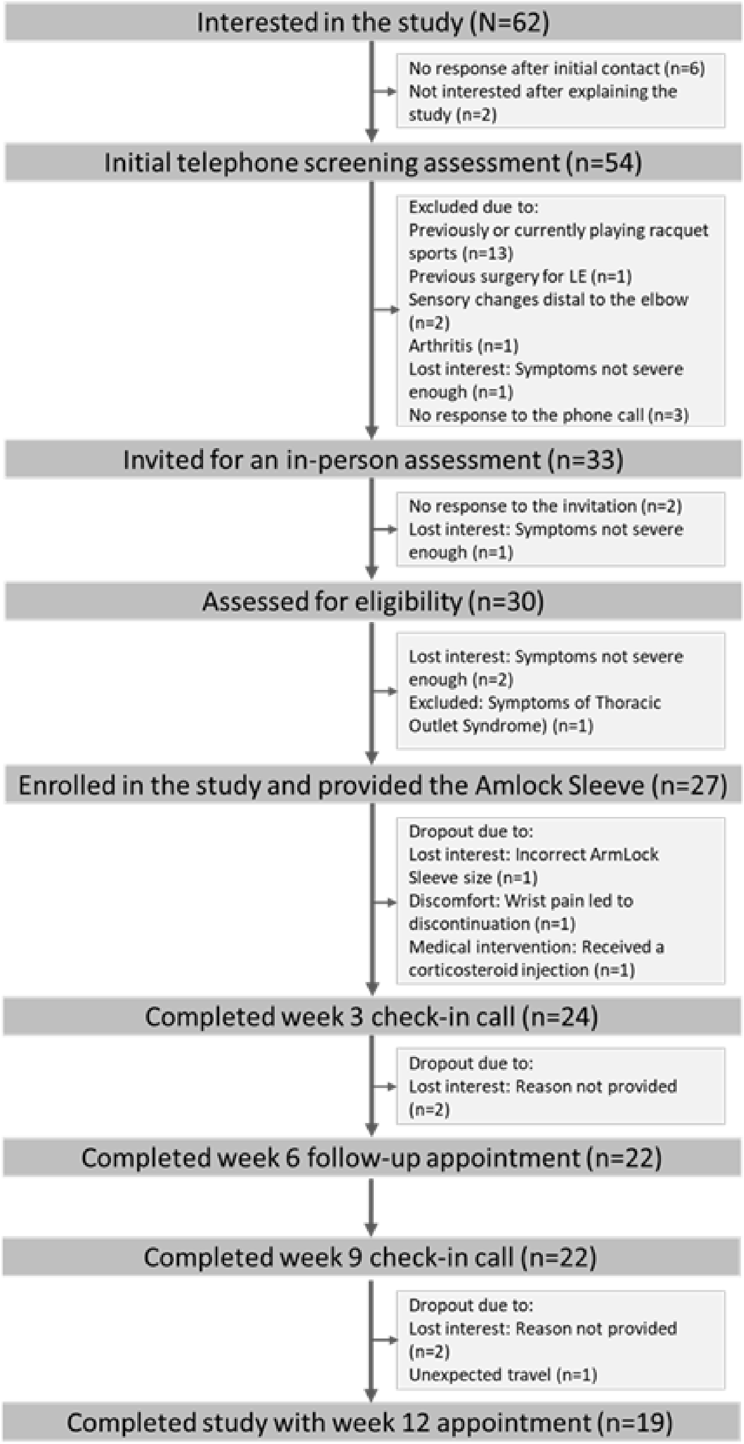
Results

Participant Flow

Figure 2 illustrates the progression of participants throughout the study, between March 23, 2020, and April 30, 2023. A total of 62 individuals from the community contacted the researchers to express interest in participating. Of these, 54 individuals took part in an initial telephone screening assessment, with 21 excluded at this stage. Subsequently, 30 participants proceeded to an in-person eligibility assessment, where 2 declined to participate and 1 was excluded. Ultimately, 27 participants enrolled in the study, completed baseline assessments and received the ArmLock sleeve. All enrolled participants were scheduled to have follow-up assessments conducted at week 6 and week 12. Throughout the study, 8 participants dropped out for various reasons, as detailed in Figure 2. In total, 19 participants completed the study, and their data were included in the final analysis. Based on this participant’s flow, the study retention was 70% (19/27). Compliance with the intervention was measured based on the number of days the device was used and the duration of daily use. On average, participants used the device for 62 days, representing 74% of the recommended 84 days. The average daily usage time was 30 (SD 2.62) minutes.

XSL•FO
RenderX

Figure 2. Study flowchart. LE: lateral epicondylitis.



Demographics of Participants

Table 1 presents the demographic characteristics of participants. The majority of participants were middle-aged adults (mean age 44.11, SD 10.57 years), male, and employed. The most commonly reported treatment strategies used in the past to manage tennis elbow included performing ROM exercises and using a brace or elbow strap.

At the time of data collection, nearly half of the participants reported engaging in ROM exercises, approximately one-quarter were taking nonsteroidal anti-inflammatory drugs such as ibuprofen, aspirin, or naproxen, and slightly more than one-quarter reported no nonsurgical concomitant treatments. Additionally, slightly more than half of the participants reported that LE symptoms lasted for 1 hour or more.

XSL•FO
RenderX

Table 1. Demographics of participants (N=19).

Variable	Participants, n (%)
Sex	
Male	13 (68.4)
Female	6 (31.6)
Employment status	
Employed	9 (47.4)
Self-employed	2 (10.5)
Out of work and looking	3 (15.8)
Out of work but not looking	1 (5.3)
Student	1 (5.3)
Retired	2 (10.5)
Unable to work	1 (5.2)
Previous treatment	
Range of motion exercises	13 (68.4)
Brace or elbow strap	13 (68.4)
Nonsteroidal anti-inflammatory drugs	10 (52.6)
Ice	10 (52.6)
Physiotherapy or massage therapy	9 (47.4)
Steroid injections	1 (5.3)
Present treatment	
Range of motion exercises	9 (47.4)
Nonsteroidal anti-inflammatory drugs	5 (26.3)
Ice	3 (15.8)
Physiotherapy or massage therapy	3 (15.8)
Brace or elbow strap	3 (15.8)
Steroid injections	0 (0)
Duration of symptoms	
One hour or more	11 (57.9)
Less than 1 hour	8 (42.1)

Effectiveness of the ArmLock Sleeve on Reducing LE Symptoms

Table 2 presents the descriptive statistics and results of the Friedman tests for the outcome variables across measurement time points. The analysis revealed statistically significant changes in 85.7% (6/7) of the outcome variables, suggesting that the ArmLock sleeve was effective in improving several key measures. Specifically, the intervention resulted in increased composite extensibility of the wrist and finger extensor muscles, enhanced PFG, and reduced pain intensity when lifting a 2-lb weight. Additionally, participants reported decreases in self-reported pain and functional disability, as measured by the PRTEE, with significant improvements observed in the pain subscale, the functional disability subscale, and the total score.

XSL•FO
RenderX

Table 2. Descriptive statistics and Friedman test for the outcome variables across measurement time points (week 0, 6, and 12; N=19)[a].

Outcome variable	Descriptive statistics						Friedman test results		
	Week 0		Week 6		Week 12		Chi-square (*df*)	*P* value	W[b]
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)			
Ext-AA[c] (degrees)	45.0 (8.6)	45.0 (40.0-54.0)	48.6 (5.1)	49.0 (46.0-53.0)	48.0 (8.8)	51.0 (43.0-55.0)	6.08 (2)	*.048*	0.160
PFG-AA[d] (lb)	48.9 (29.9)	43.1 (20.9-76.4)	64.2 (25.5)	59.0 (42.7-88.3)	73.2 (22.9)	76.0 (65.6-88.9)	27.26 (2)	*<.001*	0.717
PPT-AA[e] (lb)	4.1 (1.1)	3.9 (3.2-5.1)	4.3 (1.1)	4.9 (3.4-4.6)	4.7 (1.7)	4.2 (3.5-6.0)	4.89 (2)	.09	0.129
PWL 2lb-AA[f]	3.16 (2.2)	3.00 (1.0-4.0)	1.7 (2.5)	1.0 (0.0-3.0)	0.84 (1.8)	0.0 (0.0.-1.0)	24.00 (2)	*<.001*	0.632
PAINS_SUB[g]	22.0 (8.2)	20.00 (16.0-30.0)	12.5 (9.0)	11.00 (5.0-17.0)	7.9 (9.5)	5.0 (2.0-8.0)	27.26 (2)	*<.001*	0.717
FUNC_SUB[h]	19.6 (10.8)	24.0 (8.5-30.5)	9.7 (9.9)	5.0 (3.0-15.0)	6.0 (9.5)	3.0 (1.0-5.5)	29.95 (2)	*<.001*	0.788
TOTAL_W[i]	41.6 (17.7)	46.0 (27.0-59.0)	22.2 (18.6)	9.0 (9.0-32.5)	14.0 (18.7)	8.0 (4.0-13.0)	28.74 (2)	*<.001*	0.756

[a] Italicized *P* values indicate statistical significance at *P*<.05.

[b] W: Kendall’s coefficient of concordance (considered a measure of effect size in the context of the Friedman Test). [c] Ext-AA: composite extensibility of wrist and finger extensor muscles–affected arm.

[d] PFG-AA: pain-free grip strength–affected arm.

[e] PPT-AA: pressure pain threshold–affected arm.

[f] PWL 2lb-AA: pain with lifting a 2-lb weight–affected arm.

[g] PAINS_SUB: Patient-Rated Tennis Elbow Evaluation pain subscale.

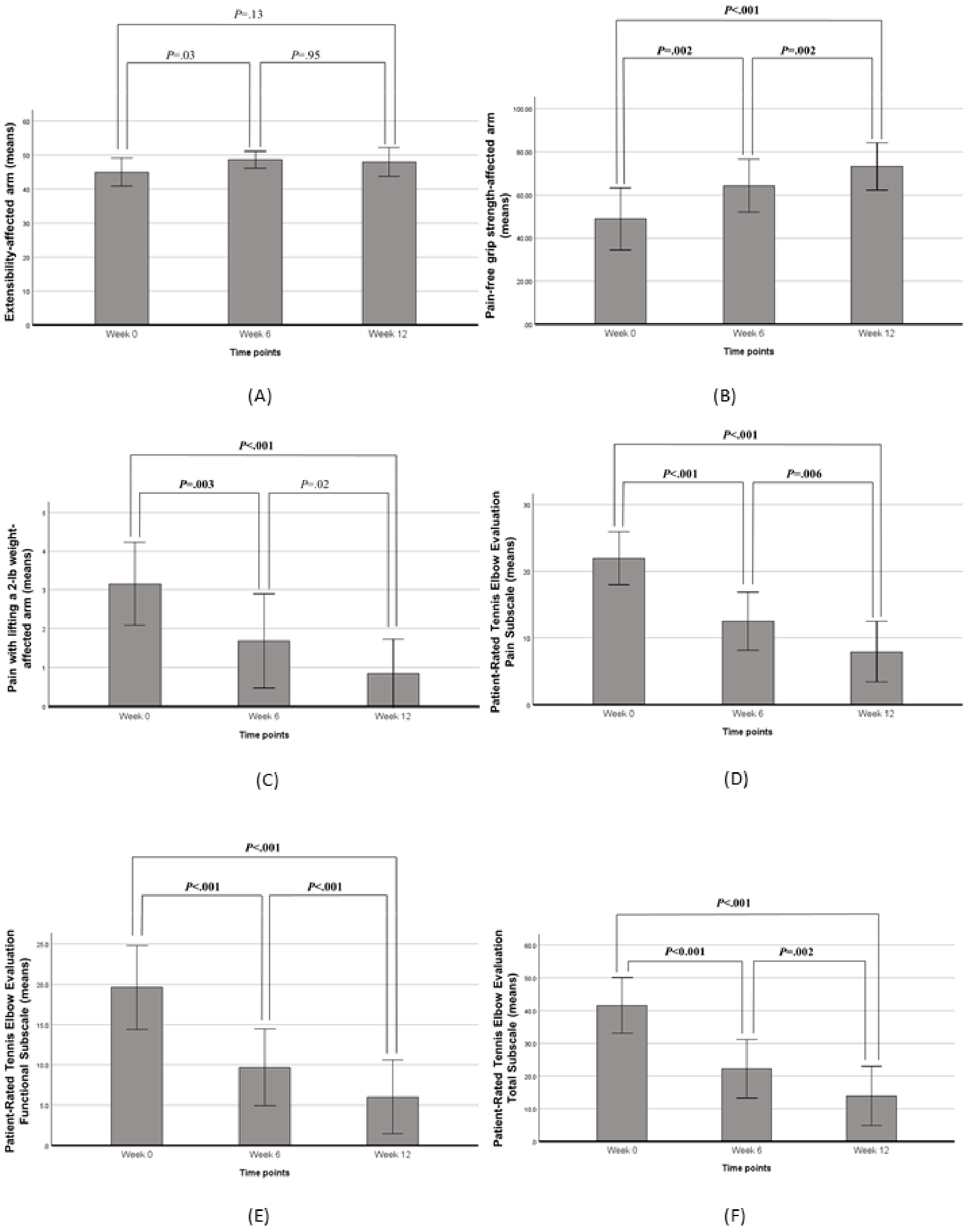
[h] FUNC_SUB: Patient-Rated Tennis Elbow Evaluation functional disability subscale.

[i] TOTAL_W: Patient-Rated Tennis Elbow Evaluation total scale.

Figures 3A-3F illustrates the mean values of each outcome variable for the affected arm across the 3 time points (week 0, week 6, and week 12), along with post hoc comparisons between time points using Bonferroni correction. The pain pressure threshold was excluded from the post hoc comparisons, as the Friedman test did not indicate statistically significant changes for this variable. Error bars represent the SEM at each time point.

XSL•FO
RenderX

Figure 3. Post hoc comparisons between time points. The *P* values in the figure are obtained from the Wilcoxon signed-rank tests.



In Figure 3, Bonferroni correction: α_{adjusted} = .017. The α_{adjusted} was calculated as follows: α_{adjusted} =/m, where m is the number of tests per outcome variable (ie, m=3) and =.05. All the *P* values in bold are statistically significant using the new α_{adjusted}.

For composite extensibility of the wrist and finger extensor muscles, the results suggest an initial improvement between week 0 and week 6, followed by stabilization from week 6 to week 12 (Figure 3A). However, despite the observed improvement at week 6, the changes from week 0 to week 6 and week 0 to week 12 were not statistically significant.

PFG showed a progressive improvement over time (Figure 3B), with statistically significant changes at each stage. These findings reinforce the effectiveness of the ArmLock sleeve in enabling participants to achieve a stronger, pain-free grip. The decreasing *P* values indicate that the cumulative improvement from week 0 to week 12 was greater than the improvement observed over shorter intervals.

XSL•FO
RenderX

Pain during resisted wrist extension, measured by the pain with lifting a 2-lb weight test, significantly decreased at each consecutive time point, suggesting a consistent reduction in pain throughout the 12-week period (Figure 3C). The lowest *P* value (*P*<.001) for the comparison between week 0 and week 12 confirms that the overall reduction in pain across the study period was statistically significant.

The subjective assessment of tennis elbow aligned with the objective measures. As shown in Figure 3E and F, both the pain subscale and functional disability subscale of the PRTEE decreased over time, indicating that the ArmLock sleeve had a positive impact on pain reduction and functional limitations in daily activities. Consequently, the total PRTEE score also declined over time, with all observed changes being statistically significant.

Because participants were not restricted from engaging in nonsurgical concomitant care, potential confounding effects were further examined analytically by including concomitant treatment status (yes or no) as a between-subjects factor in a repeated-measures ANOVA. This analysis evaluated whether changes in pain and functional outcomes over the intervention period differed between participants who did (n=14) and did not (n=5) receive concomitant treatments. The results indicated that pain pressure threshold was the only outcome significantly influenced by concomitant treatment status (β=17.073; *P*=0.03; partial η^2=0.250). Notably, pain pressure threshold was also the only outcome that did not demonstrate a significant change over time in response to the ArmLock sleeve intervention. Taken together, these findings suggest that concomitant treatments did not meaningfully confound the observed intervention-related improvements in the remaining outcomes.

Discussion

Principal Findings

This study aimed to evaluate the effects of sustained tension provided by the ArmLock sleeve on biomechanical and functional outcomes in adults diagnosed with LE. The results of this within-subjects repeated-measures design study provide evidence of the potential clinical benefits of the ArmLock sleeve in reducing pain and improving function. Despite the small sample size, both subjective and objective measures of pain and functional disability outcomes demonstrated statistically significant reductions, with large effect sizes (W≥0.5). Improvements were observed in composite extensibility of the wrist and finger extensor muscles, PFG, pain with lifting a 2-lb weight, and the pain and functional disability subscales of the PRTEE. These findings suggest that the sustained muscle stretching provided by the ArmLock sleeve contributed to a substantial reduction in symptoms and enhanced functional performance in individuals with LE.

PFG refers to the maximum force an individual can exert during a gripping action without experiencing pain. It is a widely recommended measure for assessing function and treatment progress in individuals with LE [31]. PFG has also been found to be highly correlated with self-reported pain and has demonstrated high accuracy in detecting clinically important changes over time in individuals with LE [32]. Therefore, the significant increase in PFG observed in our study for the 12-week period is supported by a valid and reliable measure capable of detecting meaningful changes in patients with LE.

Hill et al [33] identified a minimal detectable change for PFG in individuals with LE of 9.2 kg force, which is approximately 20.28 lb. In our study, the observed increase in PFG for 12 weeks was 24.3 lb, exceeding the minimal detectable change threshold of 20.28 lb. This suggests that the measured improvement is clinically meaningful and unlikely to be attributed to measurement error.

The results from objective measures align with subjective assessments, reinforcing the intervention’s effectiveness. Findings from the PRTEE indicate that participants experienced notable improvements, including reduced pain and discomfort and enhanced arm functionality in daily activities. During the exit interview, participants reported resuming various activities due to the ArmLock sleeve intervention, such as playing hockey and golf, opening jars, typing, picking up a cup of coffee, putting on pants, opening doors, snow shoveling, paddling a canoe, climbing, and carrying heavy bags or groceries. However, 2 (11%) participants reported no perceived improvements with device use. Variations in the device’s effectiveness may be influenced by lifestyle factors such as smoking and alcohol consumption, which have been associated with nonsurgical treatment failure [34].

The majority of participants found the ArmLock sleeve easy to put on and remove, indicating good usability. However, some participants faced challenges in adjusting the optimal tightness and correctly using the straps. Among those who completed the study, 3 (16%) participants reported numbness, tingling, or soreness while wearing the device. One participant noted that these symptoms resolved after loosening the wrist strap. Additionally, 1 participant withdrew from the study due to wrist pain. These findings highlight the need for future research to explore potential adverse effects associated with wearing the device and to optimize its design and usage guidelines.

The high retention rate and compliance with the intervention indicate that participants accepted the ArmLock device and tolerated the recommended daily wearing time of 30 minutes. Participants provided feedback on how they adjusted their schedules to accommodate device use, and many appreciated the ability to wear it privately at home. Several participants expressed satisfaction with the outcomes of using the device daily. However, not all were able to do so consistently due to various challenges, including time constraints, forgetfulness, evening fatigue, or a diminished perception of the stretching effect after several weeks of use. These insights highlight the importance of considering individual adherence factors in future implementations of the device.

Finally, the ArmLock sleeve controlled and gentle sustained stretch may promote collagen fiber alignment and increase blood flow, thereby facilitating tendon healing. This mechanism is consistent with the observed reductions in pain and improvements in function demonstrated in this study.

XSL•FO
RenderX

Study Limitations

This study has some limitations. First, the small sample size was influenced by recruitment challenges, particularly in a community setting where potential participants were not referred by health care professionals. Additionally, 15 individuals who initially expressed interest did not respond to follow-up calls or invitations. Future studies should consider recruiting from hospitals and clinics to improve participant retention and sample size. The COVID-19 pandemic further impacted recruitment, delaying the study for 1 year due to restrictions on in-person screening. Second, the lack of a control group limits the ability to attribute improvements solely to the ArmLock sleeve intervention. A placebo, sham device, or standard treatment group would help distinguish the intervention's effects from natural recovery or placebo effects. Third, self-reported compliance may be affected by recall bias or reporting inaccuracies, as participants tracked device usage in a diary. Future studies could incorporate wear-time sensors for more reliable adherence data. Fourth, lifestyle factors such as smoking and alcohol consumption were not measured in our study, highlighting a limitation. Future research on this device should consider assessing these variables to better understand their potential impact on outcomes. Fifth, although LE typically presents with an equal sex distribution [2], the number of male participants in our sample was twice that of female participants. Future research should consider refining recruitment strategies and materials to better engage and include female participants. Finally, the lack of long-term follow-up means it is unclear whether the observed benefits were sustained beyond 12 weeks. Building on the effect estimates generated by this study, future research with larger sample sizes should incorporate extended follow-up assessments at 3 to 6 months postintervention to evaluate the durability of intervention effects.

Conclusion

This within-subjects repeated-measures study provides preliminary evidence supporting the ArmLock sleeve as a promising nonsurgical treatment for LE. Wearing the device daily for 30 minutes across 12 weeks led to statistically significant improvements in more than 80% of the outcome variables. Specifically, the intervention resulted in increased composite extensibility of the wrist and finger extensor muscles, enhanced PFG, and reduced pain intensity when lifting a 2-lb weight. Additionally, participants reported lower self-reported pain and functional disability. These findings suggest that the sustained muscle stretching provided by the ArmLock sleeve contributes to symptom reduction and improved functional performance in individuals with LE. However, future studies with a control group and a larger sample size are needed to provide more conclusive evidence regarding the device's effectiveness.

Acknowledgments

We would like to extend our special thanks to Mykaela Belter, Takeisha Wang, and Carly De Moissac for their valuable assistance with data collection.

Funding

This work was supported by the Mathematics of Information Technology and Complex Systems (Mitacs) Accelerate Program (grant IT15917).

Conflicts of Interest

None declared.

References

1. Sims SEG, Miller K, Elfar JC, Hammert WC. Non-surgical treatment of lateral epicondylitis: a systematic review of randomized controlled trials. Hand (N Y) 2014;9(4):419-446 [FREE Full text] [doi: 10.1007/s11552-014-9642-x] [Medline: 25414603]
2. Ahmad Z, Siddiqui N, Malik SS, Abdus-Samee M, Tytherleigh-Strong G, Rushton N. Lateral epicondylitis: a review of pathology and management. Bone Joint J 2013;95-B(9):1158-1164. [doi: 10.1302/0301-620X.95B9.29285] [Medline: 23997125]
3. Fan ZJ, Silverstein BA, Bao S, Bonauto DK, Howard NL, Spielholz PO, et al. Quantitative exposure-response relations between physical workload and prevalence of lateral epicondylitis in a working population. Am J Ind Med 2009;52(6):479-490. [doi: 10.1002/ajim.20700] [Medline: 19347903]
4. Cho Y, Hsu W, Lin L, Lin Y. Kinesio taping reduces elbow pain during resisted wrist extension in patients with chronic lateral epicondylitis: a randomized, double-blinded, cross-over study. BMC Musculoskelet Disord 2018;19(1):193 [FREE Full text] [doi: 10.1186/s12891-018-2118-3] [Medline: 29921250]
5. Boden AL, Scott MT, Dalwadi PP, Mautner K, Mason RA, Gottschalk MB. Platelet-rich plasma versus tenex in the treatment of medial and lateral epicondylitis. J Shoulder Elbow Surg 2019;28(1):112-119. [doi: 10.1016/j.jse.2018.08.032] [Medline: 30551782]

XSL•FO
RenderX

6. Silverstein B, Viikari-Juntura E, Kalat J. Use of a prevention index to identify industries at high risk for work-related musculoskeletal disorders of the neck, back, and upper extremity in Washington state, 1990-1998. Am J Ind Med 2002;41(3):149-169. [doi: 10.1002/ajim.10054] [Medline: 11920960]
7. Seiça EC, Buruian A, Gameiro D, Peixoto D, Pereira C. Lateral epicondylitis: current concepts in pathology, investigation, and management. SN Compr Clin Med 2023;5. [doi: 10.1007/s42399-023-01410-6]
8. Kroslak M, Pirapakaran K, Murrell GAC. Counterforce bracing of lateral epicondylitis: a prospective, randomized, double-blinded, placebo-controlled clinical trial. J Shoulder Elbow Surg 2019;28(2):288-295. [doi: 10.1016/j.jse.2018.10.002] [Medline: 30658774]
9. Kim JH, Hoy JF, Smith SR, Sabet A, Fernandez JJ, Cohen MS, et al. Recalcitrant lateral epicondylitis: a systematic review on current nonoperative and operative treatment modalities. JBJS Rev 2024;12(8). [doi: 10.2106/JBJS.RVW.24.00059] [Medline: 39106325]
10. Nirschl RP, Ashman ES. Elbow tendinopathy: tennis elbow. Clin Sports Med 2003;22(4):813-836. [doi: 10.1016/s0278-5919(03)00051-6] [Medline: 14560549]
11. Kim CH, Park YB, Lee JS, Jung HS. Platelet-rich plasma injection vs. operative treatment for lateral elbow tendinosis: a systematic review and meta-analysis. J Shoulder Elbow Surg 2022;31(2):428-436. [doi: 10.1016/j.jse.2021.09.008] [Medline: 34656779]
12. Li S, Yang G, Zhang H, Li X, Lu Y. A systematic review on the efficacy of different types of platelet-rich plasma in the management of lateral epicondylitis. J Shoulder Elbow Surg 2022;31(7):1533-1544. [doi: 10.1016/j.jse.2022.02.017] [Medline: 35337955]
13. Riff AJ, Saltzman BM, Cvetanovich G, Frank JM, Hemu MR, Wysocki RW. Open vs percutaneous vs arthroscopic surgical treatment of lateral epicondylitis: an updated systematic review. Am J Orthop (Belle Mead NJ) 2018;47(6). [doi: 10.12788/ajo.2018.0043] [Medline: 29979804]
14. Day JM, Lucado AM, Uhl TL. A Comprehensive rehabilitation program for treating lateral elbow tendinopathy. Intl J Sports Phys Ther 2019;14(5):818-829. [doi: 10.26603/ijspt20190818]
15. Taylor L, Wiebusch M, Bisset LM, Coombes BK. Adherence to exercise in lateral elbow tendinopathy, a scoping review. Musculoskelet Sci Pract 2024;72:102978 [FREE Full text] [doi: 10.1016/j.msksp.2024.102978] [Medline: 38820868]
16. MacDermid JC, Wojkowski S, Kargus C, Marley M, Stevenson E. Hand therapist management of the lateral epicondylosis: a survey of expert opinion and practice patterns. J Hand Ther 2010;23(1):18-30. [doi: 10.1016/j.jht.2009.09.009] [Medline: 19959328]
17. Nagrale AV, Herd CR, Ganvir S, Ramteke G. Cyriax physiotherapy versus phonophoresis with supervised exercise in subjects with lateral epicondylalgia: a randomized clinical trial. J Man Manip Ther 2009;17(3):171-178 [FREE Full text] [doi: 10.1179/jmt.2009.17.3.171] [Medline: 20046624]
18. Shirato R, Aoki M, Iba K, Wada T, Hidaka E, Fujimiya M, et al. Effect of wrist and finger flexion in relation to strain on the tendon origin of the extensor carpi radialis brevis: a cadaveric study simulating stretching exercises. Clin Biomech (Bristol) 2017;49:1-7. [doi: 10.1016/j.clinbiomech.2017.08.008] [Medline: 28826010]
19. Wang RH, Kenyon LK, McGilton KS, Miller WC, Hovanec N, Boger J, et al. The time is now: a FASTER approach to generate research evidence for technology-based interventions in the field of disability and rehabilitation. Arch Phys Med Rehabil 2021;102(9):1848-1859. [doi: 10.1016/j.apmr.2021.04.009] [Medline: 33992634]
20. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. Hoboken, NJ: Pearson/Prentice Hall; 2008.
21. Nowotny J, El-Zayat B, Goronzy J, Biewener A, Bausenhart F, Greiner S, et al. Prospective randomized controlled trial in the treatment of lateral epicondylitis with a new dynamic wrist orthosis. Eur J Med Res 2018;23(1):43 [FREE Full text] [doi: 10.1186/s40001-018-0342-9] [Medline: 30219102]
22. Karbach LE, Elfar J. Elbow instability: anatomy, biomechanics, diagnostic maneuvers, and testing. J Hand Surg Am 2017;42(2):118-126 [FREE Full text] [doi: 10.1016/j.jhsa.2016.11.025] [Medline: 28160902]
23. Jones SJ, Miller JMM. Spurling test. Treasure Island (FL): StatPearls Publishing; 2025.
24. Henni S, Hersant J, Ammi M, Mortaki F, Picquet J, Feuilloy M, et al. Microvascular response to the roos test has excellent feasibility and good reliability in patients with suspected thoracic outlet syndrome. Front Physiol 2019;10:136 [FREE Full text] [doi: 10.3389/fphys.2019.00136] [Medline: 30846945]
25. Miller TA, Pardo R, Yaworski R. Clinical utility of reflex studies in assessing cervical radiculopathy. Muscle Nerve 1999;22(8):1075-1079. [doi: 10.1002/(sici)1097-4598(199908)22:8<1075::aid-mus11>3.0.co;2-u]
26. Valdes K, LaStayo P. The value of provocative tests for the wrist and elbow: a literature review. J Hand Ther 2013;26(1):32-42; quiz 43. [doi: 10.1016/j.jht.2012.08.005] [Medline: 23062797]
27. Akkurt HE, Kocabaş H, Yılmaz H, Eser C, Şen Z, Erol K, et al. Comparison of an epicondylitis bandage with a wrist orthosis in patients with lateral epicondylitis. Prosthet Orthot Int 2018;42(6):599-605. [doi: 10.1177/0309364618774193] [Medline: 29806569]
28. Rompe JD, Overend TJ, MacDermid JC. Validation of the patient-rated tennis elbow evaluation questionnaire. J Hand Ther 2007;20(1):3-10; quiz 11. [doi: 10.1197/j.jht.2006.10.003] [Medline: 17254903]

XSL•FO
RenderX

29. Farzad M, MacDermid JC, Shafiee E, Beygi AS, Vafaei A, Varahra A, et al. Clinimetric testing of the persian version of the patient-rated tennis elbow evaluation (PRTEE) and the Disabilities of the Arm, Shoulder, and Hand (DASH) questionnaires in patients with lateral elbow tendinopathy. Disabil Rehabil 2022;44(12):2902-2907. [doi: 10.1080/09638288.2020.1844318] [Medline: 33180555]
30. IQWiG (Institute for Quality and Efficiency in Health Care). Tennis Elbow: Learn More – Tennis Elbow: Strengthening and Stretching Exercises. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2025.
31. Chourasia AO, Buhr KA, Rabago DP, Kijowski R, Lee KS, Ryan MP, et al. Relationships between biomechanics, tendon pathology, and function in individuals with lateral epicondylosis. J Orthop Sports Phys Ther 2013;43(6):368-378 [FREE Full text] [doi: 10.2519/jospt.2013.4411] [Medline: 23508267]
32. Stratford PW, Levy DR. Assessing valid change over time in patients with lateral epicondylitis at the elbow. Clin J Sport Med 1994;4(2):88-91. [doi: 10.1097/00042752-199404000-00004]
33. Hill CE, Heales LJ, Stanton R, Kean CO. Pain-free grip strength in individuals with lateral elbow tendinopathy: between- and within-session reliability of one versus three trials. Physiother Theory Pract 2023;39(5):1007-1015. [doi: 10.1080/09593985.2022.2030445] [Medline: 35114892]
34. Zhang C, Jia Z, Li J, Wang X, Yang S. Impact of lifestyle and clinical factors on the prognosis of tennis elbow. Sci Rep 2024;14(1):3063 [FREE Full text] [doi: 10.1038/s41598-024-53669-x] [Medline: 38321129]

Abbreviations

ECRB: extensor carpi radialis brevis
FASTER: Framework for Accelerated and Systematic Technology-based Intervention Development and Evaluation Research
LE: lateral epicondylitis
NRS: Numerical Rating Scale
PFG: pain-free grip strength
PPT: pressure pain threshold
PRP: platelet-rich plasma
PRTEE: Patient-Rated Tennis Elbow Evaluation
RCT: randomized controlled trial
ROM: range of motion

Edited by S Munce; submitted 09.Jul.2025; peer-reviewed by K Songur, J Blasco-Abadía; comments to author 03.Oct.2025; revised version received 29.Dec.2025; accepted 13.Jan.2026; published 04.Mar.2026.

Please cite as:
Ríos Rincón AM, Guptill C, Tran A, Kamran R, Alshammari S, Miguel Cruz A
Evaluation of a Stretching Forearm Sleeve for Lateral Epicondylitis: Repeated Measures Study
JMIR Rehabil Assist Technol 2026;13:e80400
URL: https://rehab.jmir.org/2026/1/e80400
doi:10.2196/80400
PMID:41780923

©Adriana M Ríos Rincón, Christine Guptill, Ann Tran, Rija Kamran, Salamah Alshammari, Antonio Miguel Cruz. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 04.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone-Supported Vestibular Rehabilitation in Individuals With Vestibular Dysfunction: Pilot Randomized Crossover Trial Assessing Functional Clinical Outcomes and Anxiety

Azriel Kaplan[1], MSc, PT; Liran Kalderon[1], PT, PhD; Amit Wolfovitz[2], MD; Yoav Gimmon[2,3*], PT, PhD; Shelly Levy-Tzedek[1,4,5*], PhD

[1]Department of Physical Therapy, Recanati School for Community Health Professions, Ben-Gurion University of the Negev, Beersheva, Israel

[2]Department of Otolaryngology—Head and Neck Surgery, Sheba Tel-HaShomer Medical Centre, Ramat Gan, Israel

[3]Department of Physical Therapy, Faculty of Social Welfare & Health Sciences, University of Haifa, Eshkol Tower, Department of Physical Therapy, Haifa, Israel

[4]Zelman Center for Neuroscience, Ben-Gurion University of the Negev, Beersheva, Israel

[5]Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Freiburg, Germany

[*]these authors contributed equally

Corresponding Author:
Yoav Gimmon, PT, PhD
Department of Otolaryngology—Head and Neck Surgery, Sheba Tel-HaShomer Medical Centre, Ramat Gan, Israel

Abstract

Background: Vestibular disorders impair balance, increase fall risk, and reduce quality of life due to dizziness and vertigo. They are frequently accompanied by heightened anxiety, which may further limit daily functioning and contribute to avoidance behaviors. Although vestibular rehabilitation has been extensively studied and shown to be effective in managing vestibular disorders, adherence to home-based exercises remains low for many dizzy patients. This is often attributed to uncertainty about correct performance, lack of feedback, or difficulty maintaining a structured routine. To help address these barriers, Vestibulon, a smartphone app co-designed with clinicians and patients, was developed to support rehabilitation practice by providing guided exercise, structured scheduling, progress monitoring, and clear instructions intended to promote confidence and engagement.

Objective: We aim to evaluate the potential contribution of a smartphone-based app to vestibular rehabilitation outcomes and to explore the relationship between dizziness-related disability and anxiety during the intervention.

Methods: This randomized 2-period crossover pilot trial included 20 adults with vestibular dysfunction (mean age 52, SD 12 y) who completed 6 weeks of rehabilitation. Participants were randomized to begin with either app-supported or conventional treatment before crossing over to the alternate condition after 3 weeks. This design enabled each participant to experience both modes of rehabilitation. Assessments were conducted at baseline (T0), mid-study (after 3 wk; T1), and end of study (T2). Outcome measures included the Dizziness Handicap Inventory (DHI) to assess dizziness-related disability, the State-Trait Anxiety Inventory to evaluate anxiety, and the Instrumented Timed Up and Go test to quantify functional mobility. The primary outcome was the change in DHI scores across time points.

Results: Significant improvements in DHI were observed between T0 and T2 in both groups (median change: app first=34, IQR 6-35 points, P=.006; conventional first=18, IQR 10-19 points, P=.009). Improvement in Instrumented Timed Up and Go performance was observed only when the app-supported phase occurred first ($Z=-2.45$, P=.01), suggesting a potential early benefit of structured smartphone guidance. State-Trait Anxiety Inventory scores did not change significantly in either sequence. Across all time points, dizziness-related disability and state anxiety demonstrated a consistent and significant moderate correlation (r=0.64, $P<.001$), emphasizing the strong interplay between physical and psychological symptoms in individuals with vestibular disorders.

Conclusions: This pilot study indicates that smartphone-supported vestibular rehabilitation has the potential to enhance functional outcomes for some patients. The consistent association between dizziness and anxiety underscores the relevance of considering psychological factors in vestibular rehabilitation. Given the preliminary nature of this study and the small sample size, these findings should be interpreted cautiously, and further research is needed to determine the app's effectiveness in larger randomized controlled trials.

Trial Registration: ClinicalTrials.gov NCT05959278; https://clinicaltrials.gov/study/NCT05959278

(JMIR Rehabil Assist Technol 2026;13:e84207) doi:10.2196/84207

XSL•FO
RenderX

KEYWORDS

vestibular rehabilitation; dizziness; vestibular application; vestibular anxiety; vestibular technology; rehabilitation technology; mobile phone

Introduction

Central or peripheral vestibular dysfunction may cause dizziness, vertigo, and unsteadiness [1]. The peripheral injury can be classified as either unilateral or bilateral vestibular hypofunction [2]. In 2014, a total of 53 - 95 million people across the United States and Europe endured vestibular hypofunction [3]. The common symptoms are vertigo, dizziness, instability, ataxia, nystagmus, and impaired navigational abilities [1,4,5]. A decrease in balance caused by vestibular hypofunction significantly increases the risk of falling in active older adults [6]. The symptoms of vestibular hypofunction may cause impairment in participation in social gatherings and events, quality of life, and a feeling of anxiety [7,8].

Alongside its physical and functional effects, dizziness has been associated with psychological factors, particularly anxiety [7,9,10], although the mechanisms underlying this association are still being investigated. Recent evidence suggests that cerebellar involvement in vestibular processing may play a role in this relationship [11]. Accordingly, interactions between the vestibular system and the cerebellum could be key to understanding the relationship between balance issues and anxiety disorders [11]. In addition, several studies have shown that the scores of the Dizziness Handicap Inventory (DHI) questionnaire are correlated with various anxiety questionnaires, but most of these studies examined specific vestibular populations, such as people with vestibular migraine, or measured the correlation at only one point in time [12-14].

Over the years, several studies [15-18] have shown the safety and effectiveness of vestibular rehabilitation. In vestibular rehabilitation, exercises are performed to alleviate symptoms with habituation exercises, improve gaze stability with adaptation and substitution exercises, and improve balance [15]. According to the clinical practice guidelines for vestibular rehabilitation published by the American Physical Therapy Association and updated in 2022, it is recommended to practice 3 to 5 times per day for a total of 20 - 40 minutes daily over 4 to 6 weeks [15].

Despite vestibular rehabilitation's effectiveness and safety, patients report difficulty during the process, including time commitment and difficulty getting to many face-to-face meetings with the therapist [19]. Moreover, access to vestibular rehabilitation is limited (eg, in terms of distance from the clinic or clinician availability); in a European survey among vestibular therapists, 48% described accessibility as difficult or very difficult [20].

In recent years, studies have examined several ways to make vestibular practice more accessible, such as through practicing tai chi or using virtual reality in vestibular rehabilitation [21]. These methods have shown efficacy in some studies [22,23]. Despite this, these methods are rarely used in practice, possibly because it can be challenging to combine short periods of exercise several times a day [24], as recommended in clinical practice guidelines. In addition, the high cost of robotic [25] or virtual-reality [22] systems may also account for the absence of these technologies from the rehabilitation landscape [22].

Several phone applications have been developed in recent years to aid in vestibular rehabilitation [26-28]. DSilva et al [26] and Nehrujee et al [27] examined the usability of the application after a 1-time experience on individuals with vestibular disorders as well as on healthy individuals. Hovareshti et al [28] validated the performance of a tablet application (which can also be adapted for mobile phones) designed for monitoring practice while identifying head movements and gaze focus using the camera. In addition, they also presented a case series examining the effect on eye-gaze accuracy of 3 patients and 1 healthy individual who used the application [29]. However, these studies did not examine the effects of the phone applications' impact on patients' clinical outcomes over the course of the vestibular rehabilitation program.

To meet the emerging need and address the existing difficulties in the accessibility of vestibular rehabilitation [19,20,24], we developed the "Vestibulon" phone app, aimed to help patients perform their vestibular exercises; the app provides feedback for the user, gives a clear explanation of each exercise, helps manage practice time, and provides information and feedback to the therapist and patient.

The first goal of this research is to examine the effect of practice with the "Vestibulon" app on functional balance and anxiety indices. These were measured using the following parameters: (1) Instrumented Timed Up and Go Test (iTUG), (2) DHI, and (3) State-Trait Anxiety Inventory (STAI). The second goal is to assess the association between balance function (iTUG) and anxiety (STAI). We hypothesized that practice supported by the app would increase compliance with the practice, thereby improving indicators of balance and anxiety.

Methods

Study Design

This study was conducted using a randomized crossover design. Each participant was randomly assigned to 1 of 2 groups, while gender variables were stratified to ensure an equal distribution. A computer program determined the randomization into groups. Due to the nature of the intervention, outcome assessors were not blinded; no blinding of participants or therapists was possible, and allocation was not concealed. A total of 6 weeks of vestibular rehabilitation were given to each participant. During the first period, which lasted 3 weeks, group 1 used the app for their vestibular rehabilitation, and group 2 did their vestibular rehabilitation exercises without the app. At the crossover point, they switched, such that in the second period, which also lasted 3 weeks, group 1 did their rehabilitation exercises without the app, and group 2 used the app. Measurement sessions were held at Sheba Hospital's otolaryngology clinic between June 2023 and July 2024. This

XSL•FO
RenderX

study was concluded in July 2024, following completion of planned recruitment. All study participants gave their written informed consent. This trial is reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) 2010 statement: extension to randomized crossover trials (Checklist 1).

Participants

A total of 26 participants were randomly assigned to the study. Further, 6 participants dropped out during the study, resulting in 20 participants completing this pilot study. As this is a preliminary investigation, a formal power analysis was not performed. Instead, our sample size aligns with literature suggesting a minimum of 12 participants per arm to adequately gauge feasibility and variance for future calculations [30]. In this crossover design, participants received both interventions, ensuring that the final sample exceeded this minimum requirement per arm. A broad age range (18 - 75 y) was intentionally included to allow an initial feasibility assessment across diverse adult users, in line with the aims of a pilot study.

Inclusion and Exclusion Criteria

The inclusion criteria were (1) participants' age of 18 - 75 years, (2) diagnosis of vestibular dysfunction, and (3) being fluent in Hebrew. The exclusion criteria were (1) a medical condition that prevents the participant from performing home vestibular rehabilitation practice, such as orthopedic, neurological, cardiac, or visual impairment; (2) individuals with dizziness who have already been given a home vestibular rehabilitation exercise program; and (3) diagnosis of central vestibular disorder due to brain structural damage.

Recruitment

Participants were recruited through the Sheba Hospital Dizziness Clinic. Clinicians informed eligible patients about this study during routine visits and provided them with a study information sheet. Those interested provided their contact details, and a research team member subsequently contacted them to explain this study further, confirm eligibility, and schedule the baseline assessment.

Ethical Considerations

All study procedures were approved by the Helsinki Committee of Sheba Medical Center (protocol number 9576 - 22-SMC), and the trial was prospectively registered on ClinicalTrials.gov (NCT05959278; registration date: 07/06/2023). All participants provided written informed consent before enrollment and were informed of their right to withdraw from this study at any time without consequence. No secondary use of data from prior studies was involved, and thus no additional waivers or exemptions were required.

To ensure privacy and confidentiality, all data collected for this study were anonymized using coded identifiers. No personally identifying information was stored within the research database or within the investigational app. Paper-based documents (eg, consent forms) were stored in a locked cabinet inside a locked research laboratory, and all digital data were stored on secure, access-restricted institutional servers. Only authorized study personnel had access to coded datasets. Participants did not receive financial compensation for participation in this study.

Outcome Measures and Data Collection

The investigational smartphone app, Vestibulon, was designed to support the delivery of conventional vestibular rehabilitation remotely, with an emphasis on accessibility, clarity, and adherence. Its development was informed by a participatory design process, incorporating insights from patients and clinicians regarding usability, clarity of instructions, accessibility, and barriers to home-based vestibular rehabilitation [25]. These stakeholder inputs directly shaped core features of the app, including its simple interface, structured reminders, and emphasis on supporting daily adherence to standard vestibular exercises. All participants used the same version of the app during this study.

Vestibulon enables the clinician to predefine the rehabilitation parameters for each user, including the type of exercise (adaptation, habituation, and substitution), frequency, intensity, and recommended duration, based on the patient's diagnosis and clinical characteristics. The exercises included in the app correspond to standard, widely accepted components of vestibular rehabilitation. Before each exercise, the app provides step-by-step written and visual instructions to ensure correct performance.

To promote adherence, the app delivers push-notification reminders for exercise sessions. Each user selects the preferred times of day for receiving these reminders and can modify these settings at any time during the intervention period. When a reminder appears, users can choose to begin the exercise, postpone it ("snooze"), or dismiss it. After completing an exercise session, the user is prompted to provide brief feedback (eg, whether the exercise was completed and whether they wish to continue). These features are intended to help patients structure their daily practice and maintain consistency across sessions. Participants received a brief standardized onboarding session of approximately 10 minutes, during which the app was installed and demonstrated, and users were instructed on selecting reminder times, navigating the exercise interface, and submitting feedback. No advanced technical skills were required to operate the app. The app functioned reliably throughout this trial, and no major technical issues or crashes were reported.

Outcome Measures

Instrumented Timed Up and Go Test

The iTUG examines functional balance [31]. The test measures the time it takes for a participant to stand up from a seated position on a chair, walk for 3 meters at their own pace, turn around, return for 3 meters, turn again, and return to their seated position [31]. We used the iTUG for a quantitative assessment for this test [32]. For the measurement, we used Opal (APDM) sensors and the instructions in the APDM "mobility lab" application.

State-Trait Anxiety Inventory

The STAI is a self-report questionnaire designed to examine a person's general and current anxiety levels [33]. The questionnaire is divided into 2 parts, each consisting of 20

XSL•FO
RenderX

statements, which the participant should rate on a scale of 1 - 4 according to their degree of agreement [34]. This questionnaire has been used among patients who endure vestibular disorders [35]. We used this questionnaire to better distinguish between the participant's general and momentary anxiety to assess the effect of the vestibular rehabilitation on anxiety levels.

Dizziness Handicap Inventory

The DHI is a 25-item questionnaire designed to assess the impact of dizziness on everyday life in people with vestibular disorders [36]. The questionnaire has 3 parts: physical, emotional, and functional [36]. The minimal clinically important difference (MCID) of the DHI is an 18-point difference [36]. Participants with a baseline DHI score of ≤17 were excluded from the MCID analysis because, given the 18-point MCID threshold, these individuals could not mathematically achieve an MCID improvement. Including them would introduce a floor effect and underestimate the proportion achieving clinically meaningful change. In 2010, a Hebrew translation of the questionnaire was published [37], and this is the version we administered to participants. No changes to outcome measures were made after trial commencement.

Procedures

This was a 2-period, 2-sequence randomized crossover trial. A washout period was not implemented to avoid interrupting rehabilitation and potentially compromising clinical progress, although this may allow residual effects to influence the second phase. Each participant was administered home exercises for gaze stabilization and habituation for a total of 20 - 30 minutes each day, across 7 to 10 sessions (2 - 4 min per session). The exercises were administered by a vestibular physical therapist. Upon entry to this study (T0), we took the baseline measurements of the iTUG, DHI, and STAI. Participants were then randomized into 1 of 2 groups: one that practiced with the app first (AF), or one that started with conventional practice first (without an app; conventional first, CF). At the end of the first intervention period (T1), we performed the second set of measurements of the same parameters. This was the crossover point at which participants from each group switched to the alternative treatment condition. That is, during the second intervention period, the AF group received conventional treatment, and the CF group used the app to perform their rehabilitation exercises. Each intervention period lasted approximately 3 weeks, for a total of 6 weeks of rehabilitation per participant.

After the completion of both intervention periods (T2), we performed the endpoint measurement of the iTUG, DHI, and STAI. The 6-week period was chosen based on the clinical practice guidelines for vestibular rehabilitation [15]. Participants were aware of the intervention type. The app-based and conventional rehabilitation provided the same therapeutic content but differed in their mode of delivery.

Data Analysis

We used the Kolmogorov-Smirnov test to examine the normal distribution of quantitative variables on the ratio scale.

For evaluating differences in the STAI, DHI, and iTUG indices between groups, we used the Mann-Whitney *U* test. To assess changes in STAI, DHI, and iTUG across the 3 time points (T0, T1, and T2), we applied the Friedman test. When significant differences were identified, post hoc analyses were conducted using the Wilcoxon signed-rank test, with Bonferroni correction applied to account for multiple comparisons. Finally, to explore potential correlations between the iTUG indices, DHI scores, and the STAI questionnaire, we calculated Spearman correlation coefficient. Statistical analysis was performed with SPSS (version 28; IBM Corp) software, and results were considered significant when *P* values were <.05.

Results

Participant Characteristics

A total of 26 participants were recruited for this study. Further, 20 participants (mean age 52.4, SD 12.6 y; 8 men and 12 women) completed this study (Figure 1, Table 1). There was no statistically significant difference in the baseline data between the 2 groups. Participants who completed this study presented with the following vestibular disorders: bilateral vestibular hypofunction (n=6), unilateral vestibular hypofunction (n=4), Meniere disease (n=3), vestibular migraine (n=2), persistent postural perceptual dizziness (PPPD; n=3), labyrinthitis (n=1), and postconcussive syndrome (n=1). Furthermore, 6 participants dropped out during this study (mean age 46.7, SD 14.0 y; 2 men and 4 women).

XSL•FO
RenderX

Figure 1. Study overview. The flowchart illustrates the participant enrollment, intervention groups (one that started with app assistance [AF] and one that started with conventional rehabilitation [CF]), and assessment time points. The app icon indicates the app-assisted rehabilitation period. AF: app first; CF: conventional first; DHI: Dizziness Handicap Inventory; iTUG: Instrumented Timed Up and Go; STAI: State Trait Anxiety Inventory.

XSL•FO
RenderX

Table . Demographics and baseline questionnaire scores of each group. Baseline data of each group and the baseline data of all groups together. No statistically significant difference was found between the groups in any of the characteristics.

Characteristic	AF[a] (n=11)	CF[b] (n=9)	Total (N=20)	*P* value
Age (years), mean (SD)	48.5 (12.8)	57.3 (11.1)	52.4 (12.6)	.16[c]
Gender (woman/man), n (%)	7 (63.6) / 4 (36.4)	5 (55.6) / 4 (44.4)	12 (60) / 8 (40)	.53[d]
DHI[e] total score, median (IQR)	54.0 (36-64)	44.0 (32-74)	49.0 (33-68)	.88[c]
DHI physical, median (IQR)	14.0 (9-19)	14.0 (12-18)	14.0 (9-18)	.91[c]
DHI functional, median (IQR)	18.0 (8-28)	20.0 (10-30)	19.0 (9-30)	.73[c]
DHI emotional, median (IQR)	20.0 (11-29)	14.0 (10-30)	17.0 (10-30)	.73[c]
STAI[f] total, median (IQR)	83.0 (64-92)	92.0 (78-93)	87.0 (64-94)	.62[c]
STAI state, median (IQR)	42.0 (32-49)	44.0 (33-51)	43.0 (32-50)	.94[c]
STAI trait, median (IQR)	34.0 (30-50)	45.0 (34-55)	36.5 (31-54)	.67[c]
iTUG[g] total time, mean (SD)	11.4 (2.2)	11.1 (1.5)	11.1 (1.8)	.79[c]

[a] AF: app first.

[b] CF: conventional first.

[c] Mann-Whitney test.

[d] Chi-square test (Fisher Exact Test).

[e] DHI: Dizziness Handicap Inventory.

[f] STAI: State Trait Anxiety Inventory.

[g] iTUG: Instrumented Timed Up and Go.

Effects of the Phone App on Disability, Anxiety, and Functional Balance Measures

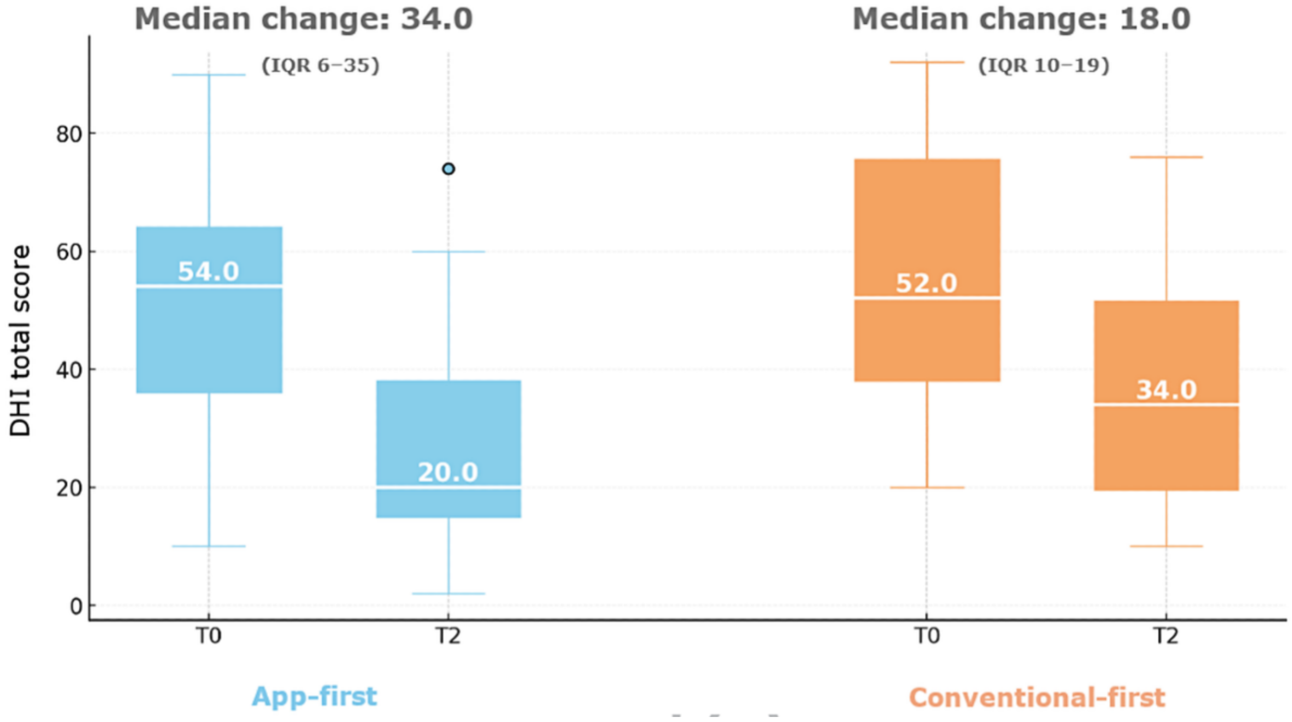
DHI Total

Both groups (AF and CF) show statistically significant differences between T0-T2 in the DHI total (AF: Z=−2.80, *P*=.006; CF: Z=−2.67, *P*=.009).

The total DHI score decreased by 62.9% in the AF group and by 34.6% in the CF group between baseline and end-of-study; however, the difference between the groups was not statistically significant (Z=−0.80, *P*=.42). Figure 2 summarizes the change in DHI total.

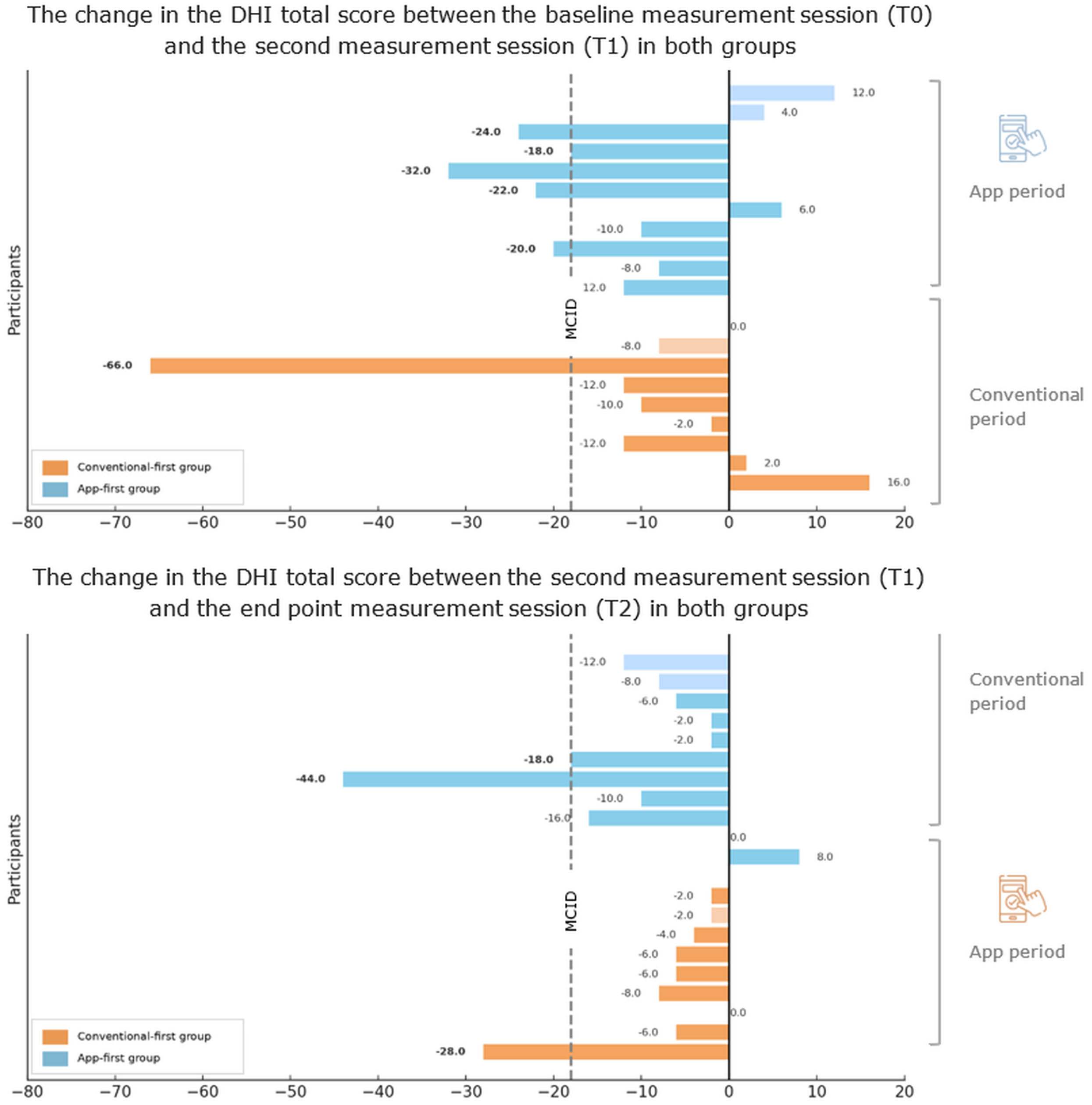
Figure 3 illustrates the changes in the DHI total score during both practice periods, as well as the number of participants who reached or exceeded the MCID.

Figure 2. The change in median DHI total between T0-T2. Within each group (AF and CF), we found a statistically significant difference between T0 and T2. A blue dot indicates an outlier. Blue color represents the app-first group, and orange color represents the conventional-first group. The values in the box plot represent the median values at T0 (first evaluation when beginning the rehabilitation period) and at T2 (last evaluation after 6 weeks of rehabilitation). AF: app first; CF: conventional first; DHI: Dizziness Handicap Inventory.



XSL•FO
RenderX

Figure 3. Change in DHI total score and number of patients achieving the minimal clinically important difference score. Top: the change in the DHI total score between the baseline measurement (first evaluation before the rehabilitation period [T0]) and the second measurement session (second evaluation after 3 weeks of rehabilitation [T1]) in both groups. Bottom: the change in the DHI total score between the second measurement session (T1) and the end point measurement session (last evaluation after 6 weeks of rehabilitation [T2]) in both groups. The results of the app-first group are marked in light blue; those of the conventional-first group are marked in orange; a phone-app icon indicates that the app was used during this period. The bolded numbers indicate that these participants' score changes met or exceeded the MCID threshold. The lighter bars indicate participants with a floor effect, that is, whose initial DHI score was 17 or below. DHI: Dizziness Handicap Inventory; MCID: minimal clinically important difference.



By the end of this study, 7 of 9 (77%) participants in the AF group exceeded the MCID threshold, compared to 2 of 8 (25%) participants in the CF group. This difference between the groups did not quite reach statistical significance (χ^2_1=4.97, *P*=.06).

When comparing the with-app practice period to the no-app practice period, regardless of the group (AF/CF), we found that 6 of 17 (35.2%) participants met or superseded the MCID threshold during the period with the app (T0-T1 for the AF group and T1-T2 for the CF group), while 3 of 17 (17.6%) participants met or superseded the MCID threshold without the app (T1-T2 for the AF group and T0-T1 for the CF group).

No statistically significant change was found between the 3 measurement sessions in the STAI state (χ^2_2=0.80, *P*=.67) or in the STAI trait (χ^2_2=1.71, *P*=.42) in the AF+CF together, and in each group separately.

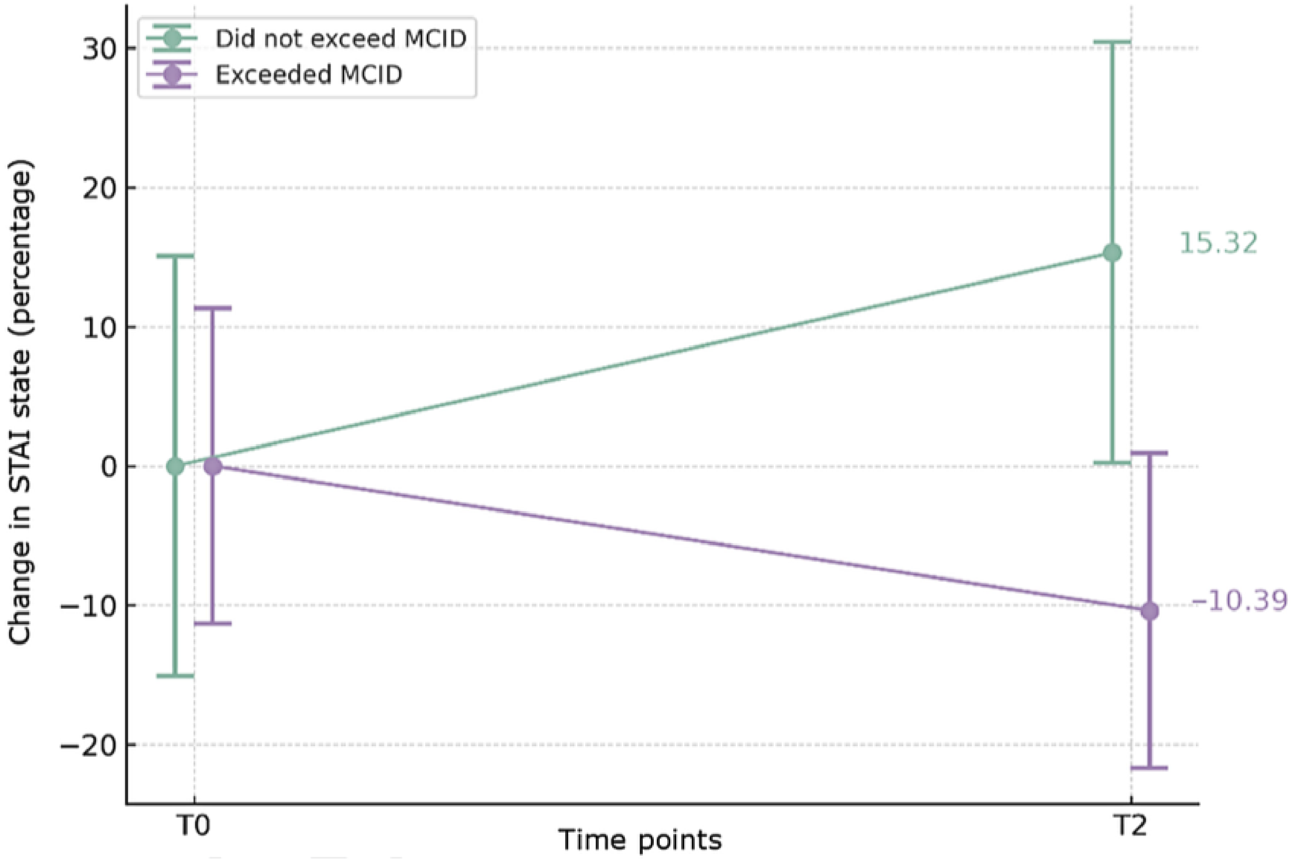
When we divided the data of all the participants into 2 groups: those who met or superseded the MCID threshold in the DHI total (a change of 18 points or more) and those who did not (a change of fewer than 18 points), we found opposing trends: the group that superseded the MCID showed a 10.3% decrease in the STAI state score between T0 and T2, while the group that did not pass the MCID showed an increase of 15.3% (Figure

XSL•FO
RenderX

4). This difference is not statistically significant (Z=−1.73, *P*=.09).

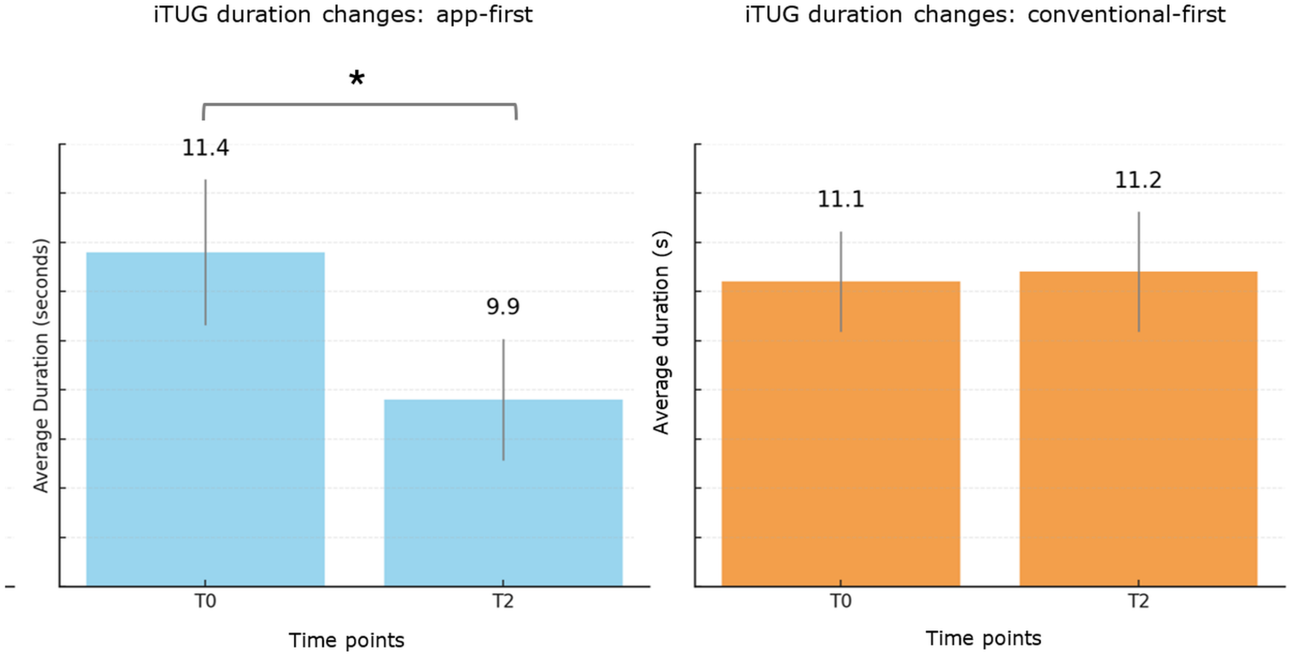
In the AF group, a statistically significant change was found across the 3 iTUG measurements (χ^2_2=9.45, *P*=.009; Figure 5). We found a statistically significant change in the post hoc test only between T0-T2 (Z=−2.45, *P*=.01). In the CF group, no statistically significant change was found across the 3 measurements (χ^2_2=1.75, *P*=.42).

Figure 4. Contrasting trends in STAI State anxiety changes relative to DHI improvement. Participants who achieved or exceeded the MCID in DHI scores between T0 and T2 showed a trend towards lower STAI State scores. Conversely, those who did not reach the MCID in DHI improvement demonstrated a trend towards higher STAI State scores. DHI: Dizziness Handicap Inventory; MCID: minimal clinically important difference; STAI: State Trait Anxiety Inventory.



XSL•FO
RenderX

Figure 5. Change in iTUG total duration between baseline (first evaluation at the beginning of the rehabilitation period [T0]) and the end of this study (last evaluation after 6 weeks of rehabilitation [T2]) * indicates a statistically significant difference between the time points. iTUG: Instrumented Timed Up and Go.



Association Between Anxiety (STAI) to Balance Function (iTUG) and Dizziness Handicap (DHI)

There was no statistically significant correlation between STAI state and iTUG duration at any time point. However, we observed a statistically significant moderate-strong positive correlation between STAI state and DHI total in all 3 measurement sessions (Table 2). Table 2 shows the Spearman correlation between the STAI state and iTUG duration, and between the STAI state and DHI total.

Table . Correlations between anxiety scores and DHI[a] and iTUG[b] for all participants.

	T0[c]	T1[d]	T2[e]
STAI[f] state and iTUG duration			
r	0.12	−0.36	−0.16
P value	.59	.19	.51
STAI state and DHI total			
r	0.59	0.65	0.68
P value	.005	.002	<.001

[a] DHI: Dizziness Handicap Inventory.

[b] iTUG: Instrumented Timed Up and Go.

[c] T0: first evaluation when beginning the rehabilitation period.

[d] T1: second evaluation after 3 weeks of rehabilitation.

[e] T2: last evaluation after 6 weeks of rehabilitation.

[f] STAI: State Trait Anxiety Inventory.

Discussion

Principal Findings

Twenty participants completed this 6-week study, with each subject undergoing 2 phases in random order: a period of conventional vestibular rehabilitation and a period of app-assisted vestibular rehabilitation. We examined how mobile-assisted vestibular rehabilitation affects functional balance measure (iTUG), dizziness disability (DHI), and anxiety (STAI). In addition, we examined the association between the measures (STAI with DHI and iTUG). We found that more individuals (7/9, 77%) in the AF group reached an improvement of 18 points or more on the DHI scale, indicating a clinically meaningful improvement, compared to the conventional-first group (2/8, 25%). Given the heterogeneity of vestibular diagnoses represented in our sample, it is possible that different conditions may respond differently to rehabilitation. For example, individuals with Ménière disease often show fluctuating symptom patterns [38,39], whereas patients with PPPD may require longer or psychologically oriented interventions [40]. As each diagnostic subgroup in our study

XSL•FO
RenderX

included only a few participants, subgroup analyses were not feasible.

The main contributions of this paper are (1) we found that using the “Vestibulon” phone app, as part of the vestibular rehabilitation process, was associated with improvements in functional clinical indices in vestibular rehabilitation (DHI and iTUG), with a trend suggesting a possible advantage of the app-assisted condition over conventional rehabilitation; (2) we found that the correlation between the DHI questionnaire and the STAI questionnaire remained relatively constant throughout the rehabilitation period and was moderate (r=0.64, P<.001), highlighting the importance of addressing both physical and psychological aspects of vestibular disorders.

The Effect of Using the Phone App on Functional Outcomes

The finding that DHI scores improved following vestibular rehabilitation in both groups is echoed in the literature [15,23,41,42]. Some of those studies showed greater improvement in DHI than shown in the current study; this difference may be due to the diverse population participating in our study, which included a relatively high percentage of patients (45%) with bilateral vestibular dysfunction, or PPPD, who sometimes require a more extended period of rehabilitation or a combination of additional treatments [15].

Our study revealed no significant changes in STAI scores in either group or in both groups together. This aligns with a report by Johansson et al [43], who found no effect on STAI score after vestibular rehabilitation based on home exercises combined with cognitive behavioral therapy, in a diverse cohort of vestibular patients. However, Meli et al [44] found an improvement in STAI score that was preserved after 2 months of follow-up; this was after a daily vestibular rehabilitation program supervised by a therapist, coupled with home practice, in diverse chronic vestibular patients. Our results may offer a potential explanation for these conflicting findings. We noted a distinction between participants who achieved the DHI MCID and those who did not. Specifically, those who met the MCID threshold demonstrated improvements in STAI scores, whereas those below the threshold showed deterioration (Figure 4). Although a higher proportion of participants reached the MCID during the app-assisted phase, this difference was not statistically significant (P=.057). Therefore, this pattern should be interpreted cautiously and viewed as preliminary, given the small sample size. Similarly, this observation is further supported by the study by Johansson et al [43], which reported no change in STAI scores and a relatively modest average change in the DHI of less than 8 points (whereas the MCID is 18). Combined, these findings suggest that vestibular rehabilitation’s impact on anxiety may be closely linked to the degree of improvement in dizziness-related symptoms, highlighting the complex interplay between physical symptoms and psychological well-being in vestibular disorders.

The results of our study suggest potential advantages of the app-assisted vestibular rehabilitation over conventional methods, although the small sample size limited our ability to demonstrate statistically significant differences. Notably, the AF group showed promising trends in several key areas. First, for DHI MCID attainment, 3 times as many participants in the AF group (7/9, 77%) met or surpassed the DHI MCID threshold compared to the CF group (2/8, 25%). Although these differences in MCID attainment did not reach statistical significance when directly comparing the 2 conditions (P=.057), the pattern may still indicate a potential benefit of the app-assisted approach and warrants further investigation in a larger sample. Second, for iTUG performance, only the AF group demonstrated a statistically significant improvement in the iTUG test (Figure 5), indicating a trend toward enhanced functional mobility. Third, for DHI improvement, the AF group achieved a median change of −34 (IQR 6 - 35) points on the DHI, compared to −18 (IQR 10 - 19) points in the CF group, potentially suggesting a more substantial reduction in dizziness-related handicap.

Furthermore, twice as many participants met or surpassed the MCID threshold of DHI while using the app (6/17, 35.2%) as participants during conventional treatment (3/17, 17.6%). While this difference did not reach statistical significance, the directionality of the findings suggests a potential advantage of the app-assisted condition. These findings, which suggest a positive trend in favor of rehabilitation with the app, need further investigation. A larger-scale study with an expanded sample size and additional outcome measures is necessary to validate the observed trends and conclusively establish whether indeed the app-assisted vestibular rehabilitation is advantageous over conventional rehabilitation.

The "Vestibulon" App Compared to Other Apps

Several applications have been developed to enhance vestibular rehabilitation, each with unique features and limitations.

For example, D’Silva et al [26] and Meldrum et al [45] developed tablet-based apps integrated with an inertial measurement unit (IMU) sensor placed on the body. D’Silva et al [26] demonstrated high usability and improved accuracy in exercise execution compared to conventional exercises among individuals with vestibular disorders. However, their findings were based on a single session rather than a full vestibular rehabilitation program, and clinical changes were not assessed. In contrast, Meldrum at el [45] evaluated their app over an average of 8 weeks of vestibular rehabilitation and reported high usability, strong adherence, and reductions in dizziness, nausea, and anxiety. Nevertheless, they did not compare their treatment to conventional rehabilitation methods, and symptom improvement was assessed by subjective ratings rather than validated clinical tests or questionnaires typically used in vestibular rehabilitation. Hall et al [46], on the other hand, developed a vestibular rehabilitation training system incorporating a computer application and an external sensor, based on feedback from clinicians and patients. In a pilot study with individuals with vestibular hypofunction, their system demonstrated feasibility and effectiveness over a 4-week rehabilitation period. Unlike the systems mentioned above, the solution by Hall et al [46] also included functional outcomes, which provide more robust evidence of clinical effectiveness. However, similar to the systems in studies by D’Silva et al [26] and Meldrum et al [45], the system requires external hardware, such as a sensor and a computer or tablet, limiting its portability and accessibility.

XSL•FO
RenderX

A main strength of these systems is the ability to provide accurate real-time feedback through the sensors, increasing exercise precision. However, the reliance on multiple components, such as sensors, laptops, or tablets, constitutes a barrier to portability and accessibility.

Hovareshti et al [28] developed a tablet application that detects eye movement without using IMUs, using the tablet's camera itself, and compared its capabilities with IMUs, which showed good accuracy in head angle error and mean interpeak time errors. They presented a case series examining the impact of the practice with the app on eye-gaze accuracy (the ability to maintain one's gaze focused on a target while the head moves) in 3 patients and 1 healthy individual. They found poor eye-gaze accuracy in a participant exposed to directed energy [29]. Their significant advantage is their simplicity of use and portability, which means they do not need external sensors while providing accurate feedback based on movement detection. On the other hand, they have not yet demonstrated the app's efficacy within the long vestibular rehabilitation process or compared it with conventional rehabilitation.

Our smartphone app was developed through a participatory-design process involving focus groups with clinicians and patients [25]. They emphasized the need for ease of use, clear instructions, feedback, cost-effectiveness, and the portability of mobile technology such as phone apps [25]. These insights guided the development of our app. We then tested the app throughout 6 weeks of vestibular rehabilitation with patients, using accepted clinical questions and functional tests.

Association Between Dizziness Disability and Anxiety

Our study revealed a moderately positive correlation between the DHI and the STAI state scores across all 3 measurement points (*r*=0.59, *r*=0.65, and *r*=0.68 for T0, T1, and T2, respectively). These findings align with previous research on the association between DHI and STAI state. For instance, Hong et al [47] found that DHI was significantly correlated with the STAI questionnaire in a multicenter study on 407 individuals with various vestibular disorders, which examined the psychological effects of the vestibular condition. In addition, Kim et al [48] examined 544 individuals with various disorders and examined the effects of anxiety and depression on the intensity of dizziness; they found that compared to individuals with low STAI scores, individuals with high STAI scores also demonstrated higher DHI scores. Both studies examined the association between STAI and DHI at a single time point, regardless of the effect of vestibular rehabilitation. Our study extends these findings by demonstrating the correlation between DHI and STAI state at 3 points throughout the process of vestibular rehabilitation: before treatment initiation, at the midpoint, and at the conclusion of the 6-week treatment. Additionally, we observed that participants who exceeded the MCID threshold on the DHI also showed improvement on the STAI state, in contrast with those who did not reach the MCID threshold and exhibited an increase in their STAI state scores (Figure 4). These findings suggest that one possible contributor to the positive trend observed with the app-assisted condition may relate to psychological elements, such as participants' perception of support or feedback, which have been highlighted as important by patients and clinicians [24]. However, because these factors were not directly measured in the present study, this explanation should be considered a hypothesis for future investigation. The results underscore the relationship between dizziness symptoms, functional limitations, and anxiety. The incorporation of these psychological components appears to address not only the physical aspects of vestibular rehabilitation but also the psychological factors that can impact recovery and adherence to treatment.

Interestingly, we found no correlation between STAI trait and DHI. Schuhbeck et al [49] did find an association between high STAI trait and DHI scores in the acute phase after vestibular stroke and high DHI scores at long-term follow-up. The discrepancy in the findings may be attributed to differences in study populations, as we excluded individuals with vestibular strokes.

We did not observe a statistically significant correlation between the iTUG test and the DHI scores at any measurement point (Table 2). Previous studies have reported conflicting results on the correlation between TUG and DHI, possibly due to variations in study populations and test instructions (eg, walking as quickly as possible vs at a comfortable pace). Further research is needed to clarify the relationship between DHI and iTUG, with careful consideration of population characteristics and standardized instructions.

Limitations

This is a pilot study with a small sample size, which limits the external validity and generalizability of the findings. This study included participants with a diverse range of vestibular disorders and a broad adult age range, which offers the advantage of reflecting the clinical heterogeneity seen in practice. However, this diagnostic and age diversity may also influence the results, as certain vestibular conditions or age groups could have a disproportionate impact on specific outcomes, particularly in a small pilot sample.

Furthermore, crossover study designs present inherent challenges, as it can be difficult to ascertain which phase of the intervention contributed to the observed changes. In addition, the lack of a washout period may have introduced carryover effects between phases, and this should be considered when interpreting the comparative findings. The absence of blinding for participants, therapists, and assessors represents an additional limitation. This may have contributed to performance or detection bias, especially for subjective outcomes (DHI and STAI) and for functional assessments such as the iTUG. Although standardized assessment procedures were used, the influence of expectancy or assessor awareness cannot be fully excluded. To address these limitations, future research should implement a randomized controlled trial with a larger sample size to enhance the robustness and applicability of the results.

Conclusions

Our findings contribute to the understanding of the interplay between dizziness, anxiety, and functional mobility in vestibular rehabilitation. The consistent correlation between DHI and STAI state scores throughout the rehabilitation process underscores

XSL•FO
RenderX

the importance of addressing both physical and psychological aspects in vestibular disorders.

The “Vestibulon” smartphone app, when integrated into vestibular rehabilitation, showed improvements across several outcomes, with patterns suggesting a possible trend favoring the app-assisted phase. Taken together with the significant within-participant improvement in iTUG performance, the observation that more participants exceeded the DHI MCID threshold during the app-assisted phase suggests that integrating the “Vestibulon” app into vestibular rehabilitation may offer clinical benefits. However, these trends did not reach statistical significance and should therefore be interpreted as preliminary. These encouraging preliminary findings support the need for larger, adequately powered trials to determine the extent of the app’s contribution to rehabilitation outcomes.

Acknowledgments

The authors would like to thank Yuliya Berdichevsky, MSc, for her assistance in data analysis and figure preparation. The authors acknowledge the use of generative AI (artificial intelligence) tools for language editing purposes only. Specifically, Grammarly (Superhuman Platform) was used to assist with grammar and style corrections. No generative AI (artificial intelligence) tools were used for data analysis, interpretation, or generation of scientific content.

Funding

The research was partially supported by the Rosetrees Trust and by the Consolidated Anti-Aging Foundation. The funders had no involvement in this study's design, data collection, analysis, interpretation, or paper writing.

Data Availability

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

Authors' Contributions

Conceptualization: YG (lead), SL-T (equal), AW (supporting), LK (supporting)
Data curation: AK
Formal analysis: AK (lead), LK (supporting)
Funding acquisition: SL-T
Investigation: AK (lead), LK (supporting)
Methodology: YG (lead), SL-T (equal), LK (supporting), AW (supporting), AK (supporting)
Project administration: YG (lead), SL-T (equal), AW (supporting)
Resources: SL-T
Supervision: AW
Validation: YG (lead), SL-T (equal), LK (supporting)
Visualization: AK (lead), LK (supporting)
Writing – original draft: AK (lead), LK (supporting)
Writing – review & editing: YG (lead), SL-T (equal), LK (supporting), AW (supporting)

Conflicts of Interest

None declared.

Checklist 1
CONSORT checklist.
[PDF File, 155 KB - rehab_v13i1e84207_app1.pdf]

References

1. Strupp M, Dlugaiczyk J, Ertl-Wagner BB, Rujescu D, Westhofen M, Dieterich M. Vestibular disorders. Dtsch Arztebl Int 2020 Apr 24;117(17):300-310. [doi: 10.3238/arztebl.2020.0300] [Medline: 32530417]
2. Starkov D, Strupp M, Pleshkov M, Kingma H, van de Berg R. Diagnosing vestibular hypofunction: an update. J Neurol 2021 Jan;268(1):377-385. [doi: 10.1007/s00415-020-10139-4] [Medline: 32767115]
3. Grill E, Heuberger M, Strobl R, et al. Prevalence, determinants, and consequences of vestibular hypofunction. Results from the KORA-FF4 survey. Front Neurol 2018;9:1076. [doi: 10.3389/fneur.2018.01076] [Medline: 30581415]
4. Whitney SL, Alghwiri AA, Alghadir A. An overview of vestibular rehabilitation. Handb Clin Neurol 2016;137:187-205. [doi: 10.1016/B978-0-444-63437-5.00013-3] [Medline: 27638071]
5. Zingler VC, Weintz E, Jahn K, et al. Causative factors, epidemiology, and follow-up of bilateral vestibulopathy. Ann N Y Acad Sci 2009 May;1164:505-508. [doi: 10.1111/j.1749-6632.2009.03765.x] [Medline: 19645958]

XSL•FO
RenderX

6. Tuunainen E, Rasku J, Jäntti P, Pyykkö I. Risk factors of falls in community dwelling active elderly. Auris Nasus Larynx 2014 Feb;41(1):10-16. [doi: 10.1016/j.anl.2013.05.002] [Medline: 23763793]
7. Bayat A, Hoseinabadi R, Saki N, Sanayi R. Disability and anxiety in vestibular diseases: a cross-sectional study. Cureus 2020 Nov 30;12(11):e11813. [doi: 10.7759/cureus.11813] [Medline: 33409058]
8. Guinand N, Boselie F, Guyot JP, Kingma H. Quality of life of patients with bilateral vestibulopathy. Ann Otol Rhinol Laryngol 2012 Jul;121(7):471-477. [doi: 10.1177/000348941212100708] [Medline: 22844867]
9. Elyoseph Z, Geisinger D, Zaltzman R, Hartman TG, Gordon CR, Mintz M. The overarching effects of vestibular deficit: imbalance, anxiety, and spatial disorientation. J Neurol Sci 2023 Aug 15;451:120723. [doi: 10.1016/j.jns.2023.120723] [Medline: 37393737]
10. Jacob RG, Furman JM. Psychiatric consequences of vestibular dysfunction. Curr Opin Neurol 2001 Feb;14(1):41-46. [doi: 10.1097/00019052-200102000-00007] [Medline: 11176216]
11. Hilber P, Cendelin J, Le Gall A, Machado ML, Tuma J, Besnard S. Cooperation of the vestibular and cerebellar networks in anxiety disorders and depression. Prog Neuropsychopharmacol Biol Psychiatry 2019 Mar 8;89:310-321. [doi: 10.1016/j.pnpbp.2018.10.004] [Medline: 30292730]
12. Zamyslowska-Szmytke E, Politanski P, Jozefowicz-Korczynska M. Dizziness handicap inventory in clinical evaluation of dizzy patients. Int J Environ Res Public Health 2021 Feb 24;18(5):2210. [doi: 10.3390/ijerph18052210] [Medline: 33668099]
13. Kim EK, Hum M, Sharon JD. Correlating vestibular migraine patient assessment tool and handicap inventory to daily dizziness symptoms. Otol Neurotol 2023 Dec 1;44(10):1052-1056. [doi: 10.1097/MAO.0000000000004014] [Medline: 37733970]
14. Zhu C, Li Y, Ju Y, Zhao X. Dizziness handicap and anxiety depression among patients with benign paroxysmal positional vertigo and vestibular migraine. Medicine (Baltimore) 2020 Dec 24;99(52):e23752. [doi: 10.1097/MD.0000000000023752] [Medline: 33350759]
15. Hall CD, Herdman SJ, Whitney SL, et al. Vestibular rehabilitation for peripheral vestibular hypofunction: an updated clinical practice guideline from the Academy of Neurologic Physical Therapy of the American Physical Therapy Association. J Neurol Phys Ther 2022 Apr 1;46(2):118-177. [doi: 10.1097/NPT.0000000000000382] [Medline: 34864777]
16. Herdman SJ, Hall CD, Maloney B, Knight S, Ebert M, Lowe J. Variables associated with outcome in patients with bilateral vestibular hypofunction: preliminary study. J Vestib Res 2015;25(3-4):185-194. [doi: 10.3233/VES-150556] [Medline: 26756134]
17. Horak FB, Jones-Rycewicz C, Black FO, Shumway-Cook A. Effects of vestibular rehabilitation on dizziness and imbalance. Otolaryngol Head Neck Surg 1992 Feb;106(2):175-180. [doi: 10.1177/019459989210600220] [Medline: 1738550]
18. McDonnell MN, Hillier SL. Vestibular rehabilitation for unilateral peripheral vestibular dysfunction. Cochrane Database Syst Rev 2015 Jan 13;1(1):CD005397. [doi: 10.1002/14651858.CD005397.pub4] [Medline: 25581507]
19. Cohen HS. Assessment of functional outcomes in patients with vestibular disorders after rehabilitation. NeuroRehabilitation 2011;29(2):173-178. [doi: 10.3233/NRE-2011-0692] [Medline: 22027079]
20. Meldrum D, Burrows L, Cakrt O, et al. Vestibular rehabilitation in Europe: a survey of clinical and research practice. J Neurol 2020 Dec;267(Suppl 1):24-35. [doi: 10.1007/s00415-020-10228-4] [Medline: 33048219]
21. Deveze A, Bernard-Demanze L, Xavier F, Lavieille JP, Elziere M. Vestibular compensation and vestibular rehabilitation. Current concepts and new trends. Neurophysiol Clin 2014 Jan;44(1):49-57. [doi: 10.1016/j.neucli.2013.10.138] [Medline: 24502905]
22. Rosiak O, Krajewski K, Woszczak M, Jozefowicz-Korczynska M. Evaluation of the effectiveness of a virtual reality-based exercise program for unilateral peripheral vestibular deficit. J Vestib Res 2018;28(5-6):409-415. [doi: 10.3233/VES-180647] [Medline: 30714985]
23. Xie M, Zhou K, Patro N, et al. Virtual reality for vestibular rehabilitation: a systematic review. Otol Neurotol 2021 Aug 1;42(7):967-977. [doi: 10.1097/MAO.0000000000003155] [Medline: 33782257]
24. Kalderon L, Kaplan A, Wolfovitz A, Levy-Tzedek S, Gimmon Y. Barriers and facilitators of vestibular rehabilitation: patients and physiotherapists’ perspectives. J Neurol Phys Ther 2024 Jul 1;48(3):140-150. [doi: 10.1097/NPT.0000000000000470] [Medline: 38426842]
25. Kalderon L, Kaplan A, Wolfovitz A, Gimmon Y, Levy-Tzedek S. Do we really need this robot? Technology requirements for vestibular rehabilitation: input from patients and clinicians. Int J Hum Comput Stud 2024 Dec;192:103356. [doi: 10.1016/j.ijhcs.2024.103356]
26. DSilva LJ, Skop KM, Pickle NT, et al. Use of stakeholder feedback to develop an app for vestibular rehabilitation-input from clinicians and healthy older adults. Front Neurol 2022;13:836571. [doi: 10.3389/fneur.2022.836571] [Medline: 35280295]
27. Nehrujee A, Vasanthan L, Lepcha A, Balasubramanian S. A smartphone-based gaming system for vestibular rehabilitation: a usability study. J Vestib Res 2019;29(2-3):147-160. [doi: 10.3233/VES-190660] [Medline: 31356221]
28. Hovareshti P, Roeder S, Holt LS, et al. VestAid: a tablet-based technology for objective exercise monitoring in vestibular rehabilitation. Sensors (Basel) 2021 Dec 15;21(24):8388. [doi: 10.3390/s21248388] [Medline: 34960480]
29. Whitney SL, Ou V, Hovareshti P, et al. Utility of VestAid to detect eye-gaze accuracy in a participant exposed to directed energy. Mil Med 2023 Jul 22;188(7-8):e1795-e1801. [doi: 10.1093/milmed/usac294] [Medline: 36208334]

XSL•FO
RenderX

30. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
31. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
32. Kim KJ, Gimmon Y, Millar J, Brewer K, Serrador J, Schubert MC. The instrumented Timed “Up & Go” test distinguishes turning characteristics in vestibular hypofunction. Phys Ther 2021 Jul 1;101(7):pzab103. [doi: 10.1093/ptj/pzab103] [Medline: 33774661]
33. Spielberger CD. State-trait anxiety inventory for adults. APA PsycTests 1983. [doi: 10.1037/t06496-000]
34. Julian LJ. Measures of anxiety. Arthritis Care Res 2011 Nov;63(S11):S467-S472. [doi: 10.1002/acr.20561] [Medline: 22588767]
35. Kalderon L, Chaimoff M, Katz-Leurer M. The distinction between state and trait anxiety levels in patients with BPPV in comparison with healthy controls. Front Psychol 2022;13:1055467. [doi: 10.3389/fpsyg.2022.1055467] [Medline: 36533063]
36. Jacobson GP, Newman CW. The development of the Dizziness Handicap Inventory. Arch Otolaryngol Head Neck Surg 1990 Apr;116(4):424-427. [doi: 10.1001/archotol.1990.01870040046011] [Medline: 2317323]
37. Kaplan DM, Friger M, Racover NK, Peleg A, Kraus M, Puterman M. The Hebrew dizziness handicap inventory. Harefuah 2010 Nov;149(11):697-700. [Medline: 21250408]
38. Wright T. Menière’s disease. BMJ Clin Evid 2015 Nov 5;2015:0505. [Medline: 26545070]
39. Hoskin JL. Ménière’s disease: new guidelines, subtypes, imaging, and more. Curr Opin Neurol 2022 Feb 1;35(1):90-97. [doi: 10.1097/WCO.0000000000001021] [Medline: 34864755]
40. Popkirov S, Staab JP, Stone J. Persistent postural-perceptual dizziness (PPPD): a common, characteristic and treatable cause of chronic dizziness. Pract Neurol 2018 Feb;18(1):5-13. [doi: 10.1136/practneurol-2017-001809] [Medline: 29208729]
41. Millar JL, Gimmon Y, Roberts D, Schubert MC. Improvement after vestibular rehabilitation not explained by improved passive VOR gain. Front Neurol 2020;11:79. [doi: 10.3389/fneur.2020.00079] [Medline: 32153490]
42. Tokle G, Mørkved S, Bråthen G, et al. Efficacy of vestibular rehabilitation following acute vestibular neuritis: a randomized controlled trial. Otol Neurotol 2020 Jan;41(1):78-85. [doi: 10.1097/MAO.0000000000002443] [Medline: 31789800]
43. Johansson M, Akerlund D, Larsen HC, Andersson G. Randomized controlled trial of vestibular rehabilitation combined with cognitive-behavioral therapy for dizziness in older people. Otolaryngol Head Neck Surg 2001 Sep;125(3):151-156. [doi: 10.1067/mhn.2001.118127] [Medline: 11555746]
44. Meli A, Zimatore G, Badaracco C, De Angelis E, Tufarelli D. Effects of vestibular rehabilitation therapy on emotional aspects in chronic vestibular patients. J Psychosom Res 2007 Aug;63(2):185-190. [doi: 10.1016/j.jpsychores.2007.02.007] [Medline: 17662755]
45. Meldrum D, Murray D, Vance R, et al. Toward a digital health intervention for vestibular rehabilitation: usability and subjective outcomes of a novel platform. Front Neurol 2022;13:836796. [doi: 10.3389/fneur.2022.836796] [Medline: 35422750]
46. Hall CD, Flynn S, Clendaniel RA, et al. Remote assessment and management of patients with dizziness: development, validation, and feasibility of a gamified vestibular rehabilitation therapy platform. Front Neurol 2024;15:1367582. [doi: 10.3389/fneur.2024.1367582] [Medline: 38872821]
47. Hong SM, Lee HJ, Lee B, et al. Influence of vestibular disease on psychological distress: a multicenter study. Otolaryngol Head Neck Surg 2013 May;148(5):810-814. [doi: 10.1177/0194599813476476] [Medline: 23380760]
48. Kim SK, Kim YB, Park IS, Hong SJ, Kim H, Hong SM. Clinical analysis of dizzy patients with high levels of depression and anxiety. J Audiol Otol 2016 Dec;20(3):174-178. [doi: 10.7874/jao.2016.20.3.174] [Medline: 27942604]
49. Schuhbeck F, Strobl R, Conrad J, et al. Determinants of functioning and health-related quality of life after vestibular stroke. Front Neurol 2022;13:957283. [doi: 10.3389/fneur.2022.957283] [Medline: 36158947]

Abbreviations

AF: app first
CF: conventional first
CONSORT: Consolidated Standards of Reporting Trials
DHI: Dizziness Handicap Inventory
IMU: inertial measurement unit
iTUG: Instrumented Timed Up and Go
MCID: minimal clinically important difference
PPPD: persistent postural perceptual dizziness
STAI: State-Trait Anxiety Inventory

XSL•FO
RenderX

Edited by S Munce; submitted 16.Sep.2025; peer-reviewed by FN Ardic, ZW Chen; revised version received 04.Jan.2026; accepted 26.Jan.2026; published 24.Mar.2026.

Please cite as:
Kaplan A, Kalderon L, Wolfovitz A, Gimmon Y, Levy-Tzedek S
Smartphone-Supported Vestibular Rehabilitation in Individuals With Vestibular Dysfunction: Pilot Randomized Crossover Trial Assessing Functional Clinical Outcomes and Anxiety
JMIR Rehabil Assist Technol 2026;13:e84207
URL: https://rehab.jmir.org/2026/1/e84207
doi:10.2196/84207

© Azriel Kaplan, Liran Kalderon, Amit Wolfovitz, Yoav Gimmon, Shelly Levy-Tzedek. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Reliability and Discriminant Ability of an Instrumented Timed Up and Go Test in People With Postsurgical Orthopedic Conditions: Quantitative Study

Marica Giardini[1], PT; Ilaria Arcolin[1], PT, MSc; Valerio Antonio Arcobelli[2], PhD; Michela Picardi[3], PT, PhD; Sabato Mellone[2], PhD; Marco Godi[1], PT, MSc

[1]Department of Physical Medicine and Rehabilitation, Institute of Veruno, Istituti Clinici Scientifici Maugeri IRCCS, Gattico-Veruno, Italy

[2]Department of Electrical, Electronic, and Information Engineering–Guglielmo Marconi (DEI), University of Bologna, Viale del Risorgimento 2, Bologna, Italy

[3]Department of Neurorehabilitation Sciences, Casa di Cura Igea, Milano, Italy

Corresponding Author:
Valerio Antonio Arcobelli, PhD
Department of Electrical, Electronic, and Information Engineering–Guglielmo Marconi (DEI), University of Bologna, Viale del Risorgimento 2, Bologna, Italy

Abstract

Background: The Timed Up and Go (TUG) test is widely used to assess mobility and fall risk in older adults and orthopedic patients. Its instrumented variant (iTUG), based on inertial measurement units, enables an objective quantification of motor performance and can even be implemented using smartphone technology. However, its broader clinical adoption remains limited by concerns about reliability, feasibility, and the interpretability of the extracted parameters.

Objective: This study aimed to evaluate the test-retest reliability of variables derived from a single-sensor iTUG in orthopedic inpatients undergoing rehabilitation and to determine whether a subset of reliable sensor-based metrics can support a multidimensional assessment of functional mobility and discriminate among common orthopedic conditions.

Methods: We recruited 104 inpatients at discharge from a rehabilitation ward after total hip arthroplasty, total knee arthroplasty, or femur fracture. Each participant performed the iTUG test on 2 consecutive days using a smartphone-based solution consisting of an inertial measurement unit placed on the lower back. From 100 extracted variables, those with excellent test-retest reliability (intraclass correlation coefficient ≥0.75) were retained. Exploratory factor analysis was used to identify underlying mobility domains, and linear discriminant analysis with 10-fold cross-validation tested their ability to classify diagnostic groups.

Results: Out of 100 iTUG-derived variables, 36 demonstrated excellent test-retest reliability, and 25 were retained for multivariate analysis. Exploratory factor analysis identified 5 factors—walking ability, pace or rhythm, sit-to-walk smoothness, turning ability, and mediolateral angular stability—explaining 80.8% of the total variance. These factors showed good classification accuracy (68%) and achieved an area under the curve of 0.86 and an overall mean accuracy of 0.68 (SD 0.14) for distinguishing among total hip arthroplasty, total knee arthroplasty, and femur fracture. In contrast, total iTUG duration alone yielded an area under the curve of 0.62. All patients used walking aids, and gait variables were more reliable than jerk-based or coordination metrics.

Conclusions: The single-sensor iTUG provides reliable and clinically informative metrics that go beyond traditional stopwatch timing, enabling a multidimensional view of functional mobility in orthopedic patients. The approach is feasible, interpretable, and compatible with real-world mobile health apps, supporting personalized rehabilitation monitoring and future integration into digital decision support systems.

(JMIR Rehabil Assist Technol 2026;13:e82632) doi:10.2196/82632

KEYWORDS

instrumented Timed Up and Go; iTUG test; inertial measurement unit; IMU; reliability; orthopedic patient; wearable sensor

Introduction

The Timed Up and Go (TUG) test is widely recognized for its simplicity and clinical applicability [1]. It is considered a fundamental tool for assessing functional mobility, including elements of balance, walking ability, and transitional movements in older adults [2-4]. The test involves rising from a chair, walking 3 m, performing a 180° turn, returning to the chair, and sitting down, with the total time recorded using a stopwatch [5].

Its widespread clinical use stems from its practicality and robust clinimetric properties. Within neurological practice, the TUG test is useful in identifying fall risk among people with

XSL•FO
RenderX

conditions such as stroke and Parkinson disease [6]. In orthopedics, the test has been recommended for evaluating patients with hip fractures [7] and for distinguishing between future recurrent fallers and nonfallers in frail older adults post hip fracture [8]. Additionally, the TUG has been demonstrated as a valuable test in predicting recovery trajectories: shorter recovery stays in patients with total hip arthroplasty (THA) [9] and early identification of postsurgical rehabilitation needs in patients with total knee arthroplasty (TKA) [10].

The advent of inertial measurement units (IMUs) has expanded the TUG’s potential through detailed and objective kinematic analysis of its subcomponents. The instrumented TUG (iTUG) enables the detection of specific functional impairments [11], improves fall risk sensitivity [12], supports rehabilitation monitoring, and predicts clinical outcomes [13-15].

Indeed, a recent review has examined different methods for analyzing the iTUG, focusing on specific aspects such as key variables, clinical applications, and predictive value [1]. While this review underscores the clinical potential of the iTUG, its psychometric properties remain insufficiently explored. Existing research has primarily addressed the reliability of parameters in healthy older adults or individuals with neurological conditions [11,16,17].

To the best of our knowledge, no studies have analyzed iTUG reliability in patients with orthopedic conditions. This study aimed to evaluate the test-retest reliability of the iTUG in orthopedic inpatients in a rehabilitative setting. Specifically, we targeted 3 common orthopedic conditions: THA, TKA, and femur fractures. Additionally, we sought to identify the most reliable iTUG variables and assess their ability to discriminate between these patient groups using exploratory factor analysis (EFA) and linear discriminant analysis (LDA). We hypothesize that iTUG-derived variables can effectively distinguish between the 3 orthopedic groups and that domains related to dynamic mobility tasks, such as walking, will provide the highest discriminative value.

The objective is not to enable the diagnosis of those orthopedic conditions; rather, the hypothesis is that, based on measures derived from the iTUG, it may be possible to discriminate among these conditions in the phases immediately prior to discharge, given the distinct biomechanical implications underlying each. This initial evidence would support the use of the iTUG as a promising tool for patient follow-up, whether at home or during subsequent check-up visits.

Methods

Patients

Between April 2023 and December 2023, all patients admitted to the musculoskeletal ward of the Department of Physical Medicine and Rehabilitation at the Istituti Clinici Scientifici Maugeri Istituto di Ricovero e Cura a Carattere Scientifico (Gattico-Veruno, Piedmont, Italy) for rehabilitation following orthopedic surgery were screened for inclusion during the last 3 to 4 days of hospitalization. Patients were eligible if their primary admitting diagnosis was unilateral THA, TKA, or a femur fracture surgically repaired with internal fixation. Inclusion criteria included being aged 18 years or older, the ability to walk at least 10 meters without assistance (Functional Ambulation Category score ≥3), using the walking aid used in the final days of hospitalization, and the cognitive ability to understand instructions (Functional Independence Measure cognitive subscale ≥29).

Exclusion criteria included the presence of major associated injuries of the lower or upper limbs, contraindications to mobilization and weight-bearing on the lower extremities, clinical instability, disorders of bone metabolism other than osteoporosis, neurological or major cardiac diseases, pathological fractures, a prior history of overt dementia, terminal illness (life expectancy <6 mo), and prehospitalization dependence on nursing home care or inability to walk independently.

Ethical Considerations

The study was approved by the local ethics committee (2678 CE) and was conducted in accordance with the World Medical Association’s Declaration of Helsinki. All participants provided informed consent for study participation. To ensure participants’ privacy and confidentiality, all data were anonymized and coded using unique identification numbers. No personally identifiable information was stored in the research database or in the investigational application, and access to the data was restricted to authorized study personnel. Participants did not receive any financial compensation for their participation.

Setting and Procedures

iTUG Test

All evaluations were conducted in the Laboratory of Posture and Movement at the ICS Maugeri of Veruno. The laboratory provided a spacious, quiet area with adequate lighting, suitable for performing iTUG tests. The iTUG test was recorded following established protocols developed by Caronni et al [18,19]. During the test, participants wore a commercial IMU secured with a belt around their lower back at the level of the third lumbar vertebra (L3).

All participants were instructed to perform the iTUG test, which consisted of standing up from a chair, walking 3 m, pivoting, returning to the chair, and sitting down. An ordinary chair with a seat height of 44 cm was used. To mark the turning point, a yellow and black tape was placed on the floor at a distance of 3 m from the chair.

Participants initiated the task upon receiving a start signal from the iTUG commercial software and were instructed to maintain a comfortable and safe walking pace throughout the test.

Instrumentation

The iTUG system was supplied by mHealth Technologies srl (Monte San Pietro, Italy) in a portable briefcase containing the following components: a motion sensor with its charging station and power supply; 2 fabric belts of different sizes, equipped with Velcro closures, to secure the sensor to the patient; an Android smartphone with its charging cable; a user manual; a quick guide; and a declaration of conformity. The motion sensor is a lightweight IMU (22 g) with dimensions of 54 mm×33

XSL•FO
RenderX

mm×14 mm. It incorporates a triaxial accelerometer, a triaxial gyroscope, and a triaxial magnetometer, offering a maximum sampling frequency of 200 Hz. The sensor is positioned at the third lumbar vertebra (L3) using a Velcro-secured belt, with the buttons facing downward. Proper positioning is critical, as incorrect placement results in unreliable data.

The system connects the IMU to a smartphone via Bluetooth, with the smartphone acting as the central processing and control unit. During the iTUG test, the IMU captures movement signals, which are processed by the smartphone app "mTUG," developed by mHealth Technologies srl. The app manages the system and generates a report on the patient's kinematic performance during the test. The IMU signals provide the total duration of the TUG test and, through validated algorithms [16,17], automatically segment the test into four distinct phases: (1) sit-to-walk, (2) walk, (3) turn 180°, and (4) turn-to-sit. From each test, a total of 100 metrics is obtained (Multimedia Appendix 1).

Procedure

The iTUG procedures were conducted by 4 licensed physical therapists who underwent a dedicated half-day training session focused on standardized test administration, including sensor placement, task instructions, and use of the data acquisition and analysis software. Each participant was evaluated twice: first on the third-to-last day of hospitalization and then 24 hours later. A rater, randomly selected from the pool of 4 trained raters, conducted the evaluation. The rater provided instructions, demonstrated the procedure, secured the belt with the sensor, and initiated the measurement by activating the software to give the start signal. To ensure familiarity with the test, participants performed an initial trial that was not recorded and excluded from the analysis, following established guidelines [11]. After the familiarization trial, participants completed 2 recorded trials. During each trial, the rater manually recorded the time to complete the TUG test using a stopwatch in addition to the automated measurements provided by the iTUG system. The entire procedure, including the familiarization trial and the 2 recorded trials, was repeated 24 hours later by a different randomly selected trained rater.

Data Analysis

Reliability Measures

The best iTUG trial, that is, the trial with the shortest time recorded by the stopwatch, recorded on the first day was compared with the best trial recorded on the second day. These iTUG tests were used to calculate the test-retest reliability of the 100 variables collected for each participant, expressed as intraclass correlation coefficients (ICCs). In particular, we used an ICC 2,1 (model 2, form 1) [20] because the rater was not fixed and the data used to calculate the ICC were the measures of a single walking trial. An ICC less than 0.40 was considered to indicate poor reliability, between 0.40 and 0.59 indicated moderate reliability, between 0.60 and 0.74 indicated good reliability, and 0.75 or greater indicated excellent reliability [21]. All variables that obtained an ICC value less than 0.75 during reliability evaluation were excluded from the subsequent analysis.

Only after completing the variable reduction process, and thus after analyzing the correlation matrix, was it decided to calculate other measures of agreement for the resulting variables. Indeed, the SE of measurement (SEM), to assess the absolute error of the instrument, and the minimum detectable change (MDC), which represents the smallest change considered significant beyond the measurement error of an individual, were employed [20,22]. The SEM was calculated by multiplying the SD of the measurements by the square root of 1 minus the ICC (SD × √1-ICC). The MDC at 90% was derived using the formula MDC=1.65×2×SEM.

Correlation Matrix

For simplicity, we chose to perform all analyses on the variables collected during the first trial on the first day of testing. The variables included in the correlation matrix were tested for normal distribution using the Shapiro-Wilk test and Q-Q plot. For variables with nonnormal distributions, a logarithmic transformation was applied, followed by a reassessment of normality. Subsequent analyses were performed using the transformed variables where appropriate.

The data were screened for factorability using several well-described criteria, including the Kaiser-Meyer-Olkin (KMO) measure for sampling adequacy, the Bartlett test of sphericity, an anti-image correlation matrix [23-26], and the Pearson correlation and determinant score.

- The KMO statistic ranges from 0 to 1, with values above 0.50 considered acceptable for individual variables and a total value above 0.70 deemed adequate for the overall correlation matrix [20].
- The Bartlett test of sphericity evaluates the null hypothesis that the original correlation matrix is equal to an identity matrix (one with all 1's on the diagonal). A significant Bartlett test (*P*<.05) suggests that the variables are suitable for factor analysis. This condition prevents the application of factor analysis because it seeks items that are sufficiently correlated for the extraction of the factors or dimensions [20].
- The anti-image correlation matrix was inspected to test whether partial correlations among variables were small; we calculated the percentage of elements with values exceeding 0.30.
- From the correlation matrix, created to explore the relationships between variables, only variables with Pearson correlation coefficients ≥0.20 but <0.90 were considered, in order to avoid the presence of both isolated variables and multicollinearity [20].
- The determinant of a correlation matrix assesses issues related to extreme multicollinearity (ie, questionnaire items with high correlations) and singularity (questionnaire items with perfect correlations). These issues can hinder the application of exploratory factor analysis, indicating the need to exclude problematic items. Since the determinant of a correlation matrix can range from 0 to 1, it should be greater than 0.00001 to mitigate the potential problems associated with multicollinearity and singularity [27].

XSL•FO
RenderX

Exploratory Factor Analysis

We conducted an EFA using principal component extraction with varimax rotation. While principal component analysis is not inherently a latent variable model, it is commonly employed in EFA as a factor extraction method when the objective is to identify underlying constructs [23]. The resulting factors were interpreted as latent dimensions of iTUG performance. iTUG variables were first standardized and mean-centered. The number of factors to retain was determined using the following standard criteria [28]:

- Kaiser criterion [28,29], which retains only factors with eigenvalues ≥1
- Cattell scree test [30], based on the visual inspection of the plot of eigenvalues
- Horn parallel analysis [31], where random eigenvalues are compared to the observed data to identify meaningful factors

Only variables with factor loadings ≥|0.40| were considered relevant contributors to each factor. The factor scores for each individual were then computed and used in subsequent analyses, representing the extent to which each participant expressed the corresponding latent dimension [32].

Linear Discriminant Analysis

The factor scores derived from the EFA were entered into an LDA with 10-fold cross-validation to evaluate the ability of iTUG variables to classify patients into the 3 diagnostic groups (THA, TKA, and femur fracture). Model performance was assessed using receiver operating characteristic (ROC) analysis, and discrimination was quantified using the area under the curve (AUC). The AUC of the proposed model was compared with that obtained from manually timed TUG performance (stopwatch) and with AUC values derived from an LDA model, including all iTUG variables with good reliability (ICC>0.75). AUC values were computed and reported separately for the THA, TKA, and femur fracture groups. AUC values were interpreted as follows: poor discrimination, AUC<0.70; moderate discrimination, AUC between 0.70 and 0.89; and high discrimination, AUC≥0.90. Classification accuracy, calculated from the confusion matrix, was also reported separately for each diagnostic group.

Statistical Analysis

Mean (SD) values were used for descriptive statistics, and mean (SE) values were used for the figures. One-way ANOVA was used for the direct comparison between 2 variables (“Total Duration till initial contact with the chair” vs TUG test duration by chronometer). The level of significance was set at *P*<.05. Statistical analysis was conducted using R software (R Foundation for Statistical Computing), with the packages *blandr*, *EFAtools*, *FactoMineR*, *graphics*, *irr*, *lattice*, *MASS*, *methods*, *multiUS*, *psych*, and *stats*.

Sample Size Calculation

The sample size required for the reliability study was calculated based on the ICC method [33]. With 2 assessors using the iTUG system and assuming a reliability of at least 0.7, a sample of 104 patients was determined to be adequate.

Results

Participants

Of the 266 inpatients screened, 112 (42%) met the inclusion criteria and were enrolled in the study. The remaining patients were excluded from the study for the following reasons: no walking ability, cognitive impairments, clinical instability, other types of fractures besides femur fractures, or severe comorbidities. Among the 112 recruited participants, 3 did not undergo the second-day trial due to clinical reasons, 2 did not return for reassessment, and for 3 others, iTUG data could not be extracted due to software malfunctions. Thus, the final sample comprised 104 participants. Of these, 47% were hospitalized for femur fractures, 36% for THA, and 17% for TKA. Table 1 presents the clinical characteristics of the assessed patients.

XSL•FO
RenderX

Table . Characteristics of the recruited sample (N=104).

Characteristics	Value
Demographic	
Age (y), mean (SD)	73.6 (10.9)
Sex, n	
Female	69
Male	35
Clinical	
BMI, mean (SD)	22.8 (3.5)
Interval between surgery and assessment (d), mean (SD)	26.4 (6.8)
Injury side, n	
Right	54
Left	50
Type of surgical intervention, n	
TKA[a]	18
THA[b]	37
Femur fracture by osteosynthesis	20
Femur fracture by THA	29

[a] TKA: total knee arthroplasty.

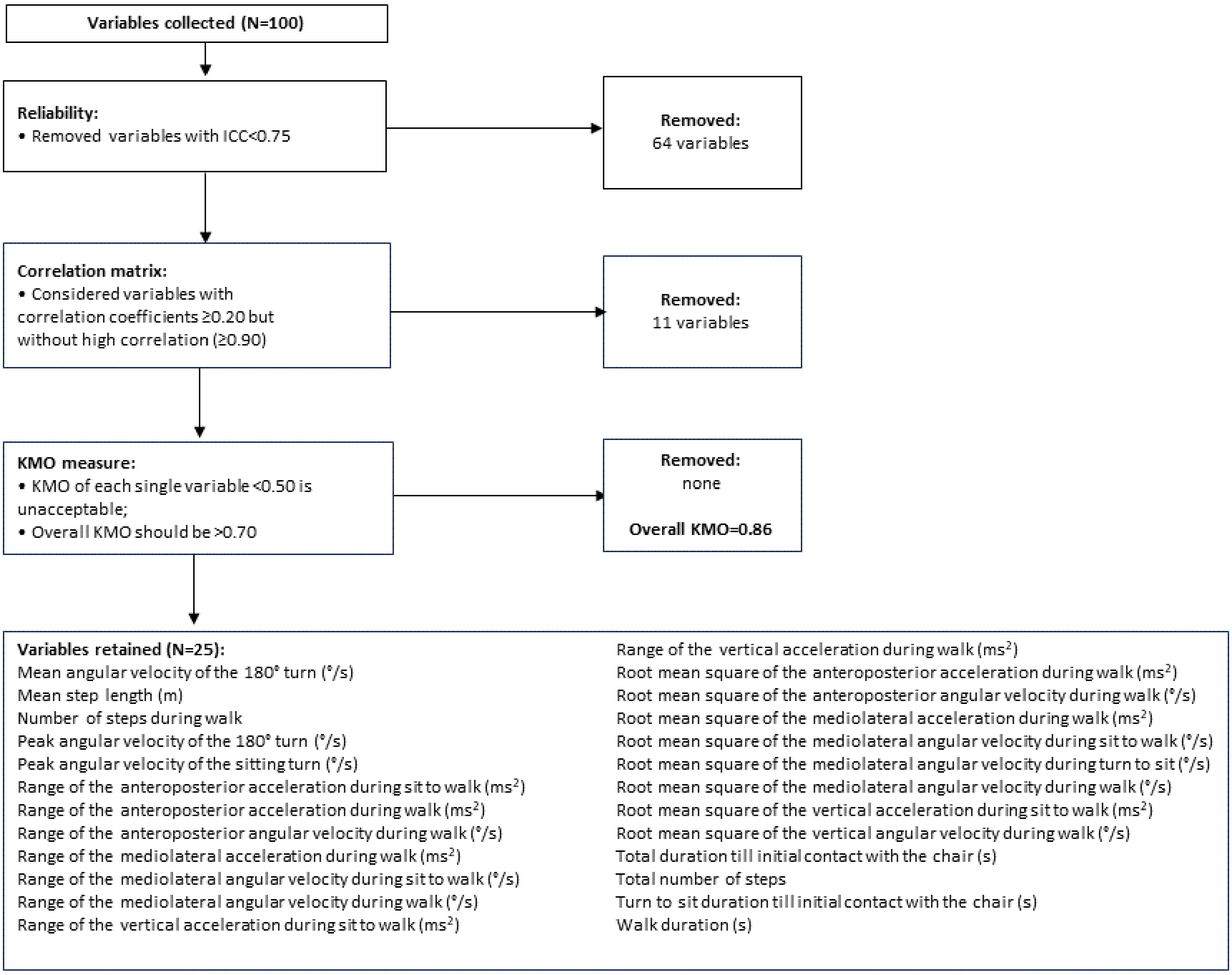
[b] THA: total hip arthroplasty.

Reliability Measures

The total iTUG duration recorded using the stopwatch was highly comparable to the iTUG-derived variable “Total Duration till initial contact with the chair” (correlation: $y = 0.99x + 0.24$; R^2=0.99). The mean time was 17.5 (SD 6.3) seconds for manual versus 17.6 (SD 6.3) seconds for iTUG, with no significant differences among orthopedic groups (ANOVA; P=.09). The ICC for this variable was 0.90 (95% CI 0.849 ‐ 0.928), and it was therefore included in the subsequent EFA. Figure 1 illustrates the stepwise selection process for reliable iTUG variables.

XSL•FO
RenderX

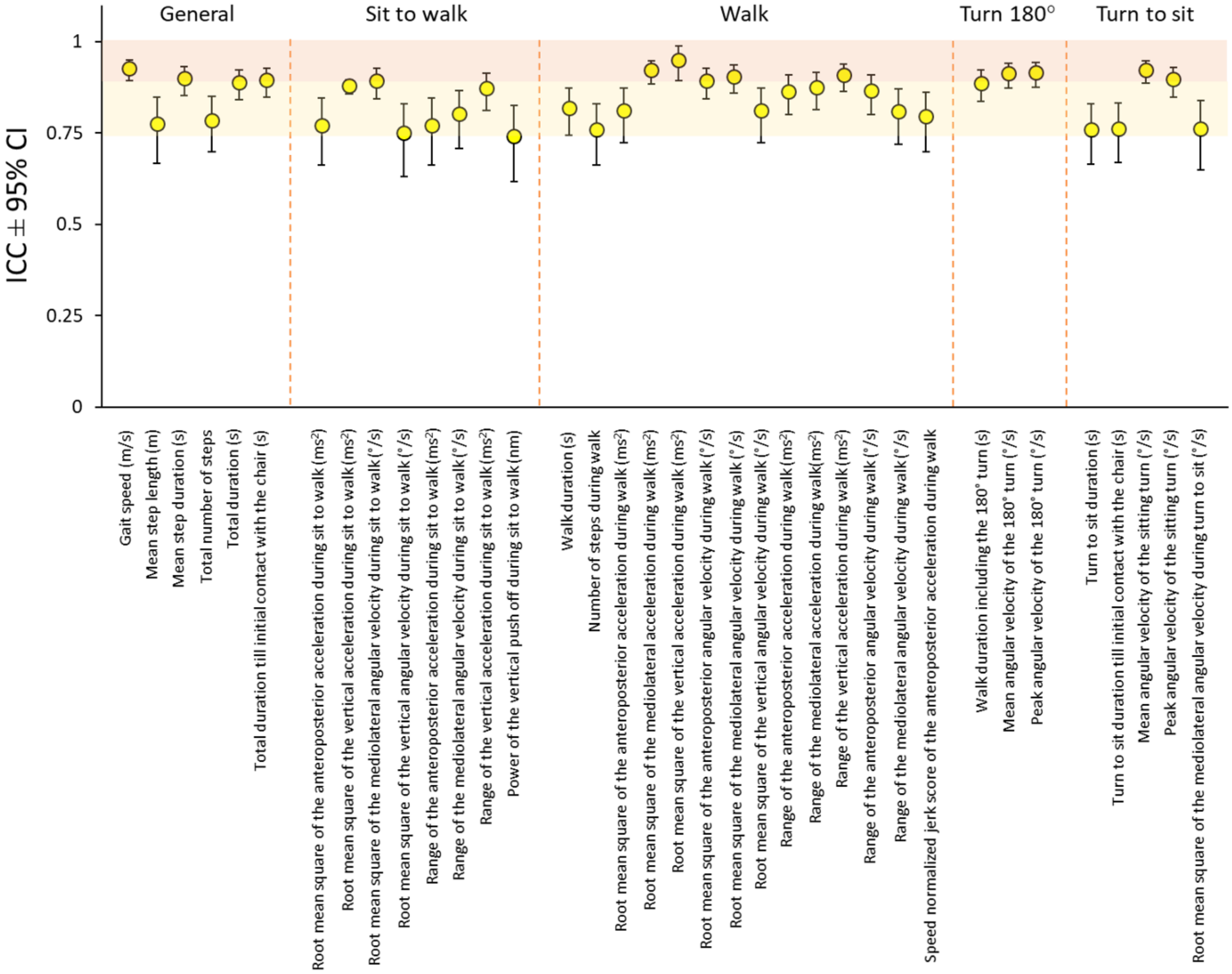
Figure 1. Flowchart describing the initial selection of variables from the instrumented Timed Up and Go (iTUG) test. ICC: intraclass correlation coefficient; KMO: Kaiser-Meyer-Olkin.



ICC values for all 100 iTUG variables initially considered were calculated. Subsequently, only variables with high ICC values were retained, resulting in a total of 36 variables. Figure 2 illustrates the ICC of these variables, providing an assessment of test-retest reliability: each point represents the ICC value for a variable, with error bars indicating the 95% CIs around the estimate. These intervals reflect the precision of the ICC measurement; narrower intervals denote higher confidence in the reliability estimate. With these 36 reliable variables, a correlation matrix was calculated.

XSL•FO
RenderX

Figure 2. Representation of the test-retest reliability of the 36 variables, divided into the different tasks of the instrumented Timed Up and Go (iTUG) test, with intraclass correlation coefficient (ICC) values >0.75. The top orange bar indicates excellent ICCs, meaning >0.90, while the yellow bar indicates ICCs >0.75, used as a cutoff to identify highly reliable variables.

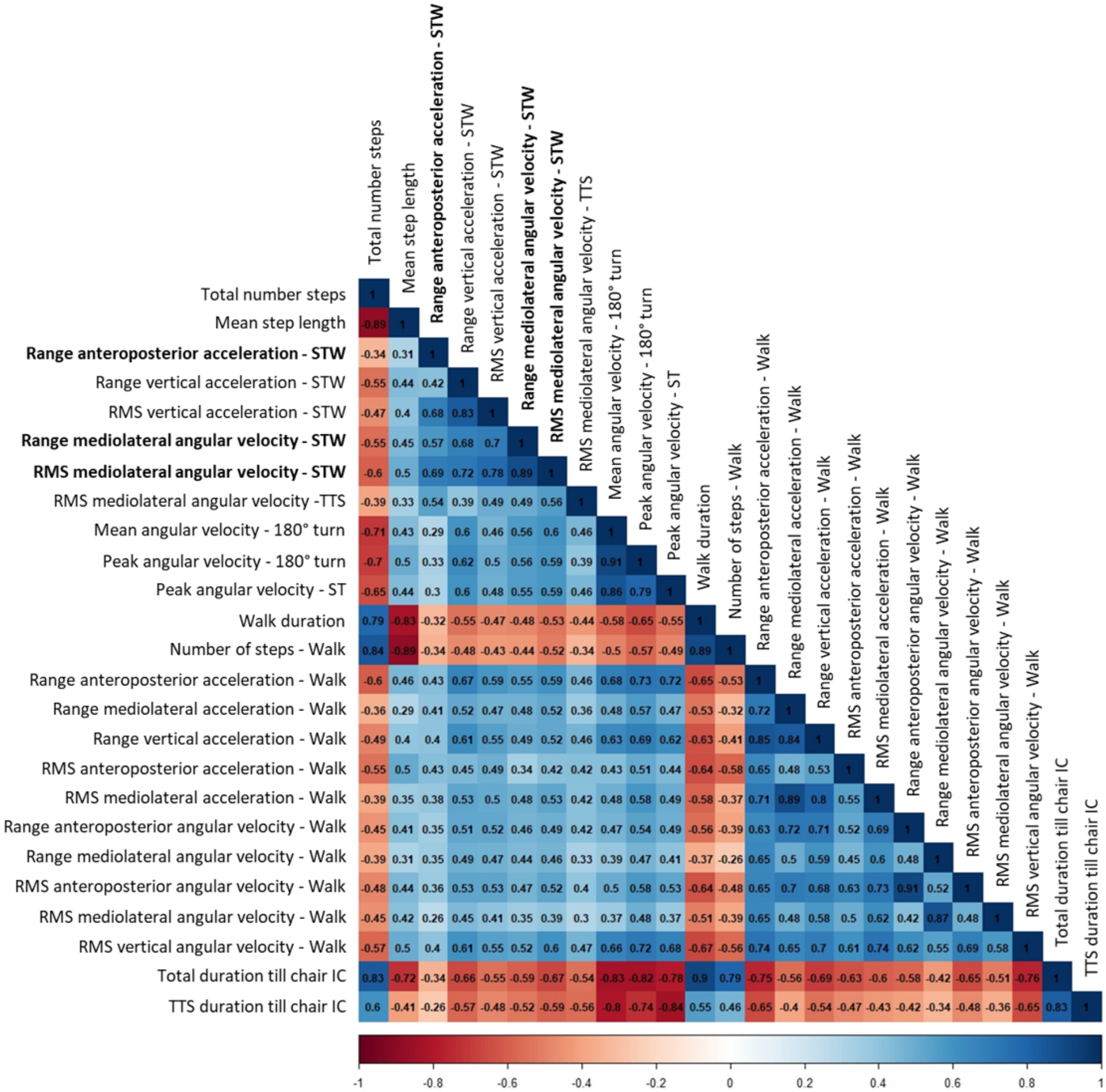


Correlation Matrix and Factor Analysis

Of the 36 retained variables, 6 were normally distributed; the remaining variables were log-transformed, and their distributions were then evaluated again for normality. Based on the predefined criteria for the correlation matrix (see “Methods” section), 11 variables were eliminated from the initial group of 36 variables (Figure 3).

XSL•FO
RenderX

Figure 3. Correlation matrix, shown as heatmap, between the final 25 variables of the instrumented Timed Up and Go (iTUG) test. Data from all participants were used to prepare the matrix (N=104). Variables shown in bold were already normally distributed and did not require logarithmic transformation. IC: initial contact; RMS: root mean square; ST: sitting turn; STW: sit-to-walk; TTS: turn-to-sit.



	Total number steps	Mean step length	**Range anteroposterior acceleration - STW**	Range vertical acceleration - STW	RMS vertical acceleration - STW	**Range mediolateral angular velocity - STW**	**RMS mediolateral angular velocity - STW**	RMS mediolateral angular velocity - TTS	Mean angular velocity - 180° turn	Peak angular velocity - 180° turn	Peak angular velocity - ST	Walk duration	Number of steps - Walk	Range anteroposterior acceleration - Walk	Range mediolateral acceleration - Walk	Range vertical acceleration - Walk	RMS anteroposterior acceleration - Walk	RMS mediolateral acceleration - Walk	Range anteroposterior angular velocity - Walk	Range mediolateral angular velocity - Walk	RMS anteroposterior angular velocity - Walk	RMS mediolateral angular velocity - Walk	RMS vertical angular velocity - Walk	Total duration till chair IC	TTS duration till chair IC
Total number steps	1																								
Mean step length	-0.89	1																							
Range anteroposterior acceleration - STW	-0.34	0.31	1																						
Range vertical acceleration - STW	-0.55	0.44	0.42	1																					
RMS vertical acceleration - STW	-0.47	0.4	0.68	0.83	1																				
Range mediolateral angular velocity - STW	-0.55	0.45	0.57	0.68	0.7	1																			
RMS mediolateral angular velocity - STW	-0.6	0.5	0.69	0.72	0.78	0.89	1																		
RMS mediolateral angular velocity -TTS	-0.39	0.33	0.54	0.39	0.49	0.49	0.56	1																	
Mean angular velocity - 180° turn	-0.71	0.43	0.29	0.6	0.46	0.56	0.6	0.46	1																
Peak angular velocity - 180° turn	-0.7	0.5	0.33	0.62	0.5	0.56	0.59	0.39	0.91	1															
Peak angular velocity - ST	-0.65	0.44	0.3	0.6	0.48	0.55	0.59	0.46	0.86	0.79	1														
Walk duration	0.79	-0.83	-0.32	-0.55	-0.47	-0.48	-0.53	-0.44	-0.58	-0.65	-0.55	1													
Number of steps - Walk	0.84	-0.89	-0.34	-0.48	-0.43	-0.44	-0.52	-0.34	-0.5	-0.57	-0.49	0.89	1												
Range anteroposterior acceleration - Walk	-0.6	0.46	0.43	0.67	0.59	0.55	0.59	0.46	0.68	0.73	0.72	-0.65	-0.53	1											
Range mediolateral acceleration - Walk	-0.36	0.29	0.41	0.52	0.47	0.48	0.52	0.36	0.48	0.57	0.47	-0.53	-0.32	0.72	1										
Range vertical acceleration - Walk	-0.49	0.4	0.4	0.61	0.55	0.49	0.52	0.46	0.63	0.69	0.62	-0.63	-0.41	0.85	0.84	1									
RMS anteroposterior acceleration - Walk	-0.55	0.5	0.43	0.45	0.49	0.34	0.42	0.42	0.43	0.51	0.44	-0.64	-0.58	0.65	0.48	0.53	1								
RMS mediolateral acceleration - Walk	-0.39	0.35	0.38	0.53	0.5	0.48	0.53	0.42	0.48	0.58	0.49	-0.58	-0.37	0.71	0.89	0.8	0.55	1							
Range anteroposterior angular velocity - Walk	-0.45	0.41	0.35	0.51	0.52	0.46	0.49	0.42	0.47	0.54	0.49	-0.56	-0.39	0.63	0.72	0.71	0.52	0.69	1						
Range mediolateral angular velocity - Walk	-0.39	0.31	0.35	0.49	0.47	0.44	0.46	0.33	0.39	0.47	0.41	-0.37	-0.26	0.65	0.5	0.59	0.45	0.6	0.48	1					
RMS anteroposterior angular velocity - Walk	-0.48	0.44	0.36	0.53	0.53	0.47	0.52	0.4	0.5	0.58	0.53	-0.64	-0.48	0.65	0.7	0.68	0.63	0.73	0.91	0.52	1				
RMS mediolateral angular velocity - Walk	-0.45	0.42	0.26	0.45	0.41	0.35	0.39	0.3	0.37	0.48	0.37	-0.51	-0.39	0.65	0.48	0.58	0.5	0.62	0.42	0.87	0.48	1			
RMS vertical angular velocity - Walk	-0.57	0.5	0.4	0.61	0.55	0.52	0.6	0.47	0.66	0.72	0.68	-0.67	-0.56	0.74	0.65	0.7	0.61	0.74	0.62	0.55	0.69	0.58	1		
Total duration till chair IC	0.83	-0.72	-0.34	-0.66	-0.55	-0.59	-0.67	-0.54	-0.83	-0.82	-0.78	0.9	0.79	-0.75	-0.56	-0.69	-0.63	-0.6	-0.58	-0.42	-0.65	-0.51	-0.76	1	
TTS duration till chair IC	0.6	-0.41	-0.26	-0.57	-0.48	-0.52	-0.59	-0.56	-0.8	-0.74	-0.84	0.55	0.46	-0.65	-0.4	-0.54	-0.47	-0.43	-0.42	-0.34	-0.48	-0.36	-0.65	0.83	1

Sampling adequacy was confirmed (KMO=0.86), and the Bartlett test indicated significant correlations among variables (χ^2_{300}=3199.92; *P*<.001). The determinant of the correlation matrix was 7.36×10^{-16}, indicating no issues with multicollinearity or singularity. The percentage of partial correlations below the threshold of 0.3 in the anti-image correlation matrix was 16%, a value that can be considered adequate for proceeding with factor analysis.

Table 2 shows the correlations among the 25 selected variables, with all pairwise correlations exceeding the minimum acceptable threshold of 0.2. The table also reports the factor loadings for each variable on the 5 extracted factors, which were labeled according to their conceptual meaning and the naming convention proposed by Coni et al [34]: "walking ability," "pace or rhythm," "sit-to-walk smoothness," "turning ability," and "mediolateral angular stability." The factor loadings were derived from a PCA with varimax rotation, which preserves orthogonality and retains the full variance of the original data. PCA identified 5 components with eigenvalues ≥1 (Kaiser criterion), and this was confirmed by the Horn parallel analysis. Together, these 5 components explained 80.8% of the total variance in the iTUG dataset, with individual contributions of 0.54, 0.09, 0.07, 0.06, and 0.05 of the variance, respectively.

XSL•FO
RenderX

Table . Factor loadings of the instrumented Timed Up and Go (iTUG) variables after principal component analysis[a].

	Walking ability	Pace or rhythm	Sit-to-walk smoothness	Turning ability	Mediolateral angular stability
Range of the anteroposterior acceleration during walk (ms^2)	*0.529*	—[b]	—	0.489	0.410
Range of the anteroposterior angular velocity during walk (°/s)	*0.811*	—	—	—	—
Range of the mediolateral acceleration during walk (ms^2)	*0.828*	—	—	—	—
Range of the vertical acceleration during walk (ms^2)	*0.704*	—	—	0.403	—
Root mean square of the anteroposterior acceleration during walk (ms^2)	*0.474*	0.492	—	—	—
Root mean square of the anteroposterior angular velocity during walk (°/s)	*0.791*	—	—	—	—
Root mean square of the mediolateral acceleration during walk (ms^2)	*0.782*	—	—	—	—
Root mean square of the vertical angular velocity during walk (°/s)	*0.530*	—	—	0.469	—
Mean step length (m)	—	*0.903*	—	—	—
Number of steps during walk	—	*–0.900*	—	—	—
Total duration till initial contact with the chair (s)	—	*–0.598*	—	–0.630	—
Total number of steps	—	*–0.771*	—	–0.451	—
Walk duration (s)	–0.405	*–0.786*	—	—	—
Range of the anteroposterior acceleration during sit-to-walk (ms^2)	—	—	*0.823*	—	—
Range of the mediolateral angular velocity during sit-to-walk (°/s)	—	—	*0.749*	—	—
Range of the vertical acceleration during sit-to-walk (ms^2)	—	—	*0.576*	0.422	—
Root mean square of the mediolateral angular velocity during sit-to-walk (°/s)	—	—	*0.801*	—	—
Root mean square of the mediolateral angular velocity during turn-to-sit (°/s)	—	—	*0.547*	—	—

XSL•FO
RenderX

	Walking ability	Pace or rhythm	Sit-to-walk smoothness	Turning ability	Mediolateral angular stability
Root mean square of the vertical acceleration during sit-to-walk (ms^2)	—	—	*0.795*	—	—
Mean angular velocity of the 180° turn (°/s)	—	—	—	*0.858*	—
Peak angular velocity of the 180° turn (°/s)	—	—	—	*0.743*	—
Peak angular velocity of the sitting turn (°/s)	—	—	—	*0.827*	—
Turn to sit duration till initial contact with the chair (s)	—	—	—	*–0.823*	—
Range of the mediolateral angular velocity during walk (°/s)	—	—	—	—	*0.845*
Root mean square of the mediolateral angular velocity during walk (°/s)	—	—	—	—	*0.857*

[a] Italicized values indicate the highest loading for each variable, representing the primary factor association. Only values greater than 0.4 are reported in the table. Variables shown in italics were already normally distributed and did not require logarithmic transformation.

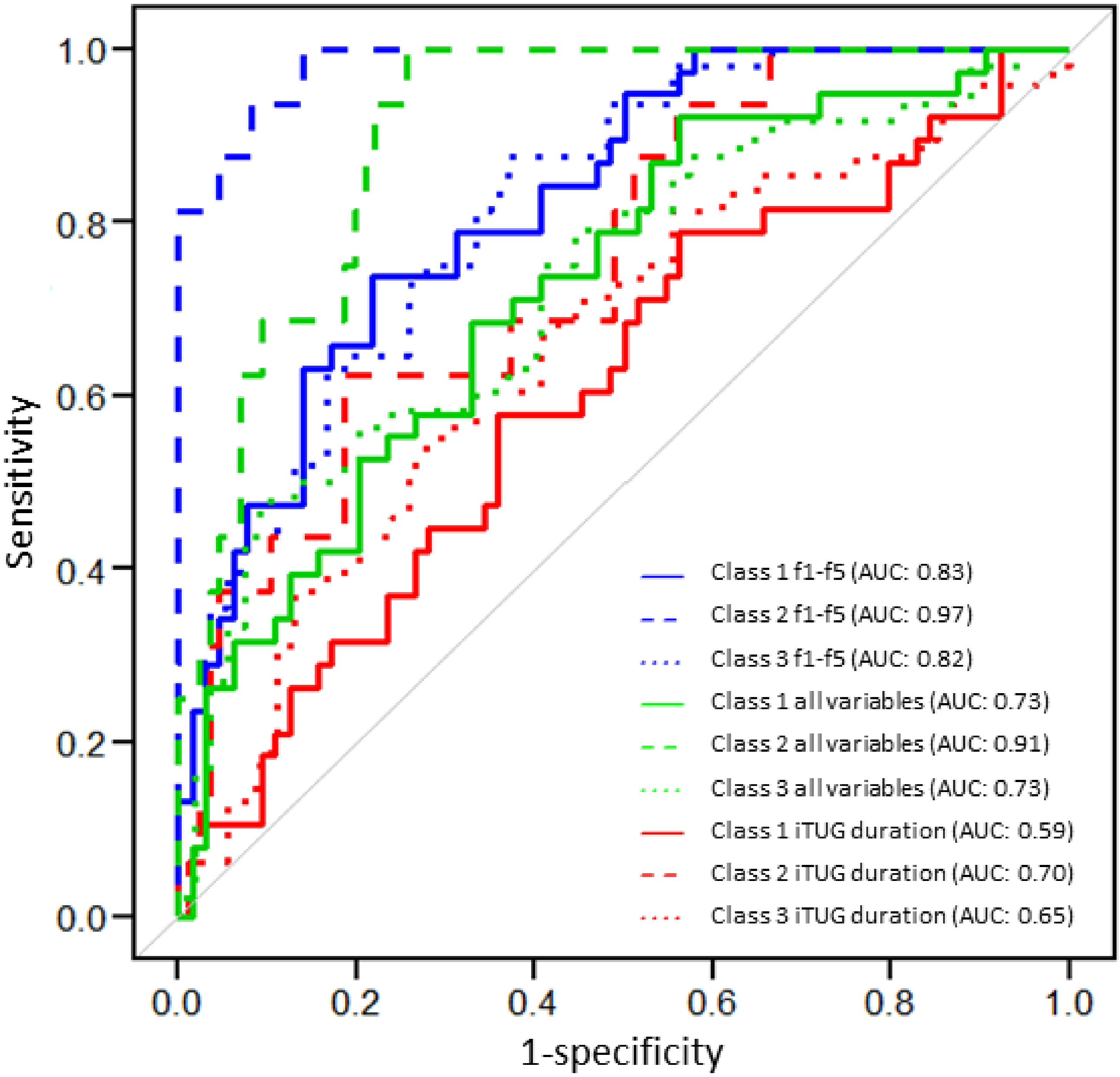
[b] Not applicable.

Linear Discriminant Analysis

LDA revealed that the first 5 factor scores, corresponding to the factors walking ability, pace or rhythm, sit-to-walk smoothness, turning ability, and mediolateral angular stability, successfully discriminated between the 3 diagnoses with an accuracy of 0.68. These 5 factor scores demonstrated a total AUC of 0.87, with an AUC of 0.83 for THA, 0.97 for TKA, and 0.82 for femur fractures (Figure 4).

XSL•FO
RenderX

Figure 4. Representation of receiver operating characteristic (ROC) curves. ROC compares the classification performance of instrumented Timed Up and Go (iTUG) duration (variable labeled as iTUG "total duration till initial contact with the chair"; red), factors (f1–f5; blue), and 5 randomly selected variables (green). Sensitivity is plotted against 1–specificity, illustrating the discriminative ability of each variable. The curves demonstrate the ability of different predictors to distinguish between patient diagnoses: total hip arthroplasty (THA; solid lines), total knee arthroplasty (TKA; long dashed lines), and femur fracture (short dashed lines). AUC: area under the curve.



The accuracy of the discriminant analysis model using the variable "total duration" was 0.50, with a total AUC of 0.65. In contrast, the model including all iTUG variables with good reliability achieved an accuracy of 0.54 and an AUC of 0.79.

Measures of Agreement

Table 3 shows the descriptive statistics, test-retest reliability indices, and measurement error parameters for the retained variables (n=25) across 2 sessions (day 1 and day 2). Each variable is accompanied by its mean and SD for both days, reflecting the overall consistency of the measures across time. The SEM provides an estimate of the measurement error attributable to random variations, offering insight into the precision of each variable. Lower SEM values reflect greater measurement precision. The MDC, derived from the SEM, represents the smallest change in a variable that can be interpreted as a true difference beyond measurement error. Larger MDC values indicate greater variability in detecting meaningful changes over time, whereas smaller values suggest higher sensitivity.

XSL•FO
RenderX

Table . Descriptive statistics and agreement measures of the retained variables (n=25).

Variable	Day 1, mean (SD)	Day 2		ICC[a]		
		Mean (SD)	ICC	95% CI	SEM[b]	MDC[c] (90%)
Total number of steps	15.298 (4.952)	14.644 (4.110)	0.779	0.689-0.846	2.138	4.973
Mean step length (m)	0.573 (0.192)	0.589 (0.167)	0.775	0.669-0.847	0.085	0.198
Range of the anteroposterior acceleration during sit-to-walk (ms^2)	9.842 (2.058)	9.702 (2.217)	0.784	0.681-0.853	0.993	2.310
Range of the vertical acceleration during sit-to-walk (ms^2)	4.724 (2.091)	4.733 (1.812)	0.805	0.712-0.868	0.862	2.005
Root mean square of the vertical acceleration during sit-to-walk (ms^2)	1.337 (0.512)	1.373 (0.505)	0.882	0.826-0.920	0.174	0.406
Range of the mediolateral angular velocity during sit-to-walk (°/s)	143.058 (53.462)	150.363 (52.583)	0.871	0.808-0.913	19.064	44.351
Root mean square of the mediolateral angular velocity during sit-to-walk (°/s)	52.204 (15.502)	54.974 (15.589)	0.890	0.830-0.927	5.173	12.035
Root mean square of the mediolateral angular velocity during turn-to-sit (°/s)	0.700 (0.305)	0.726 (0.340)	0.482	0.236-0.649	0.232	0.539
Mean angular velocity of the 180° turn (°/s)	66.099 (27.070)	72.208 (29.286)	0.901	0.834-0.938	8.911	20.730
Peak angular velocity of the 180° turn (°/s)	124.330 (42.056)	131.032 (45.224)	0.911	0.863-0.941	13.07	30.405
Peak angular velocity of the sitting turn (°/s)	130.800 (46.552)	140.803 (51.975)	0.889	0.823-0.928	16.489	38.359
Walk duration (s)	9.134 (3.804)	8.181 (3.073)	0.790	0.648-0.869	1.597	3.714

XSL•FO
RenderX

Variable	Day 1, mean (SD)	Day 2		ICC[a]	SEM[b]	MDC[c] (90%)
		Mean (SD)	ICC	95% CI		
Number of steps during walk	10.019 (4.098)	9.404 (3.287)	0.855	0.783-0.903	1.417	3.296
Range of the anteroposterior acceleration during walk (ms^2)	5.369 (2.001)	5.697 (2.093)	0.862	0.794-0.908	0.76	1.769
Range of the mediolateral acceleration during walk (ms^2)	5.868 (2.192)	6.338 (2.704)	0.868	0.797-0.913	0.896	2.085
Range of the vertical acceleration during walk (ms^2)	6.810 (2.722)	7.349 (2.670)	0.901	0.840-0.937	0.849	1.974
Root mean square of the anteroposterior acceleration during walk (ms^2)	1.438 (0.490)	1.553 (0.557)	0.810	0.711-0.874	0.23	0.534
Root mean square of the mediolateral acceleration during walk (ms^2)	0.915 (0.287)	0.982 (0.325)	0.912	0.846-0.946	0.091	0.212
Range of the anteroposterior angular velocity during walk (°/s)	54.028 (19.465)	56.768 (20.359)	0.863	0.797-0.907	7.372	17.150
Range of the mediolateral angular velocity during walk (°/s)	72.046 (27.695)	73.475 (27.123)	0.812	0.723-0.872	11.862	27.595
Root mean square of the anteroposterior angular velocity during walk (°/s)	9.390 (3.202)	9.997 (3.766)	0.888	0.828-0.926	1.172	2.726
Root mean square of the mediolateral angular velocity during walk (°/s)	13.355 (4.248)	14.005 (5.014)	0.901	0.852-0.934	1.46	3.396

XSL•FO
RenderX

Variable	Day 1, mean (SD)	Day 2		ICC[a]		
		Mean (SD)	ICC	95% CI	SEM[b]	MDC[c] (90%)
Root mean square of the vertical angular velocity during walk (°/s)	20.331 (6.604)	21.725 (6.561)	0.803	0.702-0.868	2.934	6.826
Total duration till initial contact with the chair (s)	17.671 (6.248)	15.838 (5.291)	0.853	0.619-0.929	2.241	5.214
Turn-to-sit duration till initial contact with the chair (s)	4.165 (1.931)	3.683 (1.655)	0.738	0.603-0.826	0.926	2.155

[a] ICC: intraclass correlation coefficient.

[b] SEM: SE of measurement.

[c] MDC: minimal detectable change.

Discussion

This study demonstrated that, in a sample of heterogeneous orthopedic patients discharged from a rehabilitative ward, only 25 out of the 100 original or exported iTUG variables were both reliable and capable of fitting into a robust assessment model with good indices. EFA on this model identified 5 domains that contribute to motor performance: walking ability, pace or rhythm, sit-to-walk smoothness, turning ability, and mediolateral angular stability. The ability of these factors to differentiate between the 3 orthopedic diagnoses is significantly higher than the total duration of the test. This finding underscores the clinical value of moving beyond single-duration metrics toward a multidimensional, sensor-based assessment framework.

A total of 36 out of the 100 variables derived from the iTUG exhibited excellent test-retest reliability. The fact that only one-third of the variables are reliable is consistent with findings previously reported in a population of elderly patients with femoral fractures [35]. In recent years, several studies have investigated the reliability of the iTUG in different populations [36-39], yet none have conducted a detailed analysis of individual variable reliability in patients with knee or hip replacements.

In our study, “Total Duration till initial contact with the chair” was found to be more reliable than the durations of the individual subphases. The most reliable variables were those related to the walking phase, whereas jerk-based metrics did not reach adequate levels of reliability. This finding is consistent with previous literature suggesting that the limited reliability of jerk-based features may be due to multiple concurrent factors, particularly during postural transitions such as sit-to-stand [39]. As reported by Weiss et al [40], outcome measures like jerk and range, which are highly sensitive to subtle variations and signal noise, may show greater variability and therefore reduced reliability, despite standardized acquisition protocols.

As expected, the measures that can be obtained using a stopwatch—that is, total duration and the gait variables derived from it (ie, gait speed, mean step duration, and cadence)—showed high reliability. It is worth noting that, although these variables can be derived from manual timing, the instrumented test additionally offers the advantage of being operator-independent [41].

Measures of stability, regularity, and coordination exhibited low reliability, likely due to the limited number of steps performed during the iTUG test. It is well established that reliable stability and coordination metrics typically require more than 20 to 25 steps [42,43]. In our sample, participants performed, on average, fewer than 15 steps, highlighting a limitation inherent to the test itself rather than to the measures.

Furthermore, all participants in our orthopedic patient sample used a walking aid, which appears to compromise jerk-related measures by disrupting the natural flow of gait [44]. As a result, only 1 jerk variable during the walking phase demonstrated good reliability.

A recent review has confirmed the reliability of single-point IMUs for gait metric analysis and their potential to support clinical applications, especially when placed over the L3-L4 vertebrae [45]. This approach has been shown to be valid for the iTUG as well [16]. Despite the simplicity of use and the demonstrated validity of using a single IMU placed over the lumbar region, the accuracy of spatiotemporal gait parameter estimation is significantly influenced by sensor placement [46]. For instance, although the average error across multiple strides remained consistent across different sensor positions, significant variations were observed in single-stride errors and variability parameters [47]. Practical considerations, such as ease of execution and time efficiency associated with different IMU

XSL•FO
RenderX

configurations, also play an important role. The single-sensor approach offers a clear advantage in terms of usability, as starting data collection is as straightforward as pressing "start" on a smartphone, comparable to using a stopwatch. On the other hand, while adding more sensors can enhance data richness and improve the accuracy of the estimation of some features, it might notably increase the time required to complete the evaluation, and its reliability would still be influenced by sensor placement [47,48]. Therefore, the iTUG test with only a single-sensor setup provides a good balance between sensor count and wearability, which is an optimal tradeoff between accuracy and practicality in real-world applications.

Furthermore, a single sensor for mobility assessment facilitates the transition of clinical evaluations from traditional supervised clinical settings to unsupervised home environments, thereby promoting and enabling continuity of care scenarios where the patient receives integrated, coordinated, and person-centered care, both during acute episodes and in the management of chronic conditions, reducing fragmentation in health care delivery [45].

Previous studies have demonstrated the use of iTUG measures in distinguishing preoperative functional status across various orthopedic conditions [49]. For instance, Bloomfield [50] categorized patients undergoing total knee replacement into moderate-functioning and low-functioning groups based on preoperative test duration. Similarly, Gasparutto et al [51] conducted a biomechanical analysis of the TUG test in total hip replacement patients, comparing functional deficits in each phase of the test in patients with those of a control group. However, these discriminative capabilities within relatively homogeneous orthopedic populations appear to be primarily driven by large differences in total test duration, rather than by the intrinsic discriminative power of the extracted sensor-based variables. In fact, both studies reported differences exceeding the 4-second minimal clinically important difference for orthopedic populations [52], potentially limiting the added value of sensor-based analysis.

In contrast, given the lack of significant differences in total TUG duration among the 3 patient groups, our findings suggest that this measure alone is insufficient to differentiate between groups in a sample of relatively high-functioning individuals. Instead, our iTUG-based model, which incorporates the 5 principal components, provided a significantly superior discriminative ability compared to total duration alone. Specifically, the iTUG model achieved an AUC greater than 0.8, whereas the model based solely on total duration yielded an AUC less than 0.7, indicating limited predictive power.

Overall, our results reinforce the added value and sensitivity of iTUG over the traditional stopwatch-based TUG. At the same time, they show that when sensor technology is not used and only standard variables are considered, the total duration recorded by the iTUG remains comparable to conventional measurements, as recently reported by Dos Santos et al [53]. This underscores the importance of integrating sensor-based analytics to refine mobility assessments and optimize rehabilitation strategies, ultimately improving clinical decision-making and patient care.

Although the LDA conducted in this study demonstrated good discriminant ability among groups (mean overall accuracy 0.68, SD 0.14), more advanced predictive models, such as machine learning algorithms, could perform even better. Machine learning approaches have already been applied to iTUG assessments in various populations and for different purposes, including evaluating levodopa responsiveness in individuals with Parkinson disease [54] and predicting falls across different pathologies [1]. In this perspective, the present study proposes a reliability-driven and interpretable framework for variable selection and dimensionality reduction, which may provide a preliminary layer of feature selection for the development of predictive models not only in orthopedic rehabilitation but also in other populations characterized by mobility impairments, such as neurological disorders. While the discriminative patterns identified here are specific to postsurgical orthopedic patients, the latent mobility domains derived from the iTUG (eg, walking ability, turning, sit-to-walk smoothness, and mediolateral stability) consist of fundamental components of functional mobility, which are also affected in other conditions, including neurological disorders.

This study has several clinical implications. First, the use of multidimensional, sensor-derived metrics, particularly those related to walking dynamics, offers a standardized and operator-independent approach to assessing mobility in orthopedic patients. This supports consistent tracking of rehabilitation progress across settings. Second, the factor structure identified through EFA provides clinicians with interpretable domains (eg, sit-to-walk smoothness and mediolateral angular stability) with strong discriminative performance (AUC>0.8) that can inform targeted therapy planning beyond total time. Third, the iTUG model's ability to discriminate between patients with similar test durations suggests that it can uncover latent performance differences that might otherwise remain undetected, especially in high-functioning individuals.

Selected variables showed stable and reproducible measurements in the present dataset. This suggests their potential suitability not only for cross-sectional group differentiation but also for longitudinal monitoring, where consistency of measurement is required to detect true change beyond measurement noise. From a clinical perspective, this capability is particularly relevant for stratifying and monitoring patients within the same orthopedic condition, such as those undergoing THA or TKA. The assessment and selection of sensor-based features represent an initial step toward the development of an evaluation model aimed at quantifying motor impairments that may distinguish the categories and reflect rehabilitation targets. The underlying objective is to capture functional dimensions that could inform treatment decisions; accordingly, these variables may serve as candidate outcome measures to track patient progress, evaluate treatment effectiveness, and support early adjustment of rehabilitation strategies. They may also contribute to finer personalization within each category by highlighting individual functional profiles.

Finally, once a reliable and clinically meaningful feature set is identified, the same variables could provide a basis for additional predictive models, such as fall-risk estimation or forecasting

XSL•FO
RenderX

functional decline. However, these findings should be considered exploratory. Validation in larger and more diverse cohorts will be necessary to assess the robustness, generalizability, and actual performance of both the assessment and predictive models and to determine their potential clinical use. Future work will therefore focus on increasing the sample size and conducting within-diagnosis analyses to identify distinct functional recovery profiles, enabling the distinction between patients who are recovering as expected and those showing subtle but clinically meaningful deviations from typical recovery trajectories. Such early identification may support timely, personalized rehabilitation adjustments before functional decline or plateau becomes evident using conventional time-based measures.

Moreover, the feasibility of a single-sensor setup reinforces the iTUG's potential for remote monitoring and continuity of care postdischarge. Reliable domains such as walking ability and pace or rhythm may represent valuable longitudinal indicators of recovery and could be integrated into digital tools for monitoring rehabilitation outcomes.

In conclusion, this study confirms the reliability of selected iTUG variables in orthopedic patients and reinforces their clinical value for a multidimensional assessment of motor performance. Factor analysis identified 5 meaningful components capable of distinguishing between common orthopedic conditions, addressing the limitations of relying solely on total test duration. These findings support the hypothesis that the development of predictive models and clinical decision support tools based on sensor-based functional tests, such as the iTUG, has the potential to improve clinical decision-making and optimize rehabilitation interventions.

Acknowledgments

The authors would like to express their sincere gratitude to all the people, patients, and physiotherapists who participated in this study.

Funding

This work was partially supported by the Ricerca Corrente funding scheme of the Ministry of Health, Italy.

Data Availability

The dataset used for this study is not publicly available due to patient privacy concerns but may be made available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: M Godi
Data curation: M Giardini, M Godi
Formal analysis: M Giardini, IA, VAA, M Godi
Investigation: M Giardini, IA, M Godi
Methodology: MP, SM
Writing–original draft: M Giardini, M Godi
Writing–review and editing: M Giardini, IA, VAA, MP, SM, M Godi
All authors participated in discussions, read and approved the final manuscript.

Conflicts of Interest

SM owns a share of mHealth Technologies srl, Monte San Pietro, Italy. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multimedia Appendix 1
The 100 parameters extracted from the instrumented Timed Up and Go (iTUG) test, categorized according to the specific functional phases: sit-to-walk, walking, turning, and turn-to-sit.
[PDF File, 143 KB - rehab_v13i1e82632_app1.pdf]

References

1. Ortega-Bastidas P, Gómez B, Aqueveque P, Luarte-Martínez S, Cano-de-la-Cuerda R. Instrumented Timed Up and Go Test (iTUG)-more than assessing time to predict falls: a systematic review. Sensors (Basel) 2023 Mar 24;23(7):3426. [doi: 10.3390/s23073426] [Medline: 37050485]
2. Podsiadlo D, Richardson S. The Timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
3. Herman T, Giladi N, Hausdorff JM. Properties of the "Timed Up and Go" test: more than meets the eye. Gerontology 2011;57(3):203-210. [doi: 10.1159/000314963] [Medline: 20484884]

XSL•FO
RenderX

4. Nightingale CJ, Mitchell SN, Butterfield SA. Validation of the Timed Up and Go test for assessing balance variables in adults aged 65 and older. J Aging Phys Act 2019 Apr 1;27(2):230-233. [doi: 10.1123/japa.2018-0049] [Medline: 30117359]
5. Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the “Get-Up and Go” test. Arch Phys Med Rehabil 1986 Jun;67(6):387-389. [Medline: 3487300]
6. Caronni A, Picardi M, Scarano S, et al. Pay attention: you can fall! The Mini-BesTest scale and the turning duration of the TUG test provide valid balance measures in neurological patients: a prospective study with falls as the balance criterion. Front Neurol 2023;14:1228302. [doi: 10.3389/fneur.2023.1228302] [Medline: 37745667]
7. McDonough CM, Harris-Hayes M, Kristensen MT, et al. Physical therapy management of older adults with hip fracture. J Orthop Sports Phys Ther 2021 Feb;51(2):CPG1-CPG81. [doi: 10.2519/jospt.2021.0301] [Medline: 33522384]
8. Wald P, Chocano-Bedoya PO, Meyer U, et al. Comparative effectiveness of functional tests in fall prediction after Hip Fracture. J Am Med Dir Assoc 2020 Sep;21(9):1327-1330. [doi: 10.1016/j.jamda.2020.02.008] [Medline: 32276783]
9. Oosting E, Kapitein PJC, de Vries SV, Breedveld E. Predicting short stay total hip arthroplasty by use of the Timed Up and Go-test. BMC Musculoskelet Disord 2021 Apr 16;22(1):361. [doi: 10.1186/s12891-021-04240-6] [Medline: 33863323]
10. Givens DL, Eskildsen S, Taylor KE, Faldowski RA, Del Gaizo DJ. Timed Up and Go test is predictive of patient-reported outcomes measurement information system physical function in patients awaiting total knee arthroplasty. Arthroplast Today 2018 Dec;4(4):505-509. [doi: 10.1016/j.artd.2018.07.010] [Medline: 30560183]
11. van Lummel RC, Walgaard S, Hobert MA, et al. Intra-rater, inter-rater and test-retest reliability of an instrumented Timed Up and Go (iTUG) test in patients with Parkinson’s disease. PLoS ONE 2016;11(3):e0151881. [doi: 10.1371/journal.pone.0151881] [Medline: 26999051]
12. Williams JM, Nyman SR. Association between the instrumented Timed Up and Go test and cognitive function, fear of falling and quality of life in community dwelling people with dementia. J Frailty Sarcopenia Falls 2018 Dec;3(4):185-193. [doi: 10.22540/JFSF-03-185] [Medline: 32300707]
13. Tan D, Pua YH, Balakrishnan S, et al. Automated analysis of gait and modified timed up and go using the Microsoft Kinect in people with Parkinson’s disease: associations with physical outcome measures. Med Biol Eng Comput 2019 Feb;57(2):369-377. [doi: 10.1007/s11517-018-1868-2] [Medline: 30123947]
14. Mangano GRA, Valle MS, Casabona A, Vagnini A, Cioni M. Age-related changes in mobility evaluated by the Timed Up and Go Test instrumented through a single sensor. Sensors (Basel) 2020 Jan 28;20(3):719. [doi: 10.3390/s20030719] [Medline: 32012930]
15. Bergquist R, Nerz C, Taraldsen K, et al. Predicting advanced balance ability and mobility with an instrumented Timed Up and Go Test. Sensors (Basel) 2020 Sep 3;20(17):4987. [doi: 10.3390/s20174987] [Medline: 32899143]
16. Mellone S, Tacconi C, Chiari L. Validity of a smartphone-based instrumented Timed Up and Go. Gait Posture 2012 May;36(1):163-165. [doi: 10.1016/j.gaitpost.2012.02.006] [Medline: 22421189]
17. Picardi M, Redaelli V, Antoniotti P, et al. Turning and sit-to-walk measures from the instrumented Timed Up and Go test return valid and responsive measures of dynamic balance in Parkinson’s disease. Clin Biomech (Bristol) 2020 Dec;80:105177. [doi: 10.1016/j.clinbiomech.2020.105177] [Medline: 32979787]
18. Caronni A, Sterpi I, Antoniotti P, et al. Criterion validity of the Instrumented Timed Up and Go test: a partial least square regression study. Gait Posture 2018 Mar;61:287-293. [doi: 10.1016/j.gaitpost.2018.01.015] [Medline: 29413799]
19. Caronni A, Picardi M, Aristidou E, et al. How do patients improve their Timed Up and Go test? Responsiveness to rehabilitation of the TUG test in elderly neurological patients. Gait Posture 2019 May;70:33-38. [doi: 10.1016/j.gaitpost.2019.02.010] [Medline: 30802642]
20. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice, 2nd edition: Prentice Hall; 2000.
21. McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires, 3rd edition: Oxford University Press; 2006.
22. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007 Jan;60(1):34-42. [doi: 10.1016/j.jclinepi.2006.03.012] [Medline: 17161752]
23. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Prac Assess Res Eval 2005;10(1):7. [doi: 10.7275/jyj1-4868]
24. Spicer J. Making Sense of Multivariate Data Analysis: An Intuitive Approach: SAGE Publications Inc; 2005. URL: https://books.google.co.in/books/about/Making_Sense_of_Multivariate_Data_Analys.html?id=ao_8EaoUuSwC&redir_esc=y [accessed 2026-03-28]
25. Matsunaga M. How to factor-analyze your data right: do’s, don’ts, and how-to’s. Int J psychol res 2010;3(1):97-110. [doi: 10.21500/20112084.854]
26. Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL. Practical considerations for using exploratory factor analysis in educational research. Prac Assess Res Eval 2013;18(1):6. [doi: 10.7275/qv2q-rk76]
27. Can S, van de Schoot R, Hox J. Collinear latent variables in multilevel confirmatory factor analysis: a comparison of maximum likelihood and Bayesian estimations. Educ Psychol Meas 2015 Jun;75(3):406-427. [doi: 10.1177/0013164414547959] [Medline: 29795827]
28. Williams B, Onsman A, Brown T. Exploratory factor analysis: a five-step guide for novices. Australas J Paramed 2010 Jan;8(3):1-13. [doi: 10.33151/ajp.8.3.93]

XSL•FO
RenderX

29. Gaskin CJ, Happell B. On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. Int J Nurs Stud 2014 Mar;51(3):511-521. [doi: 10.1016/j.ijnurstu.2013.10.005] [Medline: 24183474]
30. Cattell RB. The scree test for the number of factors. Multivariate Behav Res 1966 Apr 1;1(2):245-276. [doi: 10.1207/s15327906mbr0102_10] [Medline: 26828106]
31. Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika 1965 Jun;30:179-185. [doi: 10.1007/BF02289447] [Medline: 14306381]
32. Astephen JL, Deluzio KJ. Changes in frontal plane dynamics and the loading response phase of the gait cycle are characteristic of severe knee osteoarthritis application of a multidimensional analysis technique. Clin Biomech (Bristol) 2005 Feb;20(2):209-217. [doi: 10.1016/j.clinbiomech.2004.09.007] [Medline: 15621327]
33. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. Stat Med 2002 May 15;21(9):1331-1335. [doi: 10.1002/sim.1108] [Medline: 12111881]
34. Coni A, Mellone S, Colpo M, Bandinelli S, Chiari L. A factor analysis model of the instrumented Timed Up and Go test for physical capability assessment. Gait Posture 2018 Oct;66:S11-S12. [doi: 10.1016/j.gaitpost.2018.07.117]
35. Baracco S, Arcolin I, Corna S, Godi M, Giardini M. Instrumented Timed Up and Go test: a reliable tool for elderly with femur fracture. Gait Posture 2024 Oct;114:S4-S5. [doi: 10.1016/j.gaitpost.2024.08.018]
36. Wüest S, Massé F, Aminian K, Gonzenbach R, de Bruin ED. Reliability and validity of the inertial sensor-based Timed “Up and Go” test in individuals affected by Stroke. J Rehabil Res Dev 2016;53(5):599-610. [doi: 10.1682/JRRD.2015.04.0065] [Medline: 27898161]
37. Craig JJ, Bruetsch AP, Lynch SG, Horak FB, Huisinga JM. Instrumented balance and walking assessments in persons with multiple sclerosis show strong test-retest reliability. J Neuroeng Rehabil 2017 May 22;14(1):43. [doi: 10.1186/s12984-017-0251-0] [Medline: 28532417]
38. Moreno-Verdú M, Ferreira-Sánchez MDR, Martín-Casas P, Atín-Arratibel M. Imagined Timed Up and Go test (iTUG) in people with Parkinson’s disease: test-retest reliability and validity. Disabil Rehabil 2023 Mar 8;2023:1-11. [doi: 10.1080/09638288.2023.2185688] [Medline: 36890615]
39. Zhou J, Yao Q, Han R, et al. Reliability and validity of Instrumented Timed Up and Go test in typical adults and elderly: a systematic review. Arch Phys Med Rehabil 2025 Jul;106(7):1092-1107. [doi: 10.1016/j.apmr.2025.03.001] [Medline: 40054550]
40. Weiss A, Herman T, Plotnik M, Brozgol M, Giladi N, Hausdorff JM. An instrumented Timed Up and Go: the added value of an accelerometer for identifying fall risk in Idiopathic Fallers. Physiol Meas 2011 Dec;32(12):2003-2018. [doi: 10.1088/0967-3334/32/12/009] [Medline: 22094550]
41. Greene BR, Rutledge S, McGurgan I, et al. Assessment and classification of early-stage multiple sclerosis with inertial sensors: comparison against clinical measures of disease state. IEEE J Biomed Health Inform 2015 Jul;19(4):1356-1361. [doi: 10.1109/JBHI.2015.2435057] [Medline: 26087505]
42. Riva F, Bisi MC, Stagni R. Gait variability and stability measures: minimum number of strides and within-session reliability. Comput Biol Med 2014 Jul;50:9-13. [doi: 10.1016/j.compbiomed.2014.04.001] [Medline: 24792493]
43. Pasciuto I, Bergamini E, Iosa M, Vannozzi G, Cappozzo A. Overcoming the limitations of the harmonic ratio for the reliable assessment of gait symmetry. J Biomech 2017 Feb 28;53:84-89. [doi: 10.1016/j.jbiomech.2017.01.005] [Medline: 28104246]
44. Bateni H, Maki BE. Assistive devices for balance and mobility: benefits, demands, and adverse consequences. Arch Phys Med Rehabil 2005 Jan;86(1):134-145. [doi: 10.1016/j.apmr.2004.04.023] [Medline: 15641004]
45. Mobbs RJ, Perring J, Raj SM, et al. Gait metrics analysis utilizing single-point inertial measurement units: a systematic review. Mhealth 2022;8:9. [doi: 10.21037/mhealth-21-17] [Medline: 35178440]
46. Seo K, Jung J, Hwang J, Kim K, Kim SH. Assessing the impact of imu sensor location on spatio-temporal gait parameter estimation. Annu Int Conf IEEE Eng Med Biol Soc 2024 Jul;2024:1-4. [doi: 10.1109/EMBC53108.2024.10781929] [Medline: 40039154]
47. Küderle A, Roth N, Zlatanovic J, Zrenner M, Eskofier B, Kluge F. The placement of foot-mounted IMU sensors does affect the accuracy of spatial parameters during regular walking. PLoS ONE 2022;17(6):e0269567. [doi: 10.1371/journal.pone.0269567] [Medline: 35679231]
48. Höglund G, Grip H, Öhberg F. The importance of inertial measurement unit placement in assessing upper limb motion. Med Eng Phys 2021 Jun;92:1-9. [doi: 10.1016/j.medengphy.2021.03.010] [Medline: 34167702]
49. Böttinger MJ, Labudek S, Schoene D, et al. “TiC-TUG”: Technology in clinical practice using the instrumented Timed Up and Go test—a scoping review. Aging Clin Exp Res 2024 Apr 27;36(1):100. [doi: 10.1007/s40520-024-02733-7] [Medline: 38676844]
50. Bloomfield RA. In-Clinic Functional Measurement and Analysis of Knee Osteoarthritis Patients Undergoing Total Knee Replacement [Doctoral Dissertation]. : The University of Western Ontario (Canada); 2021 URL: https://ir.lib.uwo.ca/etd/7808/ [accessed 2026-03-18]
51. Gasparutto X, Gueugnon M, Laroche D, Martz P, Hannouche D, Armand S. Which functional tasks present the largest deficits for patients with Total Hip Arthroplasty before and six months after surgery? A study of the Timed Up-And-Go Test phases. PLoS ONE 2021;16(9):e0255037. [doi: 10.1371/journal.pone.0255037] [Medline: 34506498]

XSL•FO
RenderX

52. Overgaard JA, Kristensen MT. Minimal clinical important difference and responsiveness of performance-based and self-reported measures used in older adults following outpatient rehabilitation programme Post-Hip Fracture Surgery. Geriatr Orthop Surg Rehabil 2021:37. [doi: 10.1177/21514593211058764]
53. Santos ACAD, Matheus D, Battistella LR. Reference values for instrumented Timed Up and Go—outcomes of vertical accelerations and angular velocities in healthy subjects. J Int Soc Phys Rehabil Med 2025;8(1):1-5. [doi: 10.1097/ph9.0000000000000050]
54. He J, Wu L, Du W, et al. Instrumented Timed Up and Go test and machine learning–based levodopa response evaluation: a pilot study. J Neuroeng Rehabil 2024 Sep 18;21(1):163. [doi: 10.1186/s12984-024-01452-4] [Medline: 39294708]

Abbreviations

AUC: area under the curve
EFA: exploratory factor analysis
ICC: intraclass correlation coefficient
IMU: inertial measurement unit
iTUG: instrumented Timed Up and Go
KMO: Kaiser-Meyer-Olkin
LDA: linear discriminant analysis
MDC: minimum detectable change
PCA: principal component analysis
ROC: receiver operating characteristic
SEM: SE of measurement
THA: total hip arthroplasty
TKA: total knee arthroplasty
TUG: Timed Up and Go

Edited by S Munce; submitted 19.Aug.2025; peer-reviewed by M Tramontano, O Vickers, S Erel; accepted 26.Jan.2026; published 01.Apr.2026.

Please cite as:
Giardini M, Arcolin I, Arcobelli VA, Picardi M, Mellone S, Godi M
Reliability and Discriminant Ability of an Instrumented Timed Up and Go Test in People With Postsurgical Orthopedic Conditions: Quantitative Study
JMIR Rehabil Assist Technol 2026;13:e82632
URL: https://rehab.jmir.org/2026/1/e82632
doi:10.2196/82632

© Marica Giardini, Ilaria Arcolin, Valerio Antonio Arcobelli, Michela Picardi, Sabato Mellone, Marco Godi. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 1.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using a Co-Designed Digital Self-Management Program to Prepare Patients for Hip or Knee Replacement Surgery: Pragmatic Pilot Study

Elizabeth Horton[1], PhD; Hayley Wright[1], PhD; Andy Turner[1], PhD; Louise Moody[2], PhD; Lucy Aphramor[3], PhD; Anna Carlson[4], PhD; Hesam Ghiasvand[5], PhD; Shea Palmer[6], PhD

[1]Centre for Intelligent Healthcare, Coventry University, Priory Street, Coventry, United Kingdom
[2]Centre for Arts, Memory and Communities, Coventry University, Coventry, United Kingdom
[3]Centre for Agroecology, Water and Resilience, Coventry University, Coventry, United Kingdom
[4]Hope For The Community Community Interest Company, Enterprise Hub, Coventry, United Kingdom
[5]Centre for Healthcare and Communities, Coventry University, Coventry, United Kingdom
[6]School of Healthcare Sciences, Cardiff University, Cardiff, United Kingdom

Corresponding Author:
Elizabeth Horton, PhD
Centre for Intelligent Healthcare, Coventry University, Priory Street, Coventry, United Kingdom

Abstract

Background: The aging population has resulted in more people living longer with musculoskeletal conditions who require hip and knee replacement surgery. Lengthening waiting lists continue to be a challenge for patients and health care services.

Objective: This pragmatic study aimed to develop and test a digital self-management intervention (the HOPE [Help Overcome Problems Effectively] program) to better prepare patients waiting for hip and knee replacement surgery.

Methods: The study used a pragmatic, pre-post with follow-up, single-arm design. All intervention and data collection components were delivered online. Patients were recruited from those on the waiting list for hip or knee surgery. Following iterative co-development of the intervention, the content was refined and optimized into a final version for testing. The resulting program was an 8-week intervention delivered via the HOPE 4 The Community (H4C) digital platform (powered by H4C). Data were collected at baseline (pre-HOPE program), 8 weeks (post-HOPE program), and 6-month follow-up. Patient-reported outcome measures related to preparation for surgery, quality of life, physical function, pain, mental well-being, self-efficacy, and physical activity. Resource usage data were collected to calculate health and social care costs. System Usability Scale data were collected post-HOPE program.

Results: One hundred participants enrolled in the HOPE program. Of these, 57 (57%) consented to take part in the evaluation and returned the baseline questionnaire. Thirty-nine participants completed ≥5 of the 8 sessions and all surveys. Among the 25 participants who had surgery at 6 months, 23 (92%) felt better prepared due to the HOPE program. Median improvements in most outcomes were observed at 8 weeks, with several continuing to improve at 6 months. The Friedman test showed significant improvements over 6 months in self-efficacy (pain: *P*=.002; other symptoms: *P*<.01), pain (*P*=.04), health status (*P*=.02), and mental well-being (*P*=.01). No significant changes were noted in physical activity. While the early cost analysis did not reach statistical significance, it indicated potential cost savings from reduced patient interactions with health care professionals. Sixty-four percent (25/39) of participants had surgery, and this likely contributed in part to improvements in outcomes. System usability was rated above average (mean score 70.1, SD 15.9).

Conclusions: The results are promising in relation to participants attending the HOPE program feeling better prepared for surgery. A fully powered efficacy and cost-effectiveness trial is needed to determine the contribution of the HOPE program to outcomes, over and above the contribution of surgery.

(JMIR Rehabil Assist Technol 2026;13:e68286) doi:10.2196/68286

KEYWORDS

prehabilitation; self-management; peer support; digital platform; hip and knee arthroplasty; musculoskeletal

XSL•FO
RenderX

Introduction

In the United Kingdom, an estimated 20.3 million people are affected by musculoskeletal conditions. These conditions account for 21% of the years lived with illness and disability [1]. The global prevalence of osteoarthritis is increasing, and if the trend continues, osteoarthritis will become one of the most prevalent diseases in populations from high-income countries in the coming decades [2]. The aging UK population is living longer with complex musculoskeletal conditions and comorbidities, causing increased demand on National Health Service (NHS) health and social care services [1], accounting for up to 30% of general practice consultations in England [3].

Lengthening waiting lists are particularly problematic in musculoskeletal medicine. A 2019 report found that in England alone, 726,000 people had severe hip osteoarthritis and 1.4 million people had severe knee osteoarthritis [4]. For those whose condition is severe, joint replacement surgery is the only option to alleviate pain and improve mobility and the ability to self-manage. Under the NHS constitution, 92% of patients should be treated within 18 weeks as part of the referral-to-treatment scheme. However, in 2019, nearly 4000 patients had been waiting for over 2 years for surgery [4], and more than 690,000 were on waiting lists in 2021 [5]. The COVID-19 pandemic had an unprecedented impact on secondary care orthopedic services, with a significant increase in waiting times for the majority of patients [5]. While on the waiting list, patients are likely to experience worsening pain, reduced mobility, increased anxiety, and deteriorating health, leading to greater demand for health and care services. In recognition of wait times, Versus Arthritis [5] and Arthritis Action [6] offer resources for self-management on their websites. By 2060, it is projected that the demand for hip and knee joint replacements in the United Kingdom will rise by nearly 40% from current levels, which will have significant implications for the health care system [7].

New ways of working are needed to optimize support for patients, maximize capacity, and mitigate risk. It is also important to address inequities: the COVID-19 pandemic foregrounded deep-rooted equality, diversity, and inclusion issues in relation to morbidity and mortality that are entangled with access to health care services. Inequities in treatment waiting time [8] for musculoskeletal services and in treatment outcomes [9] reflect this general picture and highlight the need for action. There is a need for holistic support among those waiting for hip and knee surgery in England. The NHS personalized care team recommends that patients on the waiting list should receive self-management support to "wait well" by undertaking prehabilitation. This support should empower patients through information, health coaching, and digital resources [10].

Prehabilitation is an effective way of improving perioperative outcomes through support to increase physical and mental resilience for surgery. Systematic reviews and meta-analyses have found some, generally low-quality evidence that prehabilitation improves a range of postoperative outcomes for patients undergoing hip and knee surgery, including function, pain, strength, and quality of life [11-13]. A more recent systematic review and meta-analysis specifically focused on the effects of digital prehabilitation in a range of musculoskeletal conditions awaiting surgery, including knee and hip replacements [14]. They found evidence that advanced technologies supported greater improvements in function pre- and post-operatively than standard care for knee and hip replacements. Greater improvements were also seen in preoperative pain, preoperative risk of falling, and postoperative stiffness. There was no evidence for spinal surgery or other conditions. However, few orthopedic prehabilitation interventions are digitally delivered, nor do they provide peer or emotional support, which is highly valued by many patients living with long-term conditions [11-13]. Indeed, a recent survey conducted in the United Kingdom [15] found that, although the vast majority of hospitals (97%) offered preoperative education, only 59% and 48% offered prehabilitation for knee and hip arthroplasty, respectively. Education was mainly delivered as a single talk supported by a booklet, and prehabilitation mainly as strengthening exercise, advice, and written information. Reported barriers included lack of facilities, funding, and staff. There was also a reported lack of robust evidence to support practice [15]. Across various surgical specialties, multimodal prehabilitation includes nutrition and psychological support alongside exercise training. There is some evidence of psychological factors improving postsurgical outcomes [13]. A systematic review and meta-analysis found low-quality evidence that psychological interventions have a positive effect on postsurgical anxiety and on mental components of quality of life [16].

In a review [17] of over 30 prehabilitation surgery schools in the United Kingdom and Ireland (these schools inform patients about what to expect and guide them on how to prepare physically and mentally to reduce postoperative risks of surgery), only 40% contained content to manage emotional well-being, and only 13% used digital apps. Further, many interventions were not underpinned by behavior change theory and techniques.

In 2022, Coventry University and its university spin-out social enterprise, H4C (HOPE [Help Overcome Problems Effectively] for The Community) interest company, developed a proof-of-concept digital intervention, called the Help Overcome Problems Effectively (HOPE) program, to help patients prepare for hip and knee surgery. The HOPE program for hip and knee patients shares the same underlying theoretical framework as other HOPE programs for long-term conditions offered by H4C, which have been taxonomized using the taxonomy of self-management support [18] and are described in detail in published papers [19-21]. All 14 digital versions of the HOPE program have been approved by the Quality Institute for Self-Management Education and Training for the provision of self-management structured education (QIS2020 and QIS2023 [22]) and certified by the Organisation for the Review of Care and Health Apps (ORCHA [23]), scoring 88% for Android and iOS (Apple Inc), and 86% for WebApp, indicating compliance with best practice in data security, professional assurance, usability, and accessibility.

XSL•FO
RenderX

The HOPE program for hip and knee patients combines evidence-based self-management content with a validated exercise program, incorporating a home exercise component tailored to individual needs and abilities, drawing from the work of Ageberg et al [24].

In 2023 H4C was awarded funding through the UK Research and Innovation Healthy Ageing Challenge Scaling Social Ventures competition to co-design and evaluate the HOPE program for hip and knee surgery patients. The funding competition was to support social enterprises in scaling products and services to support healthy aging and deliver social value.

The pragmatic, multimethod study aimed to optimize and evaluate the HOPE program to determine whether patients were better prepared for surgery. The study objectives included optimizing the HOPE program through co-design with stakeholders, implementing and testing the program with patients waiting for a joint replacement, and assessing their preparedness for surgery.

Methods

Study Design

This study used a co-design phase followed by a pragmatic, pre-post, with follow-up, single-arm intervention study. All intervention components and data collection were delivered online. This study is reported according to the CONSORT (Consolidated Standards of Reporting Trials) 2016 statement: extension for nonrandomized pilot trials [25]. CHERRIES (Checklist for Reporting Results of Internet E-Surveys) was used to guide the survey report [26]. All intervention and data collection activities took place online. All study data were collected online via questionnaires administered through Qualtrics Survey Software (Qualtrics).

Co-Design Phase to Optimize the HOPE Program

Ten participants took part in the development activities, which included 3 online workshops. One workshop was undertaken with 6 patient participants waiting for a hip (n=3) or knee replacement surgery (n=3), who had completed an earlier proof-of-concept HOPE program (5 female participants, aged 60 - 80 years). The purpose of the workshop was to explore their experiences and generate feedback on the HOPE program.

Two health professional workshops involving 4 NHS staff from our partner organizations were held to discuss referral pathways and useful resources for patients awaiting surgery. The roles of the professionals were Elective Recovery Lead, Team Lead Physiotherapist in Elective Orthopedics, Project Manager of a Musculoskeletal Clinical Program, and Senior Primary and Community Care Lead. Workshops and interviews were conducted online via Zoom (Zoom Video Communications, Inc) and MS Teams (Microsoft Corp) to allow for geographically dispersed participation.

Development of the Exercise Program

The exercise program central to the intervention was based on the The Neuromuscular Exercise training program for patients with knee or hip osteoarthritis assigned for total joint replacement the neuromuscular exercise training program for patients with knee or hip osteoarthritis assigned for total joint replacement program [24,27], which was specifically developed for older patients with severe knee and hip osteoarthritis before having total joint replacement surgery. Only the exercises from the neuromuscular exercise training program for patients with knee or hip osteoarthritis assigned for total joint replacement program were adopted within the HOPE program. Those exercises have also been incorporated into the Good Life with osteoArthritis: Denmark program [28-30]. The exercises have been demonstrated to be safe, patients can successfully progress them, and they contribute to improvements in a range of outcomes, including symptoms, function, medication use, and sick leave. A range of video and visual resources had previously been developed to support the exercise components [31]. Following feedback from the co-design phase, new video resources were developed to illustrate how the exercises could be adapted within the home environment. Forty-three videos were filmed in a home setting (living room, bedroom, and kitchen), using home furniture (sofa, chair, and bed) and both exercise equipment and everyday household items as exercise props, with volunteers representing different ages and genders, and incorporating visual prompts and voiceover instructions. The exercises target major lower limb muscle groups and can be adapted to individual capabilities, with 3 difficulty levels and encouragement to alter repetitions and sets. Participants could build their own home-based exercise program by answering 6 questions about their ability (eg, if they can easily get on and off the floor) and equipment (eg, if they have a step they can use at home). An algorithm was then built to create their personalized exercise program from the 43 videos. Participants progressed up and down levels of difficulty at their own pace, monitored progress, and set exercise reminders. Participants could download their exercise record in PDF format to keep or share with a health care professional. Tips on creating a safe exercise space, as well as important information to mitigate any worries or injuries, were part of the program.

The HOPE program: Intervention Content and Structure

The resulting program comprised 8 modules and was hosted on H4C's digital platform, powered by H4C. The content comprised text, images, videos, downloadable documents, interactive activities (eg, quizzes, self-monitoring tools, and diaries), and discussion forums and messaging facilities. The digital content was released at set times over the 8 weeks but could be accessed at any time (asynchronous). Participants had the option to "fast-track" the content if they were due to have surgery during the 8 weeks.

Once accessed and viewed, the app content could be viewed offline, reducing the need for a data plan or high-quality internet connection. An analog print booklet was produced, containing the same content as the digital version of the HOPE program, for those who were digitally excluded and/or experienced low digital literacy.

XSL•FO
RenderX

Pre-Post With Follow-Up Study

Participants

Broad eligibility criteria were used to ensure the study was as inclusive as possible and to provide ample opportunity for participation. Individuals were eligible if they were adults aged 18 years or older, lived in the South West of England in the United Kingdom, were currently on a waiting list for hip or knee replacement surgery, had access to the internet and a suitable device to engage with the intervention, and were able to interact with all materials provided as part of the intervention.

Patients interested in attending the HOPE program were referred to the study sign-up webpage through several routes. NHS South West referral sources included secondary care, primary care, and musculoskeletal clinics. Eligible participants were referred directly to H4C to enroll in the HOPE program and given the option to take part in the research study. Patients who chose to take part in the study were directed to the participant information sheet and consent form in Qualtrics Survey Software. Patients were informed that participation was voluntary and that their decision would not affect their quality of care.

We collected the following sociodemographic information: name, email address, gender, age, postcode, occupation and employment, and some details about their emotional health and their illness diagnosis, level of physical activity, health care visits, time on the waiting list, and date of surgery. Postcode data were used to calculate the English index of multiple deprivation (IMD [32]). IMD is an official measure of deprivation ranked from 1 (most deprived) to 10 (least deprived).

The questionnaire was administered through Qualtrics, using responsive and mobile-ready question formats. Adaptive questioning was used to conditionally display questions based on previous responses to reduce the number and complexity of the questions. Most pages contained between 1 and 6 items. Excluding the introduction, participant information sheet, and consent form, the survey was distributed over 14 pages. The responses were made mandatory to avoid missing data. The survey was not set up to allow participants to change their responses. The procedure, as outlined in the participant information sheet and survey structure, involved collecting identifiable information at registration—specifically, name and email address (rather than via technical means such as cookies or IP addresses)—which was then used by the research team to ensure each individual only completed the survey once per time point. Pre-HOPE program (baseline) questionnaires were completed during the period of July 6-13, 2023, for the first HOPE program and July 20-31, 2023, for the second HOPE program. Participants received a £60 (approximately US $80) electronic gift voucher for completion of all pre- and postprogram questionnaires. Participants were informed in the Patient Information Sheet how their data would be processed in accordance with the Data Protection Act 2018. Participation in the study was optional for patients who accessed the HOPE program.

The HOPE Program: Accessing and Completing the Program

Following completion of the pre-HOPE program survey, participants were given access to the HOPE program (start dates: July 13 or 27, 2023) through a personalized log-in link.

Throughout the program, participants were supported by 2 facilitators who were trained in line with Quality Institute for Self-Management Education and Training standards. The program content was organized into themed sessions across the 8 weeks, with an integrated tailored exercise program (described in the "Development of the Exercise Program" section above; Table 1 lists session content; refer to Multimedia Appendix 1 for a brief description of each session and screenshots of the intervention).

Table . Session content of the HOPE[a] program.

Session	Session content
1	Instilling HOPE
2	Managing pain and fatigue
3	Stress and shifting your thinking
4	Communication
5	Sleep and mindfulness
6	Setbacks and hospital stay
7	Happiness and strengths
8	Moving on with HOPE

[a] HOPE: Help Overcome Problems Effectively.

Patient-Reported Outcome Measures

Surgery Preparation

At 6-month follow-up, participants were asked if they felt better prepared for surgery using the following question from the Patient Preparedness for Surgery questionnaire [33]: "Overall, I feel or felt (if I had surgery) prepared for my upcoming surgery." There were 6 response options: strongly agree, agree, somewhat agree, somewhat disagree, disagree, and strongly disagree. Participants were also asked to provide reasons for their answers. Those who had surgery indicated whether they felt the HOPE program helped them prepare before surgery, after surgery, or both. Participants provided textual responses

XSL•FO
RenderX

to explain why they agreed or disagreed that the HOPE program helped them prepare for surgery.

The following validated patient-reported outcome measures (PROMs) were collected at baseline, post-HOPE program (8 weeks), and 6-month follow-up via Qualtrics.

Short Warwick-Edinburgh Mental Well-Being Scale

The Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS [34]) is a short version that assesses mental well-being within the adult population. The SWEMWBS uses 7 items from the full WEMWBS [35], which relate more to mental functioning than feelings. The 7 statements are positively worded, with 5 response categories ranging from “none of the time” to “all of the time.” Total scores range from 7 to 35, with higher scores indicating higher mental well-being. A change of one point or more on the SWEMWBS total score represents a minimally important level of change.

The EQ-5D Index and EQ-Visual Analogue Scale

The EQ-5D index [36] and the EQ-Visual Analogue Scale (EQ-VAS) are widely used measures of health status and health-related quality of life, respectively. The EQ-5D index assesses patients’ health state across 5 dimensions (self-care, mobility, anxiety and depression, usual activities, and pain and discomfort) that are weighted to provide a utility value based on a population tariff. Scores range from 0 (death) to 100 (perfect health). The EQ-VAS is a vertical rating scale for health, scored between 0 (worst imaginable health) and 100 (best imaginable health).

Western Ontario and McMaster Universities Arthritis Index

The Western Ontario and McMaster Universities Arthritis Index (WOMAC [37]) consists of 24 items divided into 3 subscales: Pain (5 items), Stiffness (2 items), and Physical Function (17 items). Items are scored on a scale of 0 - 4, which corresponds to: None (0), Mild (1), Moderate (2), Severe (3), and Extreme (4). The scores for each subscale are summed, with possible score ranges of 0 - 20 for Pain, 0 - 8 for Stiffness, and 0 - 68 for Physical Function. A sum of the scores for all 3 subscales gives a total WOMAC score (maximum 96). Higher scores indicate worse pain, stiffness, and functional limitations.

Arthritis Self-Efficacy Scale

The Arthritis Self-Efficacy Scale (ASES [38]) measures a person’s confidence to self-manage their arthritis symptoms and consists of 2 subscales: Pain (5 items) and Other Symptoms (6 items). Items are scored from 1 (very uncertain) to 10 (very certain). The scores for each subscale are summed, with a possible score range of 10 - 50 for Pain and 10 - 60 for Other Symptoms. Higher scores indicate higher self-efficacy.

International Physical Activity Questionnaire–Short Form

The International Physical Activity Questionnaire–Short Form (IPAQ-SF [39]) assesses physical activity undertaken across a comprehensive set of domains including: (1) leisure-time physical activity, (2) domestic and gardening (yard) activities, (3) work-related physical activity, and (4) transport-related physical activity. The items are structured to provide separate scores on walking, moderate-intensity, and vigorous-intensity activity, as well as a combined total score to describe the overall level of activity. Computation of the total score requires summation of the duration (in minutes) and frequency (days) of walking, moderate-intensity, and vigorous-intensity activity. The IPAQ-SF scoring protocol assigns the following metabolic equivalent of task (MET) values to walking, moderate, and vigorous-intensity activity: 3.3 METs, 4.0 METs, and 8.0 METs, respectively. Participants are considered to have met Centers for Disease Control and Prevention and American College of Sports Medicine physical activity recommendations if they reported at least 150 minutes per week of walking, moderate, or vigorous intensity physical activity.

Numerical Pain Rating Scale

The Numerical Rating Scale (NPRS)-11 [40] is an 11-point scale for self-report of pain. It is the most commonly used unidimensional pain scale. The respondent selects a whole number (integers 0 - 10) that best reflects the intensity (or other quality, if requested) of their pain. The anchors are 0=no pain and 10=worst possible pain (there are various wordings of the upper anchor).

HOPE Program Usability and User Engagement

Usability

The usability of the system was assessed by the System Usability Scale (SUS [41]), which was embedded in the last session of the HOPE program. It was optional for participants to complete. The SUS uses a 5-point Likert scale ranging from 1=strongly disagree to 5=strongly agree across 10 items. Odd-numbered questions (1, 3, 5, 7, and 9) generate a positive response. Even-numbered questions (2, 4, 6, 8, and 10) generate a negative response, which must be inverted. All the points added up together could gain a maximum of 40, thus the multiplication by 2.5 to make the scale out of 100. A total score of ≥68 is considered above-average usability.

User Engagement

The intervention platform collected user engagement data. For this study, we report the number of sessions completed, the number of participants who used the personalized exercise program, and the most commonly bookmarked content or activities.

Sample Size

This pragmatic study enrolled participants from an opportunity sample (n=100) comprising eligible candidates. Potential participants received an email containing a link to the study website hosted by Qualtrics. Here, participants were required to review the digital Participant Information Sheet, provide digital consent, and complete the digital questionnaires.

Analytical Methods

Data relating to sociodemographic characteristics and outcome measures were collated and presented descriptively at the group level. Outcome data were mostly ordinal and nonnormally distributed, so descriptive data were limited to frequencies (and proportions) and medians (and IQRs). While the study was not powered to detect statistically significant changes in outcomes between time points, nonparametric Friedman and post hoc Wilcoxon signed-rank tests were used to explore changes over

XSL•FO
RenderX

time between baseline, post-HOPE program (8 weeks), and 6-month follow-up. All analyses were performed using IBM SPSS (version 28). The level of statistical significance was set at *P*<.05. Textual responses to the question about surgery preparedness at the 6-month follow-up survey were summarized to illustrate the quantitative findings.

Given this was a feasibility study with complete-case analysis as the prespecified approach, we focused on participants who engaged with ≥5 sessions and completed follow-ups. This decision was made because (1) the primary aim was assessing intervention feasibility and acceptability under optimal conditions, (2) minimal data were available from noncompleters (only 4/15 provided follow-ups), and (3) high follow-up rates among completers (98% at 8 weeks and 93% at 6 months) reduced concerns about attrition bias. Future efficacy trials will use intention-to-treat (ITT) analysis.

Resource Usage

An early cost-impact analysis evaluated the change in costs associated with patients' appointments and visits with NHS England to understand the potential cost impact of the program and assess whether it could be expanded into a broader study. These data were captured via the Qualtrics survey at baseline, post-HOPE program (8 weeks), and 6-month follow-up.

The economic analysis focused on changes in the number of interactions patients had with NHS health and social care staff, measured by appointments and visits. A decision model was developed using parameters from a before-and-after analysis, literature review, and incorporating assumptions. The mean values, associated SEs, and assumptions populated the model, detailed in Multimedia Appendix 2. The total cost impact was calculated from the NHS personal and social care perspective, both per patient and per patient per week.

Costs associated with interaction changes were evaluated at 8 weeks and 6 months compared to baseline using unit costs from the Unit Costs of Health and Social Care report by the Personal Social Services Research Unit at the University of Kent [42] and the NHS National Tariff [43]. A probabilistic sensitivity analysis explored uncertainty around the results.

Ethical Considerations

The user requirements research undertaken by Coventry University received ethical approval from the Coventry University Research Ethics Committee (P151751). The research and evaluation activity has also received approval from Coventry University (P106036) and, as an amendment to a preexisting HOPE evaluation, from the Health Research Authority and Health and Care Research Wales (Integrated Research Applications System, project ID 283172).

Results

Co-Design Phase Adaptations

Adaptations to the intervention, as an outcome of patient and health professionals' feedback, were as follows: Adaptations suggested by patients were (1) guidance on how to adjust the exercises to meet individual needs and capabilities; (2) a broader range of additional activities to try, for example, pool-based exercises; (3) reassurance for people who may struggle to keep up with the program; (4) information to challenge misinformation, controversies, and conflicting advice; and (5) clearer guidance on how to access some features, for example, messaging functions.

Adaptations suggested by health professionals were (1) reminders and nudges to make healthy changes and prevent deconditioning, (2) long-term access to information for use postoperatively, (3) adjustment of exercises to cater to different abilities and comorbidities, (4) program certification to demonstrate credibility, and (5) reference to the expert input that informed the program content.

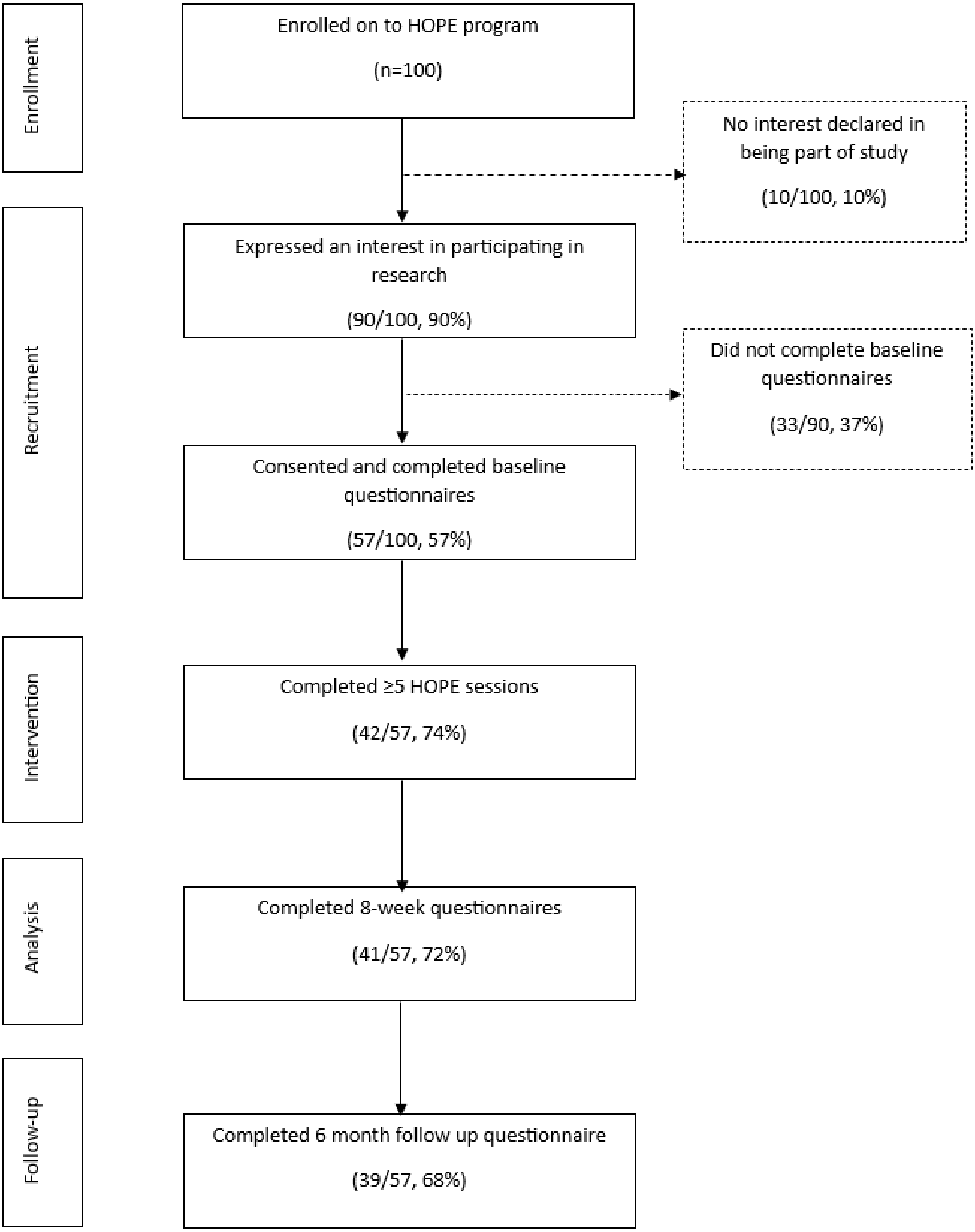
Participants

One hundred participants enrolled in the 2 HOPE programs (HOPE 1: n=59 and HOPE 2: n=41). Of these, 57 (57%) consented to take part in the evaluation and returned the baseline questionnaire (n=39, HOPE program 1 and n=18, HOPE program 2). Forty-one participants returned follow-up questionnaires at 8 weeks (41/57, 71.9%), and 39 participants returned questionnaires at 6-month follow-up (39/57, 68.4%). Forty-two participants (42/57, 73.7%) accessed ≥5 of the 8 sessions and were considered program completers.

Almost all of the HOPE program completers (41/42, 98%) returned follow-up questionnaires at 8 weeks, and 39 (93%) returned questionnaires at 6-month follow-up (Figure 1). HOPE program completers who returned both questionnaires (39/42, 93%) were included in the primary analysis. There was no missing outcome data, as these fields were required during questionnaire completion.

XSL•FO
RenderX

Figure 1. Flow of participants through the study. HOPE: Help Overcome Problems Effectively.



Participant characteristics are presented in Table 2 and are similar in the total sample (n=57) and completers (n=39). All completer participants (n=39) identified as White-English, Welsh, Scottish, Northern Irish, or British ethnicity and described English as their first language (all 39/39, 100%). One-third of participants were male (13/39, 33%) and two-thirds were female (26/39, 67%). Age was only reported by 21 (54%) participants, with a median age of 66.0 (IQR 63.0-69.5) years. The majority of participants were retired (23/39, 59%). A third of participants (13/39, 33%) were listed for hip replacement surgery, and two-thirds (26/39, 67%) for knee replacement surgery. The median IMD was 7.00 (IQR 2.5-13) and the median time on the waiting list for surgery was 6.00 (IQR 2-12) months.

XSL•FO
RenderX

Table . Participant baseline characteristics of completers (n=39) and total sample (N=57).

Characteristic	Completers (n=39)	Total sample (N=57)
Gender, n (%)		
Male	13 (33)	20 (35)
Female	26 (67)	36 (63)
Not specified	0 (0)	1 (2)
Age (years), median (IQR)	66 (63-69.5)	66 (63-69.5)
Ethnicity, n (%)		
White-English, Welsh, Scottish, Northern Irish, or British	39 (100)	56 (98)
Black, African, or Caribbean	0 (0)	1 (2)
Disability, n (%)		
Mental health condition (long-term)	5 (13)	7 (12)
Blind or partially sighted	1 (3)	1 (2)
Hard of hearing or deaf	0 (0)	1 (2)
Long-term illness or health condition (lasting more than 12 months or terminal)	4 (10)	7 (12)
Mobility impairment	24 (62)	32 (56)
Employment, n (%)		
In paid work: full-time	4 (10)	8 (14)
In paid work: part-time	4 (10)	8 (14)
Retired	23 (59)	31 (54)
Not in paid work	8 (21)	10 (18)
Not in paid work due to hip or knee condition?	8 (21)	10 (18)
Index of multiple deprivation, median (IQR)	7 (2.25)	7 (2)
Joint replacement, n (%)		
Hip	13 (33)	22 (39)
Knee	26 (67)	35 (61)
Waiting time (months), median (IQR)	6 (2-12)	7 (2.5-13)

Twenty-four out of 39 (62%) participants considered themselves to have a disability, with 9 (23%) participants reporting that day-to-day activities were “limited a little,” and 15 (39%) reporting that activities were “limited a lot.” Seven (18%) participants reported more than one specific type of disability (refer to Table 2).

User Engagement

Just over half of all participants completed all 8 sessions (30/57, 53%), with 6 participants completing 7 sessions (6/57, 11%), 1 completing 6 sessions (1/57, 2%), and a further 5 participants completing 5 sessions (5/57, 9%). Forty-nine out of 57 (86%) participants used the personalized exercise program. The top 5 bookmarked content or activities were (1) exercise program, (2) relaxed breathing, (3) mindfulness meditation, (4) compassionate approach to pain, and (5) cognitive diffusion activity.

Patient-Reported Outcomes and Estimations

By the time of the 6-month follow-up, 25 out of 39 (64%) participants had already received their surgery. Of those who had their surgery, the majority (23/25; 92%) agreed with the statement: “As a result of attending the HOPE program, overall, I felt better prepared for my surgery.” Eight out of 23 (35%) participants selected “strongly agree,” 10 (44%) selected “agree,” and 5 (22%) selected “somewhat agree.” Of the 23 participants who agreed that they were better prepared, 16 (70%) felt better prepared in the presurgery period, 3 (13%) felt better prepared postsurgery, and 5 (17%) felt better prepared pre- and postsurgery.

Of those who had not yet had surgery, the majority (13/14, 93%) agreed with the statement: “As a result of attending the HOPE program, overall, I feel better prepared for my surgery.” Of these, 1 participant (1/14, 7%) selected “strongly agree,” 7 (n=7/14, 50%) selected “agree,” and 5 (5/23, 30%) selected “somewhat agree.”

XSL•FO
RenderX

All 39 participants who completed the 6-month follow-up questionnaire provided reasons why they agreed or disagreed that the HOPE program helped them prepare for surgery. The findings are presented under 4 headings: personalized exercise, physical and mental preparation, peer support, and nothing new. Participant ID numbers 1 - 14 are those that were still waiting for surgery at 6-month follow-up, and IDs 15-39 are participants who had undergone surgery. No harms or unintended consequences were reported during the study.

Personalized Exercise

The program offered exercises that helped patients improve their physical condition and overall preparedness for surgery.

> *Better exercised and with better muscle definition.* [ID19]
>
> *It gave me some exercises to prepare for surgery.* [ID29]
>
> *It encouraged me to do the preparation exercises and helped lift my mood when needed.* [ID20]

Physical and Mental Preparation

Patients found the program beneficial for preparing both physically and mentally for surgery. It provided information and insight about what to expect before and after surgery, helping to manage pain, reduce anxiety, increase hope, and plan for the future.

> *The program gave me an insight into what to expect during and after the procedure.* [ID18]
>
> *I found the information useful and the relaxation techniques particularly helpful.* [ID38]
>
> *The information given was clear about the future after the operation.* [ID7]
>
> *I feel I manage pain better even if it becomes more painful.* [ID8]
>
> *I knew so much about what to expect, and I learned techniques to calm any anxiety.* [ID33]

Peer Support

Connecting with others who have arthritis and are waiting for surgery made patients feel less isolated. The program offered a platform for discussing shared challenges, such as surgery delays and recovery expectations, fostering a sense of community among participants. Participants valued the emotional support they received through the program. Sharing experiences with others who were undergoing similar surgeries provided comfort, while insights into the surgical process helped ease fears.

> *Hearing what other arthritis sufferers are going through made you feel that you are not alone in dealing with the pain.* [ID1]
>
> *The HOPE program gave me the opportunity to share my thoughts/fears with others who had either had their joint surgery or were waiting to undergo it.* [ID17]

Nothing New

A few participants found that the program covered what they already knew or that they already had a positive mindset.

> *I haven’t found out anything new.* [ID9]
>
> *I already had a very positive view of how to deal with the issues arising from my arthritis.* [ID10]

Usability

Only 16 participants completed the optional SUS. Participants reported a mean SUS score of 70.1 (SD 15.9; range 50 - 95). The 10-item frequency response data are provided in Multimedia Appendix 3. Compared with the 23 participants who did not complete the SUS, the 16 SUS completers were younger (median age of 64, IQR 8 vs 67, IQR 6 years; sample size n=8 and n=13, respectively), included a higher proportion of males (44% vs 26%), and more knee surgery patients (75% vs 60%) with a mobility impairment (69% vs 50%). Other patient characteristics were broadly similar. On average, the 16 completers had slightly lower disease severity: total WOMAC median 49 (IQR 23) versus 53 (IQR 17).

Table 3 summarizes the patient-reported outcomes at baseline, 8 weeks (post-HOPE program), and 6-month follow-up. Median values suggested potential improvements in many outcome measures at the end of the HOPE program (8 weeks). There were sustained improvements in median values for several outcomes at 6 months.

XSL•FO
RenderX

Table . Summary of baseline, post–Help Overcome Problems Effectively (HOPE) program (8 weeks), and 6-month follow-up outcomes (n=39).

Outcome variable	Baseline, median (IQR)	8 weeks, median (IQR)	*P* value, Wilcoxon test (baseline to 8 week)	6 months, median (IQR)	*P* value, Friedman test
Arthritis Self-Efficacy Scale (ASES)					
Confidence to manage pain (1 - 10, ↑=better)	3.8 (2.0-5.6)	4.6 (3.6-5.4)	.07	5.6 (3.8-8.2)	.002[a]
Confidence to manage other symptoms (1 - 10, ↑=better)	4.5 (2.5-5.5)	5.0 (4.2-6.5)	.001[a]	6.5 (4.2-8.3)	<.001[a]
Western Ontario and McMaster Osteoarthritis Index (WOMAC)					
Pain (0 - 20, ↑=worse)	10.0 (8.0-12.0)	10.0 (7.0-13.0)	.19	8.0 (5.0-12.0)	.0[a]4
Stiffness (0 - 8, ↑=worse)	4.0 (4.0-5.0)	4.0 (4.0-5.0)	.92	4.0 (2.0-5.0)	.11
Physical functioning (0 - 68, ↑=worse)	35.0 (26.0-41.0)	34.0 (24.0-42.0)	.44	29.0 (10.0-42.0)	.07
Total (0 - 96, ↑=worse)	49.0 (40.0-58.0)	48.0 (33.0-59.0)	.38	39.0 (19.0-60.0)	.09
Numerical Pain Rating Scale (NPRS)					
Pain (0 - 10, ↑=worse)	6.0 (5.0-7.0)	5.0 (4.0-7.0)	.19	5.0 (2.0-6.0)	.002[a]
EQ-5D					
Quality of Life (EQ-VAS[c]; 0 - 100, ↑=better)	58.0 (35.0-80.0)	60.0 (35.0-80.0)	.37	70.0 (45.0-85.0)	.05
Health status (EQ-Index; 0 - 1, ↑=better)	0.62 (0.30-0.74)	0.60 (0.23-0.78)	.54	0.75 (0.30-0.83)	.02[a]
Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS)					
Mental well-being (SWEMWBS; 5 - 35, ↑=better)	25.0 (21.0-28.0)	25.0 (22.0-28.0)	.86	27.0 (23.0-29.0)	.01[a]
International Physical Activity Questionnaire–Short Form (IPAQ-SF)					
Total (MET-min/wk, ↑=better)	2340 (393-7464)	2628 (480-6152)	.98	2079 (306-5988)	.55
Sitting time (min/d, ↑=worse)	360 (285-480)	360 (181-540)	.41	300 (240-480)	.15
Inactive (<600 MET[b]-min/wk), n (%)	12 (30.8)	10 (25.6)	—[d]	12 (30.8)	—[d]

[a] Statistically significant *P*<05. Data is for participants who participated in ≥5 sessions and completed both follow-up questionnaires (n=39).

[b] MET: metabolic equivalent.

[c] EQ-VAS: EQ-Visual Analogue Scale.

[d] Not applicable.

The Friedman test indicated several potential improvements across the 6 months. Median scores for ASES pain were 3.8 (IQR 2.0-5.6), 4.6 (IQR 3.6-5.4), and 5.6 (IQR 3.8-8.2) at baseline, 8 weeks, and 6-month follow-up, respectively (*P*=.002); ASES other symptoms: 4.5 (IQR 2.5-5.5), 5.0 (IQR 4.2-6.5), and 6.5 (IQR 4.2-8.3*; P*<.001); WOMAC pain: 10.0 (IQR 8.0-12.0), 10.0 (IQR 7.0-13.0), and 8.0 (IQR 5.0-12.0*; P*=.04); NRPS: 6.0 (IQR 5.0-7.0), 5.0 (IQR 4.0-7.0), and 5.0 (IQR 2.0-6.0*; P*=.002); EQ-index: 0.62 (IQR 0.3-0.74), 0.60 (IQR 0.23-0.78), and 0.75 (IQR 0.30-0.83*; P*=.02); and SWEMWBS: 25.0 (IQR 21.0-28.0), 25 (IQR 22.0-28.0), and 27.0 (IQR 23.0-29.0*; P*=.01). Separate Wilcoxon tests at 8 weeks (immediately following the end of the HOPE program) found that only ASES other symptoms was statistically significant (*P*=.001; refer to Table 3).

Ancillary Analyses

Assessment of Bias: Program Completers Versus Program Noncompleters at Baseline

A total of 15 participants were categorized as noncompleters of the Hope program (ie, completing <5 of 8 sessions). Only 4 (27%) of these participants returned both follow-up questionnaires, which was insufficient for meaningful analysis.

XSL•FO
RenderX

Therefore, bias assessment was conducted using baseline data only. Potential differences at baseline between noncompleters (<5 sessions; n=15) and completers (≥5 sessions; n=42) were explored using descriptive statistics. There were no obvious differences between noncompleters and completers in age (median 65.50, IQR 64-72.75 vs median 66.00, IQR 63.0-69.0 years, respectively) or IMD (median 7.00, IQR 5.0-8.0 vs median 7.00, IQR 6.0-8.25). There were slight differences between noncompleters and completers in gender (male: 50% vs 31%), ethnicity (White: 93% vs 100%), disability (yes: 53% vs 62%), employment (in paid work—full time and part time: 54% vs 20%), joint replacement (knee: 47% vs 67%), and waiting time (7.00, IQR 4.0-16.0 vs 6.00, IQR 2.0-13.0 months). On average, noncompleters had slightly greater disease severity. For example, noncompleters reported more pain (NPRS: median 7.00, IQR 6.0-8.0 vs median 6.00, IQR 5.0-7.0, respectively), had a higher total WOMAC score (median 60.00, IQR 40.0-71.0 vs median 49.00, IQR 38.75-58.25), and a lower EQ-5D index value (median 0.45, IQR 0.18-0.73 and median 0.636, IQR 0.29-0.74).

Impact of Surgery on Outcomes

To understand the potential impact of surgery on outcomes, an exploratory descriptive comparison between those who had and had not received surgery at 6 months was made (data presented in Multimedia Appendix 4). This comparison was only based on the ASES data, since improvements in this outcome were statistically significant at both 8 weeks and 6 months. The results show that those who had received surgery at 6 months had larger median improvements in self-efficacy (for both pain and other symptoms). Those who had not had surgery showed marginal improvements in self-efficacy for pain and for other symptoms at 8 weeks. These were maintained at 6 months for pain self-efficacy but not for other symptoms.

Health Care Resource Usage

The component of the study that focused on resource use for this early cost-impact analysis had a sample size of 39 patients, who completed the web-based questionnaire at all 3 time points: baseline, after 8 weeks, and after 6 months. Of these, 25 patients had their surgical intervention within the period covered (ie, within 6 months) and were therefore excluded from the analysis, leaving a total sample size of 14 analyzed. Results are provided in Table 4 (total cost-impact per patient) and Table 5 (cost-impact per patient per week). Cost-impact per patient per week evaluation revealed overall cost savings over 8 weeks as well as over 6 months, but this failed to reach statistical significance. Face-to-face general practitioner interactions at the 6-month interval showed a statistically significant change. Further details of the economic analysis are provided in Multimedia Appendix 2.

Table . Total cost impact per patient (£per week).

Cost category[a]	Cost change from baseline to 8 weeks, mean (95% CI)	Cost change from baseline to 6 months, mean (95% CI)
Face-to-face visit with a physiotherapist	12.85 (–3.80 to 36.02)	–8.64 (–76.34 to 27.29)
Remote visit with GP[b]	–7.72 (–19.55 to 0.36)	0.22 (–3.38 to 4.07)
Face-to-face visit with GP	9.66 (–12.57 to 34.00)	14.09 (–5.39 to 36.29)
Face-to-face hospital visit	–7.46 (–29.65 to 10.52)	–8.91 (–38.35 to 10.84)
Total	7.34 (46.49 to –27.61)	–3.25 (46.10 to –87.02)

[a] Positive values correspond to cost savings.

[b] GP: general practitioner.

Table . Cost impact per patient (£ per week).

Cost category[a]	Cost change from baseline to 8 weeks, mean (95% CI)	Cost change from baseline to 6 months, mean (95% CI)
Face-to-face visit with physiotherapist	1.61 (–0.48 to 4.5)	0.69 (–3.34 to 3.94)
Remote visit with [c]GP	–0.96 (–2.44 to 0.04)	0.20 (–0.1 to 0.63)
Face-to-face visit with GP	1.21 (–1.57 to 4.25)	2.45 (0.50 to 5.08)[b]
Face-to-face hospital visit	–0.93 (–3.71 to 1.31)	–0.03 (–1.85 to 1.82)
Total	0.92 (–3.45 to 5.81)	3.31 (–1.93 to 8.12)

[a] Positive values correspond to cost savings.

[b] Statistically significant (*P*<.05).

[c] GP: general pratitioner.

Sample Size Calculation for Future Trial

Data collected as part of the current evaluation were used to inform likely sample sizes for future studies in this area. This sample size calculation was based on ASES-8 data. Unfortunately, the minimum clinically important difference of the ASES-8 is unknown [44]. However, it is sensitive to change, with an effect size of 0.31 previously reported for the ASES-8

XSL•FO
RenderX

following interdisciplinary group therapy for fibromyalgia [45]. Moderate effect sizes of this magnitude are common for conservative interventions in musculoskeletal conditions. In this pilot study, the mean and SD values for ASES pain and ASES other symptoms at baseline were 3.98 (SD 1.93) and 4.29 (SD 2.08), respectively. Assuming a 1-tailed hypothesis, an effect size of 0.3, α=.05, 90% power, and a 1:1 allocation ratio, 191 participants would be required in each group (N=382) to detect a ≥0.58-point difference in ASES pain and ≥0.62-point difference in ASES other symptoms.

Discussion

Principal Findings

This study evaluated the HOPE program, a digital self-management intervention designed to support patients awaiting hip and knee replacement surgery. Results from 39 completers suggested potential improvements in self-efficacy, pain, health status, and mental well-being over 6 months. Most participants felt better prepared for surgery, and the program was rated above average for usability (mean SUS score 70.1).

Participant feedback revealed some key areas that underscore the program’s potential usefulness. Some participants appreciated the targeted exercises that improved their physical and mental readiness for surgery. The program provided comprehensive information about the surgical process, helping patients manage pain, reduce anxiety, and plan for the future. Studies have shown that patients have difficulties remembering information immediately after deciding to undergo surgery [46]. Having access to digital information, which can be regularly and quickly updated with evidence-based information, is a useful resource for patients. By fostering a sense of community, the program helped some participants connect with others facing similar challenges. However, some participants noted that the program offered nothing new, as they already enjoyed a positive mindset or previous knowledge.

The demographic profile of completers (median age 66, IQR 63-69.5 years; 100% White; and 66.7% female) was almost identical to a recent UK study, which found that digital health coaching delivered to patients waiting for lower limb arthroplasty improved patient activation and reduced length of hospital stay [47]. It should be noted that noncompleters of the program were more likely to be male, in paid employment, and awaiting a hip replacement.

Engagement with the HOPE program was high, with 73.7% (42/57) of participants attending ≥5 of 8 sessions. Follow-up and engagement rates were lower when based on the 100 participants who enrolled: 39% (39/100) completed the 6-month follow-up questionnaires, and 42% (42/100) who completed ≥5 sessions. Among those who completed all study procedures, 93% (39/42) engaged with the program.

A recent national digital attitudes and behavior survey conducted in the United Kingdom by ORCHA in 2023 described the willingness of older respondents to use digital apps for self-monitoring, symptom tracking, and managing recovery [48].

At the 6-month follow-up, nearly two-thirds (25/39, 64%) of participants had undergone surgery. More than 90% (23/25) of these participants agreed that the program helped them prepare better for surgery. Statistically significant median improvements in most PROMs were evident at the end of the HOPE program, and several scores continued to improve at 6-month follow-up, including self-efficacy, pain, health status, and mental well-being. The exercise program was the most bookmarked page, and despite the majority of participants (49/57, 86%) starting the personalized exercise program, there were no improvements in time spent sitting or in the proportion of participants classified as inactive. The exercise program may require greater input from facilitators to encourage optimum engagement. Research shows that exercise supervision involving trained physical therapists improves compliance with exercises, especially in older adults [13,49]. Alternatively, it may be that the IPAQ-SF lacks sensitivity to adequately assess physical activity [39]. More objective measures of physical activity, such as accelerometry, could be considered in future research.

The high number of participants undergoing surgery makes it challenging to attribute potential improvements in PROMs to either intervention. In their systematic review and meta-analysis, Punnoose et al [13] showed that variability in surgical procedures can influence postoperative recovery; therefore, postsurgical improvements cannot be attributable solely to prehabilitation. Owing to the often degenerative nature of musculoskeletal conditions, potential improvements in PROMs in this study were not anticipated a priori. Rather, it was hypothesized that attending the HOPE program would slow the rate of decline through the acquisition of effective self-management and coping strategies. Thus, the observed trend for median improvements across the majority of PROMs is encouraging.

Resource Usage

This early cost analysis suggests that the HOPE program may lead to a reduction in patient interactions with care professionals at both 8 weeks and 6 months. However, the small sample size results in wide CIs, which limits the reliability of these findings and affirms the need for further studies to assess the cost-effectiveness of the program. Despite this limitation, the initial results highlight the potential for the HOPE program to offer cost-saving benefits at a societal level.

Strengths and Limitations

A strength of this real-world study was the inclusion of participants with lived experience at all stages of the project, providing input into the HOPE program intervention development process and follow-up feedback to optimize it for further studies. The majority of participants started the exercise program, which is a cornerstone of prehabilitation. Other strengths include the use of validated PROMs, high levels of engagement with the intervention, and good survey completion rates at 6 months. This version of the HOPE program was rapidly developed and deployed by adding new musculoskeletal content to an existing taxonomized evidence-based intervention. Some of the health professionals involved in the co-design workshops suggested that patients needing only conservative management and not requiring surgery would also benefit from

XSL•FO
RenderX

the program. Our co-creation and intervention development process could develop and test a program for these patients and for other groups of nonorthopedic presurgery patients. The powered by H4C platform currently hosts more than 15 digital self-management and health interventions. Using a single platform to deliver multiple interventions and modules offers several advantages for funders, researchers, health care providers, and patients. Many patients live with comorbid conditions requiring diverse information and self-management techniques. Platform delivery can incorporate and streamline self-management support. Torous and Vaidyam [50] asserted that "instead of a plethora of apps, there is a need for a few that meet the needs of many." Drawing on successful examples from the automobile, space, and clean energy sector, Ansar and Flyvbjerg [51] outline the benefits of platforms over one-off designs, such as repeatability, extendibility, absorptive and adaptive capacity, resulting in "faster, better, cheaper" services and products. They concluded that sectors such as health "are ripe for a platform rethink." Another strength of this application is the partnership between a social enterprise company and an academic institution. A recent Wellcome report [52] recommended that companies, including nonprofits, can be better at developing and scaling digital health solutions than university research groups.

The limitations of the study include the small number (16/39) of participants who completed the SUS. It is possible that these participants had a more positive user experience compared to those who did not complete the scale. Among the 39 study completers, most participants (>90%) agreed that the program helped them prepare better for surgery, and the textual responses supporting this question provided limited feedback. A broader set of feedback questions and/or postprogram qualitative interviews or focus groups analyzed using rigorous and transparent methods—with participants who did not complete the program—could elicit more critical or negative experiences.

The self-selecting nature of recruitment may have resulted in participants who were inherently more inclined to seek assistance or engage in self-help efforts.

Without a control group comparator, it is not possible to directly attribute any change in the PROMs to the HOPE program. It is important to note that many improvements were not statistically significant, and the statistical analyses performed were likely to be underpowered. Furthermore, a recent systematic review of hip arthroplasty prehabilitation interventions suggested that measures such as the WOMAC may not be the most appropriate measure to detect differences and suggest alternative objective measures such as the chair rise test, gait speed, or stair climbing [53]. That review also found that more than 8 weeks of prehabilitation was associated with improved outcomes, suggesting that future trials of the HOPE program should consider extending the length of the intervention. While our completer analysis provides valuable proof-of-concept data, it limits generalizability to real-world implementation, where attrition is typically higher. The baseline differences between completers and noncompleters suggest our effect estimates may be optimistic. Future randomized controlled trials (RCTs) should combine ITT and per-protocol analyses to distinguish efficacy from effectiveness.

Separating the effects of the intervention from the effects of surgery is problematic. The ancillary analysis of the ASES data (refer to Table 4) suggests that surgery was probably a major contributor to improvements in self-efficacy at 6 months. This is not surprising, given that the excellent outcomes of hip replacement surgery have led to the procedure being described by The Lancet as the "operation of the century" [54]. Approximately 96.2% and 90.8% of patients have previously reported satisfaction with their hip and knee replacement surgery, respectively [55].

A future definitive RCT should be appropriately powered to directly compare an intervention group (ie, the HOPE program) against an appropriate control group (ie, treatment as usual). Subgroup analysis should compare PROMs in those who have had, or are still awaiting, surgery at 6-month follow-up. Such a design would help to distinguish the effects of the intervention from the effects of surgery.

The baseline data show that program noncompleters (ie, those who completed <5 sessions) had slightly greater disease severity at baseline than program completers. Owing to limited follow-up data, it is not known whether these participants could not complete the program due to factors relating to their musculoskeletal condition, their experience of the program, or random intervening factors. Nonresponders were also more likely to be male, in paid employment, and awaiting a hip replacement. Such findings raise questions about how to engage people with greater disease severity and these sociodemographic characteristics in future support programs. Further research is needed to understand individual needs and how they change as disease and pain progress, and to determine how best to support individuals through targeted interventions.

In line with the wealth of other UK health care research studies, the participant sample in this study lacked diversity in terms of ethnicity and socioeconomic characteristics. The study sample reflects the demographics of NHS waiting lists and can be understood as a manifestation of structural inequalities. People living in the most deprived areas of the United Kingdom are more likely to require replacement surgery but less likely to receive it [56] and less likely to have good outcomes [57,58] compared with those living in the least deprived areas. This recurring finding underscores the need for research into the impact of structural barriers to self-management, which may, in turn, suggest the need for more options or a new paradigm approach. Health care interventions that disproportionately meet the needs of nonmarginalized groups embed injustice by widening health inequity. The earlier statement that no harm was reported during the study holds when "harm" is understood within the parameters of evidence-based medicine and its associated framework of biomedical ethics. However, when a framing such as distributive justice is applied, the intervention may be associated with unintended adverse consequences that emerge from and perpetuate ideologies such as structural racism and classism. Lack of attention to unintended harm linked to the lack of diversity in self-management research highlights the need for an expanded ethical framework informed by disability justice scholarship [59]. Recommendations from a recent report into musculoskeletal health inequalities in the United Kingdom included prioritizing surgery and self-management support for

XSL•FO
RenderX

patients living in the most deprived areas [60]. More effort is required to understand the needs of and actively recruit these groups of participants in future self-management trials. A national digital attitudes and behavior survey conducted in the United Kingdom by ORCHA in 2021 [61] found that advocacy for digital health apps was highest among people of Black African heritage (89%), followed by Asian (80%), and then White (64%) respondents. Studies from the United States highlight the importance of recruiting low-income and ethnic minority participants, showing that these groups are more willing to attend [62] and engage more [63] with health interventions compared with White participants in higher-income groups. However, data from this study show that deprivation levels were similar between HOPE program completers and noncompleters.

Conclusion

The results are promising in relation to the acceptability of a peer-supported self-management program for people awaiting hip or knee surgery. Overall, participants felt better prepared for surgery. Textual feedback was generally positive, and participants attributed improvements in their mental and physical well-being to techniques they learned in the HOPE program. However, comparing self-efficacy in those who had and had not received surgery suggests that surgery might have been a more important agent of change than the HOPE program. Overall, the study has demonstrated potential benefit and no evidence of harm or unintended consequences. A randomized controlled efficacy and cost-effectiveness trial design, involving a socioeconomically and ethnically representative sample, is required to delineate the effects attributable to the HOPE program, as opposed to effects of having surgery or natural variation in PROMs. While these preliminary results are promising, they require confirmation in a fully powered RCT using ITT analysis to account for real-world attrition patterns.

Acknowledgments

The authors would like to thank the participants involved in the co-creation workshops, stakeholders, partners, and course participants. The HOPE program is delivered by the Coventry University spin-out social enterprise, Hope 4 The Community (H4C) CIC. We would also like to thank Health Tech Enterprise for contributing to the health economic analysis. The project, Hope and Prepare Effectively for Surgery (HoPES; 10053263), was funded by UK Research and Innovation as part of the Healthy Ageing Challenge Scaling Social Ventures competition, delivered by Innovate UK. Funding was awarded to H4C through the Small Business Research Initiative Social Ventures competition, run by UK Research and Innovation as part of the Healthy Ageing Challenge, delivered by Innovate UK. This pilot study was not registered with a clinical trials registry. There was no published protocol for this pilot study.

Funding

No external financial support or grants were received for this work.

Conflicts of Interest

AT is a co-founder and director of H4C and the co-inventor of the original HOPE program. AC is the head of Partnerships and Projects at H4C.

Multimedia Appendix 1
Intervention screenshot and session content.
[DOCX File, 560 KB - rehab_v13i1e68286_app1.docx]

Multimedia Appendix 2
Health economic evaluation analysis.
[DOCX File, 40 KB - rehab_v13i1e68286_app2.docx]

Multimedia Appendix 3
System Usability Scale scores.
[DOCX File, 22 KB - rehab_v13i1e68286_app3.docx]

Multimedia Appendix 4
Changes in Arthritis Self-Efficacy Scale.
[DOCX File, 29 KB - rehab_v13i1e68286_app4.docx]

Checklist 1
CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist.
[DOCX File, 28 KB - rehab_v13i1e68286_app5.docx]

XSL•FO
RenderX

Checklist 2
CONSORT-EHEALTH checklist.
[PDF File, 683 KB - rehab_v13i1e68286_app6.pdf]

References

1. The state of musculoskeletal health 2021: arthritis and other musculoskeletal conditions in numbers. Arthritis UK. URL: https://www.versusarthritis.org/media/24238/state-of-msk-health-2021.pdf [accessed 2025-12-12]
2. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. The Lancet 2019 Apr;393(10182):1745-1759. [doi: 10.1016/S0140-6736(19)30417-9]
3. Musculoskeletal health. NHS England. 2020. URL: https://www.england.nhs.uk/elective-care-transformation/best-practice-solutions/musculoskeletal/ [accessed 2024-08-07]
4. Versus arthritis position statement on the clinically-led review of NHS access standards. Versus Arthritis. 2019. URL: https://www.versusarthritis.org/media/21674/clinical-standards-review-nov19.pdf [accessed 2024-08-07]
5. What we’re doing to tackle waiting times for people with arthritis. Versus Arthritis. 2023 Dec. URL: https://www.versusarthritis.org/news/2023/december/what-we-re-doing-to-tackle-waiting-times-for-people-with-arthritis/ [accessed 2024-08-07]
6. Treatment waiting times. Arthritis Action. 2024. URL: http://www.arthritisaction.org.uk/living-with-arthritis/resource-centre/waiting-times/ [accessed 2024-08-07]
7. Matharu G, Culliford D, Blom A, Judge A. Projections for primary hip and knee replacement surgery up to the year 2060: an analysis based on data from The National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. Ann R Coll Physicians Surg Can 2022 Jun;104(6):443-448. [doi: 10.1308/rcsann.2021.0206]
8. People living in the poorest areas waiting longer for hospital treatment: The King’s Fund and Healthwatch England share new analysis. The King’s Fund. 2021 Sep. URL: https://www.kingsfund.org.uk/insight-and-analysis/press-releases/poorest-areas-waiting-longer-hospital-treatment#:~:text=New%20analysis%20from%20The%20King’s,in%20the%20least%2Ddeprived%20areas [accessed 2024-08-07]
9. Bergstra SA. Health inequalities across patients with early inflammatory arthritis of different ethnicities: what could be the driving factors? Rheumatology (Oxford) 2022 Dec 23;62(1):7-8. [doi: 10.1093/rheumatology/keac383] [Medline: 35786709]
10. Elective recovery planning supporting guidance. NHS England. 2022 Apr. URL: https://www.england.nhs.uk/wp-content/uploads/2021/12/B1269-elective-recovery-planning-supporting-guidance.pdf [accessed 2024-08-07]
11. Vasta S, Papalia R, Torre G, et al. The influence of preoperative physical activity on postoperative outcomes of knee and hip arthroplasty surgery in the elderly: a systematic review. J Clin Med 2020 Mar 31;9(4):969. [doi: 10.3390/jcm9040969] [Medline: 32244426]
12. Wang L, Lee M, Zhang Z, Moodie J, Cheng D, Martin J. Does preoperative rehabilitation for patients planning to undergo joint replacement surgery improve outcomes? A systematic review and meta-analysis of randomised controlled trials. BMJ Open 2016 Feb;6(2):e009857. [doi: 10.1136/bmjopen-2015-009857]
13. Punnoose A, Claydon-Mueller LS, Weiss O, Zhang J, Rushton A, Khanduja V. Prehabilitation for patients undergoing orthopedic surgery: a systematic review and meta-analysis. JAMA Netw Open 2023 Apr 3;6(4):e238050. [doi: 10.1001/jamanetworkopen.2023.8050] [Medline: 37052919]
14. Guida S, Vitale JA, Swinnen E, et al. Effects of prehabilitation with advanced technologies in patients with musculoskeletal diseases waiting for surgery: systematic review and meta-analysis. J Med Internet Res 2024 Dec 12;26:e52943. [doi: 10.2196/52943] [Medline: 39666971]
15. Omar I, Wylde V, Fogg J, Whitehouse M, Bertram W. Pre-operative education and prehabilitation provision for patients undergoing hip and knee replacement: a national survey of current NHS practice. BMC Musculoskelet Disord 2025 Apr 29;26(1):421. [doi: 10.1186/s12891-025-08637-5] [Medline: 40301783]
16. Perry R, Herbert G, Atkinson C, et al. Pre-admission interventions (prehabilitation) to improve outcome after major elective surgery: a systematic review and meta-analysis. BMJ Open 2021 Sep 30;11(9):e050806. [doi: 10.1136/bmjopen-2021-050806] [Medline: 34593498]
17. Fecher-Jones I, Grimmett C, Carter FJ, Conway DH, Levett DZH, Moore JA. Surgery school—who, what, when, and how: results of a national survey of multidisciplinary teams delivering group preoperative education. Perioper Med 2021 Dec;10(1):20. [doi: 10.1186/s13741-021-00188-2]
18. Pearce G, Parke HL, Pinnock H, et al. The PRISMS taxonomy of self-management support: derivation of a novel taxonomy and initial testing of its utility. J Health Serv Res Policy 2016 Apr;21(2):73-82. [doi: 10.1177/1355819615602725] [Medline: 26377727]
19. Martin F, Wright H, Moody L, et al. Help to overcome problems effectively for cancer survivors: development and evaluation of a digital self-management program. J Med Internet Res 2020 May 19;22(5):e17824. [doi: 10.2196/17824] [Medline: 32209529]
20. Wright H, Martin F, Clyne W, et al. A digital self-management program (help to overcome problems effectively) for people living with cancer: feasibility randomized controlled trial. J Med Internet Res 2021 Nov 5;23(11):e28322. [doi: 10.2196/28322] [Medline: 34738912]

XSL•FO
RenderX

21. Wright H, Turner A, Ennis S, et al. Digital peer-supported self-management intervention codesigned by people with long COVID: mixed methods proof-of-concept study. JMIR Form Res 2022 Oct 14;6(10):e41410. [doi: 10.2196/41410] [Medline: 36166651]
22. Accreditation register. Quality Institute for Self Management Education and Training (QISMET). 2024. URL: https://www.qismet.org.uk/accreditation/accreditation-register/ [accessed 2024-08-07]
23. Delivering safe digital health. The Organisation for the Review of Care and Health Apps (ORCHA). 2024. URL: https://orchahealth.com [accessed 2024-08-07]
24. Ageberg E, Link A, Roos EM. Feasibility of neuromuscular training in patients with severe hip or knee OA: the individualized goal-based NEMEX-TJR training program. BMC Musculoskelet Disord 2010 Jun 17;11(1):126. [doi: 10.1186/1471-2474-11-126] [Medline: 20565735]
25. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5(1):114. [doi: 10.1186/s40814-019-0499-1] [Medline: 31608150]
26. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
27. Ageberg E, Nilsdotter A, Kosek E, Roos EM. Effects of neuromuscular training (NEMEX-TJR) on patient-reported outcomes and physical function in severe primary hip or knee osteoarthritis: a controlled before-and-after study. BMC Musculoskelet Disord 2013 Dec;14(1):232. [doi: 10.1186/1471-2474-14-232]
28. Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA:D™): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. BMC Musculoskelet Disord 2017 Feb 7;18(1):72. [doi: 10.1186/s12891-017-1439-y] [Medline: 28173795]
29. Roos EM, Barton CJ, Davis AM, et al. GLA:D to have a high-value option for patients with knee and hip arthritis across four continents: Good Life with osteoArthritis from Denmark. Br J Sports Med 2018 Dec;52(24):1544-1545. [doi: 10.1136/bjsports-2017-098904]
30. Barton CJ, Kemp JL, Roos EM, et al. Program evaluation of GLA:D® Australia: physiotherapist training outcomes and effectiveness of implementation for people with knee osteoarthritis. Osteoarthr Cartil Open 2021 Sep;3(3):100175. [doi: 10.1016/j.ocarto.2021.100175] [Medline: 36474815]
31. Goff AJ, De Oliveira Silva D, Ezzat AM, et al. Co-design of the web-based “My Knee” education and self-management toolkit for people with knee osteoarthritis. Digit HEALTH 2023;9:20552076231163810. [doi: 10.1177/20552076231163810] [Medline: 37009308]
32. English indices of deprivation 2019. GOV.UK. 2019 Sep. URL: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 [accessed 2025-12-12]
33. Kenton K, Pham T, Mueller E, Brubaker L. Patient preparedness: an important predictor of surgical outcome. Am J Obstet Gynecol 2007 Dec;197(6):654. [doi: 10.1016/j.ajog.2007.08.059] [Medline: 18060968]
34. Ng Fat L, Scholes S, Boniface S, Mindell J, Stewart-Brown S. Evaluating and establishing national norms for mental wellbeing using the short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): findings from the Health Survey for England. Qual Life Res 2017 May;26(5):1129-1144. [doi: 10.1007/s11136-016-1454-8] [Medline: 27853963]
35. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. Health Qual Life Outcomes 2009 Feb 19;7(1):15. [doi: 10.1186/1477-7525-7-15] [Medline: 19228398]
36. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
37. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988 Dec;15(12):1833-1840. [Medline: 3068365]
38. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self‐efficacy in people with arthritis. Arthritis & Rheumatism 1989 Jan;32(1):37-44. [doi: 10.1002/anr.1780320107]
39. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the international physical activity questionnaire short form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8(1):115. [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
40. Stratford PW, Spadoni G. The reliability, consistency, and clinical application of a numeric pain rating scale. Physiotherapy Canada 2001;53:88-91.
41. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
42. Unit costs 2022-2027. University of Kent. URL: https://www.pssru.ac.uk/unitcostsreport/ [accessed 2024-08-07]
43. National tariff payment system documents, annexes and supporting documents. NHS England. 2023. URL: https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/ [accessed 2023-08-07]
44. Brady TJ. Measures of self-efficacy: Arthritis Self-Efficacy Scale (ASES), Arthritis Self-Efficacy Scale-8 Item (ASES-8), Children’s Arthritis Self-Efficacy Scale (CASE), Chronic Disease Self-Efficacy Scale (CDSES), Parent’s Arthritis

XSL•FO
RenderX

Self-Efficacy Scale (PASE), and Rheumatoid Arthritis Self-Efficacy Scale (RASE). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S473-S485. [doi: 10.1002/acr.20567] [Medline: 22588769]
45. Mueller A, Hartmann M, Mueller K, Eich W. Validation of the arthritis self-efficacy short-form scale in German fibromyalgia patients. Eur J Pain 2003;7(2):163-171. [doi: 10.1016/S1090-3801(02)00097-6] [Medline: 12600798]
46. Specht K, Kjaersgaard-Andersen P, Pedersen BD. Patient experience in fast-track hip and knee arthroplasty--a qualitative study. J Clin Nurs 2016 Mar;25(5-6):836-845. [doi: 10.1111/jocn.13121] [Medline: 26708610]
47. Powley N, Tew GA, Durrand J, et al. Digital health coaching to improve patient preparedness for elective lower limb arthroplasty: a quality improvement project. BMJ Open Qual 2023 Dec 7;12(4):e002244. [doi: 10.1136/bmjoq-2022-002244] [Medline: 38061840]
48. Digital health in the UK today, annual national attitudes and behaviour research. ORCA. 2023. URL: https://info.orchahealth.com/digital-health-attitudes-behaviour-2023-report [accessed 2025-12-12]
49. Shubert TE. Evidence-based exercise prescription for balance and falls prevention: a current review of the literature. J Geriatr Phys Ther 2011;34(3):100-108. [doi: 10.1519/JPT.0b013e31822938ac] [Medline: 22267151]
50. Torous J, Vaidyam A. Multiple uses of app instead of using multiple apps - a case for rethinking the digital health technology toolbox. Epidemiol Psychiatr Sci 2020 Jan 31;29:e100. [doi: 10.1017/S2045796020000013] [Medline: 32000876]
51. Ansar A, Flyvbjerg B. How to solve big problems: bespoke versus platform strategies. Oxford Review of Economic Policy 2022 May 18;38(2):338-368. [doi: 10.1093/oxrep/grac009]
52. The Wellcome Trust. Catalytic environments to drive progress in digital mental health research. Wellcome Open Res 2024;9:301. [doi: 10.21955/wellcomeopenres.1115389.1]
53. Widmer P, Oesch P, Bachmann S. Effect of prehabilitation in form of exercise and/or education in patients undergoing total hip arthroplasty on postoperative outcomes-a systematic review. Medicina (Kaunas) 2022 May 30;58(6):742. [doi: 10.3390/medicina58060742] [Medline: 35744005]
54. Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. The Lancet 2007 Oct;370(9597):1508-1519. [doi: 10.1016/S0140-6736(07)60457-7]
55. Appiah KOB, Khunti K, Kelly BM, et al. Patient - rated satisfaction and improvement following hip and knee replacements: development of prediction models. Evaluation Clinical Practice 2023 Mar;29(2):300-311. [doi: 10.1111/jep.13767]
56. Cook MJ, Lunt M, Ashcroft DM, Board T, O'Neill TW. The impact of frailty and deprivation on the likelihood of receiving primary total hip and knee arthroplasty among people with hip and knee osteoarthritis. J Frailty Aging 2023;12(4):298-304. [doi: 10.14283/jfa.2023.36] [Medline: 38008980]
57. Clement ND, Muzammil A, MacDonald D, Howie CR, Biant LC. Socioeconomic status affects the early outcome of total hip replacement. J Bone Joint Surg Br 2011 Apr;93-B(4):464-469. [doi: 10.1302/0301-620X.93B4.25717]
58. Clement ND, Jenkins PJ, MacDonald D, et al. Socioeconomic status affects the Oxford knee score and short-form 12 score following total knee replacement. Bone Joint J 2013 Jan;95-B(1):52-58. [doi: 10.1302/0301-620X.95B1.29749] [Medline: 23307673]
59. Piepzna-Samarasinha LL. Care Work: Dreaming Disability Justice: Arsenal Pulp Press; 2018:263.
60. Act now: musculoskeletal health inequalities and deprivation. : Arthritis and Musculoskeletal Alliance (ARMA); 2024 URL: https://arma.uk.net/wp-content/uploads/2024/03/Musculoskeletal-Health-Inequalities-and-Deprivation-report_v07.pdf [accessed 2024-08-07]
61. Digital health in the UK: national attitudes and behaviours. ORCHA. 2021. URL: https://us.orchahealth.com/wp-content/uploads/2021/07/2107_ICS_Research_Report_2021_National_final.pdf [accessed 2025-12-12]
62. Lorenzo-Luaces L, Wasil A, Kacmarek CN, DeRubeis R. Race and socioeconomic status as predictors of willingness to use digital mental health interventions or one-on-one psychotherapy: national survey study. JMIR Form Res 2024 Apr 11;8:e49780. [doi: 10.2196/49780] [Medline: 38602769]
63. Horrell LN, Kneipp SM, Ahn S, et al. Chronic disease self-management education courses: utilization by low-income, middle-aged participants. Int J Equity Health 2017 Jun 27;16(1):114. [doi: 10.1186/s12939-017-0604-0] [Medline: 28655319]

Abbreviations

ASES: Arthritis Self-Efficacy Scale
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
CONSORT: Consolidated Standards of Reporting Trials
EQ-VAS: EQ-Visual Analogue Scale
H4C: HOPE 4 The Community Interest Company
HOPE: Help Overcome Problems Effectively
IMD: index of multiple deprivation
IPAQ-SF: International Physical Activity Questionnaire–Short Form
ITT: intention-to-treat
MET: metabolic equivalent of task
NHS: National Health Service

XSL•FO
RenderX

NPRS: Numerical Pain Rating Scale
ORCHA: Organisation for the Review of Care and Health Apps
PROM: patient-reported outcome measure
RCT: randomized controlled trial
SUS: System Usability Scale
SWEMWBS: Short Warwick-Edinburgh Mental Well-Being Scale
WOMAC: Western Ontario and McMaster Universities Arthritis Index

---

Edited by S Munce; submitted 01.Nov.2024; peer-reviewed by G Martin-Alemany, J Bailey, P Marier-Deschenes; revised version received 23.Oct.2025; accepted 04.Nov.2025; published 07.Jan.2026.

Please cite as:
Horton E, Wright H, Turner A, Moody L, Aphramor L, Carlson A, Ghiasvand H, Palmer S
Using a Co-Designed Digital Self-Management Program to Prepare Patients for Hip or Knee Replacement Surgery: Pragmatic Pilot Study
JMIR Rehabil Assist Technol 2026;13:e68286
URL: *https://rehab.jmir.org/2026/1/e68286*
doi:*10.2196/68286*

© Elizabeth Horton, Hayley Wright, Andy Turner, Louise Moody, Lucy Aphramor, Anna Carlson, Hesam Ghiasvand, Shea Palmer. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 7.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Reducing Educational Bias in Cognitive Assessment via Dynamic Support Vector Machine Weighting: Validation Study on an Education-Stratified Dataset

Qing Liu[1*], MMed; Chi Ma[2*], BSc; Mengyuan Liu[1], MA; Suhui Chen[3], BSc; Mengting Yu[4], BSc; Lijuan Xia[5], BSc; Qi Zhang[3], BSc; Ming Wu[3], BSc

[1]School of Humanities and Social Sciences, University of Science and Technology of China, Hefei, China
[2]School of the Gifted Young, University of Science and Technology of China, Hefei, China
[3]Department of Rehabilitation Medicine, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Tian'e Hu No.1, Hefei, China
[4]Department of Rehabilitation Medicine, The Second People's Hospital, Wuhu, China
[5]Shuguang Hospital Anhui Branch Affiliated to Shanghai University of Traditional Chinese Medicine, Hefei, China
[*]these authors contributed equally

Corresponding Author:
Ming Wu, BSc
Department of Rehabilitation Medicine, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Tian'e Hu No.1, Hefei, China

Abstract

Background: The Mini-Mental State Examination (MMSE) remains widely used for cognitive screening, yet its performance varies substantially across educational backgrounds. Linear education corrections fail to capture the nonlinear interference patterns among subitems.

Objective: This study aimed to analyze how educational level shapes MMSE subitem contributions and to develop an education-adaptive optimization strategy using support vector machine–based weighting.

Methods: MMSE data from 812 participants were stratified into 4 education groups. Subitem deletion experiments quantified each subitem’s contribution (Δ). Education-specific support vector machine models were then constructed to derive dynamic weighting coefficients. Performance improvements were assessed before and after weighting.

Results: The illiterate group relied heavily on spatial orientation and memory, whereas university-educated individuals depended more on executive and calculation functions. Several education-dependent interference items were identified (eg, visuospatial construction in the primary group and basic orientation tasks in the university group). Dynamic weighting improved accuracy in all cohorts, most notably among illiterate individuals (Δ=7.25%; *P*=.06), followed by the primary school group (Δ=3.12%; *P*=.03).

Conclusions: Education-stratified weighting enhances the fairness and interpretability of MMSE-based screening. External validation confirmed generalizability, although multicenter studies are needed.

(JMIR Rehabil Assist Technol 2026;13:e79841) doi:10.2196/79841

KEYWORDS

machine learning; Mini-Mental State Examination; MMSE; support vector machine; SVM; dynamic weighted model; educational background

Introduction

The Mini-Mental State Examination (MMSE) is one of the most widely used cognitive screening tools in clinical and community settings, designed for the rapid detection of cognitive impairment [1]. Since its development by Folstein et al in 1975 [2], the MMSE has become a core instrument for assessing five cognitive domains—orientation, immediate recall, attention or calculation, language, and visuospatial ability—and can typically be completed within 10 to 15 minutes [3]. Its broad adoption reflects its operational simplicity and solid psychometric performance, with reported sensitivity of 80% to 85% and specificity of 75% to 80% for dementia screening [4].

Alzheimer disease is characterized by progressive neurodegeneration, including amyloid-β deposition and structural decline. As Alzheimer disease progresses in a relatively predictable neuroanatomical sequence, MMSE subdomains provide clinically meaningful stage markers:

XSL•FO
RenderX

orientation and memory deficits often correspond to early hippocampal involvement, while language and visuospatial impairments reflect later temporoparietal degeneration [5,6]. This multidimensional structure allows the MMSE to map aspects of the disease trajectory beyond a single total score.

Nevertheless, the diagnostic performance of the MMSE varies considerably across educational levels [7,8]. Individuals with higher education often exhibit false-negative results due to two well-recognized mechanisms: compensatory neuroplasticity that delays the manifestation of cognitive symptoms [9] and the use of test-taking strategies that allow them to maintain near-normal scores despite underlying pathology [10,11]. In contrast, individuals with limited formal education show substantially higher false-positive rates, particularly on education-dependent items, such as literacy tasks and object naming, with error rates increasing by approximately 30% to 35% compared with education-neutral instruments [12-15]. As a result, the MMSE demonstrates reduced sensitivity (68%‐72%) and specificity (65%‐70%) when a conventional cutoff score of 24 points is applied across heterogeneous educational backgrounds [16-18]. Current clinical guidelines therefore emphasize the need to incorporate educational history into cognitive assessment to minimize diagnostic inaccuracies [19,20].

Machine learning (ML) has emerged as a transformative approach in medical diagnostics, with particular relevance to cognitive assessment optimization. Support Vector Machines (SVMs) offer several advantages for psychometric refinement, including (1) robust performance in limited-sample settings (typically n<1000) [21,22], (2) effective handling and interpretability of high-dimensional feature spaces [23,24], and (3) strong compatibility with multimodal data integration frameworks that combine neuropsychological assessments with imaging- or biomarker-derived measures [25-28]. Earlier ML classifiers trained solely on MMSE total scores or subitems—most commonly logistic regression, SVM, or random forest—have demonstrated moderate diagnostic performance, with reported accuracies of 72% to 85% and area under the curve (AUC) values of 0.75 to 0.85 [29]. In contrast, multimodal models integrating magnetic resonance imaging (MRI) radiomics, positron emission tomography (PET) signatures, or speech biomarkers consistently achieve accuracies exceeding 90% to 95% [30,31]. This performance gap highlights the need for methodological innovations that enhance unimodal MMSE-based classifiers while preserving their scalability and low implementation cost.

This study introduces an education-sensitive adaptation of the MMSE based on an SVM-guided dynamic weighting framework. We analyzed cognitive screening data (n=812) collected from Chinese tertiary hospitals and community health centers and stratified participants into four educational cohorts: illiterate (0 y), primary (≤6 y), secondary (7‐12 y), and tertiary (≥13 y). The model uses systematic item-response analysis combined with SVM-guided weighting to adjust the relative contribution of each MMSE subitem in an education-specific manner.

This work provides 2 methodological innovations. First, it moves beyond conventional linear education corrections by using nonlinear SVM modeling to identify “cognitive interference” patterns unique to each education level. Second, it introduces a dynamic weighting strategy specifically designed to mitigate the disproportionate false-positive burden experienced by low-literacy populations. Together, these innovations aim to deliver a more equitable and clinically practical adaptation of the MMSE, particularly relevant for resource-limited settings.

Although unimodal or condensed classifiers derived solely from MMSE subitems are highly accessible, they are inherently constrained by limited feature richness and therefore tend to exhibit moderate predictive performance. In contrast, multimodal diagnostic systems—such as those leveraging MRI, PET, radiomics, or speech biomarkers—achieve substantially higher accuracies but remain impractical for widespread screening due to cost, technical requirements, and limited availability in primary care. Thus, methodological advances that enhance the diagnostic utility of unimodal MMSE-based approaches, while maintaining their affordability and scalability, are urgently needed. This study addresses this gap by proposing an education-stratified dynamic weighting method designed to improve predictive performance within the intrinsic limitations of unimodal cognitive classifiers.

Methods

Ethical Considerations

This study followed the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of the University of Science and Technology of China (approval number 2024-RE-431). All clinical and community datasets were obtained with authorization from the original data custodians. During data processing, only anonymized records were used, and no personally identifiable information (eg, names, addresses) was accessed, ensuring compliance with the Declaration of Helsinki.

Sample

Cognitive assessment data were obtained from two sources: the Rehabilitation Medicine Department of a tertiary hospital and a community-based cognitive screening program. A total of 812 valid records were included. Among these, 404 samples were collected from outpatients and inpatients with clinically diagnosed cognitive impairment. The inclusion and exclusion criteria are summarized in Table 1.

XSL•FO
RenderX

Table . Data inclusion and exclusion criteria.

Source	Number (n)	Inclusion criteria	Exclusion criteria	Data assessment requirements
Clinical dataset	404	• Outpatients and inpatients with clinically confirmed cognitive impairment • Complete clinical and assessment records	• Patients who underwent surgery or died during hospitalization • Missing MMSE[a] assessment data	• For clinical data: All MMSE assessors completed a standardized 12-h training program and passed certification. • Postcollection, automated missing value detection and manual verification were performed.
Community dataset	408	• Community residents with no medical consultation for cognitive complaints within the past 12 mo • Intact consciousness and ability to complete cognitive assessments	• Individuals with diagnosed mild cognitive impairment or suspected cognitive decline • Individuals with incomplete assessment procedures or severe organic neurological conditions	• For community data: Assessments followed standardized community screening procedures. Assessments were administered following standardized community screening procedures with verification of data completeness and adherence to protocol.

[a] MMSE: Mini-Mental State Examination.

Scale Selection and Scoring Criteria

The MMSE was used as the primary cognitive assessment tool. The scale comprises five cognitive domains: orientation, memory, attention and calculation, language, and visuospatial construction. Education-adjusted cutoff scores were based on Chinese normative standards [3]: illiterate group: ≤17, primary school group: ≤20, secondary school group: ≤22, and university group: ≤23.

Data Processing and Grouping

Overview

A total of 1000 assessments were collected initially. After quality control, 188 (18.8%) records were excluded due to missing data (n=107, 10.7%), nonstandardized test administration (n=35, 3.5%), or unqualified assessors (n=46, 4.6%). Before merging the hospital and community datasets, baseline demographic comparability was examined using an independent-samples *t* test (age) and chi-square test (gender) to minimize potential selection bias.

The final dataset (N=812) was stratified by educational attainment as follows: illiterate (n=108, 13.3%), primary school (n=105, 12.9%), secondary school (n=364, 44.8%), and university (n=235, 28.9%). Within each group, patient and control counts were recorded as illiterate (patients: n=60, 7.4%; and controls: n=48, 5.9%), primary (patients: n=53, 6.5%; and controls: n=52, 6.4%), secondary (patients: n=185, 22.8%; and controls: n=179, 22.0%), and university (patients: n=106, 13.1%; and controls: n=129, 15.9%). This preprocessing ensured a standardized data foundation for subsequent ML procedures.

Model Selection

A SVM classifier with a Radial Basis Function kernel was selected due to its suitability for medium-sized datasets and its capacity to model nonlinear decision boundaries. A stratified 5-fold cross-validation scheme was used to maximize data utilization while preserving the original case-control ratio within each educational group.

Hyperparameter Optimization

Hyperparameters were optimized through a grid search embedded within the cross-validation framework. The search grid included penalty parameter C: {0.1, 1, 5, 10, 100} and kernel coefficient γ: {0.001, 0.01, 0.1, 1}.

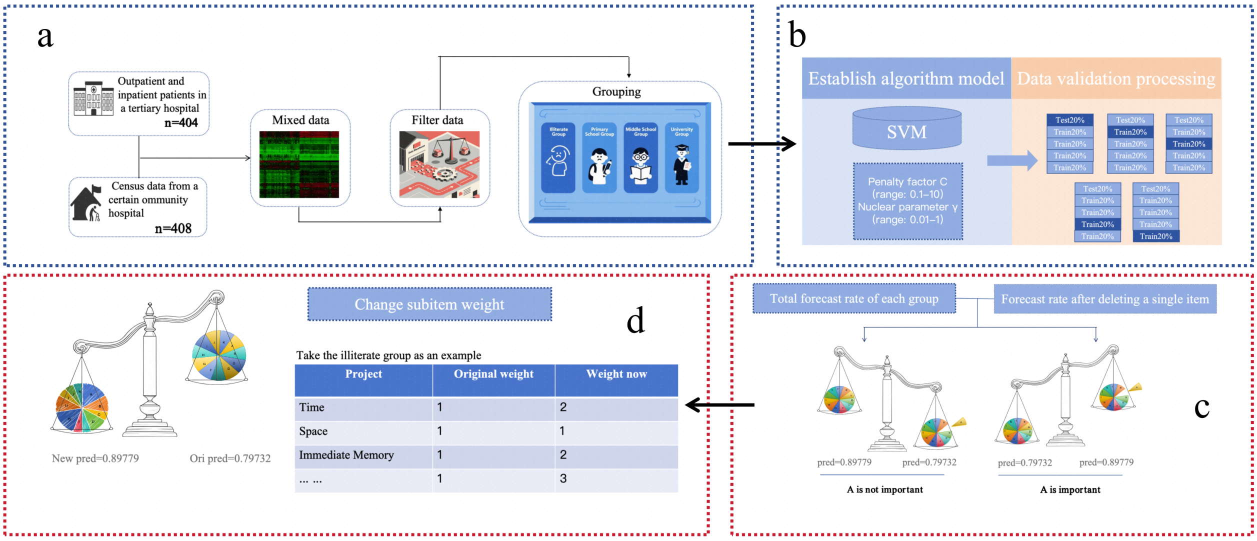
For each parameter pair, the model was trained on 4 folds and evaluated on the remaining fold, and the average accuracy across all 5 folds was used to identify the optimal configuration. Education-specific models were then retrained on the full dataset using the selected hyperparameters to ensure maximal predictive performance while avoiding information leakage.

Experimental Design

The overall experimental workflow is presented in Figure 1.

XSL•FO RenderX

Figure 1. Contribution of individual Mini-Mental State Examination (MMSE)j items across educational strata. Feature importance was quantified using the percentage change in cross-validated accuracy (Δ_i) after deleting each subitem. We operationally defined critical factors as items whose removal caused a performance decrease of $\Delta<-1.0$ and interference factors as items whose removal increased accuracy by $\Delta>+0.5$. SVM: support vector machine.



Take the illiterate group as an example

Project	Original weight	Weight now
Time	1	2
Space	1	1
Immediate Memory	1	2
... ...	1	3

The process began with data fusion and initial grouping: clinical and community data were merged to construct the initial dataset (n=812), and participants were stratified into 4 educational groups (ie, illiterate, primary school, secondary school, and university) to ensure subsequent analyses reflected education-specific cognitive characteristics. Next, a Support Vector Machine classifier with a radial basis function kernel was trained using a stratified 5-fold cross-validation scheme, preserving the original case-control ratio within each group. To quantify the contribution of each MMSE subitem, a systematic ablation procedure was performed. For each educational group, each subitem was removed in turn, and an SVM model was retrained. The feature contribution was quantified using the cross-validated change in prediction accuracy (Δ), calculated as:

$$\Delta=\text{Accuracypostdeletion}-\text{Accuracybaseline}$$

A negative Δ indicates a positive contribution (performance dropped when the item was removed), while a positive Δ suggests the item acted as noise. Finally, based on the subitem contribution profiles, an education-specific dynamic weighting scheme was constructed to adjust the relative importance of MMSE subitems before classification.

Other Statistical Techniques

Routine statistical analyses were performed using SPSS version 24.0 (IBM Corp), and ML model construction and evaluation were conducted using Python (version 3.10.10; Python Software Foundation)

For categorical variables, chi-square or Fisher exact tests were applied depending on expected frequencies. Continuous variables with a normal distribution were expressed as mean (SD) and compared using independent-samples *t* tests. Variables not following a normal distribution were summarized using median (IQR) and analyzed using nonparametric tests.

A 2-tailed $P<.05$ was considered statistically significant.

Results

Demographic Comparison

Following data cleansing, demographic characteristics of the clinical and community datasets were compared to verify baseline homogeneity before integration. As presented in Table 2, differences in age and gender distributions were not statistically significant ($P>.05$).

XSL•FO
RenderX

Table . Demographic characteristics and data composition stratified by educational background.

Characteristic	Illiterate (n=108)	Primary school (n=105)	Secondary school (n=364)	University (n=235)	*P* value
Age, mean (SD)	59.02 (8.5)	58.55 (7.8)	58.62 (8.1)	56.7 (7.4)	.05
Gender, n (%)					.37
Male	67 (62.0)	60 (57.1)	219 (60.2)	140 (59.6)	
Female	41 (38.0)	45 (42.9)	145 (39.8)	95 (40.4)	
Data source, n (%)					.12
Clinical[a] (patient)	60 (55.6)	53 (50.5)	185 (50.8)	106 (45.1)	
Community[b] (control)	48 (44.4)	52 (49.5)	179 (49.2)	129 (54.9)	

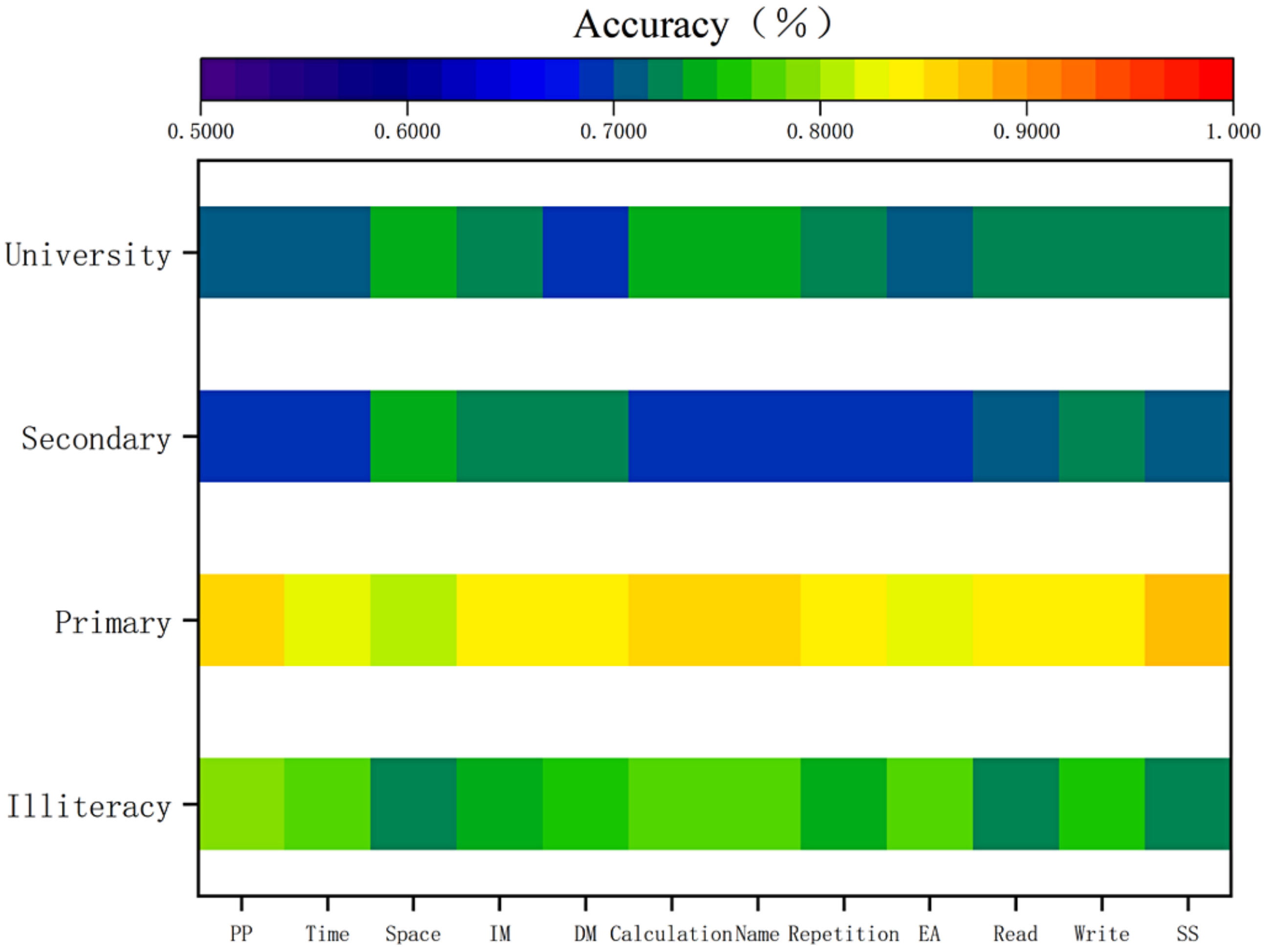
[a] Hospital dataset: patients diagnosed or evaluated in the tertiary hospital.

[b] Community dataset: participants recruited from community-based cognitive screening programs.

Item-Wise Contribution Analysis

To assess the contribution of individual MMSE items across educational groups, we conducted an item deletion experiment. The change in cross-validated accuracy (Δi) after removing each subitem was used to quantify its importance. A negative Δi (< -1.0) indicated a critical item whose removal impaired performance, while a positive Δi (> +0.5) indicated an interference item whose removal improved accuracy. The results of this analysis are visualized in Figure 2.

Figure 2. Heatmap visualization of feature contribution analysis. The color spectrum reflects prediction accuracy after item deletion. Cooler colors (blue and green) denote "critical features" (accuracy drop), while warmer colors (yellow and red) denote "neutral/interfering features" (accuracy stable/increase). The y-axis represents educational groups with the following sample sizes (N_patient/N_control): illiterate (60/48), primary school (53/52), secondary school (185/179), and university (106/129). DM: delayed memory; EA: execution ability; IM: immediate memory; PP: penmanship and praxis; SS: spatial structure.



XSL•FO
RenderX

Illiterate Group

In the illiterate cohort (baseline accuracy=78.83%), all subitems demonstrated positive contributions. The most influential items were spatial orientation (Δ=−6.58% postremoval), immediate memory (Δ=−3.85%), and writing (Δ=−2.86%). No interference items were identified.

Primary School Group

In the primary school group (baseline accuracy=85.71%), spatial construction functioned as an interference factor (Δ =+0.95%), while calculation showed a negligible effect (Δ=0). Removal of immediate memory, delayed recall, reading, or writing reduced accuracy to 83.81% (Δ=−1.90%).

Secondary School Group

In the secondary school group (baseline accuracy=69.78%), multiple interference factors were observed: spatial orientation (Δ=+3.57%), delayed recall (Δ=+3.02%), executive function (Δ=+2.19%), reading (Δ=+1.64%), immediate memory (Δ=+2.47%), writing (Δ=+0.20%), and spatial construction (Δ=+1.10%). Only temporal orientation (Δ=−0.83%) and calculation ability (Δ=−0.80%) demonstrated positive contributions.

University Group

In the university group (baseline accuracy=71.49%), five interference items were identified: spatial orientation (Δ=+2.98%), calculation (Δ=+2.98%), immediate memory (Δ=+1.70%), naming (Δ=+1.70%), and delayed recall (Δ=+1.42%). Repetition (Δ=−0.57%) and reading (Δ=−0.30%) showed small positive contributions.

Item Weight Analysis

On the basis of these contribution patterns, education-specific weighted coefficients were calculated for all MMSE subitems. The weighted diagnostic score was defined as follows:

$$S=\sum t-111\ (vi\times w1)$$

where S is the final weighted diagnostic score, v_i is the raw score of the i-th MMSE subitem, and w_i is the dynamic weighting coefficient assigned to the i-th subitem.

Classification was determined by comparing S to a group-specific classification threshold (T): S<T indicated cognitive impairment, and S≥T indicated normal cognition.

After normalization, the maximum possible weighted score was 60. Education-specific diagnostic thresholds were defined as follows: illiterate (*t*=30), primary (*t*=31), secondary (*t*=32), and university (*t*=33). As presented in Tables 3 and 4, accuracy improved across all educational strata after weight calibration.

Table . Model performance and weighting coefficients for core cognitive domains by educational group.

Group	Prediction accuracy before weighting	Prediction accuracy after weighting	Orientation (time)	Orientation (space)	Memory (immediate)	Memory (delayed)	Attention and calculation
Illiterate (%)	79.83	86.08	1	3	2	2	1
Primary school (%)	85.71	88.16	2	3	2	2	1
Secondary school (%)	69.78	69.83	2	3	2	1	2
University (%)	71.49	73.62	1	3	2	2	2

Table . Weighting coefficients for language and visuospatial domains by educational group.

Group	Naming	Repetition	Execution ability	Reading	Writing	Spatial structure
Illiterate (%)	3	3	2	3	2	3
Primary school (%)	1	2	3	2	2	1
Secondary school (%)	2	2	2	1	1	1
University (%)	2	3	1	1	2	2

Table 5 summarizes the diagnostic performance before and after dynamic weighting, including accuracy, sensitivity, specificity, AUC, and the statistical significance of performance changes. The largest improvement occurred in the illiterate group (Δ=+7.25%), whereas the secondary school group showed the smallest improvement (Δ=+0.05%).

Weight distributions differed across education levels. For basic cognitive domains, spatial orientation weights were uniformly set to 3 across all groups, and immediate memory weights remained at 2. In the illiterate group, naming, repetition, and spatial construction received the highest weights (all=3). In the primary school group, temporal orientation was weighted at 2 and spatial construction was weighted at 1. In the secondary school group, executive function and calculation were weighted at 2, while spatial construction remained at 1. In the university group, executive function received the highest weight (3), calculation was weighted at 2, and spatial construction, reading, and writing each received a weight of 1.

XSL•FO
RenderX

Table . Comparison of diagnostic performance metrics and statistical significance before and after dynamic weighting.

Education group and metric	Before weighting	After weighting	Improvement (Δ)	*P* value
Illiterate			7.25%	.06
Accuracy (%)	78.83	86.08		
Sensitivity (%)	81.45	84.92		
Specificity (%)	76.21	87.24		
AUC[a]	0.835	0.912		
Primary school			2.45%	.03
Accuracy (%)	85.71	88.16		
Sensitivity (%)	86.10	88.54		
Specificity (%)	85.32	87.78		
AUC	0.894	0.928		
Secondary school			0.05%	.46
Accuracy (%)	69.78	69.83		
Sensitivity (%)	70.25	70.31		
Specificity (%)	69.31	69.35		
AUC	0.752	0.753		
University			2.13%	.18
Accuracy (%)	71.49	73.62		
Sensitivity (%)	72.15	74.88		
Specificity (%)	70.83	72.36		
AUC	0.768	0.795		

[a] AUC: area under the curve.

Discussion

Principal Findings

This study developed an education-sensitive optimization framework for the MMSE using a dynamic SVM-based weighting strategy. Across 812 education-stratified participants, the model demonstrated that cognitive subitems contribute differently to diagnostic prediction depending on years of formal education. Dynamic weighting enhanced the performance of unimodal MMSE classifiers, particularly in low-literacy groups, and reduced education-related diagnostic bias. These findings refine the diagnostic utility of condensed cognitive screening tools and support their broader applicability in resource-limited settings.

Comparison With Prior Work

ML has been widely applied to cognitive assessment [32-35], yet existing studies typically use MMSE scores as features or labels without addressing education-driven measurement bias [36,37]. Prior condensed MMSE-only classifiers generally achieve accuracies of 72% to 85%, whereas multimodal MRI-, PET-, and biomarker-based models often exceed 90% [32,33]. Although our model remains unimodal, the dynamic weighting mechanism improved prediction by amplifying high-value signals and attenuating education-dependent interference, partially narrowing the performance gap with multimodal systems while sustaining low cost and operational simplicity.

Traditional MMSE optimization approaches—such as linear education corrections or standard ML models—lack the capacity to capture nonlinear, education-specific feature interactions [38]. In contrast, our framework integrates item deletion experiments with SVM-informed dynamic weighting to generate transparent, subgroup-adaptive coefficients. This approach offers an interpretable alternative to black box ensemble models and provides an effective mechanism for addressing item-level educational bias in cognitive assessment. A comprehensive comparison between the proposed dynamic weighted SVM model and existing cognitive assessment optimization strategies is presented in Table 6.

XSL•FO
RenderX

Table . Comparison of the proposed dynamic weighted Support Vector Machine (SVM) model with existing cognitive assessment optimization strategies.

Optimization strategy	Key methodology	Data requirements	Advantages	Limitations
Traditional linear correction (eg, MMSE[a]-E)	Linear regression and fixed point addition	MMSE scores+demographics	Simple calculation and clinically familiar	Ignored nonlinear "cognitive interference" and low precision
Standard machine learning (eg, RF[b], ANN[c])	Black box classification using raw scores	MMSE subitems or total scores	High classification accuracy	Low interpretability ("black box") and hard to explain to clinicians
Multimodal fusion models	Deep learning (CNN[d]/RNN[e]) integration	MRI[f]/PET[g] imaging+biomarkers+scale	Highest accuracy (>90%) and comprehensive pathology mapping	High cost, low accessibility in primary care, and complex deployment
This study (dynamic weighted SVM)	Nonlinear dynamic weighting via SVM	MMSE subitems only (low cost)	High interpretability (visible weights), education adaptive, and high accessibility	Accuracy is lower than multimodal models (feature limitation)

[a] MMSE: Mini-Mental State Examination.

[b] RF: random forest.

[c] ANN: artificial neural network.

[d] CNN: convolutional neural network.

[e] RNN: recurrent neural network.

[f] MRI: magnetic resonance imaging.

[g] PET: positron emission tomography.

Interpretation of Education-Specific Patterns

Education-stratified analyses revealed distinct, cognitive reliance patterns. The illiterate group depended heavily on orientation and immediate memory, consistent with basic functional domains commonly preserved in low-literacy populations. The primary school group demonstrated changes in weighting for temporal orientation, whereas secondary school participants exhibited varied interference across multiple domains and minimal model improvement (Δ=0.05), suggesting high within-group heterogeneity. University-educated participants showed stronger reliance on executive functioning and calculation. These patterns align with prior evidence that cognitive performance becomes increasingly distributed and strategy dependent with higher educational attainment [39,40].

The dynamic weighting strategy provided incremental accuracy improvements across all educational cohorts while reducing false-positive rates in lower-education groups. Subitem deletion further allowed visualization of feature importance (Figure 2), conceptually analogous to Shapley Additive Explanations–based explainability, thereby enhancing the transparency of SVM decision processes. Clinical validation by senior rehabilitation physicians confirmed that weight distributions reflected plausible neurocognitive patterns.

External validation using an independent community dataset (n=314) demonstrated stable performance (overall accuracy: 82.48% and AUC: 0.88), with the highest accuracy observed in the illiterate group (88.89%). These results support the generalizability and robustness of the proposed dynamic weighting model and mitigate concerns about overfitting.

Limitations

This study has several limitations. First, the sample was derived from a single geographic region, leading to educational imbalance across subgroups and potentially limiting generalizability. Broader multicenter sampling is needed. Second, demographic confounders beyond age and gender—such as socioeconomic status, occupational complexity, and urban-rural residence—were not available in the dataset, restricting the ability to fully disentangle education effects from related social determinants. Third, although the MMSE provides valuable screening utility, ceiling effects in highly educated individuals limit sensitivity to early cognitive decline. Integrating additional modalities, such as the Montreal Cognitive Assessment or imaging markers, may complement the proposed model. Finally, the SVM model relied on grid-search hyperparameter tuning; future research may explore AutoML-based optimization to improve efficiency and predictive performance.

Conclusions

This study proposes a dynamic SVM-based weighting framework that enhances the diagnostic fairness of MMSE-based cognitive screening across diverse educational backgrounds. By quantifying item-level contributions and adapting subitem weights for each education group, the method addresses a longstanding source of measurement bias in cognitive assessment. The approach retains the accessibility and scalability of condensed cognitive screening tools while improving prediction accuracy and interpretability. These findings provide a practical foundation for developing equitable cognitive assessment strategies, particularly in resource-limited regions.

XSL•FO
RenderX

Acknowledgments

The authors sincerely acknowledge all individuals and institutions that contributed to this work. This research was jointly supported by the Health Research Program of Anhui Province (project number AHWJ2023A10123) and the Anhui Provincial Teaching Research Project (project number 2022jyxm1851) titled "Research and Implementation of a Standardized Training Curriculum System for Rehabilitation Therapists Based on Competency-Based Education." Special thanks are extended to the funding bodies for their support.

Qi Zhang is co-corresponding author (email: ustczhangqi@163.com; phone +86 17333070277).

Funding

This work was supported by the Anhui Provincial Teaching Research Project (grant 2022jyxm1851), the Health Research Program of Anhui (grant AHWJ2023A10123), and the Anhui Provincial Graduate Academic Innovation Project “Optimization and Prediction of Alzheimer’s Disease-Related Cognitive Scales Based on Machine Learning” (grant 2024xscx009). This research received no commercial sponsorship or financial support that could compromise its objectivity. All authors declare no financial or proprietary interests relevant to this study. The funding agencies provided only financial support and had no involvement in study design, data collection and analysis, manuscript preparation, or submission for publication. The research team assumes full responsibility for the study's independence and scientific integrity.

Data Availability

The clinical dataset used in this study was obtained under ethical approval (2024-RE-431) and contains sensitive patient information; therefore, it cannot be made publicly available. Deidentified clinical records (n=404) may be shared upon reasonable request to the corresponding author under a data sharing agreement. Community screening data (n=408) are available with permission from the participating community health centers.

The complete source code for data preprocessing, Support Vector Machine modeling, dynamic weighting, and statistical analysis is openly available on GitHub [41].

Authors' Contributions

QL: Conceptualization, writing – original draft, formal analysis, data curation, methodology
CM: Conceptualization, writing – original draft, formal analysis, data curation, methodology
ML: Investigation, resources
SC: Data curation (clinical)
MY: Data curation (clinical and community)
LX: Data curation (clinical and community)
QZ: Supervision, project administration
MW: Supervision, project administration
All authors have reviewed and approved the final manuscript, agreeing to assume public responsibility for its content.
QZ is the co-corresponding author.

Conflicts of Interest

None declared.

References

1. Chun CT, Seward K, Patterson A, Melton A, MacDonald-Wicks L. Evaluation of available cognitive tools used to measure mild cognitive decline: a scoping review. Nutrients 2021 Nov 8;13(11):3974. [doi: 10.3390/nu13113974] [Medline: 34836228]
2. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
3. Jia X, Wang Z, Huang F, et al. A comparison of the Mini-Mental State Examination (MMSE) with the Montreal Cognitive Assessment (MoCA) for mild cognitive impairment screening in Chinese middle-aged and older population: a cross-sectional study. BMC Psychiatry 2021 Oct 4;21(1):485. [doi: 10.1186/s12888-021-03495-6] [Medline: 34607584]
4. Patnode CD, Perdue LA, Rossom RC, et al. Screening for cognitive impairment in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2020 Feb 25;323(8):764-785. [doi: 10.1001/jama.2019.22258] [Medline: 32096857]
5. Scheltens P, De Strooper B, Kivipelto M, et al. Alzheimer’s disease. Lancet 2021 Apr 24;397(10284):1577-1590. [doi: 10.1016/S0140-6736(20)32205-4] [Medline: 33667416]
6. Yin TT, Cao MH, Yu JC, et al. T1-weighted imaging-based hippocampal radiomics in the diagnosis of Alzheimer’s disease. Acad Radiol 2024 Dec;31(12):5183-5192. [doi: 10.1016/j.acra.2024.06.012] [Medline: 38902110]

XSL•FO
RenderX

7. Cardoso S, Barros R, Marôco J, de Mendonça A, Guerreiro M. Different MMSE domains are associated to cognitive decline and education. Appl Neuropsychol Adult 2024;31(4):533-539. [doi: 10.1080/23279095.2022.2041018] [Medline: 35234096]
8. Bird HR, Canino G, Stipec MR, Shrout P. Use of the Mini-Mental State Examination in a probability sample of a Hispanic population. J Nerv Ment Dis 1987 Dec;175(12):731-737. [doi: 10.1097/00005053-198712000-00005] [Medline: 3500273]
9. Koepsell TD, Kurland BF, Harel O, Johnson EA, Zhou XH, Kukull WA. Education, cognitive function, and severity of neuropathology in Alzheimer disease. Neurology 2008 May 6;70(19 Pt 2):1732-1739. [doi: 10.1212/01.wnl.0000284603.85621.aa] [Medline: 18160675]
10. Baker DW, Gazmararian JA, Sudano J, Patterson M. The association between age and health literacy among elderly persons. J Gerontol B Psychol Sci Soc Sci 2000 Nov;55(6):S368-S374. [doi: 10.1093/geronb/55.6.s368] [Medline: 11078114]
11. Ouvrard C, Berr C, Meillon C, et al. Norms for standard neuropsychological tests from the French CONSTANCES cohort. Eur J Neurol 2019 May;26(5):786-793. [doi: 10.1111/ene.13890] [Medline: 30575234]
12. Aprahamian I, Martinelli JE, Cecato J, Yassuda MS. Screening for Alzheimer's disease among illiterate elderly: accuracy analysis for multiple instruments. J Alzheimers Dis 2011;26(2):221-229. [doi: 10.3233/JAD-2011-110125] [Medline: 21593559]
13. Rezende GP, Cecato J, Martinelli JE. Cognitive abilities screening instrument-short form, Mini-Mental State Examination and functional activities questionnaire in the illiterate elderly. Dement Neuropsychol 2013;7(4):410-415. [doi: 10.1590/S1980-57642013DN74000009] [Medline: 29213866]
14. Katzman R, Zhang MY, et al. A Chinese version of the Mini-Mental State Examination; impact of illiteracy in a Shanghai dementia survey. J Clin Epidemiol 1988;41(10):971-978. [doi: 10.1016/0895-4356(88)90034-0] [Medline: 3193141]
15. Weiss BD, Reed R, Kligman EW, Abyad A. Literacy and performance on the Mini-Mental State Examination. J Am Geriatr Soc 1995 Jul;43(7):807-810. [doi: 10.1111/j.1532-5415.1995.tb07057.x] [Medline: 7602038]
16. Paddick SM, Gray WK, McGuire J, Richardson J, Dotchin C, Walker RW. Cognitive screening tools for identification of dementia in illiterate and low-educated older adults, a systematic review and meta-analysis. Int Psychogeriatr 2017 Jun;29(6):897-929. [doi: 10.1017/S1041610216001976] [Medline: 28274299]
17. Pellicer-Espinosa I, Díaz-Orueta U. Cognitive screening instruments for older adults with low educational and literacy levels: a systematic review. J Appl Gerontol 2022 Apr;41(4):1222-1231. [doi: 10.1177/07334648211056230] [Medline: 34856843]
18. Brucki SMD, Mansur LL, Carthery-Goulart MT, Nitrini R. Formal education, health literacy and Mini-Mental State Examination. Dement Neuropsychol 2011;5(1):26-30. [doi: 10.1590/S1980-57642011DN05010005] [Medline: 29213716]
19. Carnero-Pardo C. Should the Mini-Mental State Examination be retired? Neurologia 2014 Oct;29(8):473-481. [doi: 10.1016/j.nrl.2013.07.003] [Medline: 24140158]
20. Cai H, Shao Y, Liu XY, et al. Interpretable prognostic modeling for long-term survival of Type A aortic dissection patients using Support Vector Machine algorithm. Eur J Med Res 2025 Apr 15;30(1):277. [doi: 10.1186/s40001-025-02510-w] [Medline: 40229872]
21. Luo W. Predicting cervical cancer outcomes: statistics, images, and machine learning. Front Artif Intell 2021;4:627369. [doi: 10.3389/frai.2021.627369] [Medline: 34164615]
22. Cai F, Cherkassky V. Generalized SMO algorithm for SVM-based multitask learning. IEEE Trans Neural Netw Learn Syst 2012 Jun;23(6):997-1003. [doi: 10.1109/TNNLS.2012.2187307] [Medline: 24806769]
23. Huang C, Chung FL, Wang S. Multi-view L2-SVM and its multi-view core vector machine. Neural Netw 2016 Mar;75:110-125. [doi: 10.1016/j.neunet.2015.12.004] [Medline: 26773824]
24. Mastropietro A, Feldmann C, Bajorath J. Calculation of exact Shapley values for explaining Support Vector Machine models using the radial basis function kernel. Sci Rep 2023 Nov 10;13(1):19561. [doi: 10.1038/s41598-023-46930-2] [Medline: 37949930]
25. Jasodanand VH, Bellitti M, Kolachalama VB. An AI-first framework for multimodal data in Alzheimer's disease and related dementias. Alzheimers Dement 2025 Sep;21(9):e70719. [doi: 10.1002/alz.70719] [Medline: 40983947]
26. Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. Neuroimage 2017 Jul 15;155:530-548. [doi: 10.1016/j.neuroimage.2017.03.057] [Medline: 28414186]
27. Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage 2011 Apr 1;55(3):856-867. [doi: 10.1016/j.neuroimage.2011.01.008] [Medline: 21236349]
28. Jasodanand VH, Kowshik SS, Puducheri S, et al. AI-driven fusion of multimodal data for Alzheimer's disease biomarker assessment. Nat Commun 2025 Aug 11;16(1):7407. [doi: 10.1038/s41467-025-62590-4] [Medline: 40789853]
29. Dincer A, Gordon BA, Hari-Raj A, et al. Comparing cortical signatures of atrophy between late-onset and autosomal dominant Alzheimer disease. Neuroimage Clin 2020;28:102491. [doi: 10.1016/j.nicl.2020.102491] [Medline: 33395982]
30. Qiu S, Miller MI, Joshi PS, et al. Multimodal deep learning for Alzheimer's disease dementia assessment. Nat Commun 2022 Jun 20;13(1):3404. [doi: 10.1038/s41467-022-31037-5] [Medline: 35725739]
31. Castellano G, Esposito A, Lella E, Montanaro G, Vessio G. Automated detection of Alzheimer's disease: a multi-modal approach with 3D MRI and amyloid PET. Sci Rep 2024 Mar 3;14(1):5210. [doi: 10.1038/s41598-024-56001-9] [Medline: 38433282]

XSL•FO
RenderX

32. Vyas A, Aisopos F, Vidal ME, Garrard P, Paliouras G. Identifying the presence and severity of dementia by applying interpretable machine learning techniques on structured clinical records. BMC Med Inform Decis Mak 2022 Oct 17;22(1):271. [doi: 10.1186/s12911-022-02004-3] [Medline: 36253849]
33. Feng F, Wang P, Zhao K, et al. Radiomic features of hippocampal subregions in Alzheimer’s disease and amnestic mild cognitive impairment. Front Aging Neurosci 2018;10:290. [doi: 10.3389/fnagi.2018.00290] [Medline: 30319396]
34. Battista P, Salvatore C, Castiglioni I. Optimizing neuropsychological assessments for cognitive, behavioral, and functional impairment classification: a machine learning study. Behav Neurol 2017;2017:1850909. [doi: 10.1155/2017/1850909] [Medline: 28255200]
35. Khatun S, Morshed BI, Bidelman GM. A single-channel EEG-based approach to detect mild cognitive impairment via speech-evoked brain responses. IEEE Trans Neural Syst Rehabil Eng 2019 May;27(5):1063-1070. [doi: 10.1109/TNSRE.2019.2911970] [Medline: 30998476]
36. Wu Y, Jia M, Xiang C, Lin S, Jiang Z, Fang Y. Predicting the long-term cognitive trajectories using machine learning approaches: a Chinese nationwide longitudinal database. Psychiatry Res 2022 Apr;310:114434. [doi: 10.1016/j.psychres.2022.114434] [Medline: 35172247]
37. Martin SA, Townend FJ, Barkhof F, Cole JH. Interpretable machine learning for dementia: a systematic review. Alzheimers Dement 2023 May;19(5):2135-2149. [doi: 10.1002/alz.12948] [Medline: 36735865]
38. Kantayeva G, Lima J, Pereira AI. Application of machine learning in dementia diagnosis: a systematic literature review. Heliyon 2023 Nov;9(11):e21626. [doi: 10.1016/j.heliyon.2023.e21626] [Medline: 38027622]
39. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 8;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
40. Stern Y. Cognitive reserve in ageing and Alzheimer’s disease. Lancet Neurol 2012 Nov;11(11):1006-1012. [doi: 10.1016/S1474-4422(12)70191-6] [Medline: 23079557]
41. Dynamic-SVM-weighting-for-MMSE-optimization. GitHub. URL: https://github.com/masC1962/Dynamic-SVM-Weighting-for-MMSE-Optimization [accessed 2026-01-05]

Abbreviations

AUC: area under the curve
ML: machine learning
MMSE: Mini-Mental State Examination
MRI: magnetic resonance imaging
PET: positron emission tomography
SVM: support vector machine

Edited by S Munce; submitted 30.Jun.2025; peer-reviewed by AM Hasan, H Kan, S Neufang, ZY Jiang; accepted 23.Dec.2025; published 25.Feb.2026.

Please cite as:
Liu Q, Ma C, Liu M, Chen S, Yu M, Xia L, Zhang Q, Wu M
Reducing Educational Bias in Cognitive Assessment via Dynamic Support Vector Machine Weighting: Validation Study on an Education-Stratified Dataset
JMIR Rehabil Assist Technol 2026;13:e79841
URL: https://rehab.jmir.org/2026/1/e79841
doi:10.2196/79841

© Qing Liu, Chi Ma, Mengyuan Liu, Suhui Chen, Mengting Yu, Lijuan Xia, Qi Zhang, Ming Wu. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Design Implications of Comfort and Usability of Manual Stairclimbing Wheelchair: Ergonomic Assessment and Pilot Study Using Surface Electromyography Inputs

Abhishek Verma, PhD; Rohit Kumar, M Tech; J Ramkumar, PhD

Department of Design, Indian Institute of Technology Kanpur, DJAC Building, Kanpur, India

Corresponding Author:
J Ramkumar, PhD
Department of Design, Indian Institute of Technology Kanpur, DJAC Building, Kanpur, India

Abstract

Background: Stairclimbing wheelchairs offer enhanced mobility for users navigating multilevel environments, yet limited research addresses the ergonomics of lever propulsion-based stair climbing mechanisms. Comprehensive ergonomic assessment integrating both subjective user feedback and objective biomechanical analysis is essential for optimizing assistive device design for comfort and usability.

Objective: This pilot study aims to assess the ergonomic design of a transformable stair-climbing wheelchair through a dual-methodology approach, evaluating plane surface movement accessibility and quantifying muscle activation patterns during lever-propelled stair-climbing operations using surface electromyography (sEMG).

Methods: This 2-part study involved anthropometric measurements from 20 male participants to establish design parameters using 5th and 95th percentile values. Part A assessed plane surface movement with 9 participants (7 healthy, 2 with paraplegia) navigating a simulated urban course featuring a 5° ramp, a 90° turn, and narrow passages across 3 trials. Task completion times and subjective ride easiness ratings were recorded. Part B used a Taguchi-based fractional factorial design to evaluate 3 ergonomic factors, including torso angle (λ), lever distance (L), and lever orientation (ψ), across 7 healthy participants. Maximum voluntary contraction (MVC) was measured for 4 muscles, including biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT).

Results: In Part A, the ramp and 90° turn proved most challenging due to the wheelchair's 65 kg weight and large turning radius (~1450 mm). Driving control scored highest (6/10), while comfort scored lowest due to the tilted seat design. In Part B, a straight torso (λ=0°) consistently reduced muscle strain, particularly for brachioradialis. A lever distance of approximately 50 mm and a neutral to slightly supinated orientation (ψ=0°-30°) optimized muscle effort. Interaction effects revealed high strain configurations (λ=45°; L=100 mm; ψ=−30°) exceeding 75 MVC, while optimal settings reduced strain to approximately 50 MVC.

Conclusions: Optimal ergonomic parameters of λ=0°, L=37.5 mm, and ψ=15° are recommended to minimize fatigue and enhance user comfort. Design improvements should prioritize weight reduction, compact form factor for maneuverability, and adjustable seat tilt. The modular wheelchair design permits customization for diverse user populations. Future research should include larger, gender-diverse participant groups and real-world validation studies.

(JMIR Rehabil Assist Technol 2026;13:e78965) doi:10.2196/78965

KEYWORDS

stairclimbing wheelchair; ergonomic design; surface electromyography; taguchi design; maximum voluntary contraction; likert scale

Introduction

Ergonomic analysis of any newly designed wheelchair or assistive technology intervention is crucial for assessing user comfort and efficiency using already established methodologies. Subjective evaluations, such as Likert scales and visual analog scale (VAS), are used to evaluate user perception about usability and comfort of an assistive technology. Cowan et al [1], assessed paraplegic users, using questionnaires on maneuverability, stability, and comfort, alongside time and kinetic data, providing a comprehensive view of user experience. Van der Woude et al [2], reviewed research and innovation in manual wheelchairs, focusing on rehabilitation, sports, daily life, and health, emphasizing physiological responses during propulsion. However, most studies focus on flat surfaces, with limited research specifically addressing the ergonomics of lever-propulsion–based stair-climbing, indicating a gap that this study aims to address. Dubowsky et al [3] compared kinematics, kinetics, and electromyography in able-bodied and paraplegic

XSL•FO
RenderX

users, analyzing biceps brachii and triceps brachii during propulsion, highlighting axle placement effects.

Studies on specific design variables that may affect muscle activity are crucial for design optimization. Marrow et al [4] examined the influence of grip position on upper-extremity posture and muscle activity, linking pronation, supination, and neutral positions along with torso inclination to muscle demands. Kurup et al [5] looked at push handle height, affecting the distance from the shoulder to the lever, and its impact on muscle activity. Verma et al [6] reviewed how different ergonomic and functional considerations inform manual wheelchair design, emphasizing wheelchair dimensions based on the stair-climbing mechanisms. Among all these studies based on the wheelchair functional or comfort assessments, there is an implicit absence of pilot studies on newly developed stair-climbing technologies that use electromyography to enhance comfort and usability, highlighting the role of electromyography in designing for reduced muscle strain. Such a type of approach has been carried out in this research to discuss implications of integrating electromyography data into design processes for minimizing muscle strain and improving overall user experience in the case of stair-climbing operations.

The ergonomic evaluation of stair-climbing wheelchair functions is necessary for further iterations into the wheelchair design developed by Verma et al [7]. For an assistive technology such as this stair-climbing wheelchair, the goal is to offer assistance in activities of daily living such as mobility, nutrition, personal care, and employment and also act as an enabler of social interactions such as family or relationship responsibilities and interpersonal or community relationships, etc. Both subjective and objective information about the effectiveness of the wheelchair is required for improvement in this design of the stair-climbing wheelchair for both plane surface movement and stair-climbing function. This research focuses on understanding the wheelchair use case for a short duration. Any long-term consequences of the wheelchair design require a more detailed study spanning over years and are out of the scope of this research. This research can be broadly divided into 2 parts, the first one to collect and assess wheelchair mobility on a plane surface when the wheelchair is transformed into a plane-surface wheelchair, and the second part addresses the ergonomics of stair-climbing when the participant performs the pull-push action through a set of levers to rotate stair-climbing wheel.

This pilot study uses an integrated dual-methodology approach to comprehensively evaluate the stair-climbing wheelchair design. Part A focuses on ergonomic assessment of plane surface movement through anthropometric measurements, functional mobility testing, and subjective user feedback. Part B complements this assessment through objective biomechanical analysis using surface electromyography (sEMG) to quantify muscle activation patterns during stair-climbing operations. These 2 methodological components are mutually reinforcing; the subjective assessments from Part A identify user-perceived comfort and usability issues, while the objective sEMG measurements from Part B provide physiological validation of these subjective experiences and reveal underlying biomechanical factors that may not be immediately apparent to users. This integrated approach enables both immediate user feedback and deeper understanding of the physiological demands imposed by specific design parameters, facilitating evidence-based design optimization that addresses both perceived comfort and actual physical strain. The combination of these methodologies provides a comprehensive framework for assistive technology assessment that can inform future design iterations and clinical recommendations.

Methods

Study Design Overview

This 2-part pilot study used a dual-methodology approach to evaluate the ergonomic design of a transformable stair-climbing wheelchair. Part A assessed plane surface movement accessibility through anthropometric measurements, functional mobility testing on a simulated urban course, and subjective user feedback. Part B evaluated lever propulsion ergonomics during stair-climbing using sEMG to quantify muscle activation patterns under varying ergonomic configurations determined by a Taguchi-based fractional factorial design of experiments.

Part A: Ergonomic Assessment of Plane Surface Movement

Participants and Anthropometric Measurements

Anthropometric measurements were collected from 20 male participants (average age 27.5, SD 2.76 years) to establish design parameters for the modular wheelchair. This demographic was selected to determine conservative upper-bound dimensions, recognizing that male participants generally have larger anthropometric dimensions. Measurements were taken with the user seated in an optimal position as shown in Figure 1. The dimensions collected included hip width, buttock-popliteal length, popliteal height, subscapular height, shoulder height, elbow height, knee height, shoulder width, elbow fingertip length, upper limb length, and shoulder grip length. These measurements were used to develop a fit and reach matrix according to user comfort, with 5th and 95th percentile values calculated for design purposes (Table 1).

Figure 1. Anthropometric measurements are required to design wheelchair seating and reach dimensions.

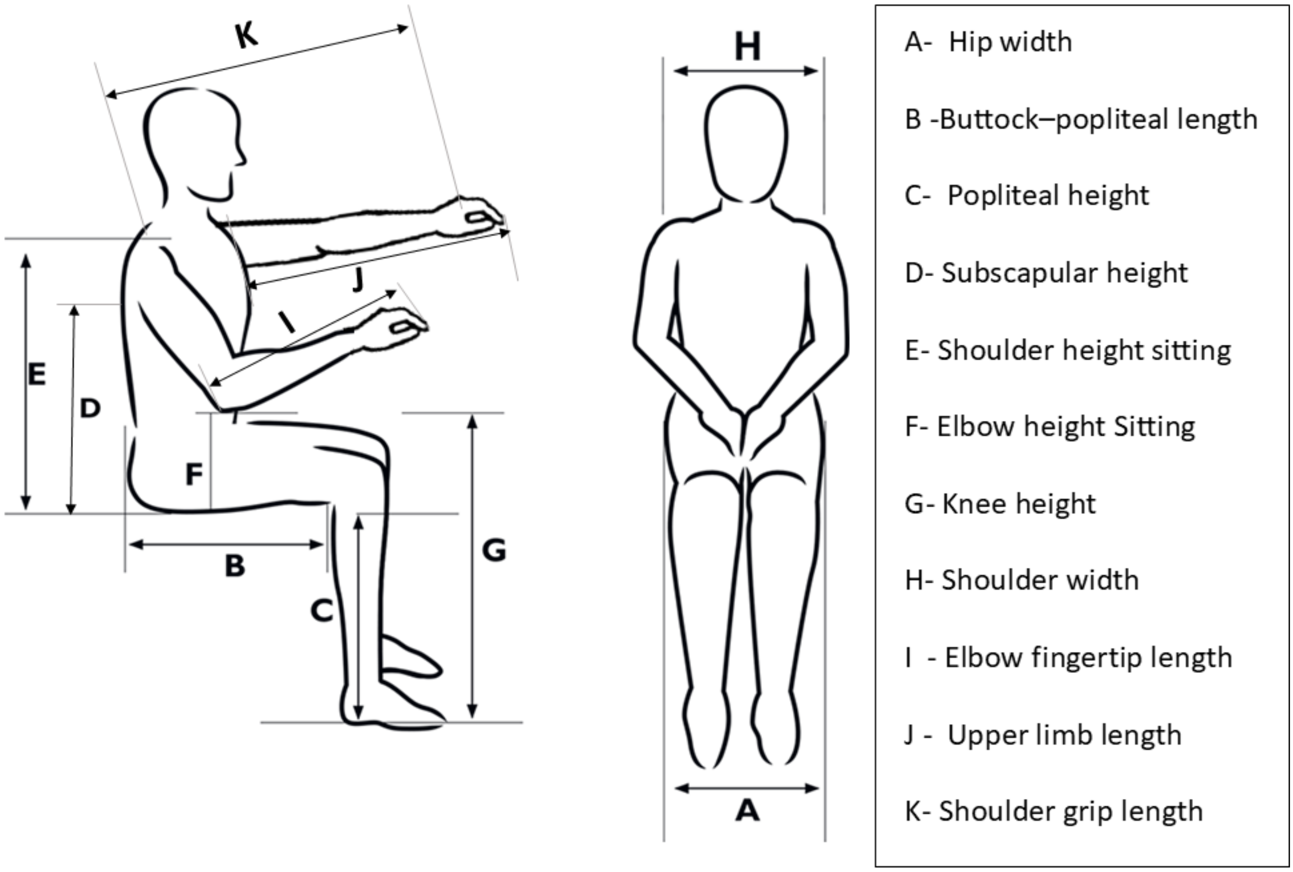


Table . Anthropometric data recorded for wheelchair design (dimensions in centimeters).

Annotation	Anthropometric feature	Mean (SD)	5th -95th percentile
—[a]	Age (years)	27.5 (2.76)	23.95-31.1
—	Height (cm)	173.25 (3.46)	168-178
—	Weight (kg)	71.75 (4.98)	63.95-78.15
A	Hip width (cm)	38.45 (3.26)	34-42.15
B	Buttock-Popliteal length	51 (4.41)	42.95-56.05
C	Popliteal height (cm)	42.1 (2.67)	38.95-47.05
D	Subscapular height (cm)	45.75 (2.22)	43-50.05
E	Shoulder height (cm)	60.85 (3.28)	54.95-60.85
F	Elbow height (cm)	24.85 (2.13)	22-28.05
G	Knee height (cm)	53.95 (4.50)	48.95-60.15
H	Shoulder width (cm)	55.45 (4.5)	50.85-63.15
I	Elbow fingertip length (cm)	36.35 (2.75)	33-41.05
J	Upper limb length (cm)	85.4 (4.28)	78.95-92.1
K	Shoulder grip length (cm)	75.8 (5.01)	67.95-82.15

[a] Not applicable.

Functional Anthropometry for Reach and Clearances

The wheelchair design dimensions were based on standards and guidelines from the Central Public Works Department (CPWD), the Office of Chief Commissioner for Persons with Disabilities, IS 4963 (building code standard), and IS 7454 (wheelchair specification standard). The clearance dimensions are summarized in Table 2.

Table . Clearance dimensions for the wheelchair based on standards and guidelines (in mm).

Reach ranges	Standards and guidelines			
	IS-7454[a] [8]	CPWD[b] [9]	CCD[c] [10]	ISO[d] [11]
Forward high	1350 - 1600	1200	1200	1200
Forward low	Not specified	400	400	400
Forward obstructed high	715 - 830	1100	1100	Not specified
Lateral high	1350 - 1770	1300	1300	Not specified
Lateral low	Not specified	250	250	Not specified
Lateral obstructed high	Not specified	1200	1200	Not specified
Maneuvering clearances				
Circular diameter-360° turn	1500	1500 - 2000	1800	1500

[a] IS: Indian Standard in architecture.

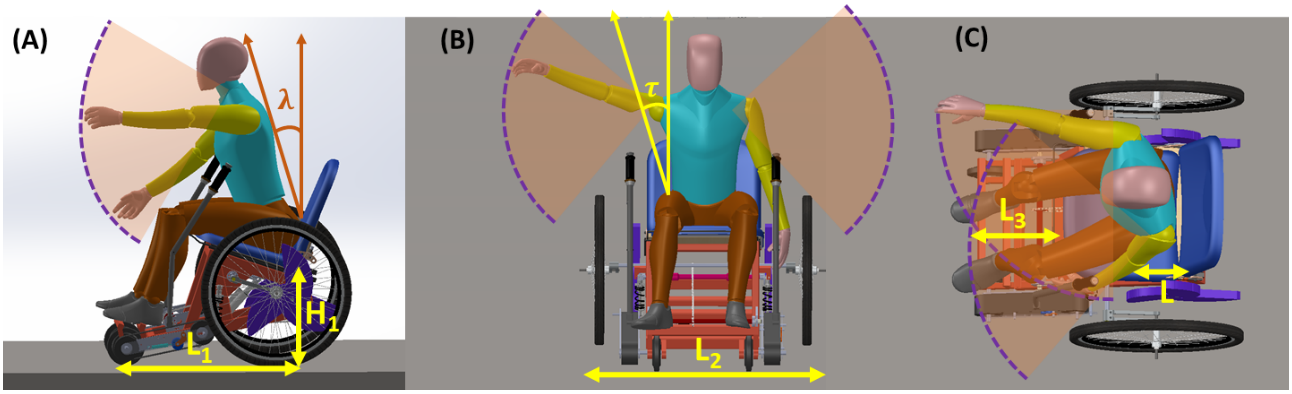
[b] CPWD: Central Public Works Department.

[c] CCD: Chief Commissioner-Disability.

[d] ISO: International Standard Organization.

The wheelchair dimensions for clearances and reach were based on the maximum of maximum (95th percentile) and minimum of minimum (5th percentile) human anthropometric dimensions, respectively. Wheelchair dimensions such as seat length and backrest angle were designed to be adjustable over the 5th and 95th percentile range using an angle-adjustable backrest and a slide-lock mechanism. The reach space design was based on wheelchair height (H1) for vertical reaches and length L1 as shown in Figure 2A-2B, width L2 (Figure 2B), and distance L3 (Figure 2C) for horizontal distances, calculated using a digital human model (DHM).

Figure 2. Maximum reach illustrated by a digital human model in (A) sagittal plane, (B) coronal plane, and (C) transverse plane.



The primary reach relationships derived from the DHM are given by equations (1–4):

Forward high: (1)

$$H1+E\times\cos\lambda+K\times\sin(\pi/4)\geq 1200mm$$

Forward low: (2)

$$H1+E\times\cos\lambda-K\times\sin(\pi/4)\geq 400mm$$

Lateral high: (3)

$$H1+E+K\times\sin(\pi/4)\geq 1300mm$$

Lateral low: (4)

$$H1+E\times\sin(\tau)-K\times\sin(\pi/4)\geq 250mm$$

Where H1 is the height of the wheelchair from the ground to the lowest point on the seat (500 mm), E is the shoulder height from the seat, λ is the torso angle the participants make with the coronal plane, K is the shoulder grip length, and τ is the angle the torso makes with the sagittal plane.

The minimum turning distance required for a 360° turn was calculated using Equation (5):

$$Tcircular=2\times\pi\times L12+L22 \quad (5)$$

Where L12+L22 is the turning radius of the wheelchair

For this prototype, L1=600 mm and L2=950 mm

XSL•FO
RenderX

Therefore, the distance required for a 360^o turn calculated using DHM will be,

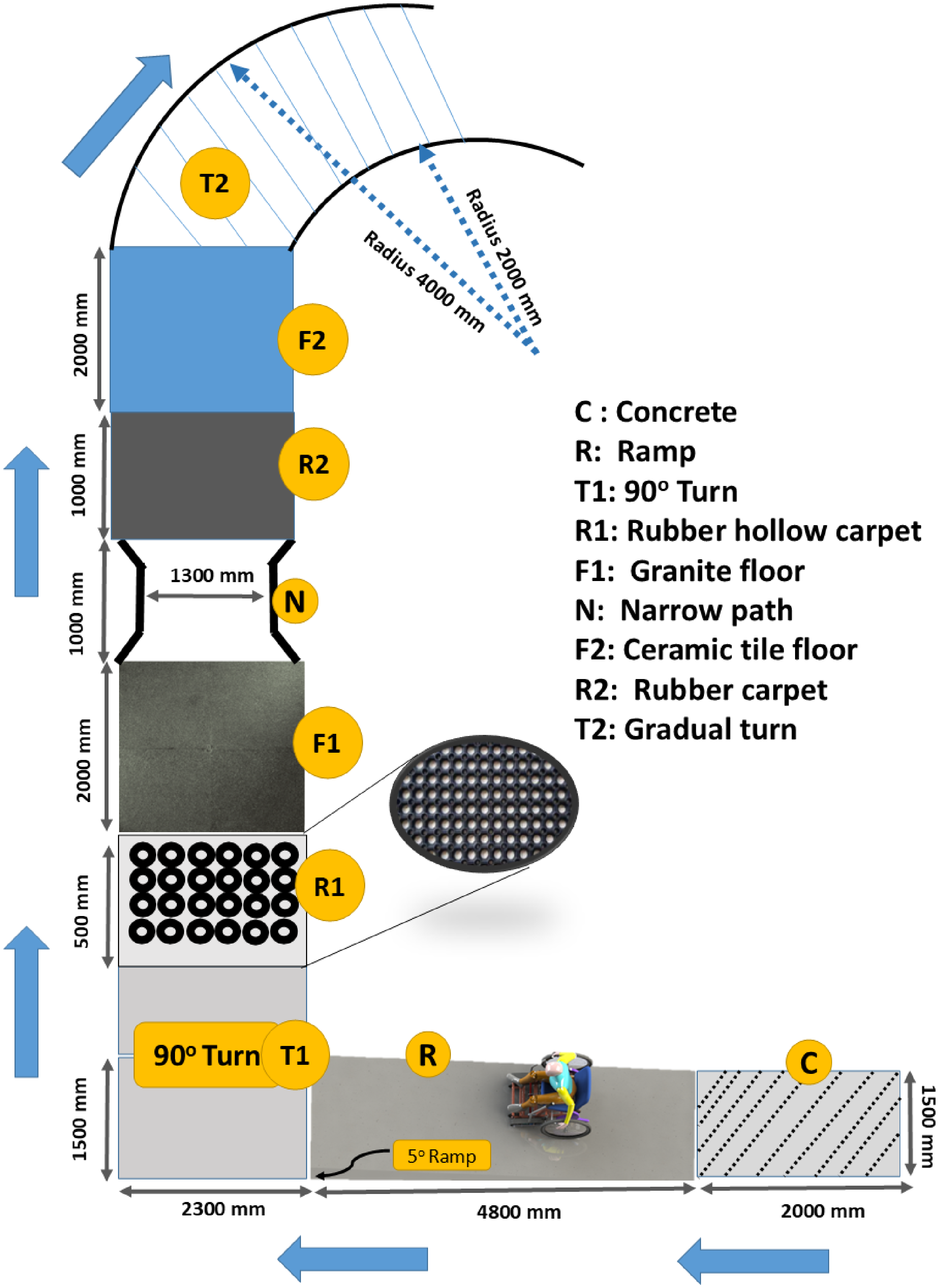
$T_{circular}$=4870 mm.

Simulated Course Design and Protocol

The manual stair-climbing wheelchair is a dual-purpose wheelchair designed to provide multidimensional accessibility on both stairs and plane surfaces. To test the efficacy of wheelchair driving on plane surfaces, a simulated course was designed to test maneuverability and control. The simulated course contained architectural hurdles that a wheelchair user faces in an urban setting, including a 5° ramp followed by a sharp 90° turn and then converging and diverging spaces. The course is illustrated in Figure 3.

XSL•FO
RenderX

Figure 3. Wheelchair course designed for ergonomic design study of plane surface movement configuration of wheelchair.



This pilot study used n=9 participants for plane surface assessment (7 healthy, 2 with paraplegia), following established guidelines for feasibility studies in assistive technology research [12,13]. Two participants with paraplegia provided critical

XSL•FO
RenderX

real-world user validation. The course was designed as a controlled experimental environment to ensure participant safety and enable systematic data collection, including multiple stair configurations with varying heights (15 - 20 cm) and depths (25 - 30 cm), curved pathways, simulated obstacles, and various surface transitions representative of typical urban environments.

The subjective descriptors used for the response for each riding task using Likert scales were (1) strongly favorable (score=4), (2) moderately favorable (score =3), (3) moderately unfavorable (score=2), and (4) strongly unfavorable (score=1). Task completion times were recorded for each of the 9 tasks across 3 trials. After the completion of all tasks, a subjective survey of the whole task was taken on a visual satisfaction scale ranging from 0 (Very negative) to 10 (Very positive) for overall ride comfort, stability, driving control, perception, and maneuverability.

Part B: Lever Propulsion sEMG Study

Muscles Under Study

The biomechanical evaluation of arm and shoulder muscles was performed to identify and mitigate shoulder pain and fatigue generated during lever propulsion. Use of sEMG is a prominent method for recording muscle activation artifacts to understand muscle effort, muscle activation timing, and muscle fatigue. The activation timing study of a muscle group was performed to understand the loading profile of muscle at a specific interval in the work cycle [14]. Surface electromyography also helps in identifying muscles under fatigue by observing an increase in amplitude and decrease in median frequency of the bioelectrical signal [15].

The upper limb and shoulder are the most common sites for any possible stress injury or musculoskeletal disorder due to long-term cyclic loading during lever-based propulsion [16,17]. The muscle groups under study for this specific research go under cyclic dynamic interactions and muscle fiber contractions during the lever-based propulsion cycle and are anisometric (concentric and eccentric) in nature. Therefore, the shoulder and arm joint angles are also required to be recorded while recording sEMG activity. The recorded electromyography signal has artifacts from the velocity- and angle-dependent muscle excitation of the shoulder and elbow extensors and flexors. The muscle activation of the extensor and flexor groups will be significant during both actions of pulling and pushing. But, since only pulling stroke is responsible for stair climbing, only those muscles will show significant dynamic electromyography activity that are associated with pulling action (elbow flexors and shoulder extensors muscle groups).

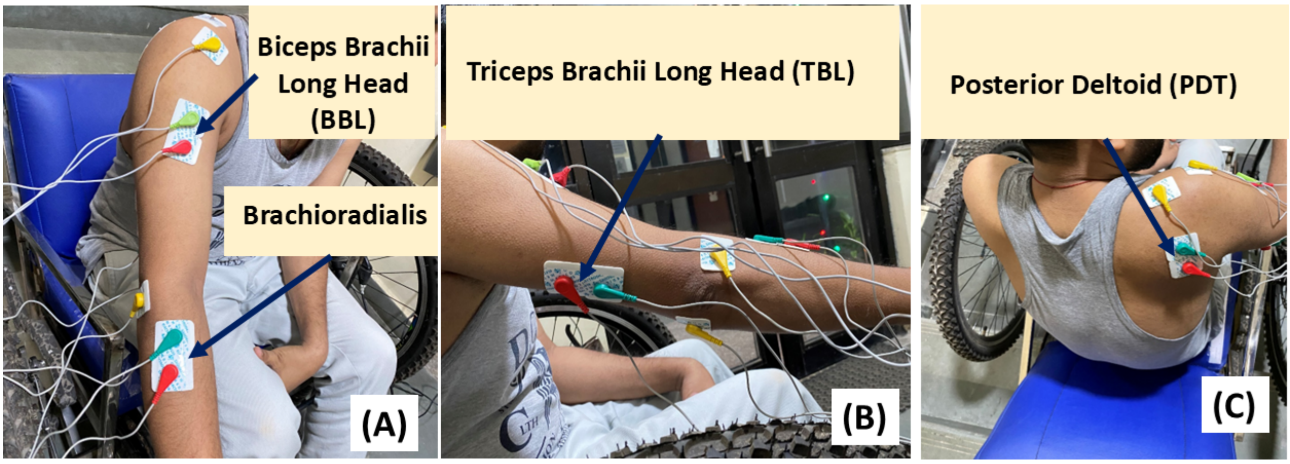
The action of lever propulsion in wheelchairs and its physiological effects are reported by only a few researchers, as the wheelchairs developed with a lever mechanism are very much limited, and in consequence, electromyography-based ergonomic evaluations based on them. There are few studies about lever-based propulsion effort, but they are limited to physiological measurements such as oxygen uptake, respiration ratio, minute ventilation, heart rate, etc [18-20]. With the advent of surface electromyography recording and processing techniques, the idea of using electromyography as a dominant physiological parameter that can influence design is increasing. Push-rim–based propulsion of a plane-surface wheelchair has been highly explored using surface electromyography studies on dominant-side muscle groups such as the deltoid group, triceps group, and bicep groups. As reported in the literature, there are some overlapping muscle groups engaged during both lever propulsion and hand-rim propulsion, such as biceps brachii, posterior deltoid, triceps, pectoralis major, etc [21-23]

The action of lever propulsion is a push-pull action that requires elbow (flexion and extension) and shoulder (flexion and extension) movement. The action of pulling requires flexion of the elbow and extension of the shoulder muscles and vice versa for pushing. The pulling stroke is responsible for rotating the stair-climbing wheel on the staircase. The pushing stroke (moving the lever away from the chest) is idle and requires less effort compared to pulling. Tomiak et al [24], in their research related to bimanual rowing movement, which mimics a lever push-pull activity, group together the pulling and pushing (returning) muscle groups as elbow and shoulder muscles. The primary goal of the ergonomic evaluation in this section is to identify extreme joint loading forces by trying different posture configurations defined by ergonomic factors. Pushing and pulling tasks can cause extreme lower back and upper joint loading, leading to musculoskeletal disorders (MSDs) [25,26].

Four muscle groups were selected for sEMG recording: (1) biceps brachii long head (BBL), a primary muscle involved in flexing the elbow, selected due to availability of higher density of motor neurons (Figure 4A); (2) triceps brachii long head (TBL), located on the back of the upper arm, responsible for extending the elbow during the pushing stroke (Figure 4B); (3) brachioradialis, responsible for forearm and elbow flexion necessary for pulling, with secondary function of supination. and pronation for holding the lever (Figure 4B); and (4) posterior deltoid (PDT), located at the shoulder, extends during shoulder extension required for pulling the lever toward the chest (Figure 4C).

XSL•FO
RenderX

Figure 4. Nickel-chromium (Ni-Cr) alloy electrodes stuck on muscle surface using silver-silver chloride electrode (Ag-Cl) hydrogel (A) location of brachioradialis and BBL muscle group recording, (B) triceps brachii long head (TBL), and (C) posterior deltoid (PDT) muscle group recording.

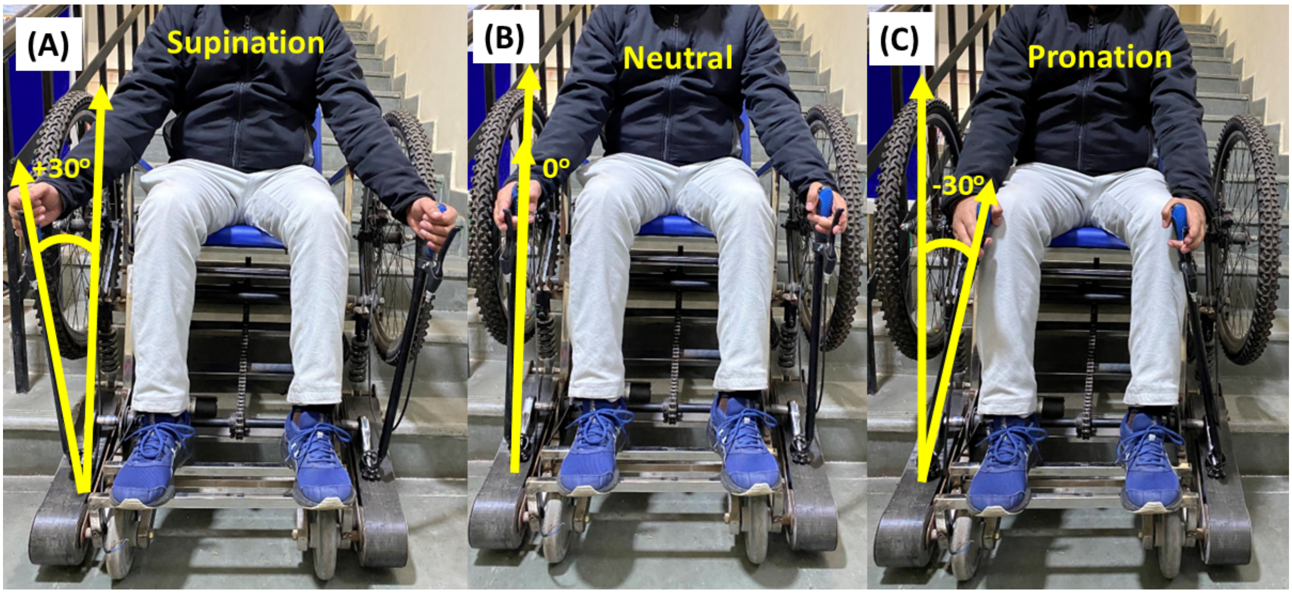


Ergonomic Factors

Three ergonomic factors were evaluated in this study:

(1) Lever orientation (ψ): the range of motion for supination and pronation was limited to ±30°, where supination is +30° (Figure 5A), the neutral position is 0° (Figure 5B), and pronation is −30° (Figure 5C). Although Sullivan et al [27] reported that up to a certain degree of pronation, subjective discomfort is lesser and higher effort can be achieved compared to supination and neutral hand orientation [28], objective evaluation using electromyography was required.

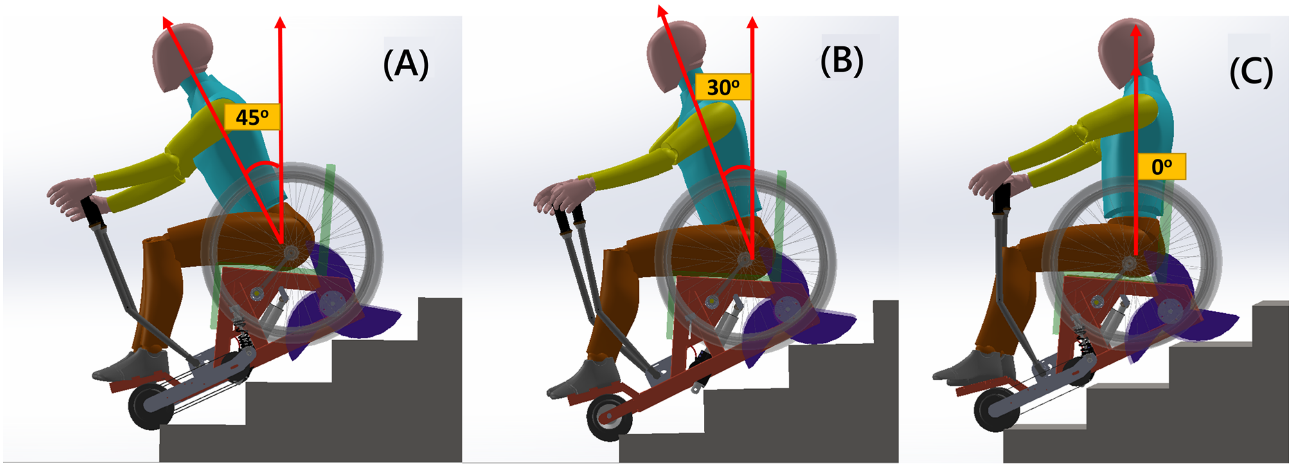
Figure 5. Lever orientation (ψ) as an ergonomic factor to study the effect of (A) supination, (B) neutral, and (C) pronation on muscle activation.



(2) Torso angle (λ): 3 trunk orientations in the coronal plane were studied [29,30], including 45° (Figure 6A), 30° (Figure 6B), and 0° upright (Figure 6C). The maximum low reach can be achieved at 45°, while maximum stability and minimum spinal loading can be attained at 0°.

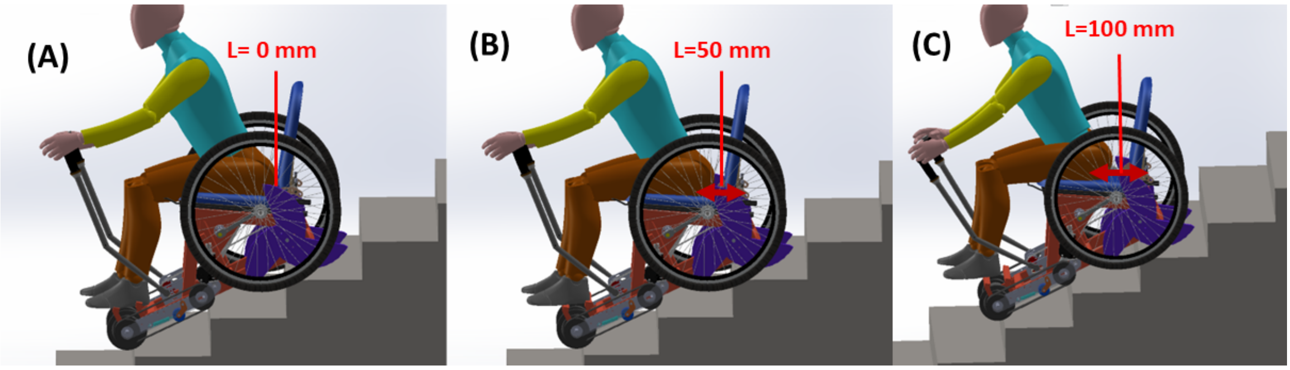
XSL•FO
RenderX

Figure 6. Torso angle made with vertical coronal plane (A) 45°, (B) 30°, and (C) 0°.



(3) Lever distance (L): the distance from the shoulder to the lever (shoulder grip length+L) was varied using a novel seat adjustment mechanism. The distance L was calculated from the lower back of the seated participant to the backrest of the wheelchair. Three levels were tested, including L=0 mm (Figure 7A), L=50 mm (Figure 7B), and L=100 mm (Figure 7C).

Figure 7. Ergonomic factor lever distance L for (A) L = 0, (B) L = 50 mm, and (C) L = 100 mm.



Participants and Experimental Design

Seven healthy participants (S1–S7) were selected for the surface electromyographic study. The anthropometric data of the participants is shown in Table 3.

Table . Anthropometric characteristics of participants in the surface electromyography (sEMG) study.

Participants	Age (years)	Upper arm length (cm)	Lower arm length (cm)	BMI
S1	28	32	24	22.6
S2	33	34	23	26
S3	30	34	26	21
S4	31	36	28	20.7
S5	31	36	27	22.5
S6	33	32	24	24
S7	31	37	29	25

A 3-factor, 3-level experimental design was used. The factors and their levels are shown in Table 4. A full factorial design would require 27 runs; this was reduced to a 1/3 fractional factorial design using the Taguchi method, requiring 9 experimental runs (Table 5). Each participant performed the lever push-pull action while seated in the configuration specified by each experiment.

XSL•FO
RenderX

Table . Factors and their respective levels for identifying ergonomic seating.

Factors	Level 1	Level 2	Level 3
Torso angle (λ)	45^o	30^o	0^o
Lever distance (L)	0 mm	50 mm	100 mm
Lever orientation (ψ)	-30^o	0^o	$+30^o$

Table . 1/3 fractional factorial design of experiment (Taguchi) for 3 factors, 3 levels.

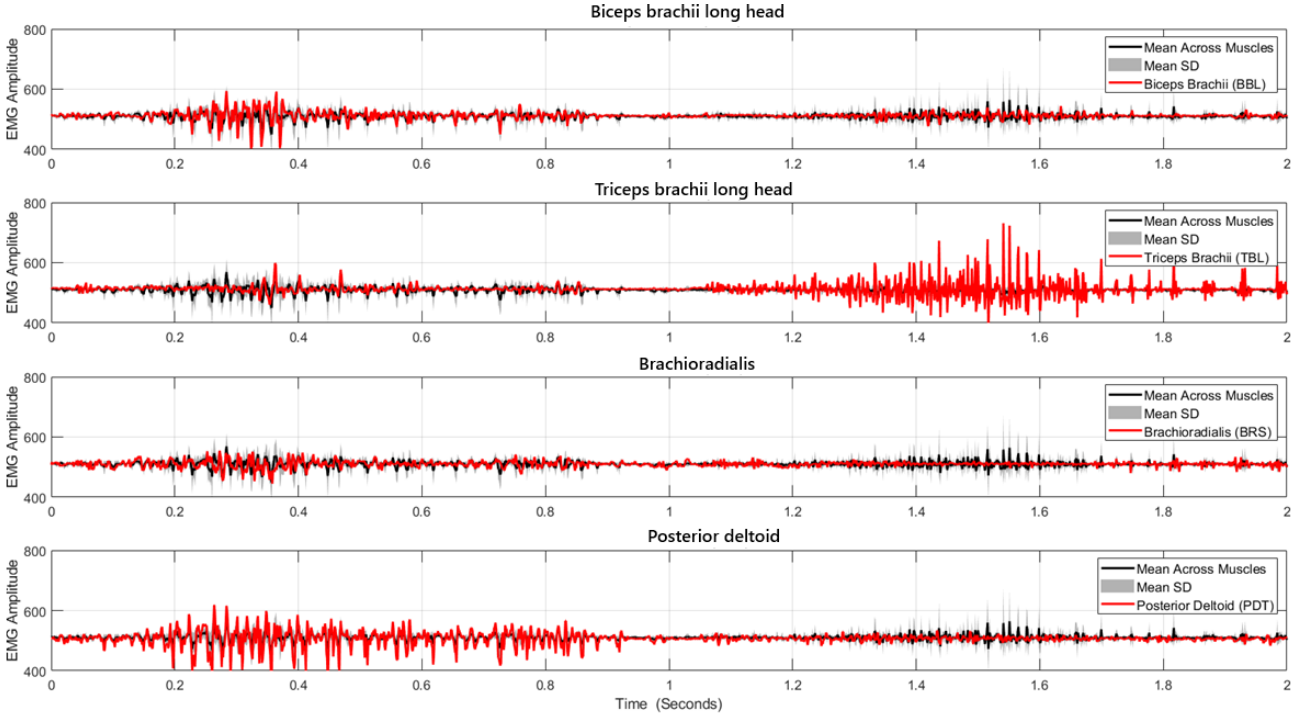
Experiment number	Torso angle level (λ)	Lever distance level (L; in mm)	Lever orientation level (ψ)
E1	45^o	0	-30^o
E2	45^o	50	0^o
E3	45^o	100	$+30^o$
E4	30^o	0	0^o
E5	30^o	50	$+30^o$
E6	30^o	100	-30^o
E7	0^o	0	$+30^o$
E8	0^o	50	-30^o
E9	0^o	100	0^o

Instruments and Electromyography Data Recording

The test wheelchair used in this study is a manually propelled stair-climbing wheelchair with a ratchet-based lever mechanism that provides a to-and-fro motion where pulling toward the chest engages the lever with the propulsion shaft, providing rotational motion to the stair-climbing wheel. For recording sEMG data, an ADS1299-based evaluation board (Texas Instruments) was used, providing 8 analog channels with a common reference, as reported in earlier sEMG studies [31-33]. These 8 channels were configured to create 4 channels with a common-mode rejection ratio for high-quality sEMG data for all 4 muscle groups as shown in Figure 8. An instrumentation gain of 11X and an overall system gain of 2420X were applied. Ultra-low impedance (<100 ohms) nickel-chromium (Ni-Cr) alloy electrodes were used with silver-silver chloride electrode (Ag-Cl) adhesive solid hydrogel at the electrode-skin interface to reduce skin impedance. The sampling frequency was 500 Hz.

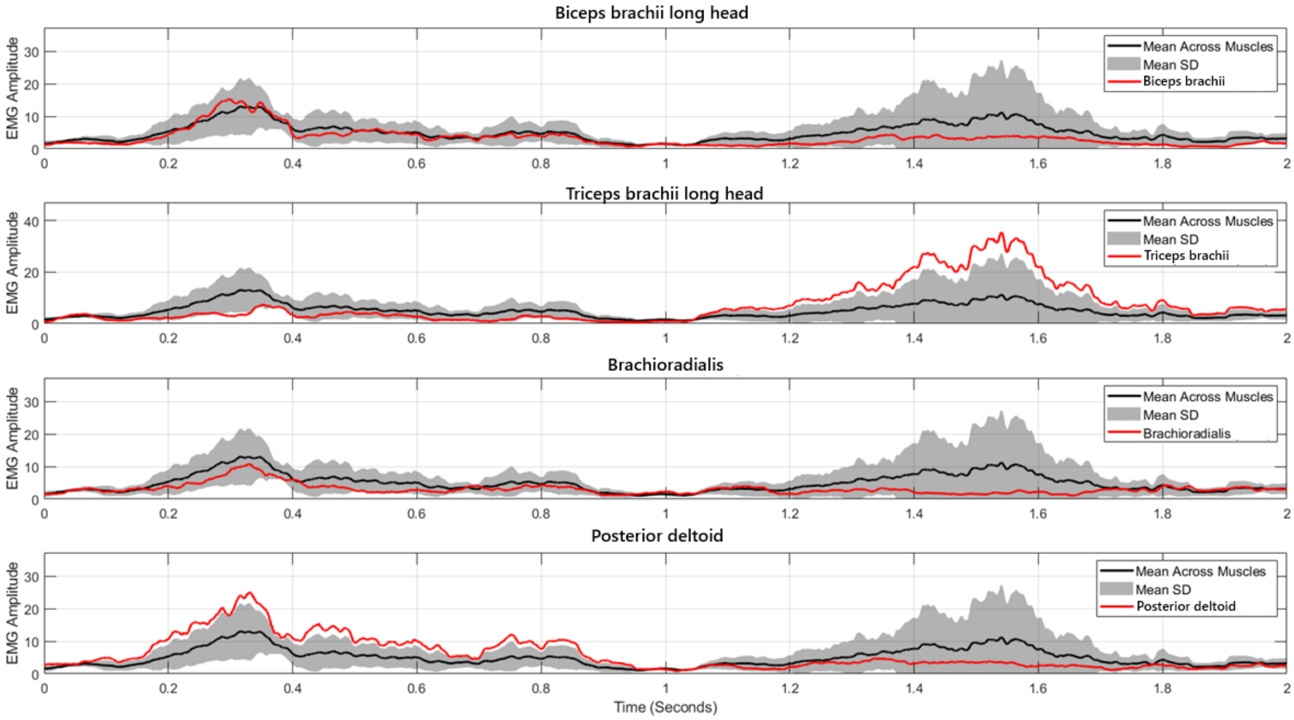
XSL•FO
RenderX

Figure 8. Surface electromyography (sEMG) data recorded on an ADS1299-based board for muscle groups biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT) for participant S1 during experiment E1.



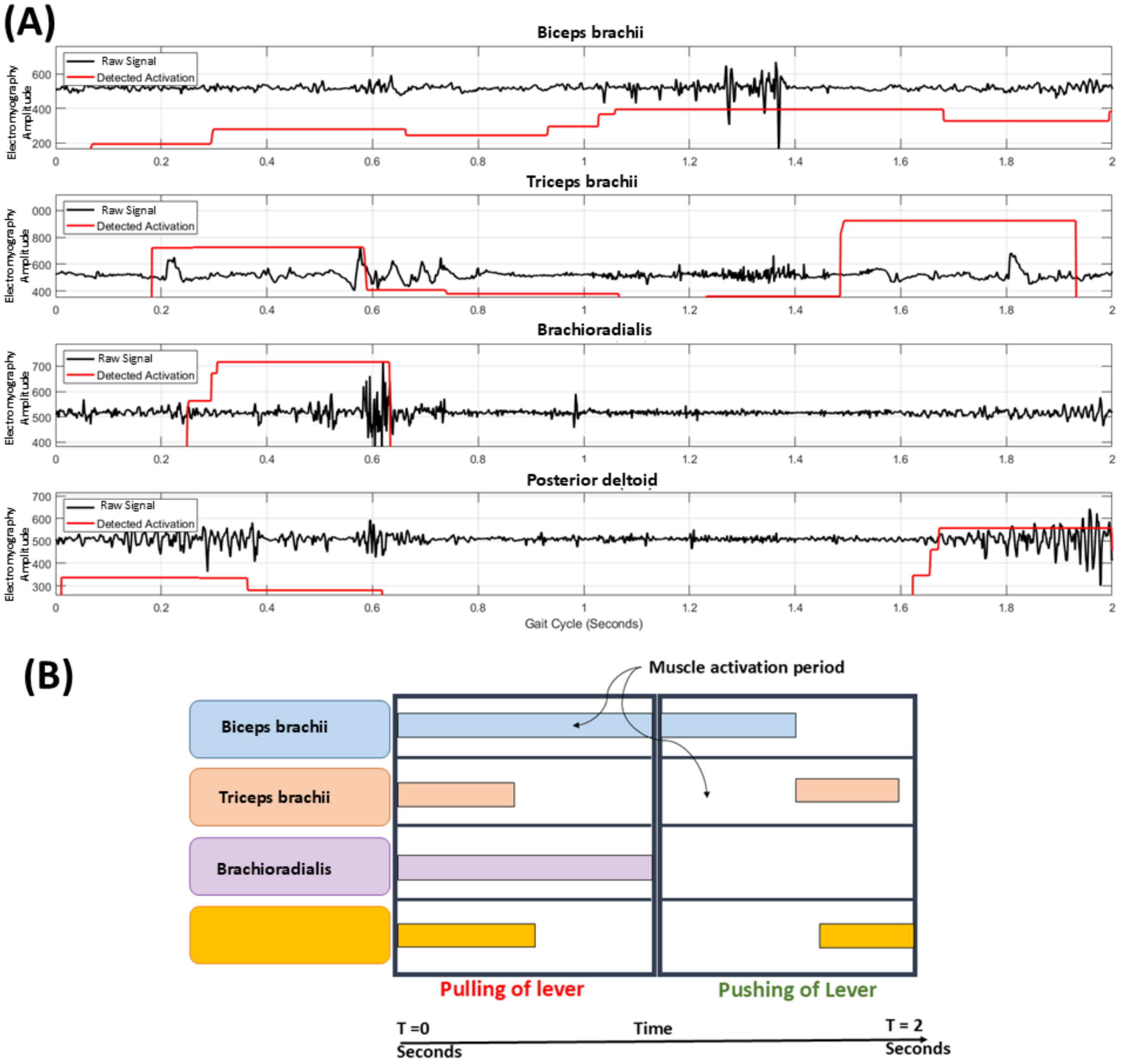
The raw electromyography signal was filtered using a band-pass Butterworth Infinite Impulse Response filter. The filtered electromyography signal was further processed by taking the absolute value and applying a moving average filter (window of 50 samples) as shown in Figure 9. For identification of muscle activation timing based on adaptive threshold, the Hilbert transform was used with a threshold duration of 25 samples to avoid influence of noise or other involuntary artifacts visible in Figure 10A. From the activation timing study, an activation window for the pull-push cycle is illustrated in Figure 10B.

Figure 9. Digitally filtered data for muscle groups BBL, TBL, brachioradialis, and PDT for participant S1 during experiment E1, showing individual activation of muscles following activation timing patterns of pushing and pulling.



XSL•FO
RenderX

Figure 10. Muscle activation detection and its representation (A) Detection of muscle activation timing and amplitude threshold on a digitally processed electromyography signal (participant S1, experiment E8) using the Hilbert transform. (B) Muscle activation window in pull-push cycle of lever propulsion. BBL: biceps brachii long head; PDT: posterior deltoid; TBL: triceps brachii long head.



Electromyography Data Processing: Maximum Voluntary Contraction

Maximum voluntary contraction (MVC) was calculated from the digitally filtered electromyography envelope with absolute values as a quantifiable parameter for normalizing the electromyography signal. The weighted average of envelope amplitude and envelope root-mean square (RMS) was used to calculate MVC [34], expressed as:

$MVC=w1 \times A+w2 \times RMS$ (6)

Where,A is the maximum amplitude of the electromyography envelope given by

$A=max (X) -min (X)$ (7)

X={x1,x2,x3……xn}; X is the dataset of filtered electromyography signal recorded over time t.

x_i represents the amplitude of the electromyography signal at i-th time point.

RMS means root mean square of dataset X.

Following Merletti and Parker (2004) [34], the weights were set as w1=0.7 and w2=0.3, prioritizing sensitivity to maximum activation for detecting full recruitment peaks while complementing it with RMS stability to mitigate outlier effects.

Muscle coordination patterns were assessed using the Pearson correlation coefficient “r” to measure the linear relationship between MVC values of muscle pairs [35]. The coefficient is given by:

$$r=\sum i=1n(xi-x-)(yi-y-)\sum i=1n(xi-x-)2\sum i=1n(yi-y-)2$$

where x_i and y_i are the MVC values for 2 muscles (eg, anterior deltoid and biceps brachii) across all observations, and x- and y- are their means.

XSL•FO
RenderX

Main Effect and Interaction Effect Analysis

Quadratic regression models were fitted for each ergonomic factor–muscle response relationship to characterize the design space. Main effects and interaction effects (AB: $\lambda\times L$; AC: $\lambda\times\psi$; BC: $L\times\psi$) were computed by averaging MVC values across paired factor combinations. Interaction effects were visualized using line plots and heatmaps. 3D surface plots were generated to delineate the ergonomic design space. Statistical significance of correlations between anthropometric dimensions and task performance was assessed at $P<.05$.

Ethical Considerations

This study was approved by the Institutional Ethics Committee (IEC) of the Indian Institute of Technology Kanpur (IEC Communication Number: IITK/IEC/2024-25/I/29, Date of Approval: June 20, 2024). The study followed the protocol titled "Design and Development of a multi-functional assistive device for restricted Mobility and Rehabilitation." All procedures involving human participants adhered to the institutional ethical standards and the 1964 Helsinki Declaration and its later amendments or comparable standards. Written informed consent was obtained from all participants, who were informed of objectives, procedures, risks, and their right to withdraw without consequences. No identifying information is included here; data were anonymized and securely stored per institutional and national privacy guidelines. Participants received no financial compensation. For those with paraplegia, extra precautions ensured comfort and safety during testing.

Results

Part A: Plane Surface Movement Results

Task Completion Times

The time taken for each of the 9 participants to complete each task on the simulated course is presented in Table 6.

Table . Time taken for each participant to complete the task (in seconds).

Task	Participants								
	S1	S2	S3	S4	S5	S6	S7	S8	S9
C[a]	5	4	3.5	3	3.5	4.2	4	3.5	5.3
R[b]	12.7	14.5	15	12.5	14	18	13.8	16.4	16.2
T1[c]	21.4	23	25.4	19.5	20.5	24	20	19	22
F1[d]	6	7.2	8.1	6.3	5	4.5	4.3	4	5.5
R1[e]	1.9	2.5	2	2	2	2	2	2.3	2.8
N[f]	10.4	9.8	6	8	7	9.5	6.5	8.3	9.9
F2[g]	2	2	2	2.8	2	2.5	2	2.2	2.3
R2[h]	5.5	3	3.4	4	4	4	5	3.2	4.4
T2[i]	9.2	7	6	7.3	8.2	9	8.5	7	7

[a] C: concrete.
[b] R: ramp.
[c] T1: 90° turn
[d] F1: granite floor.
[e] R1: rubber hollow carpet.
[f] N: narrow path.
[g] F2: ceramic tile floor.
[h] R2: rubber carpet.
[i] T2: gradual turn.

The ramp (R) and sharp 90° turn (T1) emerged as the most time-intensive tasks. The ramp required an average of approximately 14.8 seconds, while the 90° turn required approximately 21.6 seconds. Tasks on smooth floor surfaces (F1 and F2) and rubber carpets (R1 and R2) required relatively less time, averaging 2 ‐ 5 seconds each.

The time required to complete the 9 standardized wheelchair driving tasks, along with participant favorability ratings, is summarized below. All times are reported as mean (SE) in seconds.

Task C (straight corridor): Participants completed the task in 4.1 ± 0.9 s and rated it as moderately favorable. Task R (right turn): 14.8 (1.8)s, rated as moderately unfavorable. Task T1 (tight turn 1): 22.3 (2.1) s, rated as strongly unfavorable. Task F1 (forward maneuver 1): 5.7 (1.2) s, rated as moderately unfavorable. Task R1 (reverse 1): 2.4 (0.4)s, rated as moderately favorable. Task N (narrow passage): 8.5 (1.5) s, rated as strongly favorable. Task F2 (forward maneuver 2): 2.5 (0.5) s, rated as moderately favorable. Task R2 (reverse 2): 4.1 (0.8) s, rated as moderately favorable. Task T2 (tight turn 2): 7.8 (1.1) s, rated as moderately favorable.

XSL•FO
RenderX

These results indicate that tasks involving tight maneuvers (particularly T1) required substantially more time and received the least favorable ratings, whereas simpler forward and reverse tasks were completed quickly and viewed more positively. The pattern highlights the importance of adequate space and clear pathways in wheelchair-accessible environments.

Subjective Ratings and Overall Assessment

In the overall subjective response study for all 9 participants evaluating wheelchair attributes on a visual analog scale from 1 to 10, overall ride comfort received the lowest mean rating of 3.33 (SD 1.58), indicating areas for potential improvement in user experience during mobility. Wheelchair stability was rated moderately higher at a mean of 5.44 (SD 1.74), while driving control emerged as the strongest attribute with the highest average score of 6.00 (SD 1.22), suggesting effective handling and responsiveness in the device's design. Perception of the wheelchair scored a mean of 3.89 (SD 1.17), reflecting possibly mixed sensory feedback, whereas maneuverability—potentially encompassing response aspects—achieved a solid mean of 5.11 (SD 1.17), highlighting reliable navigation capabilities as depicted in the bar graph with error bars representing SDs. The correlation between task difficulty (higher time and low level of easiness) and anthropometric dimensions of the participants was not statistically significant (*P*=.04), indicating that task difficulty could be directly associated with wheelchair design variables rather than participant characteristics.

Part B: sEMG and MVC Results

Electromyography Amplitude and MVC Values

The average electromyography amplitude and RMS of all 7 partitipants for each muscle group, along with calculated MVC values, are presented in Table 7. There was no significant statistical correlation between the anthropometric features and electromyography activity in terms of MVC (*P*=.04). The maximum electromyography amplitude for each muscle averaged from all 9 experiments is illustrated in Figure 12A. Muscle coordination patterns derived from MVC data using the Pearson correlation coefficient are shown in Figure 12B.

Table . Average electromyography amplitude and root mean square (RMS) of all 7 participants for each muscle group along with their calculated maximum voluntary contraction (MVC) values.

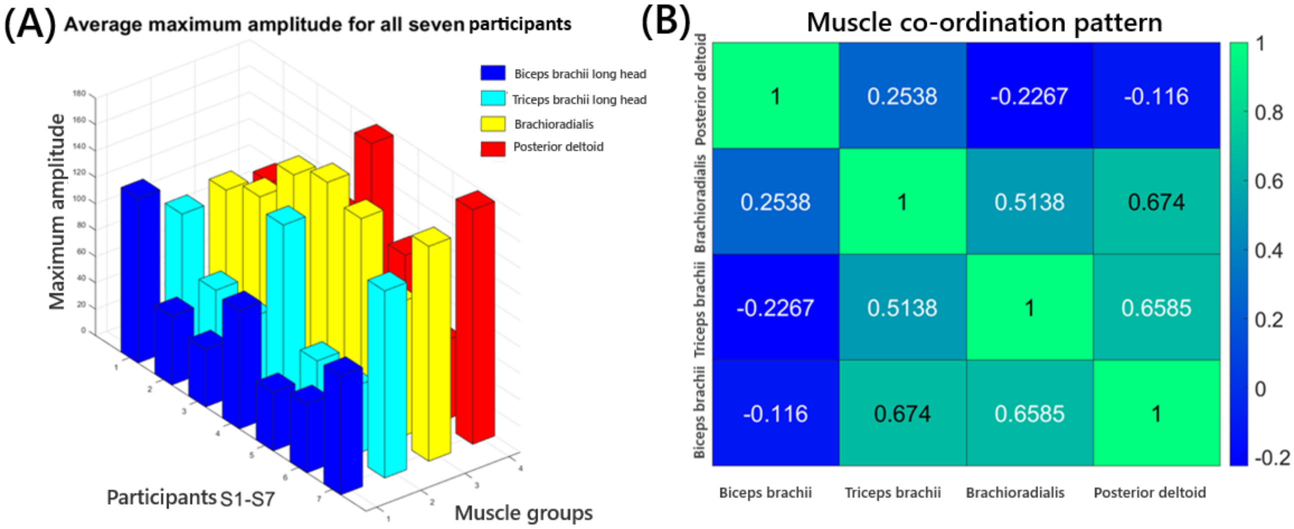
Experiment	Electromyography average amplitude				Electromyography root mean square				MVC			
	BBL[a]	TBL[b]	Brachioradialis	PDT[c]	BBL	TBL	Brachioradialis	PDT	BBL	TBL	Brachioradialis	PDT
E1	69.3	94.4	149.6	132.3	7.8	12.2	22.3	17.4	50.85	69.74	111.41	97.83
E2	55.5	74.8	179.0	95.0	7.3	10.5	13.24	12.4	41.04	55.51	129.272	70.22
E3	56.8	105.2	136.1	88.3	9	11.9	14.3	11.5	42.46	77.21	99.56	65.26
E4	52.9	75.1	128.1	82.6	8.3	12.0	13.0	10.0	39.52	56.17	93.57	60.82
E5	59.0	76.5	129.8	83.9	8.6	10.3	15.0	12.0	43.88	56.64	95.36	62.33
E6	56.7	102.5	130.9	121.3	9.0	13.9	16.8	19.0	42.39	75.92	96.67	90.61
E7	47.8	72.9	85.6	77.6	8.6	11.5	11.6	10.1	36.04	54.48	63.4	57.35
E8	48.7	55.8	75.1	62.1	8.0	9.0	11.3	10.0	36.49	41.76	55.96	46.47
E9	56.6	64.0	132.4	64.0	10.8	9.5	14.8	9.7	42.86	47.65	97.12	47.71

[a] BBL: biceps brachii long head.

[b] TBL: triceps brachii long head.

[c] PDT: posterior deltoid.

Figure 11. Electromyography voltage output for all 4 muscle groups of 7 participants. (A) Average maximum amplitude of electromyography signal at different muscle groups for all 7 participants (S1-S7) calculated from all 9 experiments (E1-E9). (B) Muscle coordination pattern calculated using MVC values. BBL: biceps brachii long head; PDT: posterior deltoid; TBL: triceps brachii long head.



Quadratic Regression Analysis

Quadratic regression models were fitted for each factor-muscle relationship. The regression equations, residual sum of squares, and coefficients of determination are presented in Table 8.

Table . Quadratic regression analysis and its parameters for muscle group response and ergonomic factors.

Relation	Equation	Residual sum of squares	Coefficient of determination (%)
BBL[a]-lever distance	$BBL=42.14-0.0710\times L+.000753\times L^{2}$	4.96	4.8
BBL-torso angle	$BBL=38.46+.0658\times +0.001659\times ^{2}$	3.98	38.7
BBL-lever orientation	$BBL=41.14-0.04083\times \psi+0.000976\times \psi^{2}$	4.91	6.8
TBL[b]-lever distance	$TBL=60.13-0.4210\times L+.004890\times L^{2}$	11.79	30.6
TBL-torso angle	$TBL=47.96+.6270\times -0.00429\times ^{2}$	9.81	52.0
TBL- lever orientation	$TBL=53.11+.0051\times \psi+0.01057\times \psi 2$	13.05	15.1
Brachioradialis- lever distance	$Brachioradialis=89.46+.0796\times L+.000036\times L^{2}$	25.41	2.6
Brachioradialis-torso angle	$Brachioradialis=72.16+.470\times +0.00992\times ^{2}$	15.35	64.5
Brachioradialis-lever orientation	$Brachioradialis=106.7-0.0318\times \psi -0.02177\times \psi 2$	23.11	19.4
PDT[c]-lever distance	$PDT=72.00-0.4517\times L+.004103\times L^{2}$	19.27	9.6
PDT-torso angle	$PDT=50.51+.863\times -0.00571\times ^{2}$	14.43	49.3
PDT-lever orientation	$PDT=59.58-0.2776\times \psi -0.01155\times \psi 2$	17.48	25.6

[a] BBL: biceps brachii long head.

[b] TBL: triceps brachii long head.

[c] PDT: posterior deltoid.

XSL•FO
RenderX

Main Effects on MVC

The main effects of the 3 ergonomic factors on MVC for each muscle group are summarized in Table 9 and visualized in the main effect plots (Figure 12).

Table . Main effects on maximum voluntary contraction values for factor A, B, and C.

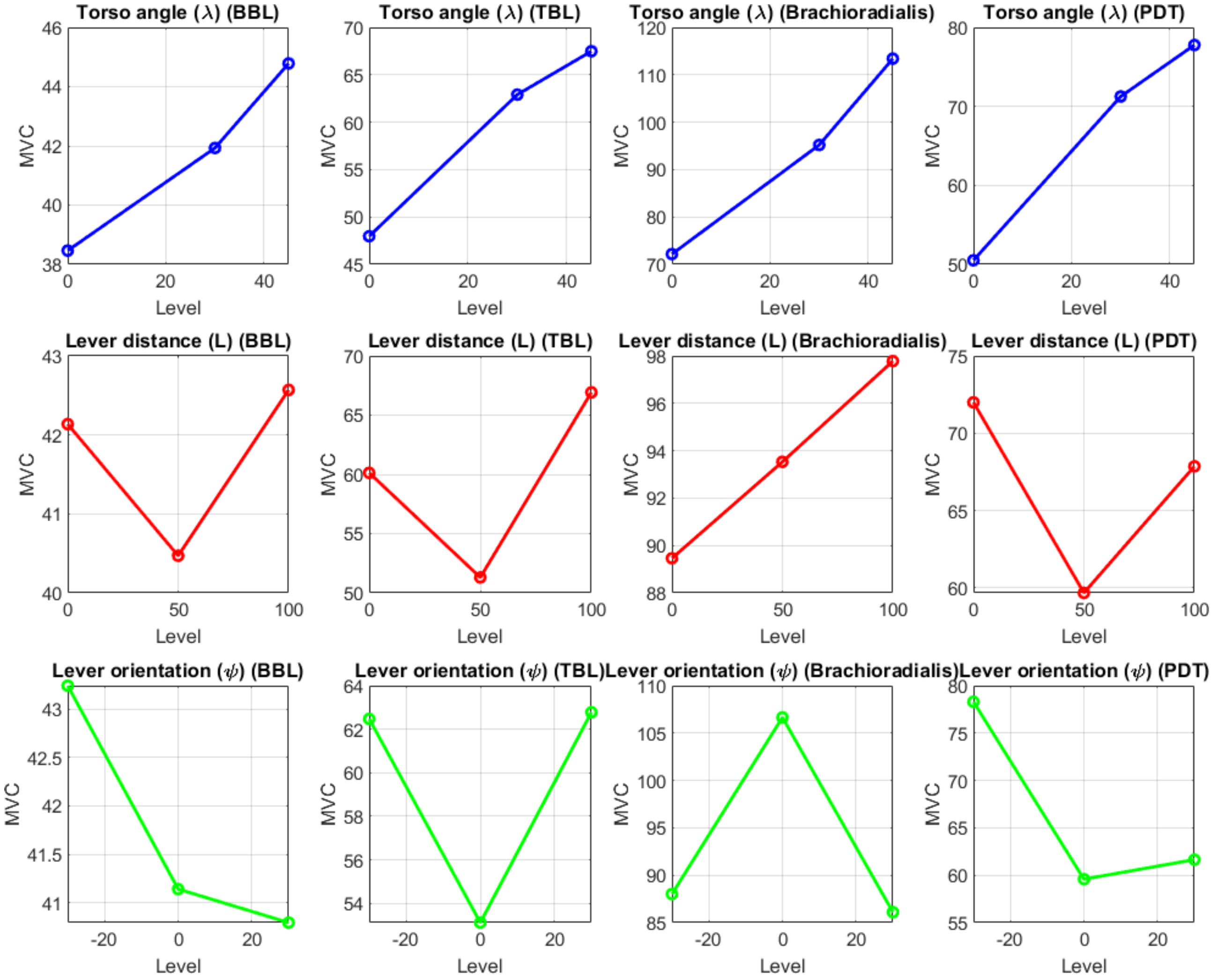
Muscle	Level 0° (straight)	Level 30° (inclined)	Level 45° (highly inclined)
Factor A: torso angle (λ)			
BBL[a]	38.46	41.93	44.78
TBL[b]	47.96	62.91	67.49
Brachioradialis	72.16	95.20	113.41
PDT[c]	50.51	71.25	77.77
Muscle	Level 0 mm (farthest)	Level 50 mm (neutral)	Level 100 mm (closest)
Factor B: lever distance (L)			
BBL	42.14	40.47	42.57
TBL	60.13	51.30	66.93
Brachioradialis	89.46	93.53	97.78
PDT	72.00	59.67	67.86
Muscle	Level –30° (pronation)	Level 0° (neutral)	Level 30° (supination)
Factor C: lever orientation (ψ)			
BBL	43.24	41.14	40.79
TBL	62.47	53.11	62.78
Brachioradialis	88.01	106.65	86.11
PDT	78.30	59.58	61.65

[a] BBL: biceps brachii long head.

[b] TBL: triceps brachii long head.

[c] PDT: posterior deltoid.

Figure 12. Main effect plot derived from the mean of maximum voluntary contraction (MVC) values with respect to ergonomic factors for MVC of biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT) changes with change in torso angle (λ), lever distance (L), and lever orientation angle (ψ). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



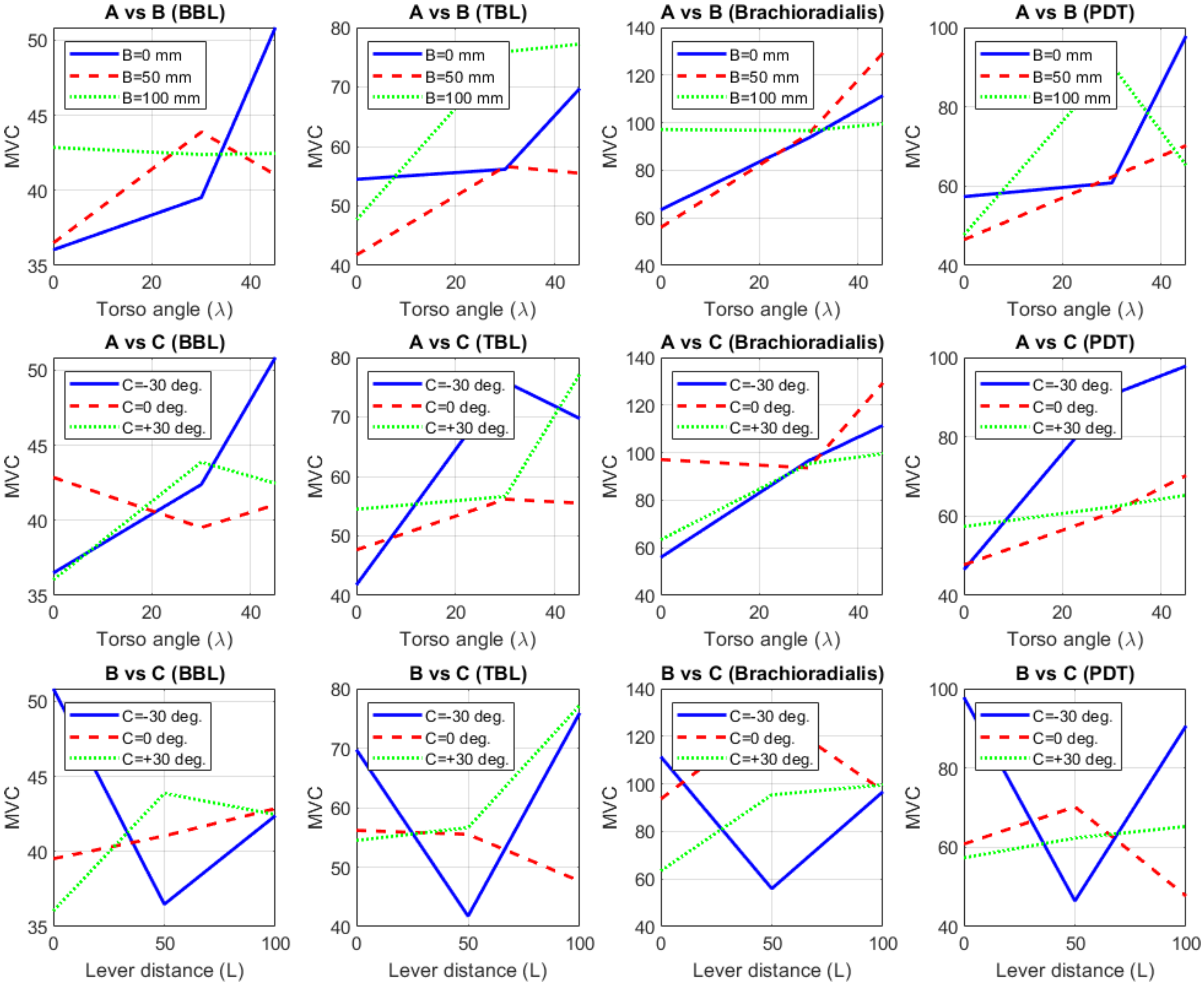
For torso angle (λ), MVC increased with inclination across all muscles, including BBL from 38.46 to 44.78, TBL from 47.96 to 67.49, brachioradialis from 72.16 to 113.41, and PDT from 50.51 to 77.77, with a straight torso (0°) yielding the lowest strain across all muscles, particularly brachioradialis. For lever distance (L), BBL remained relatively stable (40.47 - 42.57, lowest at 50 mm), TBL was optimized at 51.30 (50 mm), brachioradialis rose from 89.46 to 97.78 with decreasing L, and PDT was lowest at 59.67 (50 mm). For lever orientation (ψ), BBL slightly favored 30° (40.79); TBL and PDT were minimized at 0° (53.11 and 59.58, respectively); and brachioradialis peaked at 0° (106.65) but dropped to 86.11 at 30°.

Interaction Effects

Interaction effects (AB: λ×L; AC: λ×ψ; BC: L×ψ) were computed by averaging MVC values across paired factor combinations. These interactions are visualized in line plots (Figure 13) and heatmaps (Figure 14).

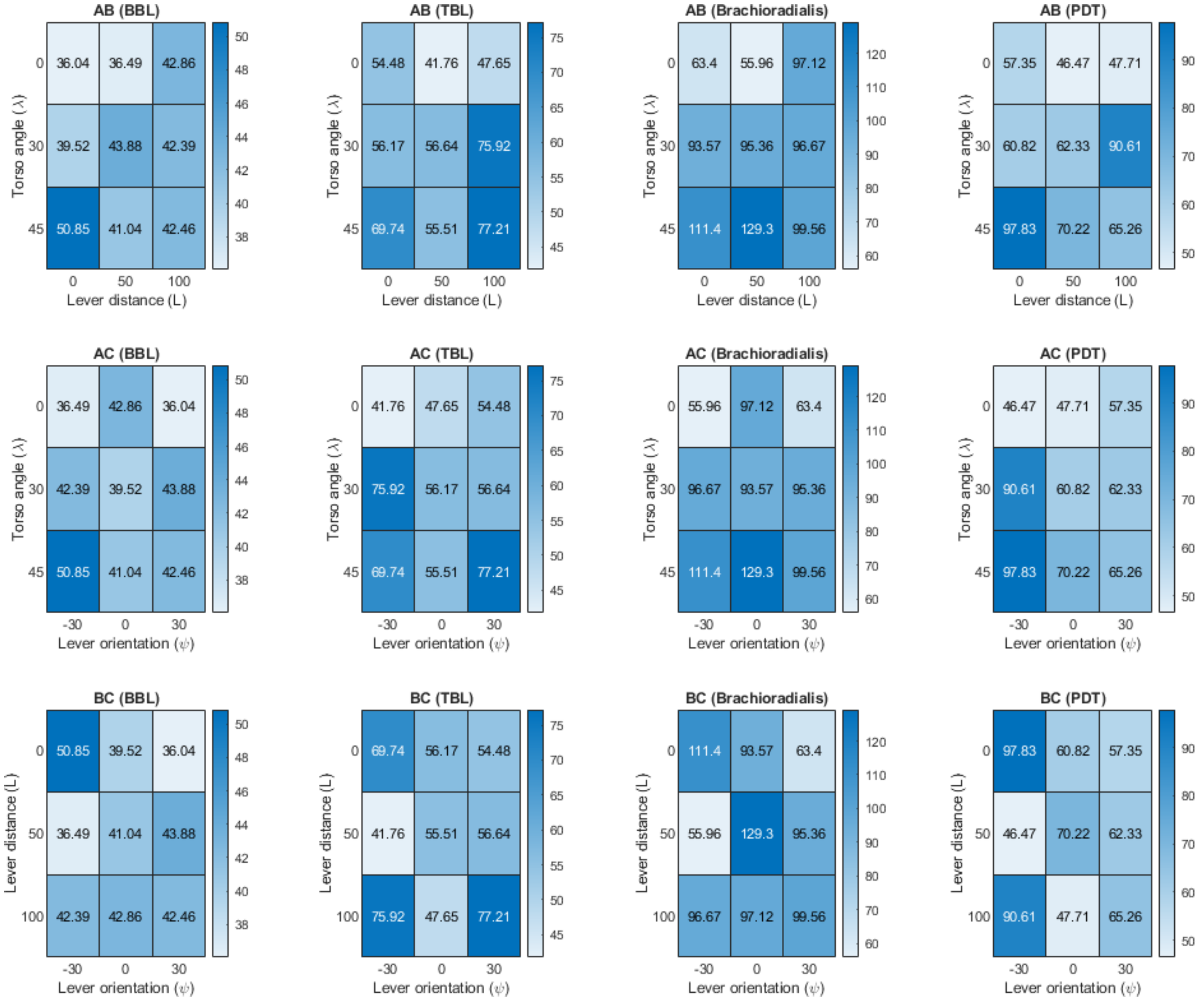
XSL•FO
RenderX

Figure 13. Interaction plot between ergonomics factors lever distance, torso angle, and lever orientation for biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



XSL•FO
RenderX

Figure 14. Interaction heat map between ergonomics factors lever distance, torso angle, and lever orientation for biceps brachii long head (BBL), triceps brachii long head (TBL), brachioradialis, and posterior deltoid (PDT). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



For the AB interaction (Torso angle ×Lever distance), forward lean (λ=45°) combined with long lever distance (L>50 mm) produced elevated effort in BBL and TBL, with lines crossing in Figure 13 and heatmaps showing high-strain zones. Upright posture (λ=0°) at L≈50 mm maintained low effort levels.

For the AC interaction (Torso angle ×Lever orientation), supination (ψ =+30°) combined with λ=0° reduced effort for brachioradialis and PDT, while pronation (ψ = –30°) at λ=45° increased effort, shown by crossing lines and heatmap contrasts. TBL was less affected by this interaction.

For the BC interaction (Lever distance ×Lever orientation), L ≈ 50 mm with ψ =+30° minimized brachioradialis and PDT strain, while a long lever distance with ψ = –30° maximized strain. TBL effort increased with pronation (see Figure 14).

Design Space Optimization

The design space heatmaps (Figure 15) and their associated 3D surface plots (Figure 16) delineated the ergonomic design space. For the AB interaction (λ×L), heatmaps revealed BBL (MVC 38 - 50) and PDT (50-95) with high strain at λ=45°, L=100 mm, and low strain at λ=0°, L=0 mm. The 3D surface plot peaked at 80 MVC where λ=45° and L=100 mm and dipped to 50 MVC at λ=0° and L=0 mm as shown in Figure 16A.

XSL•FO
RenderX

Figure 15. Design space heatmaps for ergonomic factors torso angle (λ) or A, lever distance (L) or B, and lever orientation (Ψ) interacting in AB, AC, and BC configurations. BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.

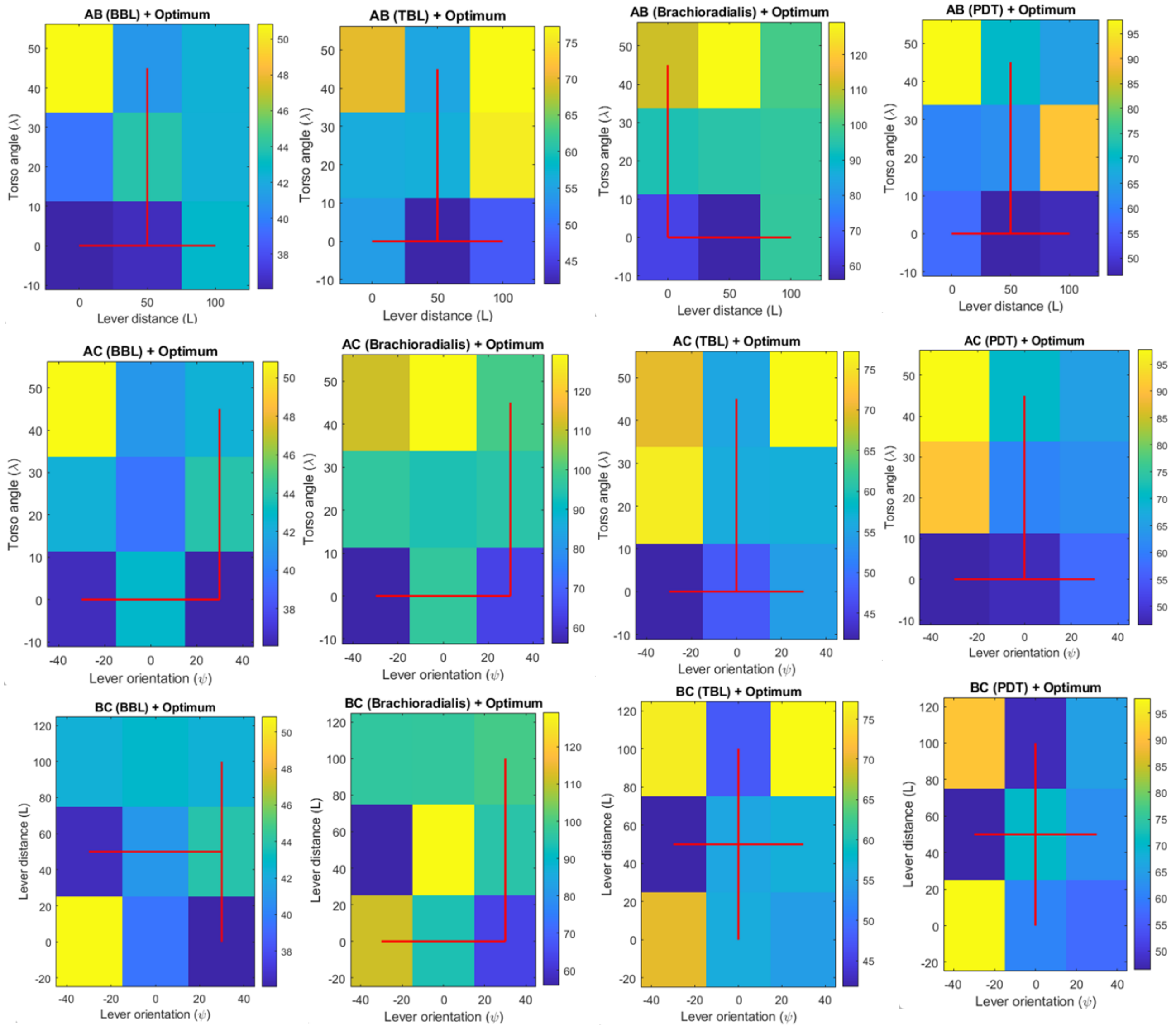
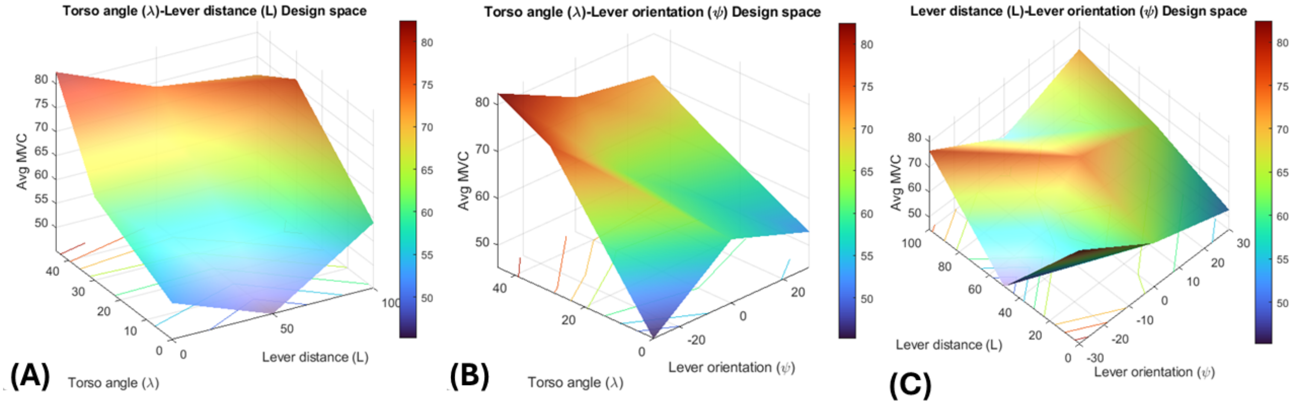


Figure 16. Design space for different ergonomic factors (A) torso angle (λ) and lever distance (L), (B) torso angle (λ) and lever orientation (Ψ), and (C) lever distance (L) and lever orientation (Ψ). BBL: biceps brachii long head; MVC: maximum voluntary contraction; PDT: posterior deltoid; TBL: triceps brachii long head.



For the AC interaction (λ × ψ), BBL minimized strain at λ=0°, ψ=30° (supination), and PDT at λ=0°, ψ=0°, with high strain at λ=45° and ψ=−30°. The 3D plot rose to 75 MVC at λ=45°, ψ=−30° and fell to 50 MVC across λ=0° and ψ=0°–30° as shown in Figure 16B. For the BC interaction (L × ψ), BBL was optimized at L=50 mm, ψ=30°, and PDT at L=50 mm, ψ=0°. The 3D plot peaked at 80 MVC for L=100 mm, ψ=−30° and reached approximately 50 MVC near L=50 mm, ψ=0°–30°, as shown in Figure 16C.

XSL•FO
RenderX

Collectively, strain exceeded 75 MVC at extreme configurations (eg, λ=45°, L=100 mm, ψ=-30°) and dropped to approximately 50 MVC at optimal settings (eg, λ=0°, L=0 - 50 mm, ψ=0°-30°).

Optimal Design Parameters

Optimal settings for minimizing individual muscle activity are shown in Table 10. These preferences vary by muscle but consistently favor a straight torso.

Table . Optimum levels for minimum muscle activity for all 4 muscle groups.

Muscle	Torso angle (λ)	Lever distance (L)	Lever orientation (ψ)
BBL[a]	0° (Straight)	50 mm (Neutral)	30° (Supination)
TBL[b]	0° (Straight)	50 mm (Neutral)	0° (Neutral)
Brachioradialis	0° (Straight)	0 mm (Farthest)	30° (Supination)
PDT[c]	0° (Straight)	50 mm (Neutral)	0° (Neutral)

[a] BBL: biceps brachii long head.

[b] TBL: triceps brachii long head.

[c] PDT: posterior deltoid.

Averaging across all muscles for a generalized design, the overall ergonomic recommendation is presented in Table 11 along with optimal wheelchair design parameters in Table 12.

Table . Overall ergonomic recommendation (averaged across muscles).

Parameter	Optimal setting
Torso angle (λ)	0.0° (Straight)
Lever distance (L)	37.5 mm (Between farthest and neutral)
Lever orientation (ψ)	15.0° (Between neutral and supination)

Table . Optimal wheelchair design parameters.

Muscle	Torso angle (in degrees)	Lever distance (mm)	Lever orientation (in degrees)	Torso angle	Lever distance	Lever orientation
BBL[a]	0	50	30	Straight	Neutral	Supination
TBL[b]	0	50	0	Straight	Neutral	Neutral
Brachioradialis	0	0	30	Straight	Farthest from lever	Supination
PDT[c]	0	50	0	Straight	Neutral	Neutral
Overall	0	37.5	15	Straight	Between farthest and neutral	Between Neutral and Supination

[a] BBL: biceps brachii long head.

[b] TBL: triceps brachii long head.

[c] PDT: posterior deltoid.

Discussion

Summary of Key Findings

This study systematically evaluated the ergonomic factors influencing muscle activity and user comfort in a stair-climbing wheelchair design through a dual-methodology approach. In Part A, the plane surface movement study revealed that the ramp and sharp 90° turn were the most challenging tasks, directly attributable to the wheelchair's weight and large dimensions. In Part B, the sEMG analysis demonstrated that a straight torso (λ=0°) consistently minimized muscle strain across all four muscle groups, with an overall ergonomic recommendation of λ=0°, L=37.5 mm, and ψ=15° to balance strain reduction across all muscles.

Design Implications of the Plane Surface Movement Study

The lack of statistical correlation between task difficulty and anthropometric dimensions (P=.04) confirms that any difficulty in performing a task can be directly associated with wheelchair design variables, which is the primary goal of this ergonomic assessment. The increase in time and effort at the ramp (R) can be directly correlated with the wheelchair weight increase due to the stair-climbing apparatus installed along with the plane-surface wheelchair. A significant amount of deadweight (net weight of 65 kg) is carried along when the plane surface

XSL•FO
RenderX

mode is active. The next iteration in design will require a detailed study of strength of materials to remove as much weight as possible without compromising strength and durability.

Similarly, the sharp 90° turn (T1) required extreme maneuverability skills at the user's side. The calculated turning distance for a 90° turn of approximately 1450 mm is a direct consequence of the overall length and width of the wheelchair, which plays a critical role in maneuverability and steering experience. The form factor must be reduced not only for addressing the sharp turn problem but also for movement through narrow spaces (Task N), which also took significant time although it had a contradictorily favorable subjective response.

The low comfort scores, particularly during plane surface movement, were attributed to the slightly tilted seat design that, while helping maintain the wheelchair user's center of gravity away from the descent direction during stair-climbing, causes considerable discomfort during plane surface movement. These findings highlight the inherent design tension between stair-climbing functionality and plane surface comfort, suggesting that future designs should incorporate a more dynamic seat tilt adjustment mechanism. Importantly, any design changes inferred from the plane surface movement study must not compromise the primary stair-climbing function.

Interpretation of sEMG Findings and Rationale for Optimal Parameters

The optimal design parameters (λ=0°, L=37.5 mm, and ψ=15°) were systematically determined through Taguchi design of experiments analysis, representing a carefully considered balance of biomechanical efficiency, user comfort, and practical manufacturability.

Torso angle (λ=0°): the neutral torso position minimizes shoulder abduction and maintains a natural arm movement trajectory, reducing rotator cuff stress. The straight torso position showed significant reduction in brachioradialis activation compared to forward-leaning postures (λ=15°, 30°), indicating lower muscular effort. This finding aligns with previous research on wheelchair propulsion biomechanics showing that neutral spine positioning reduces cumulative shoulder joint loading and decreases the risk of overuse injuries. As torso angle increases, muscle effort rises across the forearm (brachioradialis), arm (BBL and TBL), and shoulder (PDT) groups due to anterior torso displacement. The prototype's backrest design mitigates fatigue, with the 0° posture proving optimal during stair-climbing compared to forward-leaning positions.

Lever distance (L=37.5 mm): this distance provides optimal mechanical advantage while maintaining comfortable reach, derived from anthropometric data analysis and biomechanical modeling. Shorter distances resulted in cramped postures and increased BBL activation, while longer distances showed increased PDT activation, suggesting compensation through shoulder muscles. The 37.5 mm distance represents the optimal compromise minimizing total muscular activation across all monitored muscle groups. Prior literature suggests that extreme upper limb reaches elevate MSD risk, a hypothesis confirmed by these results: muscle effort increased with greater L for most muscles.

Lever orientation (ψ = 15°): this slight supination optimizes weight distribution, provides postural support, and maintains shoulder joint alignment. The brachioradialis, BBL, and PDT exhibited the lowest effort at +30° (supination), while TBL effort peaked during the pronation-to-supination transition due to flexion demands. An outward lever orientation consistently reduces muscle activity, lowering fatigue and MSD risk. The 15° angle balances stair-climbing stability requirements with plane surface comfort; neutral seating (0°) showed marginally lower muscle activation but compromised stability during stair-climbing, while greater supination angles (30°) increased discomfort during prolonged sitting and reduced forward reach capability.

The interaction effects, visualized through heatmaps and 3D surface plots, underscored the interdependence of these factors. The interaction plots helped identify critical regions where the combined effects of ergonomic factors diverge, aiding in making informed decisions about the physical dimensions of the wheelchair. For instance, the AB interaction plot showed that forward lean (λ=45°) combined with long lever distance (>50 mm) should be avoided as it spikes effort, while upright posture at L≈50 mm keeps effort low. The consistency between heatmaps and 3D surfaces highlights an ergonomic sweet spot that reduces muscle effort and fatigue, offering clear guidance for refining wheelchair design to enhance user comfort and efficiency during stair-climbing tasks.

The modular design architecture permits adjustment of all 3 parameters to accommodate users with different anthropometric characteristics, strength capabilities, and functional requirements. Individual optimization may be necessary for users at anthropometric extremes or with specific upper extremity limitations.

Weight Reduction Feasibility and Future Design Strategies

This prototype weight of 65 kg represents a significant barrier to usability, particularly for flat surface maneuverability and transportation. This weight is substantially higher than conventional manual wheelchairs (15‐25 kg) and constitutes a primary design limitation requiring immediate attention. Weight reduction strategies under investigation include the following:

1. Material substitution: replacing steel frame components with aerospace-grade aluminum alloys (estimated 8‐12 kg reduction) and carbon fiber composites for non–load-bearing elements (estimated 3‐5 kg reduction).
2. Mechanism optimization: redesigning the transformation mechanism to eliminate redundant components and implementing topology optimization for load-bearing structures (estimated 4‐6 kg reduction).
3. Component integration: combining multiple discrete parts into unified assemblies to reduce fastener count and material overlap (estimated 2‐3 kg reduction).

Preliminary engineering analysis suggests a realistic target weight of 45‐50 kg is achievable through these combined

XSL•FO
RenderX

strategies, representing a 23% - 31% reduction. This improved weight would significantly enhance maneuverability while maintaining structural integrity for stair-climbing operations. Cost-benefit analysis and finite element modeling are ongoing to validate these approaches prior to next-generation prototype development.

Safety Systems, Descending Mechanism, and Wheelchair Adjustability

While this manuscript focuses on ergonomic assessment during upward stair-climbing and plane surface movement, the wheelchair incorporates comprehensive safety and functional systems essential for complete operational capability:

1. Descending mechanism: stair descent uses a controlled ratchet-and-pawl braking system integrated with the stair-climbing wheel assembly. The mechanism permits incremental wheel rotation (step-by-step descent) while preventing uncontrolled backward movement. User-activated brake levers provide variable descent speed control (range: 1 - 4 steps per minute). Antitip sensors provide auditory feedback when the center of gravity approaches the instability threshold.
2. Braking systems: dual independent braking systems ensure redundancy and safety. Primary system lever-operated disc brakes on plane-surface wheels providing immediate stopping capability. Secondary system electromagnetic parking brakes preventing rollback during stationary periods. Emergency brake activation requires <1.5-second response time from full speed, meeting safety standards for powered mobility devices.
3. Adjustability features: the modular design permits extensive customization, including seat height adjustment (400 - 550 mm from ground, 50 mm increments), backrest angle adjustment (0 - 30° recline, 5° increments), footrest height and angle adjustment, lever position adjustment (radial distance 300 - 500 mm, angular position ±30°), and seat width adjustment (400 - 500 mm) to accommodate different pelvic widths. All adjustments use tool-free quick-release mechanisms for rapid configuration changes.

Comprehensive documentation of these systems, including mechanical specifications, safety testing protocols, and user operating procedures, will be presented in subsequent publications focused specifically on mechanical design and safety validation.

Strengths and Limitations

This study has several strengths. It represents one of the first pilot studies to integrate surface electromyography with subjective usability assessment for a stair-climbing wheelchair, providing both user-perceived comfort data and objective biomechanical evidence. The Taguchi-based experimental design efficiently identified optimal ergonomic parameters while minimizing the number of experimental runs. The inclusion of participants with paraplegia, though limited in number, provided critical real-world user validation.

However, several limitations must be acknowledged. The relatively small sample size (n=9 for plane surface movement, including 2 participants with paraplegia; n=7 for sEMG study) is appropriate for this pilot investigation but limits generalizability. Future studies should include larger, more diverse participant groups, including various disability types, severity levels, and demographic characteristics, to establish comprehensive design guidelines.

A significant limitation is the inclusion of only male participants in the anthropometric survey. Female and gender-diverse individuals have different anthropometric characteristics, muscle strength distributions, and ergonomic requirements [36]. Studies have shown that female wheelchair users often experience different pressure distribution patterns, grip strength capabilities, and shoulder biomechanics during propulsion [37]. Future iterations should conduct separate anthropometric studies for different gender groups to ensure optimal comfort, safety, and usability for all users. Additionally, consideration must be given to pregnancy-related ergonomic requirements, hormone-related strength variations, and gender-specific injury risk factors that may influence long-term wheelchair use patterns.

The controlled experimental environment, while necessary for this initial validation study, limits ecological validity. Real-world wheelchair use involves additional challenges including uneven terrain, weather conditions, unpredictable obstacles, crowded spaces, and cognitive demands from environmental distractions [38]. Future research should include (1) field testing in actual urban environments under naturalistic conditions; (2) long-term usability studies capturing daily use patterns and cumulative physical demands; (3) assessment under various weather and lighting conditions to evaluate all-season functionality; and (4) evaluation of cognitive load during real-world navigation tasks, including wayfinding, obstacle avoidance, and multitasking demands typical of community mobility. Long-term usability, durability, and cumulative physical effects require extended longitudinal studies beyond the scope of this pilot investigation, such as the impact of materials for backrest, seat, and lever handle [39-41].

Conclusion

This research provides critical design feedback for the stair-climbing wheelchair. Following a generic product design and development philosophy, the goal was to identify design variables influenced by usability testing. The most basic ergonomic assessment of a wheelchair that is, its reach and clearance, has been established through a detailed anthropometric survey of 20 participants with 5th and 95th percentile data guiding design dimensions.

The plane surface movement study demonstrated that the ramp and sharp 90° turn were the most challenging tasks, attributed to the wheelchair's 65 kg weight and large dimensions requiring approximately 1450 mm for a 90° turn. Subjective feedback indicated low satisfaction, particularly for comfort, due to the tilted seat design. These findings highlight the need for design improvements, including weight reduction, a more compact form factor, and seat tilt adjustments, all while preserving stair-climbing functionality.

The lever-propulsion sEMG study systematically evaluated three ergonomic factors. Key findings indicate that a straight torso (0°) consistently minimizes strain, particularly for

XSL•FO
RenderX

brachioradialis. A neutral lever distance (approximately 50 mm) reduced effort for most muscles, while neutral to slightly supinated orientations (0°-30°) generally optimized performance. Interaction effects guided an overall ergonomic recommendation of λ=0°, L=37.5 mm, and ψ=15°. These settings offer a balanced, practical solution for minimizing fatigue and enhancing user comfort. The modular wheelchair design permits customization for diverse user populations, and future research should include larger, gender-diverse participant groups and real-world validation studies, strengthening the foundation for ergonomic wheelchair optimization.

Funding

This research was funded by the Department of Science and Technology, Government of India.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions to protect participant privacy but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Cowan RE, Nash MS, Collinger JL, Koontz AM, Boninger ML. Impact of surface type, wheelchair weight, and axle position on wheelchair propulsion by novice older adults. Arch Phys Med Rehabil 2009 Jul;90(7):1076-1083. [doi: 10.1016/j.apmr.2008.10.034] [Medline: 19577019]
2. van der Woude LHV, de Groot S, Janssen TWJ. Manual wheelchairs: research and innovation in rehabilitation, sports, daily life and health. Med Eng Phys 2006 Nov;28(9):905-915. [doi: 10.1016/j.medengphy.2005.12.001] [Medline: 16504565]
3. Dubowsky SR, Rasmussen J, Sisto SA, Langrana NA. Validation of a musculoskeletal model of wheelchair propulsion and its application to minimizing shoulder joint forces. J Biomech 2008 Oct 20;41(14):2981-2988. [doi: 10.1016/j.jbiomech.2008.07.032] [Medline: 18804763]
4. Morrow MM, Rankin JW, Neptune RR, Kaufman KR. A comparison of static and dynamic optimization muscle force predictions during wheelchair propulsion. J Biomech 2014 Nov 7;47(14):3459-3465. [doi: 10.1016/j.jbiomech.2014.09.013] [Medline: 25282075]
5. Babu Rajendra Kurup N, Puchinger M, Gfoehler M. A preliminary muscle activity analysis: Handle based and push-rim wheelchair propulsion. J Biomech 2019 May 24;89:119-122. [doi: 10.1016/j.jbiomech.2019.04.011] [Medline: 31053474]
6. Verma A, Shrivastava S, Ramkumar J. Mapping wheelchair functions and their associated functional elements for stair climbing accessibility: a systematic review. Disabil Rehabil Assist Technol 2024 Jan;19(1):200-221. [doi: 10.1080/17483107.2022.2075476] [Medline: 35613308]
7. Verma A, Shrivastava S, Ramkumar J. Development of an optimized stair-climbing mechanism using parametric curved wheels inspired by human foot strides on flat stairs. Eng Res Express 2025 Sep 30;7(3):035436. [doi: 10.1088/2631-8695/ae086f]
8. Product manual for rehabilitation equipment—wheelchairs, folding, adult size according to IS 7454:1991. Bureau of Indian Standards. 2021. URL: https://www.bis.gov.in/wp-content/uploads/2021/08/IS-7454-Product-Manual.pdf [accessed 2026-03-24]
9. Guidelines and space standards for barrier-free built environment for disabled and elderly persons. Central Public Works Department Ministry of Urban Affairs and Employment, Government of India. 1998. URL: https://cpwd.gov.in/Publication/aged&disabled.pdf [accessed 2026-03-24]
10. Planning a barrier free environment. Office of the Chief Commissioner for Persons with Disabilities Ministry of Social Justice and Empowerment, Government of India. 2001. URL: https://www.sparsh.mp.gov.in/Public/SparshPublicPages/BarreirFreeEnvironment.aspx [accessed 2026-04-10]
11. International organization for standardization. wheelchairs — maximum overall dimensions. ISO. 1985. URL: https://www.iso.org/standard/13810.html [accessed 2026-03-25]
12. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]
13. Turner TL, Balmer DF, Coverdale JH. Methodologies and study designs relevant to medical education research. Int Rev Psychiatry 2013 Jun;25(3):301-310. [doi: 10.3109/09540261.2013.790310] [Medline: 23859093]
14. Saleeby PW. An introduction to the International Classification of Functioning, Disability and Health (ICF). Int J Disabil Hum Dev 2016;15(1):1-3. [doi: 10.1515/ijdhd-2015-0027]
15. P. w. d. act, “the persons with disabilities (equal opportunities, protection of rights and full participation) act, 1995, published in part II, section 1 of the extraordinary gazette of india. ministry of law, justice and company affairs (legislative department),”

XSL•FO
RenderX

ministry of law, justice and company affairs (legislative department). 1995 URL: https://www.indiacode.nic.in/bitstream/123456789/12890/1/the_persons_with_disabilities_act%2C_1995_no._1_of_1996_date_01.01.1996.pdf [accessed 2026-03-13]
16. Guidelines and space standards for barrier free built environment for disabled and elderly. : Central Public Works Department; 1998 URL: https://cpwd.gov.in/publication/aged&disabled.PDF [accessed 2026-03-11]
17. Chief commissioner for persons with disabilities, “planning a barrier free environment,”. 2001 URL: https://www.sparsh.mp.gov.in/Common/SJ/PDFFiles/PlanningForBarrierFreeEnvironment.PDF [accessed 2026-03-13]
18. De Luca CJ. The use of surface electromyography in biomechanics. J Appl Biomech 1997;13(2):135-163. [doi: 10.1123/jab.13.2.135]
19. Louis N, Gorce P. Surface electromyography activity of upper limb muscle during wheelchair propulsion: Influence of wheelchair configuration. Clin Biomech (Bristol) 2010 Nov;25(9):879-885. [doi: 10.1016/j.clinbiomech.2010.07.002] [Medline: 20846767]
20. Potvin JR, Bent LR. A validation of techniques using surface EMG signals from dynamic contractions to quantify muscle fatigue during repetitive tasks. J Electromyogr Kinesiol 1997 Jun;7(2):131-139. [doi: 10.1016/s1050-6411(96)00025-9] [Medline: 20719698]
21. Chow JW, Levy CE. Wheelchair propulsion biomechanics and wheelers’ quality of life: an exploratory review. Disabil Rehabil Assist Technol 2011;6(5):365-377. [doi: 10.3109/17483107.2010.525290] [Medline: 20932232]
22. Gellman H, Sie I, Waters RL. Late complications of the weight-bearing upper extremity in the paraplegic patient. Clin Orthop Relat Res 1988 Aug;Available(233):132-135. [Medline: 3402118]
23. Gangelhoff J, Cordain L, Tucker A, Sockler J. Metabolic and heart rate responses to submaximal arm lever and arm crank ergometry. Arch Phys Med Rehabil 1988 Feb;69(2):101-105. [Medline: 3341886]
24. Tomiak T, Gorkovenko AV, Tal’nov AN, et al. The averaged EMGs recorded from the arm muscles during bimanual “Rowing” movements. Front Physiol 2015;6:349. [doi: 10.3389/fphys.2015.00349] [Medline: 26640440]
25. Lui J, MacGillivray MK, Sheel AW, Jeyasurya J, Sadeghi M, Sawatzky BJ. Mechanical efficiency of two commercial lever-propulsion mechanisms for manual wheelchair locomotion. J Rehabil Res Dev 2013;50(10):1363-1372. [doi: 10.1682/JRRD.2013.02.0034] [Medline: 24699972]
26. Hoozemans MJM, Kuijer P, Kingma I, et al. Mechanical loading of the low back and shoulders during pushing and pulling activities. Ergonomics 2004 Jan 15;47(1):1-18. [doi: 10.1080/00140130310001593577] [Medline: 14660215]
27. O’Sullivan LW, Gallwey TJ. Forearm torque strengths and discomfort profiles in pronation and supination. Ergonomics 2005 May 15;48(6):703-721. [doi: 10.1080/00140130500070954] [Medline: 16087504]
28. Roman-Liu D, Tokarski T. Upper limb strength in relation to upper limb posture. Int J Ind Ergon 2005;35(1):19-31. [doi: 10.1016/j.ergon.2004.07.002]
29. Santos Requejo P, Lee S, Mulroy S, et al. Shoulder muscular demand during lever-activated vs pushrim wheelchair propulsion in persons with spinal cord injury. J Spinal Cord Med 2008 Jan;31(5):568-577. [doi: 10.1080/10790268.2008.11754604]
30. Kotajarvi BR, Sabick MB, An KN, Zhao KD, Kaufman KR, Basford JR. The effect of seat position on wheelchair propulsion biomechanics. J Rehabil Res Dev 2004 May;41(3B):403-414. [doi: 10.1682/jrrd.2003.01.0008] [Medline: 15543458]
31. Cavallone P, Vieira T, Quaglia G, Gazzoni M. Electomyographic activities of shoulder muscles during Handwheelchair.Q vs pushrim wheelchair propulsion. Med Eng Phys 2022 Aug;106(1):103833. [doi: 10.1016/j.medengphy.2022.103833] [Medline: 35926952]
32. Mastinu E, Ortiz-Catalan M, Hakansson B. Analog front-ends comparison in the way of a portable, low-power and low-cost EMG controller based on pattern recognition. 2015 Presented at: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Milan p. 2111-2114. [doi: 10.1109/EMBC.2015.7318805]
33. Jiang S, Gao Q, Liu H, Shull PB. A novel, co-located EMG-FMG-sensing wearable armband for hand gesture recognition. Sens Actuators A Phys 2020 Jan;301:111738. [doi: 10.1016/j.sna.2019.111738]
34. Merletti R, Parker PJ. Electromyography: Physiology, Engineering, and Non-Invasive Applications: John Wiley and Sons; 2004. [doi: 10.1002/0471678384]
35. Pearson K. Determination of the coefficient of correlation. Science 1909 Jul 2;30(757):23-25. [doi: 10.1126/science.30.757.23] [Medline: 17838275]
36. Steinfeld E, Maisel J, Feathers D, D’Souza C. Anthropometry and standards for wheeled mobility: an international comparison. Assist Technol 2010;22(1):51-67. [doi: 10.1080/10400430903520280] [Medline: 20402047]
37. Boninger ML, Koontz AM, Sisto SA, et al. Pushrim biomechanics and injury prevention in spinal cord injury: Recommendations based on CULP-SCI investigations. JRRD 2004;42(3sup1):9. [doi: 10.1682/JRRD.2004.08.0103]
38. Tolerico ML, Ding D, Cooper RA, et al. Assessing mobility characteristics and activity levels of manual wheelchair users. J Rehabil Res Dev 2007;44(4):561-571. [doi: 10.1682/jrrd.2006.02.0017] [Medline: 18247253]
39. Muhammad Salman Ajmal H, Muneer R, Saeed A, Tanveer M, Ahsan Saeed M. Synergistic role of green ‐ synthesized zinc oxide nanomaterials in biomedicine applications. ChemistrySelect 2024 Sep;9(36) [FREE Full text] [doi: 10.1002/slct.202402517]
40. Khan MUA, Ajmal HMS, Hassan HA, Azam A, Malik E. A soft robotic sleeve for physiotherapy: improving elbow rehabilitation in baseball pitchers. Physiother Res Int 2025 Jan;30(1):e70025. [doi: 10.1002/pri.70025] [Medline: 39764679]

XSL•FO
RenderX

41. Verma A, Singh A, Ramkumar J. Topology optimization of mechanical structures in stair-climbing assistive technology. Nanomaterials and Energy 2019 Dec 17;8(2):167-177. [doi: 10.1680/jnaen.19.00033]

Abbreviations

BBL: biceps brachii long head
CPWD: Central Public Works Department
DHM: digital human model
MSD: musculoskeletal disorder
MVC: maximum voluntary contraction
PDT: posterior deltoid
RMS: root-mean square
sEMG: surface electromyography
TBL: triceps brachii long head

Edited by S Munce; submitted 12.Jun.2025; peer-reviewed by D Menychtas, HMS Ajmal; revised version received 04.Jan.2026; accepted 03.Feb.2026; published 13.Apr.2026.

Please cite as:
Verma A, Kumar R, Ramkumar J
Design Implications of Comfort and Usability of Manual Stairclimbing Wheelchair: Ergonomic Assessment and Pilot Study Using Surface Electromyography Inputs
JMIR Rehabil Assist Technol 2026;13:e78965
URL: https://rehab.jmir.org/2026/1/e78965
doi:10.2196/78965

© Abhishek Verma, Rohit Kumar, J Ramkumar. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of Upper Body Joint and Hand Motions in Eating Solid Foods With Chopsticks and Semisolid Foods With a Spoon in Healthy Males and Females: Observational Study

Jun Nakatake[1], MMSc, PhD; Shogo Maeda[1], BHS; Shigeaki Miyazaki[1], MMSc, PhD; Hideki Arakawa[1], MD, PhD; Etsuo Chosa[2], MD, PhD

[1]Rehabilitation Unit, University of Miyazaki Hospital, 5200 Kihara Kiyotake-cho, Miyazaki-shi, Miyazaki, Japan

[2]Department of Rehabilitation Medicine, Junwakai Memorial Hospital, Miyazaki-shi, Miyazaki, Japan

Corresponding Author:
Jun Nakatake, MMSc, PhD
Rehabilitation Unit, University of Miyazaki Hospital, 5200 Kihara Kiyotake-cho, Miyazaki-shi, Miyazaki, Japan

Abstract

Background: Foods are not only masticated and swallowed but they also influence the choice of utensils and their use. Comparing the contexts in which different utensils are used with each food form could help in the assessment of individuals experiencing eating difficulties in the food culture unique to East Asian countries.

Objective: Considering East Asian rehabilitation practices, in this study, we evaluated upper body movements involved in eating pickles (solid food) and yogurt (semisolid food) using chopsticks and a spoon, respectively.

Methods: Upper body kinematics, including joint and hand spatiotemporal parameters, were quantified using a 3D inertial motion-capturing system and analyzed in healthy males (n=22; mean age 27.9, SD 5.5 years) and females (n=21; mean age 26.9, SD 4.7 years) across 4 feeding phases (reaching, picking it up, transporting, and inserting food into the mouth) by comparing utensils for eating respective food forms using the Wilcoxon signed rank test.

Results: Both sexes showed smaller maximum shoulder flexion angles with chopsticks for eating solid food across all phases (Males: reaching phase, *P*<.001; picking foods up phase, *P*=.04; transport phase, *P*<.001; and mouth phase, *P*<.001. Females: reaching phase, *P*<.001; picking foods up phase, *P*=.007; transport phase, *P*<.001; and mouth phase, *P*<.001.). Elbow, forearm, and wrist ulnar deviation angle changes were smaller using chopsticks during the “picking up” phase (in both sexes, *P*<.001) compared with using a spoon. However, greater elbow joint angle changes were found during the “reaching” phase (males, *P*=.002; females, *P*=.03) and greater forearm angle changes were found during the “transporting” phase (males, *P*=.01; females, *P*=.001) with chopsticks than with spoons. Regarding hand spatiotemporal parameters, the chopstick condition involved shorter actual distances (males, *P*<.001; females, *P*<.001), lower distance efficiency (males, *P*=.04; females, *P*=.001), and slower speed (males, *P*=.001; females, *P*<.001) during food transport.

Conclusions: The joint angle and hand spatiotemporal parameter characteristics observed during chopstick use for eating solid foods and spoon use for eating semisolid foods in healthy individuals could serve as reference movements in individuals with sensorimotor dysfunctions and inform the selection of adaptive utensils in rehabilitation practice.

(JMIR Rehabil Assist Technol 2026;13:e76239) doi:10.2196/76239

KEYWORDS

activities of daily living; biomechanical phenomena; neck; torso; upper extremity

Introduction

In East Asian countries, such as China, Korea, and Japan, chopsticks are primarily used to take food into the mouth during meals. This utensil serves multiple functions, including picking up solid foods, cutting them into manageable pieces, and securely holding them while bringing them to the mouth. Due to their versatile functions, chopsticks have become an integral part of the food customs of these regions. To compensate for the difficulty in handling semisolid foods, a spoon is typically used. Thus, chopsticks and a spoon are commonly paired with solid and semisolid foods, respectively. Furthermore, since people in Japan have the habit of consuming soup by bringing the bowl to their mouths instead of using spoons, Japanese people more often use chopsticks than do people in other countries.

When patients undergo eating activity rehabilitation, selection of appropriate feeding utensils is essential. Patients, including those recovering from strokes, often use a spoon for its adaptability to foods that are transferred to the mouth, masticated, and then swallowed. Semisolid foods, regulated for viscosity, are commonly recommended for individuals with dysphagia [1], making the spoon a preferred adaptive device. Chopsticks are not routinely selected due to the difficulty in using them for semisolid foods. As the oral and swallowing abilities improve, patients transition from using spoons for semisolid foods to using chopsticks for solid foods. The food custom in East Asian countries is also focused on the use of chopsticks for eating solid foods, which facilitates the prospect of dining together and improving well-being. However, sensorimotor dysfunctions can impair the ability to perform feeding movements involving chopsticks and spoons. For instance, patients with right hemiparesis can use chopsticks or a spoon only with their dominant right hand if their impairment is mild [2]. Therefore, the selection of chopsticks or a spoon depends on patients’ ability to transfer, masticate, and swallow foods, as well as their limb function for manipulating utensils.

Daily movements and postures can be assessed to determine whether patients encounter difficulties performing tasks. Moreover, in the context of rehabilitation practice, assessments are needed to facilitate normal or compensatory movements or to change utensils used by patients while eating [3]. Understanding typical movements and postures when using feeding utensils aids in these assessments. Previous studies have separately reported differences in upper body joint motions across feeding phases during the normal use of chopsticks [4] and spoons [5]. A comparative study on normal upper limb kinematics has shown greater joint motion during spoon use than that during chopstick use when eating pickles, a common form of solid food [6]. However, a study investigating performance while using eating utensils reported that spoon use for noodle strips and tofu cubes prolonged eating time or led to failure to properly manipulate foods [7], suggesting that the choice of utensils should match food type.

In another aspect, proximal body movements, including the neck and trunk joint angles, tilting and rotation angles, or transfer distances, occur during eating [4,5,8-10] and tend to increase when foods are prone to be dropped [8,11]. The movements and postures of these proximal segments are important as normal functions in relation to preventing food dropping and form the basis for fine upper limb function. A recent study that simulated eating movements for fictitious foods using utensils demonstrated larger neck joint motion with a spoon than with chopsticks; however, trunk flexion angles were comparable between utensils [12]. These findings perhaps imply that using a spoon requires compensatory neck joint motion for preventing food drop and trunk stability for maintaining upper extremity functions. Using chopsticks, conversely, may not need neck compensation and may need trunk stability as when using a spoon. However, the actual movement and positioning of the upper body, including the neck and trunk, during solid food intake with chopsticks, compared with those during semisolid food intake with a spoon, remain unclear.

In addition to joint parameters, spatiotemporal parameters are valuable for assessing upper limb motions [13,14]. Considering the importance of rehabilitation practice in East Asia, understanding the normal kinematics of chopstick use while eating solid foods and spoon use while eating semisolid foods not only expands knowledge about the utensils or food forms in isolation but is also crucial, yet currently lacking. Therefore, a comparative study of these 2 food type and utensil type conditions is warranted in clinical settings, as previous research found significant head motion differences in the conditions of eating food forms with these respective utensils [8]. This previous study design is thought to be validated for assessing patients with eating difficulties such as swallowing function deficit with upper limb paresis in the recovery process.

In this study, we aimed to identify the characteristics of joint motions and spatiotemporal parameters during eating solid foods with chopsticks and semisolid foods with a spoon across feeding phases. We hypothesized that upper body joint motions for chopstick use for solid food intake could be reduced compared with those for spoon use for semisolid food intake across all feeding phases; however, the time required may be similar. We also hypothesized that the distance traveled by the hand, particularly during the reaching movements, could be greater when using chopsticks than when using a spoon. Additional parameters, such as speed, could provide further insights into kinematics during clinical observations. Consequently, in this study, we investigated the characteristics of upper body joint angles and hand spatiotemporal parameters across all feeding phases by comparing the use of chopsticks for eating solid foods and of a spoon for eating semisolid foods in healthy adults.

Methods

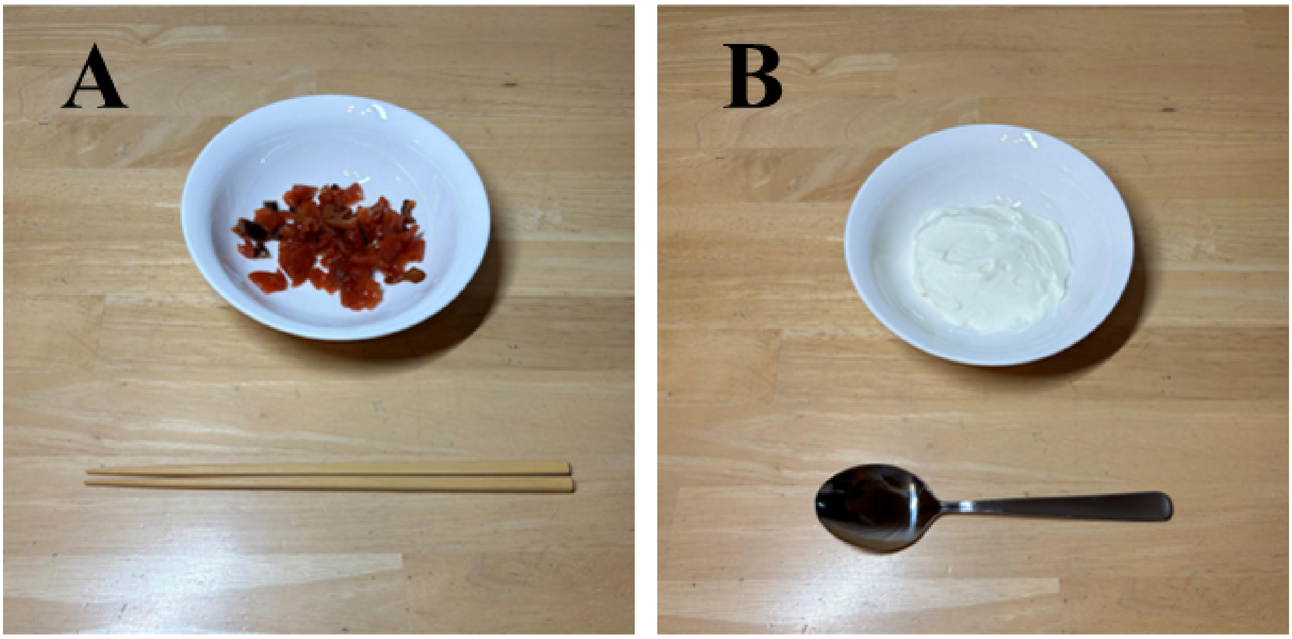
Participants

Between April 2013 and October 2017, 50 members of our institution participated per the inclusion criteria. The criteria were age 20‑39 years, right-hand dominance, and no neurological or musculoskeletal disorders. Individuals with left-hand dominance for using chopsticks or a spoon were excluded.

Eating Items and Tasks

Eating solid foods with chopsticks was performed using wooden chopsticks (22.5 centimeters in length, 8 grams in weight) and small Japanese pickled vegetables (Figure 1A). Eating semisolid foods with a spoon was performed using a stainless steel spoon (17.5 centimeters in length, 41 grams in weight) and yogurt (Figure 1B).

Figure 1. Eating items for 2 conditions. (A) Chopstick condition: small Japanese pickled vegetables in a bowl and wooden chopsticks. (B) Spoon condition: yogurt in the bowl and a stainless steel spoon.



The participants sat on a stool 40 centimeters in height, facing a table adjusted to their elbow height. The food bowl was placed in front of them at the same distance of the tip of the extended finger when their shoulder joint was in a neutral position and their elbow flexed at 90°. Neck and trunk movements were not constrained. The participants were asked to eat the foods using utensils held in their right hand while their left hand rested on their left thigh. They were instructed to eat at a comfortable pace and perform the movements sequentially 3 times. This frequency was to prevent fatigue and a full stomach. Based on the assumption that natural daily feeding movements comprise the cycle of chopsticks or a spoon traveling to and from the bowl and the mouth, starting and ending postures were not predefined to capture these natural cyclic daily movement data. The initial measurement position is illustrated in Figure 2.

XSL•FO
RenderX

Figure 2. Initial position for measurements during eating tasks using utensils with the dominant right hand while the left hand rests on the left thigh. Illustrated is an example of the task of eating pickles with chopsticks or yogurt with a spoon. The participant is shown wearing a suit, grooves, and a headband, all containing inertial sensors.

Measurement Instruments and Data Processing

Kinematic data at a frequency of 120 Hz were obtained using a 3D motion analysis system based on 17 inertial sensor units (Xsens MVN system, Xsens Technologies BV). Validation studies for this system have been previously conducted [15-19]. The inertial sensors were attached to participant body parts using a Lycra suit, gloves, and a headband and were confirmed

XSL•FO
RenderX

not to shift on the body surface or interfere with utensils and movements during feeding. The present tasks and measurements were valid; however, they might not be confirmed as sensor slipping or skin artifacts and interference with objects or bodies are recognized sources of error that could compromise the accuracy of joint angle and position data. The biomechanical model of the system consisted of 23 body parts, including the head, neck, 6 vertebrae, pelvis, scapulae, upper arms, forearms, hands, thighs, tibiae, feet, and toes. Joint angles were calculated based on the positions and orientations of the body parts. The neutral joint angle position of the model corresponded to a standing posture with the upper limbs alongside the trunk, the face directed forward, feet parallel at a distance of 1-foot width, and palms facing forward.

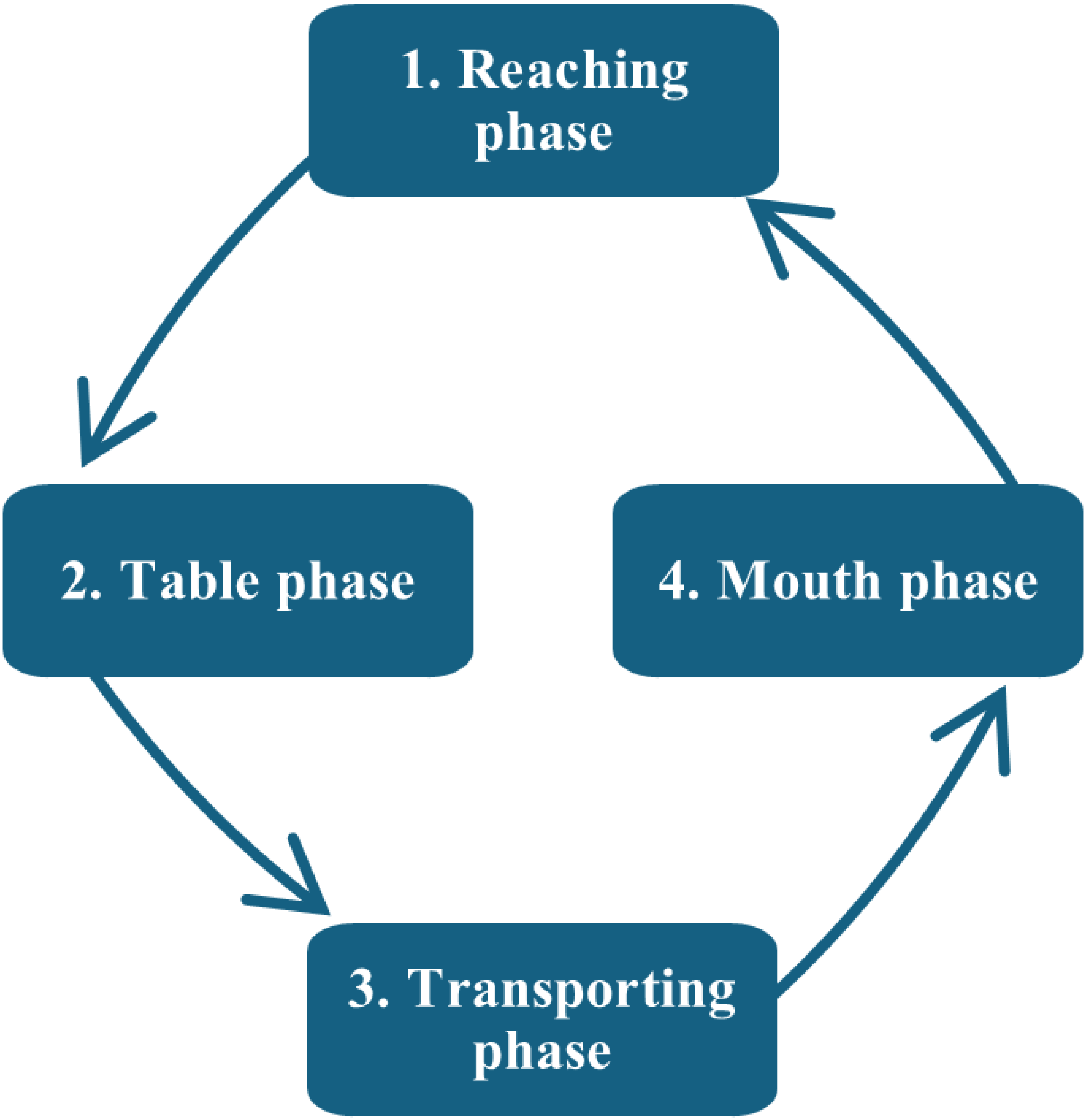
When the participants ate the food 3 times sequentially, there were 2 eating cycles. One success cycle of these 2 eating cycles (or the first cycle if both were successful) in each condition was selected for each participant for analysis. The eating cycle was defined as the interval between the moment after the utensil left the mouth and the moment immediately before it next left the mouth. This was confirmed using synchronized video recordings. To simplify the analysis, failure cycles involving the following were excluded: excessive upper limb elevation during food transport, looking away from the bowl or food, extraneous head movements, separating food, or taking more than 1 scoop of yogurt. Considering performing analysis and maintaining sample size, we decided to focus on analyzing 1 successful eating cycle, which was defined as a natural attempt of daily movements, of 2 cycles regarding each condition in each participant, excluding failed cycles that sometimes were observed.

Pauses during the reaching phase (observed in 84% of chopstick tasks and 65% of spoon tasks in the final analysis) and attempts to pick up the pickles more than once (28%) were frequently observed but not excluded from the analysis. Only participants with complete task data were included in the final analysis.

One eating cycle was divided into the following four phases: (1) reaching phase—from the moment the utensil left the mouth to the point of reaching the food, (2) table phase—picking up or scooping the food, (3) transporting phase—moving the food to the mouth, and (4) mouth phase—inserting the food into the mouth (Figure 3). These phases were defined because upper body joint motions differ during each feeding phase when using a spoon [5], which was expected to apply to chopstick use as well.

XSL•FO
RenderX

Figure 3. One eating cycle contains 4 phases. The arrows indicate the order of eating phase. The reaching phase is when the utensils move from the mouth to the food. The table phase is when the food is picked up or spooned up. The transporting phase is when the food moves from the bowl on the table to the mouth. The mouth phase is when the utensils are in the mouth.



Joint kinematic data from the biomechanical model were analyzed, including right shoulder flexion, abduction, and internal rotation; right elbow flexion and pronation; right wrist dorsiflexion and ulnar deviation (palmar flexion and radial deviation in the model); neck flexion, right lateral flexion, and right rotation (left rotation in the model); and right hip flexion. These joint kinematics were summarized as the maximum and minimum joint angles and ranges of angle change in each feeding phase. Movement time(s) was calculated in all phases.

The right-hand spatiotemporal parameters during the transporting phase included the actual distance traveled (meters), relative distance traveled, and maximum velocity (meters per second). The timing of maximum velocity (%) was calculated as motor control strategy metrics. The number of movement units was used to measure hand movement smoothness. The origin point for the right hand was defined as the midpoint between the styloid processes. A movement unit was defined as the difference between a local minimum and the next maximum velocity value that exceeded an amplitude limit of 20 millimeters per second, with a minimum interval of 150 milliseconds between peaks. This parameter was adapted from a task originally defined for drinking motions [20]. Most spatiotemporal parameters were limited to the transporting phase, as the start and end points of the reaching movement were clearly defined and undisturbed. Movements during the reaching phase showed frequent disruptions, and those during the table and mouth phases were not reaching movements. These hand spatiotemporal parameters could not be interpreted

XSL•FO
RenderX

according to the common concept of reaching movement parameters [13,14] and, consequently, were excluded from analysis.

Data Analysis

All metrics were analyzed separately for each sex because of known differences in daily movements [8,9,21-24] and fit to clinical assessments practically. The data are displayed as medians with IQRs. Since the study design compared movements between eating solid foods with chopsticks and semisolid foods with a spoon, these 2 task conditions were compared using the Wilcoxon signed rank test using IBM SPSS for Windows (version 27.0; IBM Corp), with a significance level of $P<.05$. Effect sizes (r) were calculated and categorized as small (|0.3|), medium (|0.5|), or large (|0.6|) [25].

Ethical Considerations

The protocol for this cross-sectional study, including informed consent using an opt-out method, was approved by the research ethics committee of the Faculty of Medicine of the University of Miyazaki (Miyazaki-shi, Japan; approval number O-1501). Since the sample data were derived from past studies [4,5] and participants could not be contacted directly, information related to this study was published on the local institutional website, giving past participants the opportunity to opt out of the study at any time. All participants provided written informed consent prior to participation in the previous studies [4,5]. The data were anonymized in both the previous and the present studies.

Results

Overview

Forty-three participants with 1 success cycle regarding each condition have been selected, excluding 7 participants with 2 failed cycles regarding 1 of 2 conditions. The analyzed study sample included 22 male and 21 female participants. The mean (SD) age and height of the male participants were 27.9 (5.5) years and 172.0 (5.1) cm, respectively, while those of the female participants were 26.9 (4.7) years and 158.0 (4.4) cm, respectively.

Joint Angles During Feeding Phases

In male participants (Table 1), various maximum and minimum joint angles, including shoulder flexion, abduction, and internal rotation; wrist ulnar deviation; and neck right lateral flexion, were lower when using chopsticks compared with when using spoons across many feeding phases. Additionally, during the reaching phase, the minimum wrist dorsiflexion angle with chopsticks was smaller than that with the spoon. In contrast, most elbow flexion and some neck right rotation angles were greater with chopsticks than with the spoon.

XSL•FO
RenderX

Table . Comparison of maximum and minimum joint angles between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in male participants[a].

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion										
Re	**23 (21 to 30)**	**36 (31 to 42)** [b]	−3.945	<.001	−0.595	**16 (13 to 20)**	**19 (16 to 24)**[b]	−3.1	.002	−0.467
Ta	**21 (16 to 27)**	**23 (19 to 31)**[b]	−2.029	.04	−0.306	19 (14 to 26)	22 (19 to 27)	−1.737	.08	−0.262
Tr	**25 (19 to 31)**	**36 (31 to 45)** [b]	−4.01	<.001	−0.605	**18 (15 to 24)**	**23 (19 to 30)**[b]	−3.1	.002	−0.467
Mo	**25 (19 to 32)**	**41 (33 to 49)** [b]	−4.042	<.001	−0.609	**23 (17 to 29)**	**36 (29 to 44)**[b]	−4.107	<.001	−0.619
Shoulder abduction										
Re	**32 (27 to 38)**	**38 (33 to 47)** [b]	−3.685	<.001	−0.556	**24 (20 to 29)**	**28 (22 to 33)**[b]	−2.808	.005	−0.423
Ta	30 (24 to 35)	31 (24 to 39)	−1.737	.08	−0.262	30 (24 to 34)	30 (23 to 36)	−1.315	.19	−0.198
Tr	**32 (27 to 40)**	**37 (32 to 48)**[b]	−3.847	<.001	−0.58	28 (24 to 34)	31 (24 to 39)	−1.672	.10	−0.252
Mo	**32 (26 to 40)**	**38 (33 to 48)**[b]	−3.847	<.001	−0.58	**31 (26 to 39)**	**36 (31 to 47)**[b]	−3.912	<.001	−0.59
Shoulder internal rotation										
Re	**0 (−6 to 8)**	**2 (−1 to 15)**[b]	−3.1	.002	−0.467	**−12 (−14 to −3)**	**−7 (−10 to 4)**[b]	−3.36	.001	−0.507
Ta	**−1 (−6 to 11)**	**4 (−1 to 11)**[b]	−3.036	.002	−0.458	**−4 (−8 to 5)**	**1 (−3 to 9)**[b]	−2.646	.01	−0.399
Tr	**−3 (−6 to 8)**	**4 (0 to 16)**[b]	−3.912	<.001	−0.59	**−8 (−10 to 2)**	**1 (−5 to 8)**[b]	−3.847	<.001	−0.58
Mo	**−5 (−10 to 3)**	**6 (−1 to 15)**[b]	−3.977	<.001	−0.6	**−8 (−11 to 1)**	**1 (−5 to 13)**[b]	−4.074	<.001	−0.614
Elbow flexion										
Re	**125 (117 to 132)**[b]	**121 (114 to 129)**	−3.36	.001	−0.507	97 (88 to 106)	100 (93 to 109)	−1.51	.13	−0.228
Ta	102 (97 to 112)	105 (98 to 114)	−0.568	.57	−0.086	**99 (89 to 107)**[b]	**98 (88 to 105)**	−2.062	.04	−0.311
Tr	**126 (118 to 133)**[b]	**120 (115 to 129)**	−3.393	.001	−0.512	**101 (91 to 108)**[b]	**99 (84 to 107)**	−2.451	.01	−0.37
Mo	**131 (120 to 136)**[b]	**127 (118 to 133)**	−3.295	.001	−0.497	**125 (115 to 132)**[b]	**119 (113 to 127)**	−3.523	<.001	−0.531
Forearm pronation										

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Re	90 (83 to 101)	99 (86 to 110)	−1.607	.11	−0.242	**26 (18 to 32)**	**34 (24 to 39)**[b]	−2.711	.007	−0.409
Ta	**90 (80 to 101)**	**99 (86 to 108)**[b]	−3.165	.002	−0.477	**77 (73 to 89)**[b]	**75 (55 to 83)**	−2.289	.02	−0.345
Tr	**78 (73 to 88)**[b]	**74 (55 to 83)**	−2.289	.02	−0.345	31 (23 to 37)	32 (23 to 38)	−0.828	.41	−0.125
Mo	31 (23 to 37)	34 (25 to 39)	−0.536	.59	−0.081	22 (16 to 29)	24 (16 to 32)	−0.016	.99	−0.002
Wrist dorsal flexion										
Re	26 (22 to 43)	29 (21 to 42)	−0.243	.81	−0.037	**13 (2 to 24)**	**20 (8 to 34)**[b]	−2.289	.02	−0.345
Ta	22 (9 to 37)	28 (16 to 38)	−1.217	.22	−0.183	16 (4 to 25)	21 (12 to 34)	−1.542	.12	−0.232
Tr	26 (15 to 36)	27 (18 to 39)	−0.243	.81	−0.037	19 (7 to 30)	20 (11 to 31)	−0.925	.36	−0.139
Mo	22 (11 to 33)	23 (11 to 31)	−0.016	.99	−0.002	20 (9 to 30)	20 (7 to 29)	−0.146	.88	−0.022
Wrist ulnar deviation										
Re	**5 (−3 to 10)**	**11 (4 to 16)**[b]	−3.587	<.001	−0.541	−7 (−14 to 2)	−5 (−14 to 5)	−1.347	.18	−0.203
Ta	**1 (−5 to 9)**	**7 (3 to 13)**[b]	−3.328	.001	−0.502	−2 (−10 to 7)	−4 (−12 to 5)	−0.86	.39	−0.13
Tr	**3 (−3 to 12)**	**12 (5 to 16)**[b]	−3.977	<.001	−0.6	**−1 (−7 to 5)**	**5 (1 to 12)**[b]	−3.782	<.001	−0.57
Mo	**2 (−4 to 11)**	**13 (3 to 18)**[b]	−4.074	<.001	−0.614	**1 (−5 to 8)**	**10 (3 to 15)**[b]	−4.042	<.001	−0.609
Neck flexion										
Re	27 (22 to 29)	26 (22 to 28)	−1.445	.15	−0.218	19 (14 to 25)	21 (18 to 26)	−1.25	.21	−0.188
Ta	26 (22 to 30)	25 (23 to 28)	−0.373	.71	−0.056	25 (21 to 29)	24 (22 to 27)	−0.828	.41	−0.125
Tr	25 (21 to 29)	25 (22 to 27)	−0.601	.55	−0.091	21 (18 to 25)	20 (18 to 25)	−1.542	.12	−0.232
Mo	22 (19 to 26)	22 (19 to 28)	−0.73	.47	−0.11	21 (18 to 24)	18 (17 to 24)	−1.899	.06	−0.286
Neck right lateral flexion										
Re	**−1 (−2 to 0)**	**3 (2 to 4)**[b]	−4.107	<.001	−0.619	**−4 (-5 to −3)**	**−2 (−4 to 0)**[b]	−3.036	.002	−0.458
Ta	**−3 (−4 to −2)**	**−2 (−3 to 0)**[b]	−3.036	.002	−0.458	**−3 (−5 to −2)**	**−2 (−4 to 0)**[b]	−2.678	.007	−0.404
Tr	**−2 (-3 to −1)**	**2 (−1 to 3)**[b]	−3.945	<.001	−0.595	**−3 (−4 to −3)**	**−2 (−3 to 0)**[b]	−3.393	.001	−0.512

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Mo	**−2 (−3 to 0)**	**4 (2 to 5)**[b]	−4.107	<.001	−0.619	**−3 (-4 to −1)**	**2 (−1 to 3)**[b]	−3.977	<.001	−0.6
Neck right rotation										
Re	**4 (2 to 6)**[b]	**2 (1 to 3)**	−3.263	.001	−0.492	0 (−1 to 2)	0 (−1 to 2)	−0.211	.83	−0.032
Ta	1 (0 to 2)	1 (0 to 2)	−0.016	.99	−0.002	1 (−1 to 2)	1 (−1 to 1)	−0.179	.86	−0.027
Tr	**4 (1 to 6)**[b]	**3 (0 to 3)**	−3.165	.002	−0.477	1 (−1 to 2)	1 (−1 to 2)	−0.146	.88	−0.022
Mo	**5 (3 to 7)**[b]	**3 (2 to 5)**	−3.393	.001	−0.512	**4 (1 to 5)**[b]	**2 (0 to 3)**	−3.068	.002	−0.463
Hip flexion										
Re	51 (44 to 57)	50 (45 to 54)	−0.73	.47	−0.11	43 (37 to 51)	44 (38 to 50)	−0.016	.99	−0.002
Ta	44 (38 to 51)	46 (38 to 51)	−0.341	.73	−0.051	44 (37 to 51)	45 (38 to 51)	−0.373	.71	−0.056
Tr	50 (43 to 58)	50 (44 to 54)	−0.763	.45	−0.115	44 (38 to 51)	46 (38 to 51)	−0.438	.66	−0.066
Mo	51 (45 to 58)	51 (46 to 56)	−0.179	.86	−0.027	50 (43 to 56)	51 (44 to 54)	−0.308	.76	−0.046

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test (*P*<.05), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b] Larger values from comparisons between chopstick and spoon conditions.

Forearm pronation showed differing trends according to the feeding phase. The minimum forearm pronation angle during the reaching phase and maximum angle during the table phase were smaller with chopsticks than with the spoon. Conversely, the minimum angle during the table phase and maximum angle during the transporting phase were larger with chopsticks. The differences between utensils showed small to large effect sizes. Neck and hip flexion angles did not differ significantly between utensils.

In female participants (Table 2), all shoulder joint angles, nearly all wrist dorsiflexion angles, and neck right lateral flexion angles were smaller with chopsticks than with the spoon. In contrast, most elbow flexion angles were larger when using chopsticks. Forearm pronation and wrist ulnar deviation angles varied inconsistently across feeding phases. In half the forearm pronation angles, chopsticks showed smaller values than did the spoon. However, the minimum angle during the table phase and maximum angle during the transporting phase were larger in the chopstick condition. Minimum wrist ulnar deviation angles during the reaching and table phases were larger with chopsticks, while other angles were smaller. These differences showed small to large effect sizes. Most neck flexion angles and all neck right rotation and hip flexion angles were comparable between chopsticks and spoons.

XSL•FO
RenderX

Table . Comparison of maximum and minimum joint angles between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in female participants[a].

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion										
Re	**20 (14 to 27)**	**37 (31 to 44)**[b]	-4.015	<.001	–0.62	**10 (4 to 12)**	**10 (6 to 22)**[b]	–2.45	.01	–0.378
Ta	**14 (7 to 18)**	**15 (10 to 25)**[b]	–2.694	.007	–0.416	**14 (6 to 17)**	**13 (7 to 23)**[b]	–2.485	.01	–0.383
Tr	**21 (12 to 28)**	**37 (31 to 44)**[b]	–4.015	<.001	–0.62	**13 (6 to 17)**	**14 (9 to 24)**[b]	–2.972	.003	–0.459
Mo	**22 (18 to 31)**	**40 (37 to 49)**[b]	–4.015	<.001	–0.62	**20 (15 to 27)**	**37 (31 to 44)**[b]	–4.015	<.001	–0.62
Shoulder abduction										
Re	**31 (28 to 34)**	**37 (32 to 45)**[b]	–3.98	<.001	–0.614	**21 (19 to 24)**	**23 (19 to 28)**[b]	–2.589	.01	–0.399
Ta	**24 (22 to 28)**	**28 (24 to 36)**[b]	–3.285	.001	–0.507	**23 (21 to 27)**	**28 (23 to 33)**[b]	–3.319	.001	–0.512
Tr	**30 (27 to 35)**	**38 (33 to 47)**[b]	–4.015	<.001	–0.62	**24 (21 to 28)**	**28 (24 to 33)**[b]	–3.493	<.001	–0.539
Mo	**29 (27 to 36)**	**39 (32 to 48)**[b]	–3.98	<.001	–0.614	**28 (26 to 32)**	**36 (31 to 45)**[b]	–3.945	<.001	–0.609
Shoulder internal rotation										
Re	**3 (0 to 8)**	**13 (10 to 18)**[b]	–3.98	<.001	–0.614	**-5 (-8 to -1)**	**5 (-4 to 7)**[b]	–3.493	<.001	–0.539
Ta	**3 (0 to 8)**	**11 (6 to 15)**[b]	–3.25	.001	–0.501	**1 (-2 to 5)**	**9 (5 to 13)**[b]	–3.041	.002	–0.469
Tr	**3 (0 to 6)**	**13 (9 to 17)**[b]	–4.015	<.001	–0.62	**-1 (-5 to 0)**	**7 (2 to 11)**[b]	–3.875	<.001	–0.598
Mo	**4 (-1 to 8)**	**15 (10 to 21)**[b]	–4.015	<.001	–0.62	**0 (-3 to 5)**	**13 (6 to 15)**[b]	–4.015	<.001	–0.62
Elbow flexion										
Re	**135 (130 to 138)**[b]	**128 (120 to 134)**	–3.841	<.001	–0.593	106 (101 to 111)	103 (97 to 112)	–0.504	.61	–0.078
Ta	106 (104 to 111)	103 (97 to 112)	–1.095	.27	–0.169	**103 (97 to 110)**[b]	**97 (90 to 104)**	–3.041	.002	–0.469
Tr	**134 (125 to 137)**[b]	**128 (122 to 131)**	–3.076	.002	–0.475	**104 (97 to 111)**[b]	**97 (90 to 106)**	–3.389	.001	–0.523
Mo	**138 (131 to 141)**[b]	**131 (127 to 137)**	–3.041	.002	–0.469	**134 (126 to 138)**[b]	**127 (120 to 131)**	–3.736	<.001	–0.576
Forearm pronation										
Re	**102 (80 to 116)**	**109 (99 to 123)**[b]	–2.763	.01	–0.426	**17 (12 to 43)**	**32 (19 to 49)**[b]	–3.285	.001	–0.507
Ta	**102 (81 to 119)**	**111 (99 to 122)**[b]	–2.728	.01	–0.421	**93 (68 to 108)**[b]	**80 (64 to 91)**	–3.319	.001	–0.512

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Tr	**93 (68 to 107)**[b]	**80 (64 to 91)**	–3.215	.001	–0.496	26 (16 to 44)	29 (18 to 45)	–0.122	.90	–0.019
Mo	**26 (15 to 41)**	**31 (20 to 48)**[b]	–2.346	.02	–0.362	17 (10 to 35)	27 (11 to 40)	–1.199	.23	–0.185
Wrist dorsal flexion										
Re	**22 (15 to 37)**	**34 (27 to 40)**[b]	–3.563	<.001	–0.55	**9 (-1 to 21)**	**20 (14 to 31)**[b]	–3.875	<.001	–0.598
Ta	**23 (12 to 34)**	**31 (21 to 37)**[b]	–3.18	.001	–0.491	**18 (3 to 25)**	**23 (15 to 31)**[b]	–2.694	.007	–0.416
Tr	**25 (15 to 37)**	**31 (23 to 39)**[b]	–2.381	.02	–0.367	17 (7 to 30)	23 (11 to 31)	–1.616	.11	–0.249
Mo	18 (10 to 34)	25 (15 to 33)	–1.929	.054	–0.298	**17 (7 to 31)**	**22 (12 to 31)**[b]	–2.033	.04	–0.314
Wrist ulnar deviation										
Re	**12 (5 to 17)**	**14 (9 to 18)**[b]	–2.381	.02	–0.367	**4 (-2 to 10)**[b]	**-1 (-10 to 4)**	–3.319	.001	–0.512
Ta	**10 (3 to 15)**	**10 (6 to 18)**[b]	–1.964	.05	–0.303	**6 (0 to 13)**[b]	**-3 (-11 to 4)**	–3.91	<.001	–0.603
Tr	**12 (5 to 15)**	**15 (9 to 19)**[b]	–3.424	.001	–0.528	**9 (1 to 13)**	**10 (5 to 14)**[b]	–2.346	.02	–0.362
Mo	**10 (6 to 16)**	**15 (9 to 20)**[b]	–3.806	<.001	–0.587	**9 (5 to 13)**	**12 (9 to 19)**[b]	–3.597	<.001	–0.555
Neck flexion										
Re	27 (25 to 30)	27 (25 to 29)	–1.199	.23	–0.185	19 (10 to 25)	18 (15 to 23)	–0.678	.498	–0.105
Ta	27 (25 to 30)	27 (25 to 29)	–0.33	.74	–0.051	26 (24 to 29)	26 (25 to 28)	–0.226	.82	–0.035
Tr	26 (24 to 29)	27 (25 to 28)	–0.539	.59	–0.083	18 (11 to 24)	17 (12 to 19)	–1.929	.054	–0.298
Mo	20 (11 to 25)	20 (15 to 22)	–0.678	.50	–0.105	**18 (9 to 23)**[b]	**14 (11 to 18)**	–2.485	.01	–0.383
Neck right lateral flexion										
Re	**2 (–2 to 3)**	**3 (2 to 6)**[b]	–3.736	<.001	–0.576	**-2 (-4 to 0)**	**-1 (-2 to 0)**[b]	–2.52	.01	–0.389
Ta	**–1 (–3 to 2)**	**0 (-1 to 2)**[b]	–2.242	.03	–0.346	–2 (–3 to 1)	-1 (-2 to 0)	–1.825	.07	–0.282
Tr	**2 (–1 to 3)**	**3 (1 to 5)**[b]	–3.667	<.001	–0.566	**–1 (–3 to 1)**	**0 (–1 to 1)**[b]	–3.007	.003	–0.464
Mo	**1 (–1 to 3)**	**4 (3 to 7)**[b]	–3.493	<.001	–0.539	**1 (–2 to 3)**	**2 (1 to 5)**[b]	–3.563	<.001	–0.55
Neck right rotation										
Re	3 (1 to 5)	3 (1 to 5)	–1.025	.31	–0.158	0 (–2 to 1)	0 (–2 to 0)	–0.295	.77	–0.046
Ta	1 (–1 to 1)	0 (–2 to 1)	–0.017	.99	–0.003	0 (–1 to 1)	0 (–2 to 1)	–0.226	0.82	–0.035
Tr	3 (1 to 4)	3 (2 to 4)	–0.469	.64	–0.072	1 (–1 to 1)	0 (–2 to 1)	–0.052	.96	–0.008
Mo	4 (2 to 6)	5 (2 to 6)	–0.226	.82	–0.035	3 (1 to 4)	3 (0 to 4)	–0.608	.54	–0.094

XSL•FO
RenderX

Motion direction and Phase	Maximum joint angle (degrees)					Minimum joint angle (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Hip flexion										
Re	52 (42 to 60)	52 (42 to 58)	–1.095	.27	–0.169	44 (38 to 53)	45 (38 to 52)	–0.052	.96	–0.008
Ta	45 (39 to 54)	45 (38 to 53)	–0.956	.34	–0.148	44 (38 to 54)	44 (38 to 52)	–0.504	.61	–0.078
Tr	53 (43 to 58)	49 (42 to 57)	–1.025	.31	–0.158	45 (39 to 54)	44 (38 to 53)	–1.13	.26	–0.174
Mo	54 (44 to 60)	53 (43 to 58)	–1.06	.29	–0.164	53 (43 to 58)	49 (42 to 57)	–1.616	.11	–0.249

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test (*P*<.05), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b] Larger values in comparisons between chopstick and spoon conditions.

Changes in Joint Angles

In male participants (Table 3), changes in shoulder flexion, abduction, and neck right lateral flexion angles were generally smaller in the chopstick condition than in the spoon condition across all feeding phases. Additionally, wrist ulnar deviation, neck flexion, and hip flexion showed smaller changes in some phases when using chopsticks. Conversely, neck right rotation in most phases and wrist dorsiflexion during the reaching phase showed increased changes with chopsticks. Elbow flexion showed greater changes in the reaching phase when using chopsticks; however, the changes were smaller in the table and mouth phases. Changes in forearm pronation were smaller during the table phase but larger during the transporting phase. Shoulder internal rotation changes across all phases showed no significant differences between utensils. Effect sizes for these differences were small to large.

XSL•FO
RenderX

Table . Comparison of joint angle changes between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in male participants[a].

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion					
Re	**7 (5 to 12)**	**18 (11 to 21)**[b]	–3.62	<.001	–0.546
Ta	1 (1 to 3)	2 (1 to 3)	–1.217	.22	–0.183
Tr	**6 (4 to 9)**	**13 (7 to 18)**[b]	–3.425	.001	–0.516
Mo	**2 (2 to 3)**	**3 (3 to 5)**[b]	–2.841	.01	–0.428
Shoulder abduction					
Re	**6 (4 to 12)**	**11 (7 to 13)**[b]	–2.419	.02	–0.365
Ta	**1 (0 to 1)**	**1 (1 to 2)**[b]	–2.808	.005	–0.423
Tr	**4 (2 to 6)**	**7 (5 to 11)**[b]	–3.685	<.001	–0.556
Mo	**1 (1 to 1)**	**2 (1 to 2)**[b]	–2.646	.008	–0.399
Shoulder internal rotation					
Re	9 (6 to 15)	9 (6 to 14)	–1.25	.21	–0.188
Ta	2 (1 to 4)	3 (2 to 4)	–0.958	.34	–0.144
Tr	4 (2 to 8)	4 (2 to 8)	–0.308	.76	–0.046
Mo	2 (1 to 4)	3 (2 to 4)	–0.828	.41	–0.125
Elbow flexion					
Re	**26 (22 to 36)**[b]	**22 (13 to 25)**	–3.036	.002	–0.458
Ta	**2 (1 to 4)**	**7 (4 to 13)**[b]	–3.555	<.001	–0.536
Tr	25 (21 to 30)	23 (20 to 28)	–1.055	.29	–0.159
Mo	**4 (3 to 5)**	**6 (4 to 7)**[b]	–2.354	.02	–0.355
Forearm pronation					
Re	68 (57 to 76)	67 (60 to 77)	–0.049	.96	–0.007
Ta	**11 (7 to 13)**	**26 (22 to 34)**[b]	–4.074	<.001	–0.614
Tr	**50 (44 to 55)**[b]	**42 (34 to 52)**	–2.484	.01	–0.374
Mo	8 (5 to 11)	9 (8 to 11)	–1.704	.09	–0.257
Wrist dorsal flexion					
Re	**16 (8 to 29)**[b]	**10 (7 to 14)**	–2.289	.02	–0.345
Ta	6 (4 to 9)	5 (3 to 7)	–1.412	.16	–0.213
Tr	8 (5 to 11)	4 (3 to 8)	–1.899	.06	–0.286
Mo	2 (1 to 4)	2 (1 to 3)	–0.601	.55	–0.091
Wrist ulnar deviation					
Re	**9 (5 to 14)**	**15 (9 to 22)**[b]	–2.386	.02	–0.36
Ta	**4 (2 to 6)**	**11 (6 to 18)**[b]	–3.88	<.001	–0.585
Tr	4 (3 to 6)	4 (2 to 7)	–0.86	.39	–0.13
Mo	2 (1 to 2)	1 (1 to 2)	–0.925	.36	–0.139
Neck flexion					

XSL•FO
RenderX

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Re	4 (2 to 9)	3 (2 to 6)	–1.769	.08	–0.267
Ta	0 (0 to 1)	1 (0 to 1)	–1.088	.28	–0.164
Tr	3 (1 to 6)	4 (2 to 6)	–0.99	.32	–0.149
Mo	**1 (1 to 2)**	**3 (2 to 4)**[b]	–3.1	.002	–0.467
Neck right lateral flexion					
Re	**2 (2 to 4)**	**5 (4 to 7)**[b]	–4.107	<.001	–0.619
Ta	0 (0 to 1)	1 (0 to 1)	–1.347	.18	–0.203
Tr	**1 (1 to 3)**	**2 (1 to 5)**[b]	–2.516	.01	–0.379
Mo	**1 (0 to 1)**	**2 (2 to 3)**[b]	–3.945	<.001	–0.595
Neck right rotation					
Re	**3 (2 to 5)**[b]	**2 (1 to 3)**	–3.717	<.001	–0.56
Ta	0 (0 to 1)	0 (0 to 1)	–0.016	.99	–0.002
Tr	**3 (1 to 4)**[b]	**1 (1 to 2)**	–2.971	.003	–0.448
Mo	**2 (1 to 3)**[b]	**2 (1 to 2)**	–2.354	.02	–0.355
Hip flexion					
Re	7 (3 to 9)	6 (3 to 8)	–1.412	.16	–0.213
Ta	**0 (0 to 0)**	**0 (0 to 1)**[b]	–2.354	.02	–0.355
Tr	5 (2 to 8)	5 (1 to 6)	–1.282	.20	–0.193
Mo	1 (1 to 1)	1 (0 to 2)	–0.308	.76	–0.046

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test (*P*<.05), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as "Re," the table phase as "Ta," the transporting phase as "Tr," and the mouth phase as "Mo."

[b] Larger values in comparisons between chopstick and spoon conditions.

In female participants (Table 4), changes in shoulder flexion, abduction, and neck right lateral flexion angles across most feeding phases were smaller with chopsticks than with a spoon. Additionally, changes in shoulder internal rotation, wrist ulnar deviation, and neck and hip flexion were smaller in some phases when using chopsticks. Elbow flexion changes during the reaching phase and forearm pronation changes during the transporting phase were larger in the chopstick condition, contrasting with smaller changes in elbow flexion during the table and mouth phases and smaller forearm pronation changes during the table phase. Wrist dorsiflexion and neck right rotation changes were similar across all feeding phases. Significant differences were found in small to large effect sizes.

XSL•FO
RenderX

Table . Comparison of joint angle changes between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in female participants[a].

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Shoulder flexion					
Re	**11 (7 to 17)**	**25 (15 to 33)**[b]	–3.98	<.001	–0.614
Ta	**1 (1 to 1)**	**2 (1 to 3)**[b]	–2.485	.01	–0.383
Tr	**7 (5 to 13)**	**19 (15 to 28)**[b]	–3.875	<.001	–0.598
Mo	**3 (2 to 4)**	**5 (4 to 7)**[b]	–3.528	<.001	–0.544
Shoulder abduction					
Re	**8 (6 to 14)**	**13 (6 to 18)**[b]	–2.624	.01	–0.405
Ta	1 (0 to 2)	1 (1 to 2)	–1.581	.11	–0.244
Tr	**5 (3 to 8)**	**8 (5 to 17)**[b]	–2.589	.01	–0.399
Mo	**1 (1 to 2)**	**2 (1 to 3)**[b]	–2.45	.01	–0.378
Shoulder internal rotation					
Re	9 (6 to 12)	9 (6 to 16)	–1.651	.10	–0.255
Ta	**1 (1 to 2)**	**3 (2 to 3)**[b]	to 2.311	.02	–0.357
Tr	5 (3 to 6)	5 (4 to 10)	–1.894	.06	–0.292
Mo	**2 (2 to 4)**	**4 (3 to 5)**[b]	–3.076	.002	–0.475
Elbow flexion					
Re	**29 (22 to 32)**[b]	**21 (14 to 33)**	–2.172	.03	–0.335
Ta	**3 (2 to 4)**	**8 (6 to 12)**[b]	–3.702	<.001	–0.571
Tr	28 (22 to 32)	30 (24 to 34)	–0.782	.43	–0.121
Mo	**4 (3 to 6)**	**7 (5 to 8)**[b]	–3.702	<.001	–0.571
Forearm pronation					
Re	74 (63 to 87)	78 (67 to 82)	–0.678	.50	–0.105
Ta	**8 (5 to 15)**	**33 (23 to 45)**[b]	–4.015	<.001	–0.62
Tr	**58 (52 to 70)**[b]	**49 (38 to 53)**	–3.424	.001	–0.528
Mo	7 (3 to 11)	10 (6 to 12)	–1.616	.11	–0.249
Wrist dorsal flexion					
Re	12 (8 to 20)	11 (6 to 16)	–1.86	.06	–0.287
Ta	8 (5 to 12)	8 (3 to 13)	–0.504	.61	–0.078
Tr	5 (3 to 7)	7 (4 to 11)	–1.825	.07	–0.282
Mo	3 (2 to 4)	3 (2 to 3)	–1.13	.26	–0.174
Wrist ulnar deviation					
Re	**7 (3 to 9)**	**14 (10 to 25)**[b]	–3.354	.001	–0.518
Ta	**3 (1 to 3)**	**14 (6 to 24)**[b]	–4.015	<.001	–0.62
Tr	4 (3 to 5)	5 (3 to 6)	–1.269	.21	–0.196
Mo	1 (1 to 2)	1 (1 to 3)	–1.442	.15	–0.223

XSL•FO
RenderX

Motion direction and Phases	Joint angle change (degrees)				
	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Neck flexion					
Re	8 (3 to 13)	7 (5 to 10)	–1.408	.16	–0.217
Ta	1 (0 to 1)	1 (0 to 1)	–0.365	.72	–0.056
Tr	**8 (3 to 11)**	**10 (8 to 13)**[b]	–2.416	.02	–0.373
Mo	**2 (1 to -2)**	**3 (2 to 5)**[b]	–3.98	<.001	–0.614
Neck right lateral flexion					
Re	**2 (1 to 3)**	**4 (2 to 8)**[b]	–3.215	.001	–0.496
Ta	**0 (0 to 1)**	**1 (0 to 1)**[b]	–1.964	.05	–0.303
Tr	**2 (1 to 3)**	**3 (1 to 6)**[b]	–2.589	.01	–0.399
Mo	**1 (1 to 1)**	**2 (1 to 2)**[b]	–3.18	.001	–0.491
Neck right rotation					
Re	4 (2 to 5)	3 (1 to 5)	–1.303	.19	–0.201
Ta	0 (0 to 1)	0 (0 to 1)	–1.442	.15	–0.223
Tr	3 (2 to 4)	3 (2 to 4)	–0.956	.34	–0.148
Mo	2 (1 to 2)	2 (1 to 2)	–0.713	.48	–0.11
Hip flexion					
Re	6 (5 to 8)	4 (2 to 7)	–1.408	.16	–0.217
Ta	0 (0 to 0)	0 (0 to 0)	–1.512	.13	–0.233
Tr	4 (2 to 7)	4 (1 to 7)	–0.365	.72	–0.056
Mo	**1 (0 to 1)**	**2 (1 to 2)**[b]	–3.076	.002	–0.475

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test (*P*<.05), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b] Larger values in comparisons between chopstick and spoon conditions.

Hand Spatiotemporal Parameters

In both sexes (Table 5 for males and Table 6 for females), the reaching phase duration was longer with chopsticks than with the spoon. In contrast, the mouth phase duration was shorter with chopsticks. Additionally, female participants showed a shorter transporting phase duration when using chopsticks. These differences showed small to large effect sizes.

Table . Comparison of movement times between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in male participants[a].

Variable and Phase	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Movement time (seconds)					
Re	**4.9 (2.4 to 9.6)**[b]	**2.9 (1.8 to 3.5)**	–3.393	.001	–0.512
Ta	1 (0.7 to 2.2)	1 (0.7 to 1.4)	–0.438	.66	–0.066
Tr	0.8 (0.6 to 1.4)	0.9 (0.7 to 1.2)	–1.136	.26	–0.171
Mo	**0.4 (0.3 to 0.5)**	**0.7 (0.6 to 0.8)**[b]	–3.979	<.001	–0.600

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test (*P*<.05), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b] Larger values in comparisons between chopstick and spoon conditions.

XSL•FO
RenderX

Table . Comparison of movement times between chopstick use for eating solid foods and spoon use for eating semisolid foods during each feeding phase in female participants[a].

Variable an Phase	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Movement time (seconds)					
Re	**4.8 (3.2 to 7.3)**[b]	**2.8 (1.6 to 4.7**)	–3.702	<.001	–0.571
Ta	1.1 (0.9 to 1.8)	1.5 (1.3 to 1.7)	–0.852	.39	–0.131
Tr	**0.9 (0.8 to 1.1**)	**1.1 (1 to 1.2)**[b]	–2.103	.04	–0.324
Mo	**0.5 (0.4 to 0.7**)	**0.9 (0.8 to 1.2)**[b]	–3.98	<.001	–0.614

[a] Significant differences in each phase, as determined using the Wilcoxon signed rank test (*P*<.05), are shown as values in boldface. "*r*" denotes the effect size. Each feeding phase shows the reaching phase as “Re,” the table phase as “Ta,” the transporting phase as “Tr,” and the mouth phase as “Mo.”

[b] Larger values in comparisons between chopstick and spoon conditions.

During the transporting phase in both sexes (Table 7 for males and Table 8 for females), the actual distance traveled was shorter with chopsticks; however, the relative distance traveled was greater than that with a spoon. Velocity and maximum velocity were lower with chopsticks than with spoon. The timing of maximum velocity and the number of movement units used were comparable between utensils. Small to large effect sizes were observed for significant differences between utensil conditions.

Table . Comparison of hand spatiotemporal parameters between chopstick use for eating solid foods and spoon use for eating semisolid foods during the transporting phase in male participants[a].

Variables	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Actual distance traveled (meters)	**0.11 (0.08 to 0.12**)	**0.14 (0.11 to 0.18)**[b]	–3.782	<.001	–0.570
Relative distance traveled	**1.17 (1.04 to 1.35)**[b]	**1.07 (1.04 to 1.11**)	–2.062	.04	–0.311
Velocity (meters per second)	**0.11 (0.08 to -0.15**)	**0.16 (0.14 to 0.19)**[b]	–3.393	.001	–0.512
Maximum velocity (meters per second)	**0.18 (0.16 to 0.28**)	**0.25 (0.22 to 0.32)**[b]	–3.068	.002	–0.463
Timing of maximum velocity (%)	70.6 (61.7 to 83.2)	60.2 (45.1 to 75.8)	–1.704	.09	–0.257
Number of movement units	5 (3.8 to 8)	4 (3 to 7)	–1.325	.19	–0.200

[a] Significant differences, as determined using the Wilcoxon signed rank test (*P*<.05), are shown as values in boldface. "*r*" denotes the effect size.

[b] Larger values in comparisons between chopstick and spoon conditions.

XSL•FO
RenderX

Table . Comparison of hand spatiotemporal parameters between chopstick use for eating solid foods and spoon use for eating semisolid foods during the transporting phase in female participants[a].

Variables	Chopsticks, median (IQR)	Spoon, median (IQR)	*z* value	*P* value	*r* value
Actual distance traveled (meters)	**0.12 (0.11 to 0.14)**	**0.2 (0.17 to 0.22)**[b]	–3.91	<.001	–0.603
Relative distance traveled	**1.15 (1.07 to 1.2)**[b]	**1.05 (1.04 to 1.09)**	–3.424	.001	–0.528
Velocity (meters per second)	**0.12 (0.1 to 0.16)**	**0.18 (0.14 to 0.2)**[b]	–3.702	<.001	–0.571
Maximum velocity (meters per second)	**0.22 (0.19 to 0.29)**	**0.35 (0.28 to 0.37)**[b]	–3.875	<.001	–0.598
Timing of maximum velocity (%)	70.3 (63.4 to 75.9)	63.9 (56.4 to 69.2)	–1.338	.18	–0.206
Number of movement units	5 (4 to 6)	6 (4.5 to 6.5)	–1.171	.24	–0.181

[a] Significant differences, as determined using the Wilcoxon signed rank test (P<.05), are shown as values in boldface. "*r*" denotes the effect size.

[b] Larger values in comparisons between chopstick and spoon conditions.

Discussion

Principal Findings in Joint Kinematics

The aim of this study was to determine upper body joint angles and hand spatiotemporal parameters during the use of chopsticks for solid foods and a spoon for semisolid foods across all feeding phases by comparing these conditions. In addition to these findings observed in both sexes, using subgroup analysis, male participants demonstrated greater neck rotation motions, while female participants exhibited fewer wrist dorsiflexion angles under the chopstick condition. This comparison is crucial for clinical assessments of individuals with upper body sensorimotor dysfunctions, considering food customs in East Asia.

Various joint angle positions of the upper limb and neck were lower in the chopstick condition than in the spoon condition, except for most elbow flexion angles and some forearm pronation, wrist ulnar deviation (in females), and neck right rotation (in males) angles being higher in the chopstick condition. In addition, female participants showed smaller wrist dorsiflexion positions with chopstick use than with spoon use. However, neck and hip flexion angles were generally comparable between the conditions.

The changes in upper body joint angles across phases were smaller in the chopstick condition than in the spoon condition, except for greater or fewer changes in elbow and forearm joint angles with chopsticks, and greater wrist dorsiflexion and neck right rotation angles in male participants. Based on the comparison between the conditions, chopsticks required smaller joint angle positions and changes in shoulder and neck flexion and right lateral flexion across feeding phases and sexes.

These findings align with our hypotheses and the results from previous reports on shoulder and neck joint angles [6,12]. Chopstick use requires less control of shoulder joint motions due to how food is held compared with spoon use. This rationale is supported by the fact that chopsticks may securely hold food through a “scissor-pinching” motion, which reduces trapezius muscle activity compared with “pincer pinching” [26]. The scissor-pinching technique used with chopsticks allows for a stronger grip on the food [27]. The reduction in muscle activity may explain the smaller shoulder joint motions observed in the chopstick condition. In addition, the finding of fewer neck joint angle positions and changes in our chopstick condition is supported by findings in studies showing smaller head or trunk region motions while eating solid foods with a fork or a spoon (not-easy-to-drop-foods condition), compared with eating liquid foods with a spoon (easy-to-drop-foods condition) [8,11]. The actual differences in shoulder joint motions to control utensils and compensatory neck joint motions between chopstick use for solid foods and spoon use for semisolid foods should be clinically highlighted as normative information. However, the current experimental conditions did not clarify whether food was prone to being dropped, indicating the need for further studies.

Notably, in male participants, neck right rotation angles and angle changes during some phases were greater in the chopstick condition than in the spoon condition. Additionally, hip flexion angles and changes, which indicate trunk positioning and motion, did not differ significantly between conditions in either sex group. Proximal body part stability or movements are considered the same. Although they were not anticipated at the start of the study, these postures and movements could be important for clinical assessments. Proximal body part motions are often more pronounced when eating foods prone to being dropped [8,11], which was likely the case for the spoon condition in this study. However, this remains unclear and warrants further investigation under adjusted conditions.

Contrary to our hypothesis, motion extent at the middle and distal joints using chopsticks was not consistently reduced. For instance, when transporting food to the mouth, forearm pronation angle changes were greater in the chopstick condition than in the spoon condition. However, during the table phase (when taking food from the bowl), the changes were not significant in the chopstick condition. These joint motions are

XSL•FO
RenderX

thought to be essential for approaching and manipulating food directly.

Principal Findings in Hand Spatiotemporal Parameters

Hand spatiotemporal parameters suggest that, in contrast with spoon use, chopstick use required more time during the reaching phase and less time during the mouth phase in both sex groups, with female participants also showing shorter times during the transporting phase. These utensil and food type–related differences contradicted our hypothesis. Furthermore, this finding was not recognized in a previous study [6]. A possible reason for the contrasting results is that the previous study used pickles as the food item under both the chopstick and spoon conditions, whereas this study used pickles (solid food) with chopsticks and yogurt (semisolid food) with a spoon to align with the study aim. Yogurt may be easier to consume without chewing (requiring less time during the reaching phase) than pickles [28]. However, yogurt may be more prone to spilling than pickles (requiring more time during the transporting and mouth phases). Accordingly, the present findings related to duration differences in conditions lead to the assumption that this was due to utensil use along with respective food characteristics such as viscosity and not just dependent on utensil types.

Hand spatiotemporal parameters during the reaching movement have been reported [13,14], with clearly defined start and end positions and movement not disrupted. Accordingly, this study analyzed parameters related to distance, speed, motor control strategy, and smoothness during the transporting phase, which showed no disruption during reaching or transporting. The results were interpreted in accordance with the common concept of spatiotemporal parameters to contribute to clinical assessments.

First, the actual distance traveled by the hand from the bowl to the mouth was shorter with chopsticks than with a spoon. Furthermore, the hand holding the chopsticks moved at a slower speed. This slowness might have been caused by the shorter transfer distance [29] observed in the chopstick condition than in the spoon condition. These shorter and slower hand transfers might be due to the larger forearm rotations and smaller shoulder flexion and abduction motions associated with chopstick use, as noted earlier. In contrast, when transporting food with a spoon, the larger hand transfer distance and higher speed might be due to increased shoulder motions rather than forearm movements. Thereby, the effective hand spatiotemporal transfers during the transporting phase might be mainly caused by shoulder joint motion and not forearm joint motion. Contrastingly, forearm joint motion might affect well transfer of the utensil tip rather than the hand.

The relative distance (ie, the ratio between the actual and theoretical shortest distance) for the hand with chopsticks was greater than for the hand with the spoon, indicating lower efficiency in the hand spatial transfer. This also suggests that forearm rotation might make the movement in the tip of chopsticks efficient, not in the hand holding chopsticks. However, the relationship between hand spatiotemporal parameters as well as utensil tip motion and joint angles is beyond the scope of this study and considered an area for future research.

Additionally, we analyzed the timing of maximum velocity (motor control strategy) and number of movement units (smoothness) and found no obvious differences between utensil conditions. The comparable reaching strategies and movement smoothness between conditions might indicate that the healthy participants in this study had well-learned motor control for these modes of feeding. These characteristics are generally assessed in patients [13]; this study provides valuable data on comparable motor control and smoothness between 2 utensil-food type conditions in normal hand motions regarding such assessments.

Clinical Implications

Clinicians might determine appropriate eating utensils used by patients based on their ability to use them, considering factors such as paresis and intellectual ability [2]. Rheumatoid arthritis [30] and Parkinson disease [31] may also be included in target populations as they have altered upper body kinematics during functional tasks. The assessment can refer to the findings of this study regarding joint angle positions, ranges of joint motion, and hand spatiotemporal parameters as fundamental normal movement knowledge, which could support clinical judgments when selecting chopsticks for eating solid foods or a spoon for eating semisolid foods in East Asian countries. For instance, individuals with limited shoulder flexion or abduction or increased elbow flexion angles might be suited to using chopsticks to eat solid foods, which include small-sized pickles used in this study, bite-sized meats and vegetables, or sticky rice; conversely, those who demonstrate significant forearm rotation during the table phase might benefit from using a spoon to eat semisolid foods, which include yogurt, jelly dessert, or rice porridge. Movements for using a spoon to eat semisolid foods may also be consistent with those for using a spoon to eat liquid foods such as soup or solid foods such as separated rice and finely chopped vegetables. These combinations of utensils and foods, alongside the typical movements currently revealed, would be realistic about food customs in East Asia. Additionally, differences in hip flexion positions and motions (ie, trunk inclination) may not be related to utensil selection.

In the rehabilitation of eating difficulties in East Asian countries, when changing food forms from semisolid to solid, patients' forearm supination should be well confirmed for using chopsticks. Compared with the spoon condition, chopstick use typically involves shorter and slower hand movements during the transporting phase, which can be considered normal. If the nondominant hand is trained for chopstick use with solid foods, improvements as spoon use with semisolid foods in control strategy and movement smoothness—similar to those observed in this study—could be assessed, as recently reported [32,33]. Hand spatiotemporal parameters involving nondominant or impaired hands using utensils could also be evaluated based on the current findings. Although male participants showed greater neck right rotation positions and motions and female participants showed smaller wrist dorsiflexion positions in chopstick condition than in spoon condition should be assessed as individuals with normal state in each disease, these sex-related

XSL•FO
RenderX

characteristics would be emphasized particularly in diseases with ununiform sex ratio such as collagen disease and cervical spondylotic myelopathy.

Limitations and Future Research

The present results were derived from measurements obtained during specific movement tasks using an inertial motion-capturing system, which may limit the generalizability of the findings to feeding movements observed in clinical settings. Future studies should investigate whether certain foods could be more prone to being dropped by examining the effects of different eating utensils (eg, chopsticks with supportive features, forks, and spoons) and various types of foods (eg, viscosity, size, and forms) separately, as these may be confounding factors in the current experiments; moreover, these studies should define these effects on feeding movements. This approach might help verify and expand upon the current findings. Drawing connections between joint motions and the hand spatiotemporal parameters and utensil tip motions, despite being out of this study objectives but discussed, might also be a valuable method for more deeply understanding the differences between the current 2 utensil-food type conditions; thus, eating movements would be assessed well. Moreover, future research should also involve populations with upper body dysfunctions.

Conclusions

This study investigated upper body joint angles and hand spatiotemporal parameters during feeding movements involving chopsticks for solid food and a spoon for semisolid food across 4 feeding phases and both sexes. The results indicate that chopstick use, regardless of sex, generally requires smaller shoulder and neck joint angle positions and motions across phases but involves larger elbow and forearm joint motions based on the phase. Additionally, when transporting food from the bowl to the mouth using chopsticks, the hand travels a shorter distance and moves slowly. Spoon use, in contrast, involves larger elbow, forearm, and wrist ulnar deviation motions when taking up semisolid food from the bowl. Regarding sex differences, male participants exhibit increased right neck rotation positions and motions when using chopsticks compared with using a spoon, whereas female individuals demonstrate reduced wrist dorsiflexion positions under the chopstick condition relative to the spoon condition. These movement characteristics, influenced by the combination of utensils and food types specific to East Asia countries, could facilitate clinical assessments of feeding behaviors specific to food customs in individuals with upper body dysfunction and diseases specific to each sex and aid in determining adaptive utensils.

Funding

This study was supported by a clinical research support grant from the University of Miyazaki Hospital.

Authors' Contributions

JN contributed to study conceptualization and design, data acquisition, formal analysis and interpretation of data, writing and editing the manuscript, and approval of the final version. SM and SM contributed formal analysis and interpretation of data, reviewing and editing the manuscript, and approval of the final version. HA and EC contributed to study conceptualization and design, reviewing and editing the manuscript, and approval of the final version.

Conflicts of Interest

None declared.

References

1. Clavé P, Terré R, de Kraa M, Serra M. Approaching oropharyngeal dysphagia. Rev Esp Enferm Dig 2004 Feb;96(2):119-131. [doi: 10.4321/s1130-01082004000200005] [Medline: 15255021]
2. Iokawa K, Sone T, Fujita T, Tsukada T, Kaneda M, Hasegawa K. Functional and cognitive variables predicting successful use of chopsticks or a spoon by the paretic upper extremity in patients following stroke: a cross-sectional study. Top Stroke Rehabil 2019 Jan;26(1):1-5. [doi: 10.1080/10749357.2018.1536021] [Medline: 30355062]
3. Hill SW, Mong S, Vo Q. Three-dimensional motion analysis for occupational therapy upper extremity assessment and rehabilitation: a scoping review. The Open Journal of Occupational Therapy 2022;10(4):1-14. [doi: 10.15453/2168-6408.1901]
4. Nakatake J, Totoribe K, Chosa E. Whole-body joint angles necessary for eating with chopsticks. Jpn Occup Ther Res 2019;38:163-170. [doi: 10.32178/jotr.38.2_163]
5. Nakatake J, Totoribe K, Arakawa H, Chosa E. Exploring whole-body kinematics when eating real foods with the dominant hand in healthy adults. PLoS One 2021;16(10):e0259184. [doi: 10.1371/journal.pone.0259184] [Medline: 34710151]
6. Nagao T. Joint motion analysis of the upper extremity required for eating activities. Bull Health Sci Kobe 2004;19:13-31. [doi: 10.24546/00391929]
7. Chang BC, Huang BS, Chen CK, Wang SJ. The pincer chopsticks: the investigation of a new utensil in pinching function. Appl Ergon 2007 May;38(3):385-390. [doi: 10.1016/j.apergo.2006.03.009] [Medline: 16765903]
8. Chinju K, Yamamoto Y, Inada E, Iwashita Y, Sato H. Analysis of head motions during food intake in Japanese adults using a new motion capture system. Arch Oral Biol 2024 Apr;160:105908. [doi: 10.1016/j.archoralbio.2024.105908] [Medline: 38335700]

XSL•FO
RenderX

9. Inada E, Saitoh I, Nakakura-Ohshima K, et al. Association between mouth opening and upper body movement with intake of different-size food pieces during eating. Arch Oral Biol 2012 Mar;57(3):307-313. [doi: 10.1016/j.archoralbio.2011.08.023] [Medline: 21975117]
10. Doğan M, Koçak M, Onursal Kılınç Ö, et al. Functional range of motion in the upper extremity and trunk joints: nine functional everyday tasks with inertial sensors. Gait Posture 2019 May;70:141-147. [doi: 10.1016/j.gaitpost.2019.02.024] [Medline: 30875600]
11. van der Kamp J, Steenbergen B. The kinematics of eating with a spoon: bringing the food to the mouth, or the mouth to the food? Exp Brain Res 1999 Nov;129(1):68-76. [doi: 10.1007/s002210050937] [Medline: 10550504]
12. Inagaki Y, Ishida T, Sugimori H, et al. Functional range of motion for basic seated activities of daily living tasks. Front Sports Act Living 2025;7:1646326. [doi: 10.3389/fspor.2025.1646326]
13. de los Reyes-Guzmán A, Dimbwadyo-Terrer I, Trincado-Alonso F, Monasterio-Huelin F, Torricelli D, Gil-Agudo A. Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: a review. Clin Biomech (Bristol, Avon) 2014 Aug;29(7):719-727. [doi: 10.1016/j.clinbiomech.2014.06.013]
14. Schwarz A, Kanzler CM, Lambercy O, Luft AR, Veerbeek JM. Systematic review on kinematic assessments of upper limb movements after stroke. Stroke 2019 Mar;50(3):718-727. [doi: 10.1161/STROKEAHA.118.023531] [Medline: 30776997]
15. Robert-Lachaine X, Mecheri H, Larue C, Plamondon A. Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis. Med Biol Eng Comput 2017 Apr;55(4):609-619. [doi: 10.1007/s11517-016-1537-2] [Medline: 27379397]
16. Al-Amri M, Nicholas K, Button K, Sparkes V, Sheeran L, Davies JL. Inertial measurement units for clinical movement analysis: reliability and concurrent validity. Sensors (Basel) 2018 Feb 28;18(3):719. [doi: 10.3390/s18030719] [Medline: 29495600]
17. Zhang JT, Novak AC, Brouwer B, Li Q. Concurrent validation of Xsens MVN measurement of lower limb joint angular kinematics. Physiol Meas 2013 Aug;34(8):N63-N69. [doi: 10.1088/0967-3334/34/8/N63] [Medline: 23893094]
18. Kim S, Nussbaum MA. Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks. Ergonomics 2013;56(2):314-326. [doi: 10.1080/00140139.2012.742932] [Medline: 23231730]
19. Pedro B, Cabral S, Veloso AP. Concurrent validity of an inertial measurement system in tennis forehand drive. J Biomech 2021 May 24;121:110410. [doi: 10.1016/j.jbiomech.2021.110410] [Medline: 33852942]
20. Kwakkel G, van Wegen EEH, Burridge JH, et al. Standardized measurement of quality of upper limb movement after stroke: consensus-based core recommendations from the second stroke recovery and rehabilitation roundtable. Neurorehabil Neural Repair 2019 Nov;33(11):951-958. [doi: 10.1177/1545968319886477] [Medline: 31660781]
21. Nakatake J, Totoribe K, Chosa E, Yamako G, Miyazaki S. Influence of gender differences on range of motion and joint angles during eating in young, healthy Japanese adults. Prog Rehabil Med 2017;2:20170011. [doi: 10.2490/prm.20170011] [Medline: 32789218]
22. Mesquita IA, Fonseca PFPD, Borgonovo-Santos M, et al. Comparison of upper limb kinematics in two activities of daily living with different handling requirements. Hum Mov Sci 2020 Aug;72:102632. [doi: 10.1016/j.humov.2020.102632] [Medline: 32452388]
23. Nakatake J, Arakawa H, Tajima T, Miyazaki S, Chosa E. Age- and sex-related differences in upper-body joint and endpoint kinematics during a drinking task in healthy adults. PeerJ 2023;11:e16571. [doi: 10.7717/peerj.16571] [Medline: 38144196]
24. Waslen A, Friesen KB, Lang AE. Do sex and age influence scapular and thoracohumeral kinematics during a functional task protocol? J Appl Biomech 2024 Feb 1;40(1):29-39. [doi: 10.1123/jab.2023-0085] [Medline: 37917968]
25. Zieliński G. Effect size guidelines for individual and group differences in physiotherapy. Arch Phys Med Rehabil 2025 Dec;106(12):1844-1849. [doi: 10.1016/j.apmr.2025.05.013]
26. Yoo IG, Yoo WG. Comparison of effects of pincer- and scissor-pinching modes of chopstick operation on shoulder and forearm muscle activation during a simulated eating task. J Phys Ther Sci 2012;24(10):953-954. [doi: 10.1589/jpts.24.953]
27. Chen YL. Effects of shape and operation of chopsticks on food-serving performance. Appl Ergon 1998 Aug;29(4):233-238. [doi: 10.1016/s0003-6870(97)00046-x] [Medline: 9701536]
28. Aguayo-Mendoza MG, Ketel EC, van der Linden E, Forde CG, Piqueras-Fiszman B, Stieger M. Oral processing behavior of drinkable, spoonable and chewable foods is primarily determined by rheological and mechanical food properties. Food Qual Prefer 2019 Jan;71:87-95. [doi: 10.1016/j.foodqual.2018.06.006]
29. Gordon J, Ghilardi MF, Cooper SE, Ghez C. Accuracy of planar reaching movements. II. Systematic extent errors resulting from inertial anisotropy. Exp Brain Res 1994;99(1):112-130. [doi: 10.1007/BF00241416] [Medline: 7925785]
30. Gur Kabul E, Unver F, Alptekin A, et al. The effect of rheumatoid arthritis on upper extremity functions: a kinematic perspective. Int J Rheum Dis 2022 Nov;25(11):1279-1287. [doi: 10.1111/1756-185X.14421] [Medline: 35965381]
31. Romano P, Pournajaf S, Ottaviani M, et al. Sensor network for analyzing upper body strategies in Parkinson's disease versus normative kinematic patterns. Sensors (Basel) 2021 May 31;21(11):3823. [doi: 10.3390/s21113823] [Medline: 34073123]

XSL•FO
RenderX

32. Sawamura D, Sakuraba S, Suzuki Y, et al. Acquisition of chopstick-operation skills with the non-dominant hand and concomitant changes in brain activity. Sci Rep 2019 Dec 31;9(1):20397. [doi: 10.1038/s41598-019-56956-0] [Medline: 31892724]
33. Sawamura D, Sakuraba S, Yoshida K, et al. Chopstick operation training with the left non-dominant hand. Transl Neurosci 2021 Jan 1;12(1):385-395. [doi: 10.1515/tnsci-2020-0189] [Medline: 34721894]

Edited by S Munce; submitted 19.Apr.2025; peer-reviewed by P Gawda, S Hill; revised version received 09.Nov.2025; accepted 13.Jan.2026; published 06.Mar.2026.

Please cite as:
Nakatake J, Maeda S, Miyazaki S, Arakawa H, Chosa E
Comparison of Upper Body Joint and Hand Motions in Eating Solid Foods With Chopsticks and Semisolid Foods With a Spoon in Healthy Males and Females: Observational Study
JMIR Rehabil Assist Technol 2026;13:e76239
URL: https://rehab.jmir.org/2026/1/e76239
doi:10.2196/76239

© Jun Nakatake, Shogo Maeda, Shigeaki Miyazaki, Hideki Arakawa, Etsuo Chosa. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Barriers and Enablers for the Intention to Use Assistive Robotics Among People With Spinal Cord Injury and Those Involved in Their Care: Qualitative Study

Susanne Frennert[1], PhD; Johanna Persson[1], PhD; Eva Díez-Rodríguez[2], MSc; Monica Alcobendas-Maestro[2], MSc; Fátima Villamayor Vega[2], MSc; Antonio Oliviero[2], PhD

[1]Department of Design Sciences, Division of Ergonomics and Aerosol Technology, Ingvar Kamprad Designcentrum (IKDC), Lund University, FormStråket 12, Lund, Sweden

[2]FENNSI Group, Hospital Nacional de Parapléjicos and “Instituto de Investigación Sanitaria de Castilla-La Mancha” (IDISCAM), SESCAM, Toledo, Spain

Corresponding Author:
Susanne Frennert, PhD
Department of Design Sciences, Division of Ergonomics and Aerosol Technology, Ingvar Kamprad Designcentrum (IKDC), Lund University, FormStråket 12, Lund, Sweden

Abstract

Background: Spinal cord injury (SCI) may, and often does, profoundly reshape daily life, altering physical abilities, social roles, and personal identities. While assistive technologies, including assistive robotics, are often framed as solutions to re-establish independence, their adoption is shaped by practical, emotional, and social considerations as well as functional qualities. Individuals with SCI, their relatives, and health care professionals need to navigate complex dynamics when encountering assistive robotics. Understanding how assistive technologies are perceived and positioned in everyday life may help developers and designers create assistive robotics that are meaningful and useful for intended users.

Objective: The aim of this qualitative study was to explore how individuals with SCI, relatives, and health care professionals working with patients with SCI perceive and describe the possibilities and limitations of assistive robotics. The study sought to understand the factors that influence the intention to use assistive robotics among individuals with SCIs.

Methods: We used a qualitative approach, conducting semi-structured interviews and participatory workshops in Sweden and Spain. In total, the study involved 18 interview participants with SCI, 21 workshop participants with SCI, 12 relatives, and 26 health care professionals. The interviews and workshops elicited reflections on participants’ experiences, expectations, and concerns regarding assistive robotics in general and supernumerary robotic limbs in particular. Data were analyzed using reflexive thematic analysis, with a focus on interpreting the meanings embedded in participants’ narratives.

Results: The analysis showed that participants’ engagement with assistive robotics was influenced by expectations of technological benefits and by practical constraints in everyday life. The main barriers identified were practical constraints, including the subthemes “navigating a changing reality,” “difficulties with awareness and access” and “concerns about costs”; and interaction with robots, including “doubts about meaningfulness,” “uncertainty regarding reliability and safety,” “uneasiness about competence” and “apprehension of social norms.” Participants’ visions of enhanced self-efficacy through assistive robotics were described as important enablers of the intention to use and motivation to try assistive robotics. Shared expectations and concerns about future technologies (technological imaginaries) also influenced how participants talked about assistive robotics.

Conclusions: Rather than presenting assistive robotics as an inevitable progression toward greater autonomy, this study highlights the complexities and contingencies that shape how individuals relate to assistive robotics in general and supernumerary robotic limbs in particular. Participants’ responses illustrate that robotic assistance is not merely a question of technological feasibility but is deeply entangled with embodied experiences, shifting identities, and evolving social relations. While visions of independence through assistive robotics remain compelling among participants, sociotechnical imaginaries coexist with concerns about meaningful engagement, reliability, safety, competence, and social norms, as well as challenges related to transition periods, costs, and limited awareness and access to assistive robotics.

(JMIR Rehabil Assist Technol 2026;13:e72080) doi:10.2196/72080

KEYWORDS

barriers; enabler; assistive robotics; spinal cord injury; spinal cord injury care

XSL•FO
RenderX

Introduction

Spinal cord injury (SCI) has a substantial impact on quality of life, affecting physical, psychological, social, and economic aspects of an individual's life. The severity and level of the injury influence the extent of this impact, often resulting in paralysis or loss of sensation and a need for extensive assistance in everyday life [1]. This, in turn, negatively affects the ability of individuals with SCI to independently perform everyday tasks, such as hygiene, eating, and dressing. Frequently, they depend on assistance from caregivers, relatives, and various assistive aids [2].

In robotics labs around the world, several projects are being conducted to develop valuable robotic assistance for people with SCI, such as mobile manipulators [3] and robot-assisted feeding [4]. In this paper, we use the term "assistive robotics" to refer to robotic devices that provide physical assistance for everyday activities to people with upper body-limb disabilities (eg, robotic arms, powered gloves, and exoskeletons that support feeding, reaching, dressing, or grooming).

However, despite the potential benefits of assistive technologies, abandonment rates remain high [5], often due to technology-driven design that overlooks the social and cultural factors affecting user acceptance [6]. A key factor influencing user abandonment is the intended users' perception of the technology in question [7-9].

Conceptual Framework and Research Gap

The perception of technology and its relation to the intention to use are explained by the Technology Acceptance Model (TAM) [10] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [11]. TAM focuses on personal beliefs about usefulness and ease of use [10], while UTAUT adds social influence and the broader conditions surrounding the use of technology [11]. In the context of individuals with SCI, the models suggest that the intentions to use assistive technologies are influenced not by personal beliefs but also by the views and practices of caregivers, relatives, and health care professionals (HCPs) who organize care and everyday use [12].

At the same time, TAM and UTAUT were developed for information systems and "able-bodied individuals" [10,11], and may therefore not fully capture the situated, embodied, and relational aspects of technology adoption in a disability context. Building on these models, work in assistive technologies has emphasized the importance of perceived meaningfulness, identity, and the organization of support around assistive technologies [7,9,13,14]. Concepts from Science and Technology Studies such as sociotechnical imaginaries draw attention to shared visions and narratives about how emerging technologies might reshape care and independence [15-19]. Sociotechnical imaginaries highlight how hopes, fears, and expectations about technologies are collectively produced.

Existing research on assistive robotics and people with SCI has primarily addressed rehabilitation and technical feasibility. Studies have investigated robots for rehabilitation [20-22], work participation [23,24], technology use and priorities [25,26], outcome measures [27] and user-centered design guidelines [28]. Several recent systematic reviews and state-of-the-art papers synthesize the development of upper-limb assistive and rehabilitation robots, including exoskeletons, end-effector devices, soft exosuits for people with cervical SCI and other neuromuscular conditions [29-33]. The reviews describe control strategies, clinical indications, motor outcomes, and identify challenges, such as comfort, wearability, and robustness. Furthermore, a systematic review has examined the barriers and facilitators to exoskeleton use [34] and broader acceptance of assistive technologies among people with motor disabilities [9].

Across the existing literature, user perspectives are most often operationalized in terms of usability, comfort, and short-term feasibility, focusing on the person undergoing rehabilitation [9,29,32]. While some studies included HCPs and, more rarely, relatives as respondents, they are not systematically treated as stakeholders, even though they are often responsible for procuring assistive technologies, providing training, and embedding them in care routines [34]. There are also very few empirical studies on how emerging technologies such as supernumerary robotic limbs (SRLs) are imagined before they reach the market [30,33], which leaves a gap in understanding how assistive robotics are positioned within the everyday lives of people with SCI and those involved in their care.

To address the identified gap, this study is conducted within the HARIA project, which develops SRLs to augment upper-limb function in people with substantial arm and hand impairments [35]. SRLs are wearable robotic arms and fingers, designed to enhance human sensorimotor abilities, particularly in tasks involving movement and coordination [36]. For individuals with SCI, SRLs are proposed as aids to compensate for lost functions in impaired limbs, provide necessary assistance, and substitute for impaired limbs in terms of functionality [37].

Aim and Research Question

The objective of this qualitative study was to explore how individuals with SCI, their relatives, and HCPs conceptualize, engage with, and negotiate the possibilities and limitations of assistive robotics in general, with a particular focus on SRLs. More specifically, we examined (1) intentions to use emerging assistive robotics that are not yet widely available, (2) how these intentions are shaped by experiences of both used and abandoned assistive technologies, and (3) how perspectives differ or converge across stakeholders, including individuals with SCI, relatives, and HCPs in Sweden and Spain.

We asked: What barriers and enablers shape the intention to use assistive robotics among people with SCI and those involved in their everyday care in Sweden and Spain? By exploring this question, we aimed to clarify how assistive robotics are positioned within the everyday realities of potential users and their caregivers and to inform the design and implementation of emerging assistive robotics.

Methods

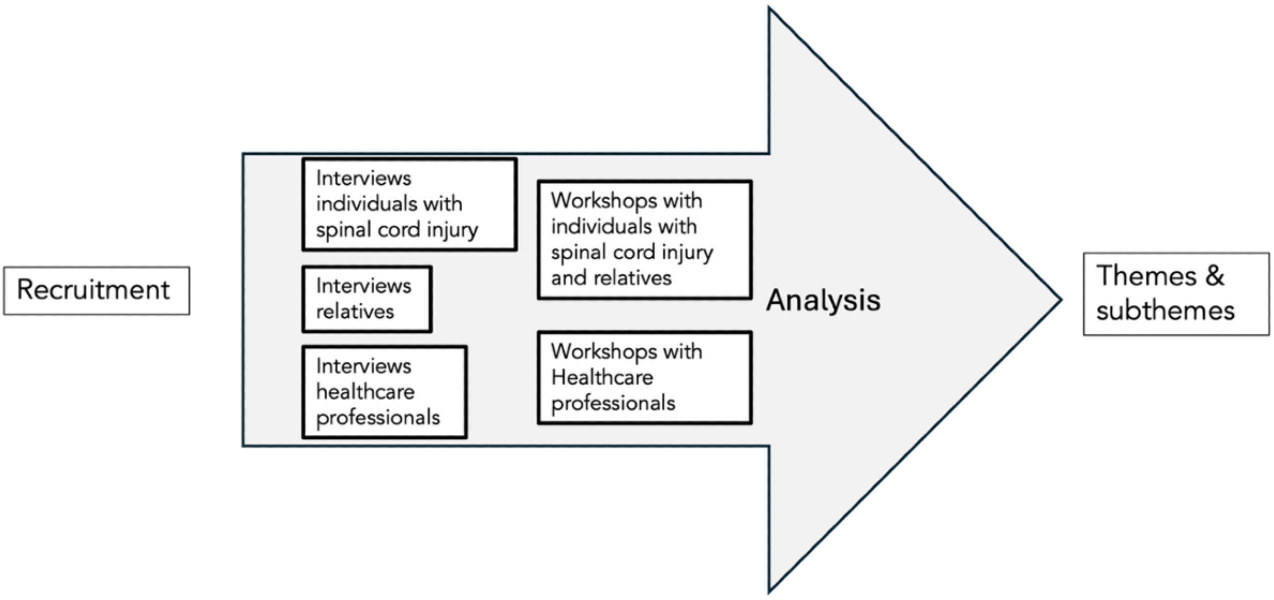
Study Design

This qualitative exploratory study is part of a broader research project centered around SRLs [35,37-39]. We adopted an

XSL•FO
RenderX

interpretive, constructivist approach [40] and used reflexive thematic analysis [41] to explore how different stakeholders make sense of assistive robotics in the context of SCI. Within this study, our focus is on individuals with SCIs and their intention to use assistive robotics, with a particular emphasis on SRLs. We combined individual semi-structured interviews [42] with participatory workshops [43] to capture both in-depth personal narratives and group-based discussions and co-imaginings of future assistive robotics (Figure 1).

Figure 1. Study design.



The multidisciplinary research team included researchers in rehabilitation engineering, design science and technology studies, and clinicians, several of whom had prior experience of SCI rehabilitation and assistive technology research. The diversity of the research team informed the development of the interview and workshop guides and the interpretation of the data.

Setting and Participants

To understand the social dynamics and contextual factors pertinent to assistive robotics, our participant pool included not only individuals with SCI but also HCPs specializing in SCIs, professionals working closely with individuals affected by SCIs, as well as relatives and caregivers. The study was conducted in both Spain and Sweden. In Sweden, participant recruitment involved collaboration with Paraplegic Associations, using their networks to disseminate information about the study to potential participants. Additionally, outreach efforts were extended to various health care facilities specializing in SCIs. In Spain, recruitment targeted HCPs, individuals with SCIs, and relatives through collaboration with Sescam - Gobierno de Castilla-La Mancha, a hospital and rehabilitation facility specializing in SCIs, which facilitated the identification and inclusion of relevant participants.

We focused on adults with cervical SCI because SRLs in the HARIA project aim to augment arm and hand function [35]. We also included relatives and HCPs, based on the assumption that the intention to use assistive robotics is affected by those who support and care for patients with SCI.

Eligibility and Sample

We contacted potential participants through email and phone. From this pool, we selected a purposive sample by applying specific criteria. To qualify for participation, individuals needed to fulfill the following criteria: (1) be 18 years or older, (2) have acquired a cervical SCI as an adult, (3) live in the community (not in long-term institutional care), and (4) be able to participate in an interview or workshop in Swedish or Spanish. Eligibility extended to those who were either a relative of an individual with a SCI or HCPs working with patients with SCI.

We aimed for heterogeneity to capture a wide range of lived experiences, needs, and expectations. Therefore, we did not set inclusion and exclusion criteria regarding time since injury, lesion level within the cervical region, functional level, or previous experience with assistive technologies or robotics. None of the participants had prior hands-on experience with SRLs, as SRLs were introduced by the researchers during the workshops (Multimedia Appendix 1).

Participants were recruited between January 2023 and September 2023. Recruitment was discontinued when we considered that data saturation was reached, indicating that further interviews produced repetitive information and no new themes or responses emerged [41].

Data Collection

Data were gathered through audio-recorded qualitative semistructured interviews [42,44] and participatory workshops [43]. The first author formulated the interview and workshop guides (Multimedia Appendices 1-5), subsequently refining them in collaboration with the rest of the research team [42,45].

The interview questions probed into the experiences of individuals living with SCIs, exploring their current use of assistive technologies and those received but not used (Multimedia Appendix 2). Such questions were included to situate participants' views on robotics within their broader histories of adopting, adapting to, and sometimes abandoning assistive technologies. Later during the interview, the focus shifted to assistive robotics technology, with questions aimed at understanding the interviewees' experiences with assistive technologies. During the latter part of the interview, the interviewer provided a brief introduction to SRLs, as the interviewees lacked prior knowledge, and SRLs are not yet commonly available on the market but are currently under development in research laboratories around the world [36,37,46]. Preferences and ideas regarding the use of assistive robotics technology were explored by asking questions about the functions requested, the required design for actual use, and how the interviewees would describe their needs and wishes concerning assistive robotics.

Interview questions and workshop activities for relatives and HCPs followed the same overall structure as for individuals with SCIs but were designed to explore their insights into the needs of the end-users (individuals with SCI) from their relational and professional standpoints (Multimedia Appendices 3-5). Our questions did not delve into the requirements of relatives and HCPs to acquire, use, or prescribe robotic devices; instead, we focused on their perceptions of the needs of end-users (individuals with SCI). Throughout the sessions, the researchers maintained a field diary. The diary served as a repository for initial analytical reflections, interpretations, and first impressions.

Qualitative semi-structured interviews were conducted either over the phone or digitally by the first [SF] and second [JP] authors in Sweden and by the third [ED], fourth [MA], and fifth [FV] authors in Spain, each lasting approximately 45 to 75 minutes. In total, 18 individuals with SCI, 10 relatives, and 13 HCPs were interviewed.

In the participatory workshops [43], the initial part focused on the lived experiences of individuals with SCIs, identifying both opportunities and challenges (Multimedia Appendix 1). The initial part of the workshop also covered currently used assistive technologies and those left unused. The subsequent part of the workshops centered on SRLs. Here, we showed both illustrations of different scenarios in which SRLs were used and videos of users interacting with SRLs (Multimedia Appendix 1). The same visual material was used across workshops to support comparability between groups of stakeholders and countries. The participants were asked to openly discuss their perceptions of the different scenarios, including perceived barriers and enablers associated with using SRLs. The participatory workshops, with an average duration of 2 hours, included one researcher as the workshop moderator and another responsible for comprehensive notetaking [43]. Upon conclusion, all materials and participant notes were collected.

Two in-person workshops were held in Sweden at Lund University by the first [SF] and second [JP] authors, each lasting around 2 hours. In Spain, 3 in-person workshops were conducted by the third [ED], fourth [MA], and fifth [FV] authors at Sescam - Gobierno de Castilla-La Mancha, a specialized hospital and rehabilitation facility for patients with SCI. Workshops were conducted in small groups and included individuals with SCI and relatives and, in separate workshops, HCPs working in SCI rehabilitation. In total, 21 individuals with SCI, 2 relatives, and 13 HCPs took part in participatory workshops in Sweden and Spain (Table 1).

Table . Participants in the qualitative semistructured interviews and participatory workshops.

Category	Qualitative semistructured interviews	Male	Female	Age (years), mean	Participatory workshops	Male	Female	Age (years), mean
Participants with spinal cord injury	18	14	4	61	21	16	5	51
Relatives	10	3	7	N/A[a]	2	1	1	N/A
HCPs	13	1	12	N/A	13	2	11	N/A

[a] Not available.

Data Analysis

The interviews were transcribed verbatim and subjected to thematic analysis following Clarke and Brown [41]. Initially, the first and the second author individually reviewed the transcripts, immersing themselves in the data through multiple readings. During the initial stage, the researchers systematically assigned codes to expressions and narratives related to barriers and enablers to the intention to use assistive robotics technology. Subsequently, the first and the second author compared codes and resolved any disagreements. The codes were organized into potential key themes and subthemes [41].

Fieldnotes and material from the workshops were analyzed using the same coding framework as the interview transcripts [41]. The shared coding framework allowed us to triangulate across data sources and explore how the same issues were discussed in one-to-one interviews and in group settings (ie, during workshops). Where workshop data introduced new nuances, the coding framework was adjusted and reapplied to interview transcripts.

During the interpretive coding phase, barriers and enabler themes were developed inductively. Initial codes captured a broad range of practical hindrances, concerns, and hopes related to assistive robotics. The codes were gradually organized into

XSL•FO
RenderX

categories that reflected patterns across the data (Table 2 presents the final set of themes and subthemes).

For the purpose of verification, workshops dedicated to member checking were organized. Member checking is a qualitative research practice used to ensure data accuracy and credibility by actively involving participants in the data analysis process [47-49]. During the member checking workshops, research findings were presented to individuals with SCIs, their relatives, and HCPs for feedback. The member checking workshops facilitated the refinement of themes to ensure close alignment with participants' lived experiences of SCI. Quotations are used to illuminate the themes.

Table . Identified barriers and enablers.

Themes	Subthemes
Barriers	
Practical constraints	• Navigating a changing reality • Difficulties with awareness and access • Concerns about costs
Interactions with assistive robotics	• Doubts about meaningfulness • Uncertainty regarding reliability and safety • Uneasiness about competence • Apprehension of social norms
Enablers	
Sociotechnical imaginaries	• Imaginings of greater self-efficacy • Imaginings of fulfillment of functional requirements

Ethical Considerations

Ethical approval for this study was obtained from the Swedish Ethical Review Authority (Dnr 2022-03154-01) and Comité Ético CHT Eic-chto-secretario at Toledo Hospital in Spain (CEIC-961). The study was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent before their participation. Participants were informed in advance that they would not receive any monetary compensation for their time and inconvenience and offered no direct personal benefits. Confidentiality and anonymity of the data provided were maintained throughout the research and publication process.

Results

Sample Characteristics

In total, 41 individuals participated in interviews and 36 in participatory workshops (Table 1). Participants were recruited from 2 countries (Sweden and Spain) and 3 stakeholders' groups (individuals with SCI, relatives, and HCPs). Among participants with SCI, most were men in their fifties and sixties (mean age 61 y in interviews and 51 y in workshops), all with cervical injuries and varying functional levels (ie, most used wheelchairs and required different degrees of assistance with everyday activities such as eating and dressing). Relatives were mainly women, typically partners and close family members. The HCPs were predominantly women and represented a range of professions working with SCI rehabilitation (eg, physiotherapists, rehabilitation physicians, and biomedical engineers).

While the national health care systems differed, for example, in Sweden, participants relied to a greater extent on personal assistants who are paid by the municipalities [50], whereas in Spain, participants with SCIs often depended on their nonpaid relatives [51], similar themes emerged across the 2 country contexts. Such contextual differences, therefore, mainly function as a backdrop for the findings rather than as a primary focus of comparison.

Overview Themes

In our interpretive reading of the material, the data revealed that participants frequently described barriers impeding their intention to use assistive robotics, both generally and specifically in the context of SRLs. Through reflexive thematic analysis, we organized the data into 2 main barrier themes, "practical constraints" and "interactions with assistive robotics," and 1 main theme capturing enablers of intention to use "imaginings of self-efficacy and functional fulfilment." Practical constraints include (1) navigating a changing reality, (2) difficulties with awareness and access, and (3) concerns about costs. Interactions with assistive robotics include (1) doubts about meaningfulness, (2) uncertainty regarding reliability and safety, (3) uneasiness about competence, and (4) apprehension of social norms. The main enabler theme encompassed two subthemes: (1) imaginings of greater self-efficacy and (2) imaginings of fulfillment of functional requirements. Key themes and subthemes are represented in Table 2.

Barrier – Practical Constraints

Overall, 3 subthemes related to barriers associated with practical constraints were identified: "navigating a changing reality," "difficulties with awareness and access," and "concerns about costs." The barrier themes were voiced across all stakeholder groups, although the specific experiences they referred to sometimes differed between individuals with SCI, relatives, and HCPs.

Navigating a Changing Reality

The subtheme, "navigating a changing reality," captures how participants described ongoing changes over time as a result of

XSL•FO
RenderX

the injury, aging, fluctuating symptoms, and evolving technologies as a barrier to engaging with assistive robotics. Most participants, including individuals with SCIs, relatives, and HCPs, described living with a SCI as a constant experience of transition periods. Transitions encompassed both the long-term adjustment to altered functionalities and needs and short-term fluctuations in pain and fatigue. Participants illustrated these experiences in expressions, such as

> *It was a brutal impact … from white to black. It shocked me in absolutely everything; it changed my life overnight. Everything has changed.* [SCI_6]
>
> *As a person with a SCI, you experience neurogenic pains that persist. Some days, it's really tough.* [SCI_15]

Those with SCIs also felt that immediately after the accident, they received considerable rehabilitation and assistance at the hospital, but upon discharge, they were largely on their own:

> *But now, after being discharged, it's difficult... no one is keeping track of me...* [SCI_4]

Such transition periods became a barrier to the intention to use assistive technology as needs, energy levels, and abilities fluctuated throughout the day and changed over time. Consequently, participants were less willing to invest time and energy in assistive technology due to the constant changes, along with fatigue and pain.

Within this subtheme, HCPs frequently cited a barrier to their intention to adopt assistive robotics. They pointed out that rapid technological developments impacted their work, for example, by requiring training in new devices, adapting clinical routines to new technology, and troubleshooting technology during already time-constrained consultations, which they perceived as increasing workload. The continuous changes and advancements placed a strain on them and their engagement with advanced technology. HCPs referred to transition periods in adapting to new technology and the continuous need to learn how to use it. As 1 HCP stated:

> *As a disadvantage, it might seem that we would have to learn and train ourselves on these new devices.* [HCP_18]

Overall, the experience of transitions and constant changes acted as barriers to the intention to use assistive robotics technology.

Difficulties With Awareness and Access

The subtheme, "difficulties with awareness and access," illustrates the barriers individuals with SCIs encounter when interacting with HCPs and others, both in understanding their injury's impact on their lives and in discovering available assistive aids. The subtheme was mainly articulated by individuals with SCI and their relatives, with HCPs offering a partly contrasting perspective on the availability of assistive robotics.

The obstacles outlined by individuals with SCIs included the lack of knowledge among municipality workers, reliance on specialized care professionals, and the challenge of self-identifying their needs:

> *Municipal workers lack knowledge about SCIs…we must have availability to direct contact with neurorehabilitation for assistive device support. We as patients typically need to identify our problems and propose potential aids, which may then be either adapted or developed for our individual needs. However, there is no predefined 'catalog' for selection.* [SCI_3]

Many participants emphasized the lack of support and assistance regarding assistive aids and knowledgeable personnel. They highlighted this by narrating

> *Above all, there is a lack of support on the issue of rehabilitation, there should be more means and more availability, both in terms of aids and personnel.* [SCI_5]
>
> *…I actually don't know who I would turn to if I wanted something [aids].* [SCI_12]
>
> *I don't know of any aid. As for support, there isn't much…* [SCI_8]

In contrast, HCPs perceived the availability of numerous assistive technologies. However, many also stressed that they did not recommend or prescribe them:

> *There are a lot of aids. For example, to eat and drink there are devices that I don't remember the name of because I have never recommended them.* [HCP_11]

In general, most participants with SCIs did not rely on advanced assistive technology but on other humans such as relatives and personal assistants. Their dependence on assistive technology mainly centered on wheelchairs and some specialized aids for the bathroom and kitchen. When it came to more advanced assistive technology and robotics, only a limited number of the participants (n=8) had experiences through participating in trials, testing at rehabilitation facilities, or using smart home technology with voice commands to control lighting, TV, and radio.

Concerns About Costs

Participants with SCI commonly faced financial pressure due to the inability to work or being limited to part-time employment. The subtheme, "concerns about costs," emphasizes a significant barrier to the intention to use assistive robotics technology tied to financial constraints. Concerns about the affordability of assistive technologies were raised by individuals with SCI, relatives, and HCPs.

In Sweden, known for subsidized health care [50], individuals with SCIs highlighted, like their Spanish counterparts, that only some assistive aids are subsidized [51]. Obtaining advanced assistive technology for free was perceived as challenging and almost impossible.

Some participants were aware of assistive robotics technology through the media but mentioned that trying it out required paying for it themselves. They perceived themselves as a disadvantaged group due to a lack of income and expressed the opinion that having money would make everything easier. For example, 1 participant stated,

XSL•FO RenderX

> *Well, the truth is that I do not have much. Around here, we have a saying that "with money" everything is easier. When you have money, you can afford everything, but if you can't pay, you will have nothing.* [SCI_11]

Another participant expressed concerns, stating,

> *In terms of support and aids, I think much is missing. Initially, you receive rehabilitation and physiotherapy support, but there comes a time when you either pay out of pocket or receive nothing.* [SCI_18]

HCPs and relatives also shared concerns about costs, viewing them as a barrier to accessing more advanced assistive technology and human support. When discussing SRLs, participants perceived them as expensive and beyond the scope of ordinary prescribed assistive technologies. As 1 HCP noted,

> *The aids must be very affordable so that patients can easily access them.*
>
> *[HCP_25]*

The cost of assistive technology was consistently perceived as a major barrier affecting the intention to use assistive robotics technology among most of the participants.

Barrier – Interaction With Assistive Robotics

While practical constraints are linked to tangible barriers in the daily lives of individuals with SCIs, impacting their intention to use assistive technology in general, a prominent theme emerged focusing on the specific challenges of interacting with assistive robotics. The main theme comprises 4 subthemes: doubts about meaningfulness, uncertainty regarding reliability and safety, uneasiness about competence, and apprehension of social norms.

Doubts About Meaningfulness

The participants reported several doubts about the meaningfulness of assistive robotics technology. Doubts about the meaningfulness of assistive robotics were expressed primarily by individuals with SCI, drawing on their own previous experiences of technologies that had promised a lot but delivered little, and were echoed by some HCPs.

A few of the participants had prior experiences of taking part in user testing of robotic devices and some had experiences with advanced technologies such as speech synthesis and Alexa. They emphasized that these technologies had not been very meaningful to use in everyday life:

> *The hospital's robotic arm I tried was large, with numerous cords and a complex setup, making it impractical for daily use. To be suitable for everyday life, I envision a more discreet, easily deployable, and user-friendly design with fewer machinery and electronic components to make it useful.* [SCI_9]

Another participant said:

> *I've tried using speech synthesis, but for two reasons, I've given up on it. Firstly, sometimes it goes well for a couple of sentences and then it turns into complete gibberish. Secondly, when constructing a text, you often don't have all the answers in your head at the beginning. You start writing, go back, and make changes, rearrange and so on. This results in a significant amount of editing work. Reversing, erasing, and adjusting are not as straightforward with voice. Dynamically building a text doesn't always mean having it fully clear in your head from the start.* [SCI_1]

These quotes illustrate that the participants questioned the usefulness of advanced technologies based on their prior experience. There were also participants who doubted the usefulness of assistive robotic technologies based on what they learned from other people and through media. They highlighted that they often hear about life-changing technologies from others but doubt the truth in these stories:

> *The media offer people like us [individuals with spinal cord injury] a thousand stories, and then 95% of the things are worthless.* [SCI_3]

They did not believe what they heard due to the lack of evidence. Also, HCPs questioned the lack of evidence regarding the benefits and usefulness of assistive robotic technologies:

> *I don't think there is much scientific evidence in this regard…* [HCP_23]

Uncertainty Regarding Reliability and Safety

Uncertainty about reliability and safety was mentioned mainly by relatives and health care professionals, while participants with SCI themselves did not express these concerns as frequently. Relatives stressed that assistive aids need not be overly sophisticated; from their perspective, low-tech solutions appeared safer. While assistive robotics technology may offer extensive functionality, it also seemed to heighten the apprehension about its reliability and safety. One relative emphasized,

> *...before it [a robot] performs a thousand functions, it has to be able to do a few basic ones in a safe and reliable way...* [REL_9]

HCPs also underscored the importance of reliability and safety, envisioning potential mishaps that could jeopardize the well-being of individuals with SCIs. One HCP illustrated,

> *Imagine it's [a robotic arm] feeding you, it turns off and leaves you halfway with the fork in your mouth, or it's combing your hair and leaves you with the comb tucked into the hair.* [HCP_13]

The uncertainty surrounding reliability and safety emerged as a significant barrier to the intention to use for both relatives and HCPs.

Uneasiness About Competence

Uneasiness about competence referred both to individuals with SCI doubting their own ability to manage complex technologies and to relatives and HCPs questioning whether potential users would be able to operate assistive robotics independently. While the participants with SCI did not raise a lot of concern about reliability and safety, as opposed to relatives and HCPs, they were instead uneasy about their own competence rather than the technologies. Almost all the participants perceived robotics

XSL•FO
RenderX

as advanced and complex technology and worried that it was not for them due to their limited technological competence but also their dependence on others for most things in life. This was illustrated through their expressions, as demonstrated by

> *I am no longer autonomous at all. I went from being the one who solved problems for others to not being able to do anything for anyone; now, they have to do everything for me.* [SCI_1]
>
> *I think that I've never used anything besides the wheelchair, which is manual and the crutches.* [SCI_12]
>
> *I don't have any experience, I don't use technology.* [SCI_7]

Relatives and health care professionals were also uneasy about the competence of individuals with SCI in using assistive robotics. They drew on their own experience of helping them to handle everyday life activities as well as technology. As one of the relatives said:

> *He will need another person to place or start the robotic arm... after all that would not be a solution because the person who is there would end up assisting him more than the robotic arm.* [REL_2]

This quote illustrates how the subthemes of doubts about meaningfulness and uneasiness about competence are interrelated. If the intended user does not have the competence to use the technology by themselves, it decreases the meaningfulness of the technology.

Apprehension of Social Norms

The subtheme, "apprehension of social norms," highlights participants' awareness of societal expectations and how others perceive them. Apprehension of social norms was voiced predominantly by individuals with SCI, with relatives and HCPs recognizing and sometimes sharing these concerns. Specifically, when discussing robotic limbs, many participants expressed concerns about being viewed as abnormal or strange.

They were already accustomed to being seen as "different" with 1 participant noting,

> *… people may see you as if you were a robot or something like that…people already look at us with a strange face.*
>
> *[SCI_13]*

The fear was that adding a robotic finger or robotic arm would further set them apart from what is considered "normal."

HCPs and relatives shared similar reservations, expressing doubts about the societal acceptance of robotics and robotic limbs. One HCP remarked,

> *I see the exoskeleton, and they look like 'Robocop.' Besides, who would put this device to work for them every day?* [HCP_8]

The comment not only points to worries about social appearance but also doubts about whether the everyday effort required to set up and use assistive robotics would be perceived as worthwhile.

As the last quotation illustrates, participants reported that thoughts about robotics and robotic limbs often evoke images from movies depicting humans becoming part-robot. However, their concerns also extend to the practical aspects of using such devices, tying back to the subtheme of the perception of meaningfulness.

Enablers

Alongside barriers, our analysis of the data revealed enablers for the intention to use assistive robotics, such as the imaginings of greater self-efficacy and independence. We interpreted the enablers as "imaginings" because they appear to be rooted in hopes and wishes for a different future. The imaginings were not specifically related to assistive robotics but rather to assistive aids in general or life in general. While the barriers described hindered participants' intentions to use assistive robotics significantly, the participants also emphasized enablers or imagined solutions. The 2 main enabler themes: "imaginings of greater self-efficacy" and "imaginings of fulfillment of functional requirements" capture how participants envisioned assistive robotics as potentially transforming everyday life, even while simultaneously doubting whether such visions would materialize.

Imaginings of Greater Self-Efficacy

A significant number of participants expressed a desire for increased independence and the ability to resume their pre-SCI lifestyle. While many participants had come to terms with their injury, its emotional impact remained, as they found themselves unable to lead the life they once did. Participants expressed a keen appreciation for any opportunity for improvement and assistance:

> *All aids that could potentially enhance my function would be very welcome.* [SCI_5]

Participants mostly discussed aids in general that could enhance their independence, rather than making specific reference to assistive robotics or robotic limbs. They did not specify the type of aids but rather conveyed their desires to carry out daily activities independently:

> *I would like to be able to use it [a robotic limb] mainly for cleaning myself, eating, drinking, putting on shoes, dressing, that kind of thing that allows me to be more autonomous. I don't know the possibilities, the more functions, the better. Give me as much autonomy as possible.* [SCI_2]

The majority of the participants with SCIs expressed imaginings of having greater self-efficacy. Such imaginings could be understood as a desire to explore all possible technologies or aids for increased independence. However, the empirical data showed that they mostly used low-tech technology and perceived much technology as lacking meaningfulness to them (as illustrated above).

Imaginings of Fulfilment of Functional Requirements

When discussing factors influencing the intention to use, "imaginings of fulfilment of requirements" emerged as one of the most significant enablers. Aids or assistive technology need to be, according to the participants, "user-friendly" and easy to

XSL•FO
RenderX

learn. Participants emphasised various requirements, with statements such as

> *..it is crucial to consider affordability, esthetic appeal and ease of application and maneuvering.* [HCP_3]
>
> *Resistance to wear and tear, absence of daily problems, ease of repair and reliable technical support are essential.* [REL_4]

If the technology met all these requirements, then it could serve as an enabler for intention to use, according to the participants.

Discussion

Principal Findings

This study explored how individuals with SCI, their relatives, and HCPs in Sweden and Spain conceptualize assistive robotics in general and SRLs in particular, and which factors shape their intention to use assistive robotics.

Our findings complement previous research that has focused on specific robotic devices and their usability [52,53] and on general patterns of assistive technology adoption and abandonment [7,13,14,54-56]. Coherent with previous research, participants in our study were interested in emerging technologies that could support activities of daily living; nevertheless, they also highlighted usability, cost, and stigma as key challenges. Our contribution lies in showing how such issues play out for emerging assistive robotics such as SRLs and in demonstrating how expectations and doubts are distributed across multiple stakeholder groups.

Relating the findings to TAM and UTAUT

Viewing our findings through the UTAUT and TAM lens [10,11], our barriers and enablers can be mapped onto core constructs while also extending them. The theme “doubts about meaningfulness” corresponds closely with performance expectancy or perceived usefulness. Thus, the participants questioned whether assistive robotics would make everyday life easier or add complexity. Subthemes of uncertainty regarding reliability, safety, and uneasiness about competence relate to effort expectancy and facilitating conditions. Hence, relatives and HCPs worried about malfunctions and safety, while individuals with SCI doubted their ability to operate assistive robotics, which were perceived as complex technologies. Apprehension of social norms directly reflects the UTUAT construct on social influence, as individuals with SCI anticipated being perceived as “too robotic” or “strange” when using visible robotic limbs in public.

The practical constraint subthemes: “navigating a changing reality,” “difficulties with awareness and access,” and “concerns about costs” speak directly to facilitating conditions. Participants highlighted limited information pathways, lack of knowledge in municipal services and financial barriers, even in subsidized health care systems like the Swedish one [50]. The findings underscore that the intention to use assistive robots is not only based on the individual attitude but shaped by broader health care system infrastructures and funding arrangements.

The enabler subtheme, “imaginaries of greater self-efficacy” and “imaginings of fulfilment of functional requirements,” align with the behavioral intention construct in TAM and UTUAT [10,11]. Participants imagined assistive robotics that were reliable, aesthetically acceptable, and well supported, and that assistive robotics would allow them to reclaim valued activities such as self-care, eating, and dressing. However, many participants also doubted that currently available or near-future assistive robotics would fully meet these requirements, suggesting a gap between positive behavioral intention and realistic expectations about actual use.

Sociotechnical Imaginaries

Building on our reference to “sociotechnical imaginaries” in the results, we recognized the interplay between expectations and broader sociotechnical narratives, echoing discussions of the myth of technology and its advancement in the participants’ perceptions of assistive robotics. As Mayor describes in her book *“Gods and Robots: Myths, Machines and Ancient Dreams of Technology”* [57] new technologies often elicit both hope and apprehension [58,59] and can be viewed as modern myths [60,61]. The imaginaries or enablers we identified in our analysis presented a paradox. We term the tension the “meaningfulness paradox,” as participants expressed doubt about the meaningfulness of assistive robotics, their competency in handling the technology, its safety, and reliability. At the same time, they also spoke about how assistive technology could enhance self-efficacy if it met all functional requirements (which they doubted).

Thus, while TAM and UTUAT focus primarily on individual beliefs and intentions [10,11], drawing on the concepts of sociotechnical imaginaries [15-19] enables us to situate the participants' beliefs within broader collective visions of what assistive robotics could or should do. Participants' imaginings of independence, reduced caregiver burden, and improved quality of life echo existing imaginaries of robotics as a solution to workforce shortages and increasing care demands [62-64]. At the same time, the participants' narratives revealed a persistent gap between idealized expectations and lived realities, including pain, fatigue, cognitive load, and bureaucratic hurdles. For example, some participants contrasted media portrayals of robots with their own experiences of bulky robotic arms that were too demanding to set up and operate.

The tension was especially visible when participants contrasted the promise of SRLs with their experiences of earlier assistive robotic devices they had tried, which turned out to be cumbersome and unreliable. As such, sociotechnical imaginaries therefore function both as enablers for adoption, motivating interest in trying out emerging technologies, and as a potential source of disappointment if the assistive robotic device cannot live up to the expectations. Our findings suggest that managing the expectation gap is important in order to avoid future abandonment of assistive robotics, echoing previous calls for more realistic and relational approaches to social robots and care technologies [22,65].

Implications for Research and Practice

Our findings suggest considerations for the future development of assistive robotics. They indicate that the participants’ perceptions and conceptualizations of assistive robotics were informed by ideas and sociotechnical imaginaries. The analysis underscores how idealized visions intersect with intention to use, where individuals may express a desire to use assistive robotics while simultaneously encountering barriers that complicate the translation of intent into action. Rather than framing identified barriers solely as obstacles to adoption, they can be understood as part of a broader negotiation between technical expectations and lived experiences [65]. This perspective shifts the focus from "non-compliant users" to assistive robotic systems and design that may not yet be well aligned with everyday realities of individuals with SCI and those who care for them.

For designers and engineers, our results highlight the importance of focusing on everyday meaningfulness rather than technical sophistication alone. Thus, prioritizing functions that users, relatives, and HCPs perceive as most valuable (such as self-care tasks), minimizing setup time and cognitive load. Furthermore, the subtheme "apprehension of social norms" indicates aesthetics and visibility matter. Hence, the participants did not want to look "different" or "too robotic." As such, the design of assistive robotic devices needs to be compatible with the user's self-perception, for example, by allowing robotic components to be concealed under clothing.

For clinicians and HCPs, our findings point to a need for clear pathways for identifying suitable candidates for assistive robotics, providing realistic information about benefits and limitations, as well as training and follow-up support. Structured introduction protocols, where assistive robotics are tried out in supervised sessions and then revisited after a trial period at home, may help individuals with SCI, relatives, and HCPs jointly assess whether the technology is useful or not. HCPs themselves also need opportunities and time to familiarize themselves with emerging technologies. Without having the time to familiarize themselves with emerging technologies, the rapid technological change of health care technologies may be perceived as additional workload instead of as support.

For policymakers and funders, ongoing concerns about costs and access underscore the need for developing reimbursement models and procurement strategies that make assistive robotics financially accessible, while also ensuring that funding covers training and maintenance. Models that support periodic reassessment may be needed to match the fluctuating needs of people with SCI. Involving people with SCI, their relatives, and HCPs in priority-setting and evaluation processes may help align investments with what is perceived as useful in everyday life.

As evidenced by numerous studies, the process of adopting new technology is far from linear [16,66,67]. Individuals' perceptions and intentions to either embrace or reject a technology are not static; rather, they are dynamic, continually shaped and negotiated as people interact with the technology [68]. Our results suggest that the ways individuals with SCI engage with assistive technologies are contingent on their lived experiences, evolving needs, and changing technological landscape. Overhyping the capabilities of assistive robotics risks fostering unrealistic expectations that could shape disengagement and disappointment [69]. Rather than focusing on promotional narratives, there is value in exploring how assistive robotics are experienced and imagined within participants' everyday lives. As researchers, we can contribute by facilitating discussions that reflect both possibilities and limitations, offering a grounded perspective on what assistive robotics can achieve and the type of support they can provide [65].

Limitations of the Study

A limitation of this study is that the inclusion and exclusion criteria did not take into account participants' time since injury, prior use of assistive technology, or whether they had completed comprehensive rehabilitation. The sampling strategy was a deliberate choice intended to invite a diverse range of perspectives and lived experiences, rather than predefining the sample based on rehabilitation history or technological exposure. However, we acknowledge that prior rehabilitation experiences and the use of assistive technologies may have shaped participants' attitudes toward adopting new assistive robotics solutions. Thus, the findings reported in the paper pertain to individuals with SCIs, constituting a highly heterogeneous group. Rather than viewing heterogeneity as a limitation alone, heterogeneity among the participants hopefully brought a more layered understanding of varying expectations, needs, and concerns surrounding assistive robotics. Nonetheless, the study has been strengthened by involving individuals from 2 different countries, as well as including relatives and HCPs. Therefore, the findings are likely to be beneficial for developers, practitioners, and researchers working on projects with similar technologies and characteristics for the target group.

Another limitation, but also an interpretive choice, was the use of the UTAUT as a lens when discussing the findings. The conceptual framings were used to help interpret and frame overarching patterns from the findings in the discussion. The data were first analyzed inductively, seeking all themes without being led by the need to fit participants' perceptions into a certain framework. As such, the interplay between emergent interpretation and structured theory did not impose a rigid framework but instead facilitated a dynamic engagement when discussing the findings, allowing both participant-driven insights and established conceptual framings to emerge. The interpretive approach was conducted with attention to maintaining the rigor of qualitative research. The recurring patterns observed in the data suggest not only coherence but also shared meanings among participants. Even so, using TAM and UTAUT may foreground certain aspects of technology acceptance while overlooking others. Future research could draw on alternative frameworks.

An additional limitation of the study was that the participants expressed gratitude for participating in the interviews and workshops, which may result in overly positive attitudes toward assistive robotics. However, as seen in the results, the data revealed multiple barriers to the intention to use, as interpreted in our analysis, which suggests that expressions of gratitude did not necessarily equate to uncritical acceptance of assistive robotics but coexisted with articulated concerns and constraints.

Conclusions

This study shows that intentions to use assistive robotics among individuals with SCI, their relatives, and HCPs are shaped by

XSL•FO
RenderX

a combination of barriers and enablers. The findings revealed that the barriers were clustered around practical constraints (such as navigating new realities, fluctuating health, limited awareness and access, and concerns about costs) and interaction with assistive robotics (including doubts about meaningfulness, uncertainty regarding reliability and safety, uneasiness about competence, and apprehension of social norms). Individuals with SCI described fluctuating pain, fatigue, and transition periods after discharge as making it difficult to invest energy in learning and using assistive robotics, while HCPs highlighted how continuous technological change increased their workload and required repeated training. Difficulties with awareness and access were mainly voiced by individuals with SCI and their relatives, who experienced fragmented information and limited support, whereas healthcare professionals perceived a wide range of available aids that were nevertheless seldom prescribed. Concerns about costs were shared by all 3 stakeholder groups and were seen as a major barrier to the intention to use assistive robotics.

The theme of "interaction with assistive robotics" revealed doubts about meaningfulness, uncertainty about reliability and safety, uneasiness about competence, and apprehension of social norms. Individuals with SCI tended to question their own ability to manage complex technologies such as assistive robotics in everyday life, while relatives and HCPs questioned whether assistive robotics would reduce dependence. Apprehension of social norms was voiced predominantly by individuals with SCI, who feared that visible robotic limbs would reinforce perceptions of them as "different." At the same time, participants' imaginings of greater self-efficacy and fulfillment of functional requirements, often expressed as hopes of regaining pre-injury abilities and reducing dependence on others, functioned as important enablers across groups. We conceptualized the pattern of ambivalence as a meaningfulness paradox, as participants expressed doubts about the meaningfulness of assistive robotics in everyday life, their own competence in handling assistive robotics, and their safety and reliability. Nonetheless, they also spoke about how assistive robotics could enhance self-efficacy if assistive robotics met all functional requirements, a circumstance they often doubted. The results suggest that assistive robotics should be understood in relation to the complexities of everyday life with SCI, caring relationships, and the imaginaries that shape how emerging technologies are perceived.

Acknowledgments

We thank all the participants.

Funding

This work was partially supported by the European Union (EU) project HARIA—Human–Robot Sensorimotor Augmentation—Wearable Sensorimotor Interfaces and Supernumerary Robotic Limbs for Humans with upper-limb disabilities (grant agreement No. 101070292). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the EU or CINEA. Neither the EU nor the granting authority can be held responsible for them.

Data Availability

Data are not publicly available due to ethical restrictions concerning participant privacy, in accordance with Swedish and Spanish regulations and ethics committee approvals but may be accessed by approved researchers after a data use agreement.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guidelines for in-depth workshop with individuals with a spinal cord injury.
[PDF File, 61 KB - rehab_v13i1e72080_app1.pdf]

Multimedia Appendix 2
Interview guide for patients.
[PDF File, 56 KB - rehab_v13i1e72080_app2.pdf]

Multimedia Appendix 3
Interview guide for relatives.
[PDF File, 57 KB - rehab_v13i1e72080_app3.pdf]

Multimedia Appendix 4
Interview guide for health care professionals.
[PDF File, 65 KB - rehab_v13i1e72080_app4.pdf]

XSL•FO
RenderX

Multimedia Appendix 5
Guidelines workshops with health care professionals.
[PDF File, 61 KB - rehab_v13i1e72080_app5.pdf]

References

1. Hauwe LVD, Sundgren PC, Flanders AE. Spinal trauma and spinal cord injury (SCI). In: Diseases of the Brain, Head and Neck, Spine 2020–2023: Diagnostic Imaging: Springer; 2020:231-240. [doi: 10.1007/978-3-030-38490-6_19] [Medline: 32119240]
2. Wagner A, Schweizer C, Ronca E, Gemperli A. The most important assistive devices for persons with spinal cord injury in Switzerland: a cross-sectional study. Disabilities 2023;3(3):367-378. [doi: 10.3390/disabilities3030024]
3. Kemp CC, Edsinger A, Clever HM, Matulevich B. The design of stretch: a compact, lightweight mobile manipulator for indoor human environments. Presented at: 2022 IEEE International Conference on Robotics and Automation (ICRA); May 23-27, 2022; Philadelphia, PA, USA p. 3150-3157. [doi: 10.1109/ICRA46639.2022.9811922]
4. Schultz JR, Slifkin AB, Yu H, Schearer EM. Proof-of-concept: a hands-free interface for robot-assisted self-feeding. Presented at: 2022 International Conference on Rehabilitation Robotics (ICORR); Jul 25-29, 2022; Rotterdam, Netherlands p. 1-6. [doi: 10.1109/ICORR55369.2022.9896535]
5. Cruz DM, Emmel MLG, Manzini MG, Braga Mendes PV. Assistive technology accessibility and abandonment: challenges for occupational therapists. Open J Occup Ther 2016;4(1):10. [doi: 10.15453/2168-6408.1166]
6. Lusardi R, Tomelleri S, Wherton J. Living with assistive robotics: exploring the everyday use of exoskeleton for persons with spinal cord injury. Front Med Technol 2021;3:747632. [doi: 10.3389/fmedt.2021.747632] [Medline: 35047959]
7. Sugawara AT, Ramos VD, Alfieri FM, Battistella LR. Abandonment of assistive products: assessing abandonment levels and factors that impact on it. Disabil Rehabil Assist Technol 2018 Oct;13(7):716-723. [doi: 10.1080/17483107.2018.1425748] [Medline: 29334475]
8. Jafar MR, Nagesh DS. Literature review on assistive devices available for quadriplegic people: Indian context. Disabil Rehabil Assist Technol 2023 Aug;18(6):929-941. [doi: 10.1080/17483107.2021.1938708] [Medline: 34176416]
9. Ventura S, Ottoboni G, Pappadà A, Tessari A. Acceptance of assistive technology by users with motor disabilities due to spinal cord or acquired brain injuries: a systematic review. J Clin Med 2023 Apr 19;12(8):2962. [doi: 10.3390/jcm12082962] [Medline: 37109297]
10. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
11. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
12. Pekkarinen S, Hennala L, Tuisku O, et al. Embedding care robots into society and practice: socio-technical considerations. Futures 2020 Sep;122:102593. [doi: 10.1016/j.futures.2020.102593]
13. Dos Santos ADP, Ferrari ALM, Medola FO, Sandnes FE. Aesthetics and the perceived stigma of assistive technology for visual impairment. Disabil Rehabil Assist Technol 2022 Feb;17(2):152-158. [doi: 10.1080/17483107.2020.1768308] [Medline: 32501732]
14. Barbareschi G, Carew MT, Johnson EA, Kopi N, Holloway C. “When They See a Wheelchair, They’ve Not Even Seen Me”-factors shaping the experience of disability stigma and discrimination in Kenya. Int J Environ Res Public Health 2021 Apr 17;18(8):4272. [doi: 10.3390/ijerph18084272] [Medline: 33920601]
15. Jasanoff S, Kim SH. Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power: University of Chicago Press; 2015. URL: https://doi.org/10.7208/chicago/9780226276663.001.0001 [accessed 2026-01-18]
16. Latour B. Reassembling the Social: An Introduction to Actor-Network-Theory: Oxford University Press; 2007. URL: https://doi.org/10.1093/oso/9780199256044.001.0001 [accessed 2026-01-18]
17. Bijker WE, Hughes TP, Pinch T. The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology: MIT press; 1987. URL: https://monoskop.org/images/1/1f/Bijker_Hughes_Pinch_eds_The_Social_Construction_of_Technological_Systems._New_Directions_in_the_Sociology_and_History_of_Technology_no_OCR.pdf [accessed 2026-01-13]
18. Akrich M. The de-scription of technical objects. In: Bijker WE, Law J, editors. Shaping Technology/Building Society: Studies in Sociotechnical Change: MIT press; 1992:205-224. [doi: 10.4000/tc.863]
19. Sand M, Schneider C. Visioneering socio-technical innovations — a missing piece of the puzzle. Nanoethics 2017 Apr;11(1):19-29. [doi: 10.1007/s11569-017-0293-6]
20. Mekki M, Delgado AD, Fry A, Putrino D, Huang V. Robotic rehabilitation and spinal cord injury: a narrative review. Neurotherapeutics 2018 Jul;15(3):604-617. [doi: 10.1007/s13311-018-0642-3] [Medline: 29987763]
21. Holanda LJ, Silva PMM, Amorim TC, Lacerda MO, Simão CR, Morya E. Robotic assisted gait as a tool for rehabilitation of individuals with spinal cord injury: a systematic review. J Neuroeng Rehabil 2017 Dec 4;14(1):126. [doi: 10.1186/s12984-017-0338-7] [Medline: 29202845]

XSL•FO
RenderX

22. Langer A, Feingold-Polak R, Mueller O, Kellmeyer P, Levy-Tzedek S. Trust in socially assistive robots: considerations for use in rehabilitation. Neurosci Biobehav Rev 2019 Sep;104:231-239. [doi: 10.1016/j.neubiorev.2019.07.014] [Medline: 31348963]
23. Trenaman LM, Miller WC, Escorpizo R, SCIRE Research Team. Interventions for improving employment outcomes among individuals with spinal cord injury: a systematic review. Spinal Cord 2014 Nov;52(11):788-794. [doi: 10.1038/sc.2014.149] [Medline: 25179659]
24. Lidal IB, Huynh TK, Biering-Sørensen F. Return to work following spinal cord injury: a review. Disabil Rehabil 2007 Sep 15;29(17):1341-1375. [doi: 10.1080/09638280701320839] [Medline: 17729082]
25. Baldassin V, Shimizu HE, Fachin-Martins E. Computer assistive technology and associations with quality of life for individuals with spinal cord injury: a systematic review. Qual Life Res 2018 Mar;27(3):597-607. [doi: 10.1007/s11136-018-1804-9] [Medline: 29417427]
26. Collinger JL, Boninger ML, Bruns TM, Curley K, Wang W, Weber DJ. Functional priorities, assistive technology, and brain-computer interfaces after spinal cord injury. J Rehabil Res Dev 2013;50(2):145-160. [doi: 10.1682/jrrd.2011.11.0213] [Medline: 23760996]
27. Scivoletto G, Galli G, Torre M, Molinari M, Pazzaglia M. The overlooked outcome measure for spinal cord injury: use of assistive devices. Front Neurol 2019;10:272. [doi: 10.3389/fneur.2019.00272] [Medline: 30967836]
28. Kabir KS, Alsaleem A, Wiese J. The Impact of Spinal Cord Injury on Participation in Human-Centered Research. 2021 Jun 28 Presented at: DIS ’21; Jun 28 to Jul 2, 2021; Virtual Event USA p. 1902-1914. [doi: 10.1145/3461778.3462122]
29. Khalid S, Alnajjar F, Gochoo M, Renawi A, Shimoda S. Robotic assistive and rehabilitation devices leading to motor recovery in upper limb: a systematic review. Disabil Rehabil Assist Technol 2023 Jul;18(5):658-672. [doi: 10.1080/17483107.2021.1906960] [Medline: 33861684]
30. Proietti T, Ambrosini E, Pedrocchi A, Micera S. Wearable robotics for impaired upper-limb assistance and rehabilitation: state of the art and future perspectives. IEEE Access 2022;10:106117-106134. [doi: 10.1109/ACCESS.2022.3210514]
31. Guatibonza A, Solaque L, Velasco A, Peñuela L. Assistive robotics for upper limb physical rehabilitation: a systematic review and future prospects. Chin J Mech Eng 2024;37(1):69. [doi: 10.1186/s10033-024-01056-y]
32. Ho JSW, Ko KSY, Law SW, Man GCW. The effectiveness of robotic-assisted upper limb rehabilitation to improve upper limb function in patients with cervical spinal cord injuries: a systematic literature review. Front Neurol 2023;14:1126755. [doi: 10.3389/fneur.2023.1126755] [Medline: 37621855]
33. Martinez-Hernandez U, Metcalfe B, Assaf T, Jabban L, Male J, Zhang D. Wearable assistive robotics: a perspective on current challenges and future trends. Sensors (Basel) 2021 Oct 12;21(20):6751. [doi: 10.3390/s21206751] [Medline: 34695964]
34. Pinelli E, Zinno R, Barone G, Bragonzoni L. Barriers and facilitators to exoskeleton use in persons with spinal cord injury: a systematic review. Disabil Rehabil Assist Technol 2024 Aug;19(6):2355-2363. [doi: 10.1080/17483107.2023.2287153] [Medline: 38009458]
35. HARIA. HARIA Human-Robot Sensorimotor Augmentation. URL: https://clem.diism.unisi.it/~haria [accessed 2026-01-13]
36. Yang B, Huang J, Chen X, Xiong C, Hasegawa Y. Supernumerary robotic limbs: a review and future outlook. IEEE Trans Med Robot Bionics 2021;3(3):623-639. [doi: 10.1109/TMRB.2021.3086016]
37. Prattichizzo D, Pozzi M, Lisini Baldi T, et al. Human augmentation by wearable supernumerary robotic limbs: review and perspectives. Prog Biomed Eng (Bristol) 2021 Sep 17;3(4):042005. [doi: 10.1088/2516-1091/ac2294] [Medline: 41367033]
38. Pozzi M, Malvezzi M, Prattichizzo D, Salvietti G. Actuated palms for soft robotic hands: review and perspectives. IEEE/ASME Trans Mechatron 2023;29(2):902-912. [doi: 10.1109/TMECH.2023.3328944]
39. Bo V, Turco E, Pozzi M, Malvezzi M, Prattichizzo D. A data-driven topology optimization framework for designing robotic grippers. Presented at: 2023 IEEE International Conference on Soft Robotics (RoboSoft); Apr 3-7, 2023; Singapore p. 1-6. [doi: 10.1109/RoboSoft55895.2023.10122000]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Clarke V, Braun V. Thematic analysis. J Posit Psychol 2017 May 4;12(3):297-298. [doi: 10.1080/17439760.2016.1262613]
42. Adams WC. Conducting semi‐structured interviews. In: Handbook of Practical Program Evaluation: Wiley; 2015:492-505. [doi: 10.1002/9781119171386]
43. Ørngreen R, Levinsen KT. Workshops as a research methodology. Electron J E-Learn 2017;15(1):70-81 [FREE Full text]
44. Kallio H, Pietilä AM, Johnson M, Kangasniemi M. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. J Adv Nurs 2016 Dec;72(12):2954-2965. [doi: 10.1111/jan.13031] [Medline: 27221824]
45. Kvale S, Brinkmann S. Interviews: Learning the Craft of Qualitative Research Interviewing: sage; 2009. URL: https://libris.kb.se/bib/10957812 [accessed 2026-01-18]
46. Hussain I, Spagnoletti G, Salvietti G, Prattichizzo D. An EMG interface for the control of motion and compliance of a supernumerary robotic finger. Front Neurorobot 2016;10:18. [doi: 10.3389/fnbot.2016.00018] [Medline: 27891088]
47. Candela AG. Exploring the function of member checking. TQR 2019;24(3):619-628. [doi: 10.46743/2160-3715/2019.3726]

XSL•FO
RenderX

48. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member checking: a tool to enhance trustworthiness or merely a nod to validation? Qual Health Res 2016 Nov;26(13):1802-1811. [doi: 10.1177/1049732316654870] [Medline: 27340178]
49. Motulsky SL. Is member checking the gold standard of quality in qualitative research? Qualitative Psychology 2021;8(3):389-406. [doi: 10.1037/qup0000215]
50. Ludvigsson JF, Bergman D, Lundgren CI, et al. The healthcare system in Sweden. Eur J Epidemiol 2025 May;40(5):563-579. [doi: 10.1007/s10654-025-01226-9] [Medline: 40383868]
51. Bernal-Delgado E, Angulo-Pueyo E, Ridao-López M, et al. Spain: health system review. Health Syst Transit 2024 Sep;26(3):1-187. [Medline: 40094303]
52. Jardón A, Gil Á, de la Peña AI, Monje CA, Balaguer C. Usability assessment of ASIBOT: a portable robot to aid patients with spinal cord injury. Disabil Rehabil Assist Technol 2011;6(4):320-330. [doi: 10.3109/17483107.2010.528144] [Medline: 20969432]
53. Laubacher M, Perret C, Hunt KJ. Work-rate-guided exercise testing in patients with incomplete spinal cord injury using a robotics-assisted tilt-table. Disabil Rehabil Assist Technol 2015;10(5):433-438. [doi: 10.3109/17483107.2014.908246] [Medline: 24712412]
54. Almenara M, Cempini M, Gómez C, et al. Usability test of a hand exoskeleton for activities of daily living: an example of user-centered design. Disabil Rehabil Assist Technol 2017 Jan;12(1):84-96. [doi: 10.3109/17483107.2015.1079653] [Medline: 26376019]
55. Riemer-Reiss ML. Applying Rogers' Diffusion of Innovations Theory to assistive technology discontinuance. J Appl Rehabil Couns 1999 Dec 1;30(4):16-21. [doi: 10.1891/0047-2220.30.4.16]
56. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: a systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
57. Mayor A. Gods and Robots: Myths, Machines, and Ancient Dreams of Technology: Princeton University Press; 2019. URL: https://journeytothewestresearch.com/wp-content/uploads/2022/09/Gods-and-Robots-Myths-Machines-and-Ancient-Dreams-of-Technology-2018.pdf [accessed 2026-01-13]
58. Corn JJ. Imagining Tomorrow: History, Technology, and the American Future 1986. URL: https://libris.kb.se/bib/4738000 [accessed 2026-01-18]
59. Natale S. Introduction: new media and the imagination of the future. : WI; 2014 URL: https://iris.unito.it/bitstream/2318/1768946/1/PUBLISHED_Natale_introduction-new-media-and-the-imagination-of-the-future.pdf [accessed 2026-01-18]
60. Mosco V. The Digital Sublime: Myth, Power and Cyberspace (Cambridge, MA and London, UK): MIT Press; 2004. URL: https://doi.org/10.7551/mitpress/2433.001.0001 [accessed 2026-01-18]
61. Dourish P, Bell G. Divining a Digital Future: Mess and Mythology in Ubiquitous Computing: Mit Press; 2011. URL: https://www.mondotheque.be/wiki/images/9/92/Paul_Dourish_Genevieve_Bell_Divining_a_Digital.pdf [accessed 2026-01-13]
62. Breuer S, Braun M, Tigard D, Buyx A, Müller R. How engineers' imaginaries of healthcare shape design and user engagement: a case study of a robotics initiative for geriatric healthcare AI applications. ACM Trans Comput-Hum Interact 2023 Apr 30;30(2):1-33. [doi: 10.1145/3577010]
63. Vallès-Peris N, Domènech M. Roboticists' imaginaries of robots for care: the radical imaginary as a tool for an ethical discussion. Engineering Studies 2020 Sep 1;12(3):157-176. [doi: 10.1080/19378629.2020.1821695]
64. Krings BJ, Weinberger N. Assistant without master? Some conceptual implications of assistive robotics in health care. Technologies (Basel) 2018;6(1):13. [doi: 10.3390/technologies6010013]
65. Rosén J. What Did You Expect?: A Human-Centered Approach to Investigating and Reducing the Social Robot Expectation Gap: University of Skövde; 2024. [doi: 10.13140/RG.2.2.36413.44001]
66. Shove E, Pantzar M, Watson M. The Dynamics of Social Practice: Everyday Life and How It Changes: SAGE Publications Ltd; 2012. [doi: 10.4135/9781446250655]
67. Hirsch E, Silverstone R. Consuming Technologies: Media and Information in Domestic: Routledge (Taylor & Francis Group); 2003. [doi: 10.4324/9780203401491]
68. Borup M, Brown N, Konrad K, Van Lente H. The sociology of expectations in science and technology. Technol Anal Strateg Manag 2006 Jul;18(3-4):285-298. [doi: 10.1080/09537320600777002]
69. Geels FW, Schot J. Typology of sociotechnical transition pathways. Res Policy 2007 Apr;36(3):399-417. [doi: 10.1016/j.respol.2007.01.003]

Abbreviations

SCI: spinal cord injury
SRL: supernumerary robotic limb
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

XSL•FO
RenderX

Edited by A Scano; submitted 03.Feb.2025; peer-reviewed by J Wangdell, J Chu; revised version received 11.Dec.2025; accepted 02.Jan.2026; published 17.Feb.2026.

Please cite as:
Frennert S, Persson J, Díez-Rodríguez E, Alcobendas-Maestro M, Villamayor Vega F, Oliviero A
Exploring Barriers and Enablers for the Intention to Use Assistive Robotics Among People With Spinal Cord Injury and Those Involved in Their Care: Qualitative Study
JMIR Rehabil Assist Technol 2026;13:e72080
URL: https://rehab.jmir.org/2026/1/e72080
doi:10.2196/72080

© Susanne Frennert, Johanna Persson, Eva Díez-Rodríguez, Monica Alcobendas-Maestro, Fátima Villamayor Vega, Antonio Oliviero. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Blockchain-Based Mobile App for Digital Identification of Older Adults in Rural Peru: Design and Usability Evaluation Study

Wilver Arana-Ramos*, BSc; Aldo Francisco Pastrana-Leon*, BSc; Juan Carlos Morales-Arevalo*, MBA, MSc

Department of Software Engineering, Faculty of Engineering, Peruvian University of Applied Sciences, Av. Alonso de Molina Nro. 1611, Lima Polo And Hunt Club (Av. Primavera 2390), Lima, Peru

*all authors contributed equally

Corresponding Author:
Juan Carlos Morales-Arevalo, MBA, MSc
Department of Software Engineering, Faculty of Engineering, Peruvian University of Applied Sciences, Av. Alonso de Molina Nro. 1611, Lima Polo And Hunt Club (Av. Primavera 2390), Lima, Peru

Abstract

Background: Older adults in rural areas of Peru encounter many challenges in accessing critical public services, such as health care, education, and social assistance, due to low levels of digital literacy, limited access to technology, and the lack of formalized, secure ID. This inhibits entry into digital health, education, and social assistance systems and increases their risk of vulnerability and social exclusion.

Objective: This study aimed to design a blockchain technology–based mobile app architecture that helps facilitate secure and inclusive digital ID for older adults in rural areas of Peru, enabling access to vital services through a decentralized, privacy-preserving solution.

Methods: This study followed the design thinking process, which consists of five phases: empathize, define, ideate, prototype, and evaluate. A total of 16 adults (aged 61-85 years) were interviewed to determine the usability barriers and security and privacy concerns with mobile technology, which was used to define functional and nonfunctional requirements. These requirements were developed based on the interviews. The primary features the target population valued included blockchain authentication, assisted registration, multilingual functionality, and a user-friendly interface. The features were prioritized and prototyped using the Figma web-based app. The architecture of the app was developed using the C4 model and accounted for sequential development while ensuring scalability, modularity, and decentralization. Usability was assessed quantitatively by administering the System Usability Scale to the same 16 participants after they had interacted with the prototype.

Results: The mean System Usability Scale score was 60.78 (SD 13.68), indicating acceptable usability. The main issues identified were a lack of skills to navigate digital interfaces, concerns regarding data security, and accessibility challenges for people with disabilities. Participants provided high ratings for the assisted registration system and notifications. The modular, blockchain-based system architecture showed substantial potential for scalability and broader inclusion. The prioritization matrix identified that, for adoption, features must incorporate good design, be multilingual, and require secure authentication.

Conclusions: The proposed blockchain-based mobile app offers a viable technical and socially inclusive model for secure digital ID of older adults in underserved contexts. Usability testing suggested that the solution was perceived as secure, usable, and appropriate for the target population. Although not fully deployed, our prototypes and system architecture provide a good starting point for future implementation. The findings in this study can contribute to efforts to facilitate digital inclusion, access to services, and respect for people’s autonomy in identity management systems for vulnerable people.

(JMIR Rehabil Assist Technol 2026;13:e79553) doi:10.2196/79553

KEYWORDS

blockchain technology; data privacy; digital divide; essential public service; portable digital identity; older people; user-centric design

Introduction

Digital technologies have become increasingly important in identity management across all sectors of the economy. The proliferation of mobile apps designed to improve access to services has further deepened this relevance [1]. Older adults living in rural areas of Peru face unique challenges, such as higher levels of digital illiteracy and limited public service coverage. These challenges hinder access to essential services such as health care, education, and social support and mirror broader barriers reported for digital health innovations in low- and middle-income health care systems in South and Southeast Asia [2-4]. Advances in digital identity models, such as self-sovereign identities (SSIs), have opened up new possibilities

XSL•FO
RenderX

for improving security, privacy, and access to services [5]. The emergence of identity verification systems that leverage blockchain infrastructure and protocols exemplifies this change, as they ensure there are no gaps in identity verification and authentication processes and allow users to exercise holistic control over their personal information [6].

Nonetheless, traditional identity management systems that typically use centralized data formats and 2-factor authentication often fail older adults in rural settings [7]. Centralized databases are limited, as they rely on dependable connectivity, technical skills, and trust in institutions to function as intended, even though there isn't adequate infrastructure in many rural contexts. In addition, centralized database models present vulnerabilities by creating a single point of failure that makes any district-based database susceptible to unauthorized access or misuse of information. Furthermore, the use of passwords, tokens, and SMS codes presents usability challenges, as older adults may be less versed in the use of mobile devices and may have difficulty remembering and entering credentials [8]. Conversely, blockchain protocols decentralize validation across identifiable participants and enable user control over personal data by creating a single entry point to the data. Collectively, these properties make blockchain protocols safer, more transparent, and more empowering for this population.

These advances in technology go well beyond data protection. They are fundamental to improving efficiency in the delivery of public and private services. For example, health credential verification and the interoperability of electronic health records are relevant in the health care context. The same is true in education and public participation [9], where the use of mobile apps that integrate SSIs could help reduce the digital divide and provide access to vital resources for all [2,3]. However, their implementation remains limited due to the fragmentation of existing platforms and the lack of a widely accepted interoperable standard. In addition, challenges remain related to centralized data sets, transparency in verification, adherence to strong data privacy policies, and the centralized control of information [10].

In this context, the focus of this study is on designing a blockchain-based mobile app architecture to support digital identity verification and access to necessary services. The goal is to create a system that can be implemented quickly and efficiently while respecting appropriate privacy, security, and decentralization practices in relation to structural data [11].

This study follows a user-centered design thinking (DT) approach to ensure that the proposed solutions are accessible, secure, and appropriate for the target population, rural older adults [12]. The research will include a comprehensive review of current solutions, such as blockchain and smart contracts applied to identity systems, in order to identify the most effective methods for improving user interaction and building trust in digital service platforms [13]. This research is qualitative and will combine a detailed literature review with the design and creation of system architecture models that will eventually be applied in the real world. The goal will be to establish a robust framework for mobile services that use advanced identity verification technologies, highlighting the potential of blockchain to ensure data permanence and integrity. The results of this study provide useful approaches for building mobile systems that prioritize accessibility, security, and transparency, with a special focus on improving essential services [13]. This strategy is even more significant for rural communities, as it allows older adults in these communities to access digital ID solutions that respect their privacy and give them control of their personal information.

Therefore, the objective of this study was to design and evaluate the usability of a blockchain-based mobile app architecture that supports secure digital ID and access to essential services for older adults living in rural areas of Peru.

Methods

Ethical Considerations

This research involved 16 older adults (aged 61-85 years) living in rural areas of Peru who voluntarily took part in interviews and usability tests. All participants signed an informed consent form. No sensitive data were collected, and personal information was anonymized and managed in accordance with the principles of confidentiality and data protection.

Formal ethics board (IRB) approval was not required according to institutional guidelines [14], as the study focused on the design and usability evaluation of a digital prototype. The research was non-interventional in nature, involved voluntary adult participants, and did not include clinical procedures or the collection of sensitive personal or health data. All participants provided informed consent prior to participation. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Research Methodology

This research adopted the DT methodology to develop a mobile app. The user-centered design approach is highly effective in digital product development because it supports users' needs at every stage while also ensuring a product is functional and accessible to the target audience [15-18]. Among the several established design frameworks, such as user-centered design, participatory design, and agile user experience, DT was selected because it provides a structured yet flexible process that begins with empathy as its core principle. This characteristic aligns with the study's focus on understanding the lived experiences of older adults in rural Peru, whose limited digital literacy and accessibility challenges require solutions grounded in human context rather than purely technical efficiency. By explicitly situating this work within a lineage of research that views empathy as the foundation for inquiry and intervention, the DT framework allows the research team to translate qualitative insights into practical, socially responsive design decisions [15-17].

The design process was broken into five phases: empathize, define, ideate, prototype, and test, and each phase was important for our design of a blockchain-based mobile product for digital ID and safe access to services. A complete overview of these activities is provided below.

XSL•FO
RenderX

Empathize Phase

A total of 16 older adults aged 61 to 85 years were interviewed to identify some of the key barriers when trying to access digital services or engage with new technologies. The interviews revealed a number of key issues directly related to participants' comfort levels and hesitancy with digital technology, and their anxiety surrounding cybersecurity, which often impaired their use of multiple digital platforms. These findings illustrate that solutions must be technically feasible, but also usable, secure, and easy to navigate [15].

Define Phase

The research team identified important findings from the interviews that influenced the project [19]. Based on the interview findings, the research team identified key user priorities, specifically the importance of a user-friendly interface, transparent privacy control, and user management (input). The study also highlighted the need for ID systems that are achievable and meaningful to users (at all digital literacy levels) [17].

Ideate Phase

At this stage of the research and development process, the team specified some core features of the app, such as authentication methods incorporating blockchain technologies, a security alerts feature, and personalization options for service access that the user sets according to their needs and preferences. The goal was to create a well-integrated solution in which users could access and manage strategic services relevant to them in a secure and efficient manner [18].

Prototype Phase

Prototypes and low- and medium-fidelity mockups were developed on the Figma platform. The prototypes helped to visualize the interface design of the app and observe users' interactions with the app's functions [19]. The technical system architecture was developed in accordance with the C4 model (this study included only "context," "containers," and "components," excluding "code," as we did not model the code level) to allow the system to be scalable and adaptable within a blockchain structure [20-22].

Testing Phase

To quantitatively evaluate prototype app usability, the study applied the System Usability Scale (SUS) developed by Brooke [23] in 1996. The SUS is a well-respected instrument used in usability research, due in part to its simplicity and its provision of a straightforward numerical metric of users' perceptions of system usability [23].

A survey was created using the standard 10-item SUS, in which respondents indicated whether they agreed or disagreed with statements using a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree"). The full set of questions used is in Textbox 1.

The launch survey was conducted on 16 older adults that were representative of the targeted users, all of whom had previous interactions with the prototype. The sample size matched the guidelines recommended for initial usability testing using the SUS. Scoring was based on subtracting 1 from the participants' response for the odd-numbered items and subtracting the even-numbered item responses from 5. The scores were summed and multiplied by 2.5 to create a final score with a range of 0 to 100, with higher scores representing better usability. User feedback was an important part of the process for improving interface accessibility and ease of use [24]. This cyclical, user-centered approach ensures that the app meets technical specifications while being accessible, safe, and easy to use for older adults with various levels of technological capabilities [16,17].

Textbox 1. Questions on the System Usability Scale.

1. I would like to use this application frequently.
2. I found the application unnecessarily complex.
3. I found the application easy to use.
4. I would need the help of a person with technical knowledge to be able to use this application.
5. The various functions of the application are well integrated.
6. I found too much inconsistency in the application.
7. I imagine most people could learn to use this application quickly.
8. I found the application cumbersome to use.
9. I felt very confident using this application.
10. I learned to use the application quickly.

Results

This section elaborates on the results from the various design phases and the process of developing a mobile app designed for digital identity and secure access to vital services based on blockchain technology, which was achieved using the DT approach. This allowed the team to better understand the needs of the targeted end users, older adults, and ensured that the prototype genuinely reflects the expectations and needs of this group. The major findings and responses presented in each phase are described below.

Empathize Phase: Interview Findings

During the empathize phase, interviews were conducted with 16 older adults, ranging in age from 61 to 85 years, to identify

XSL•FO
RenderX

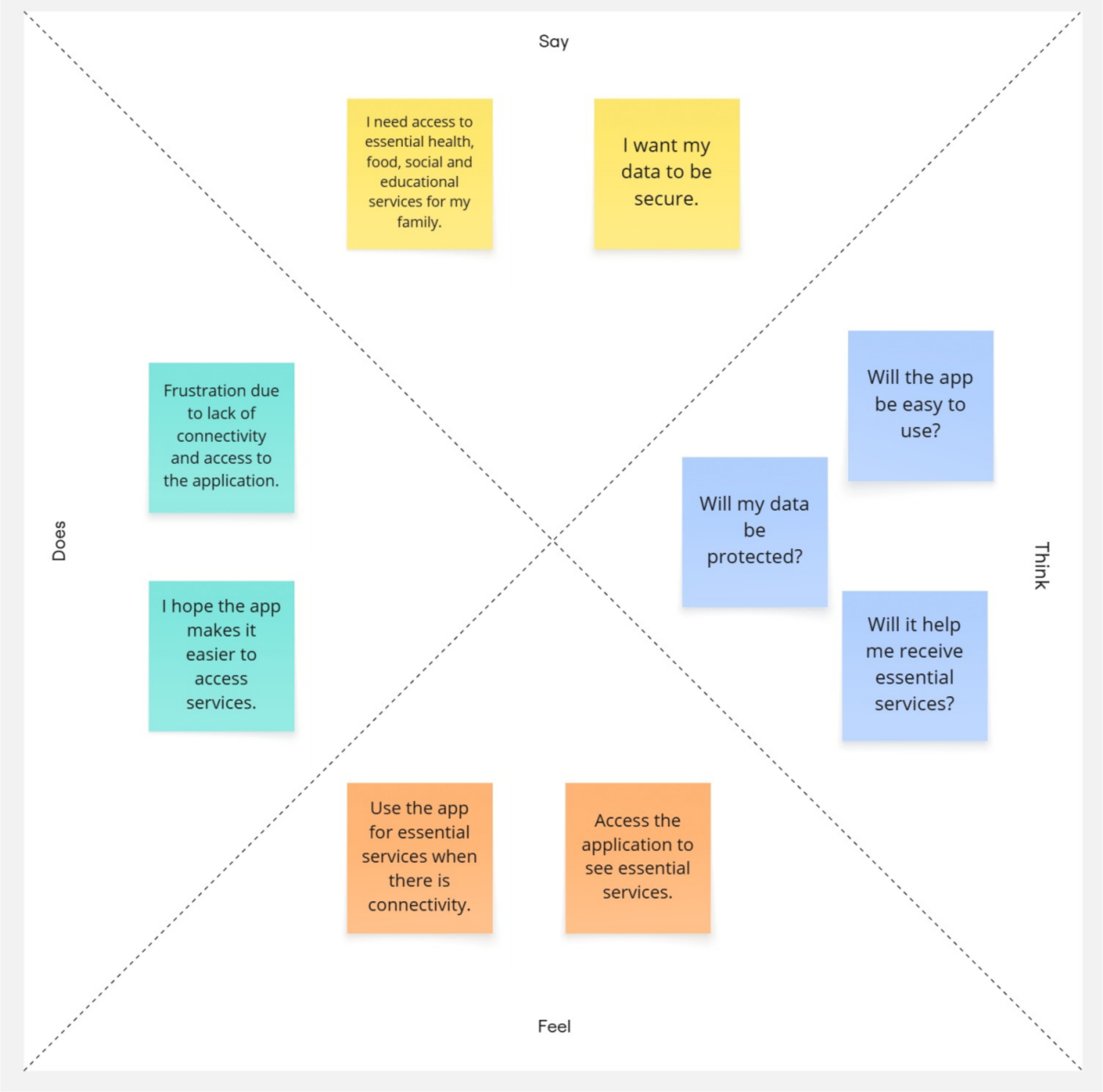
the barriers and difficulties they face when interacting with mobile apps. The most common problems encountered are presented in Textbox 2.

Based on the findings obtained, two empathy maps were developed to reflect the needs and emotions of users in relation to technology [25,26]. These are presented in Figures 1 and 2, where the first corresponds to users and the second to the local authorities.

Textbox 2. Identified problems.

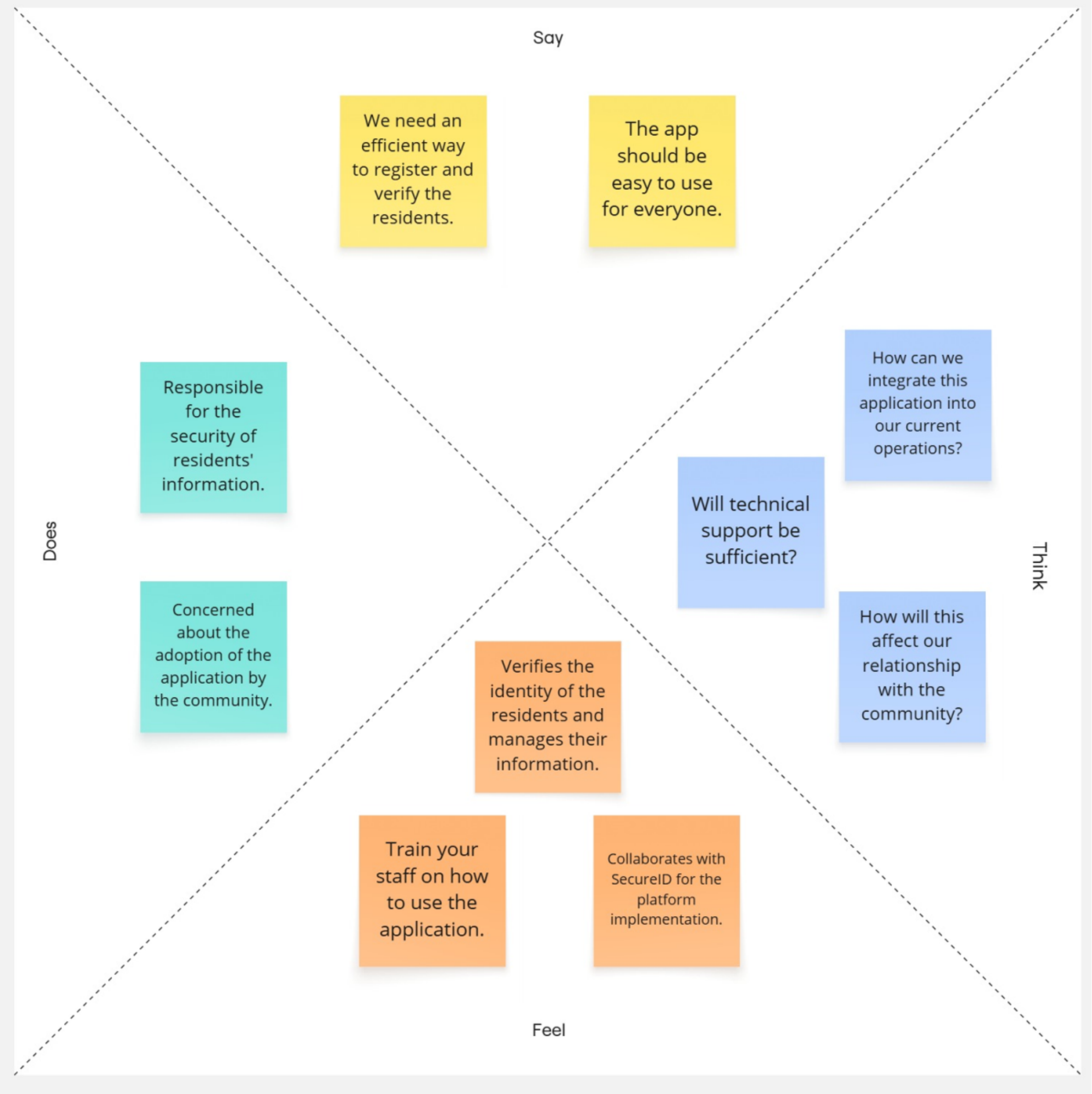
- Difficulty navigating apps: users indicated that the app interfaces were complex and difficult to understand.
- Data security concerns: older adults expressed distrust about the handling of their personal data on digital platforms.
- Lack of interface customization: some users reported that the apps were not tailored to their individual needs, making adoption difficult.
- Lack of digital literacy: many older adults have difficulty understanding apps due to their unfamiliarity with digital technologies.
- Limited accessibility: the lack of accessibility options for people with visual or motor disabilities made interaction complicated for a segment of the population.

Figure 1. Empathy map: user perspective.



XSL•FO
RenderX

Figure 2. Empathy map: local authorities' perspective.



Define Phase: Functional and Nonfunctional Requirements

Based on the findings obtained in the empathize phase, the functional and nonfunctional requirements for the mobile app were defined. These were established based on the specific needs of the users. The most relevant requirements are presented in Table 1 below.

The following are the most relevant user statements, selected from those collected for their significant contribution to understanding users' expectations regarding interaction with the app and ensuring that their needs are adequately addressed. These user stories are presented in Table 2.

XSL•FO
RenderX

Table . Functional and nonfunctional requirements.

Requirement	Description	Type
Registration in the app	Allows the user to register in the mobile app with the support of a caregiver, enabling access to their digital identification and essential services.	Functional
Digital identification display	Allows the user to securely access their digital identification based on blockchain technology, guaranteeing the authenticity and protection of their personal data.	Functional
Selection of essential health service	Provides the user with the ability to view and select health services available in their locality, facilitating timely access to vital information.	Functional
Information security	The app must ensure the integrity, confidentiality, and authenticity of user data through the use of blockchain technology and robust security protocols.	Nonfunctional
Multilingual support	The app must support multiple languages to facilitate access to services for users who do not speak the platform's main language.	Nonfunctional
Notification system	The system must have a notification controller that manages reminders, alerts, and updates, sending them to users and local authorities.	Functional

Table . Relevant user stories.

User ID	User statement
US-01	As an older adult user, I would like to view my digital identification in order to access my personal data securely.
US-02	As a local authority, I want to filter the list of registered seniors by name or digital identification to quickly locate a specific user and facilitate records management.
US-03	As an older adult user, I would like to register in the application in order to access the available services.
US-04	As an older adult user, I wish to attach and upload images such as a signature, face, or identification to be registered in my account.
US-05	As a local authority, I wish to modify the distribution of essential services.
US-06	As an older adult user, I want to visualize the services provided by the application to easily access the available options.
US-07	As a local authority, I want to view the list of registered seniors to properly track their enrollment and access to available services.
US-08	As a local authority, I want to visualize the services provided by the application to easily access the available options and also have the possibility to add new broadcasts to each selected service.

Ideate Phase: Key Functionalities and Prioritization Matrix

In this phase, the key functionalities needed for the mobile app were identified. The functionalities were prioritized according to their impact on user experience and the effort required for their implementation. These results are presented in Table 3.

Table 4 shows the prioritization matrix, which ranks the key functionalities according to their impact and implementation effort.

Table . Selected key functionalities.

Key functionality	Description
Blockchain authentication	A decentralized authentication system to ensure the security and privacy of user data.
Intuitive interface	Simple and accessible design, with customization options to enhance the experience for older adults.
Assisted registration system	Provides personalized assistance during the registration process to ensure the inclusion of older adults who are not fully technologically autonomous.
Multilingual support	The app must support multiple languages to facilitate access to services for users who do not speak the platform’s main language.
Notification system	The system should send notifications to remind users about important events or updates related to services, such as health care, education, and social assistance.

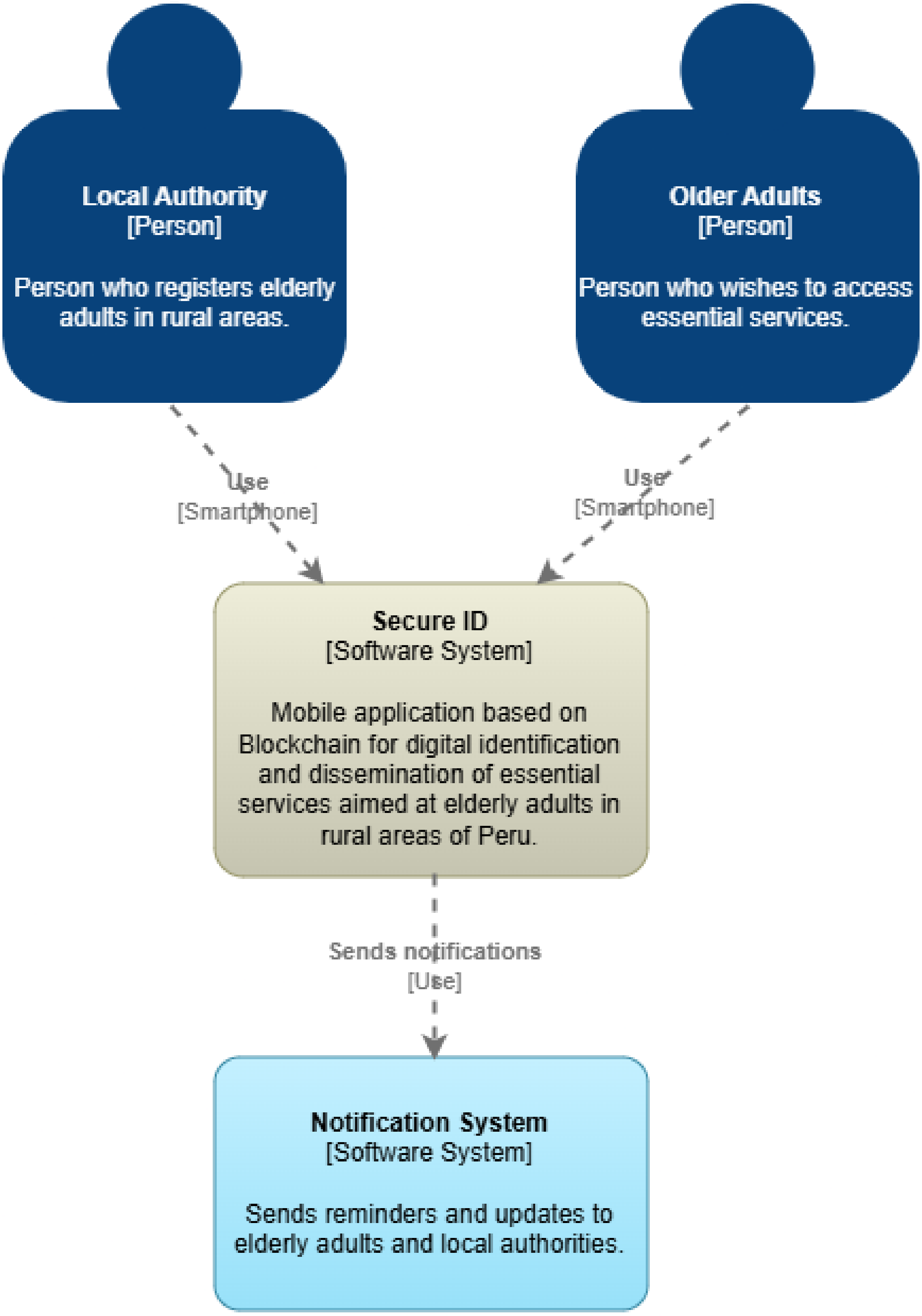
Table . Functionality prioritization matrix.

Functionality	Impact	Effort	Priority
Blockchain authentication	High	Medium	High
Intuitive Interface	High	High	Medium
Assisted registration system	High	Medium	High
Multilingual support	High	Medium	High
Notification system	High	Medium	High

Prototyping Phase: C4 Model, User Flow, and Mockups

During the prototyping phase, a comprehensive representation of the system was developed, covering both the technical architecture and the user experience. For this purpose, the C4 model was applied, which allows for structuring the system architecture at different levels of abstraction. Figure 3 shows the context diagram, which illustrates the general interaction between the mobile app, the users (older adults and local authorities), and the external identity verification services. Figure 4 presents the container diagram, which decomposes the system into its main modules: mobile interface, cloud backend, and decentralized database, along with the technologies used. Multimedia Appendix 1 presents the component diagram, which describes the internal elements of the back end, such as the authentication manager, the service controller, and the notification module.

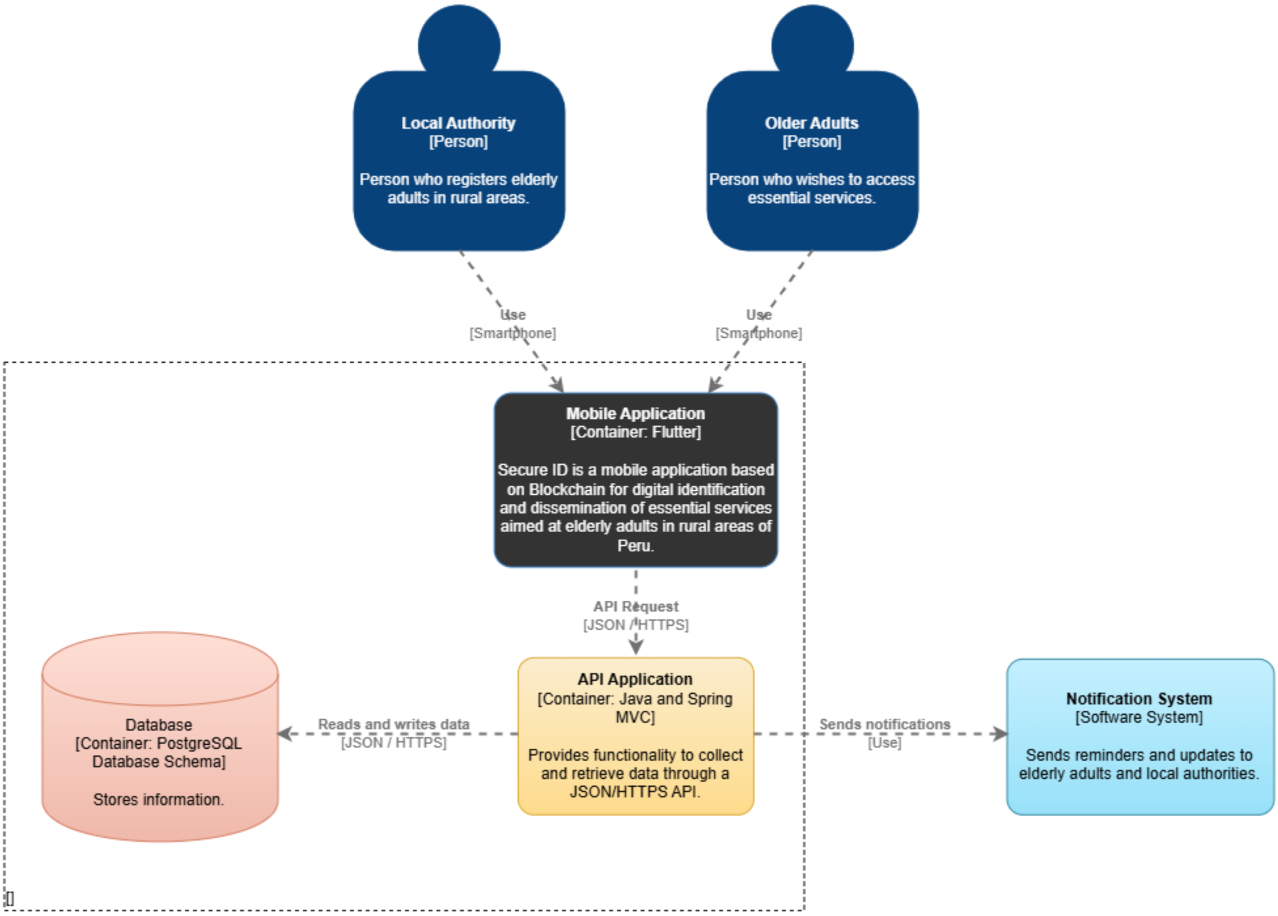
Figure 3. C4 model context diagram.



[]

XSL•FO
RenderX

Figure 4. C4 model container diagram.



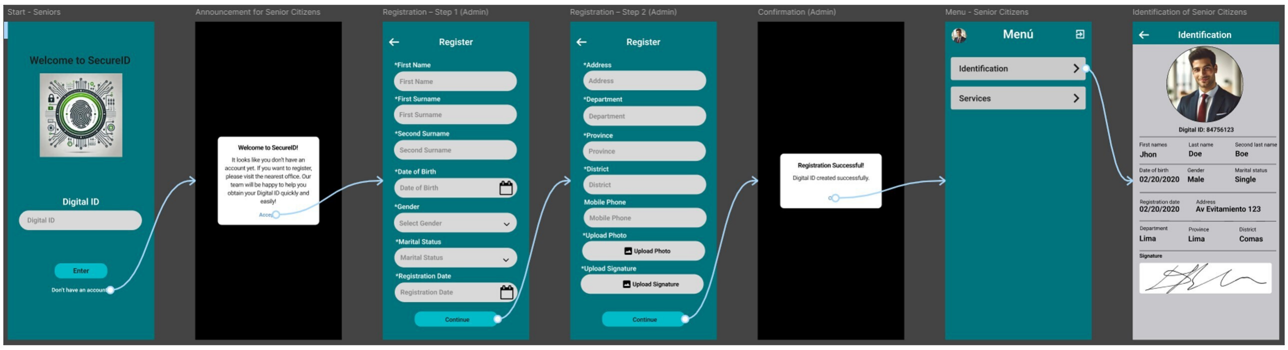
To validate the user experience, navigation flows were developed that represent common user steps from login to service access. Figure 5 shows one of these flows, focusing on digital ID.

Finally, low and medium fidelity mockups were designed in Figma. Figure 6 illustrates the most representative screens of the prototype: user registration, ID display, and service dissemination. These prototypes were essential to validate the suitability of the design to the capabilities and preferences of the target audience.

Figure 5 illustrates the proposed navigation flow for the digital ID process of older adults. This sequential design optimizes the user experience through a clear interface and guided steps from the initial registration to the display of the digital ID, ensuring accessibility, usability, and efficiency in identity management.

Figure 6 presents a series of mockups that illustrate key system functionalities from the administrator's perspective. These include user registration, display of ID information, and management of essential services such as health care, social support, education, and food. The design promotes efficient administration and clear dissemination of campaigns targeting vulnerable populations.

Figure 5. Ideal navigation flow for identification.



Welcome to SecureID!

It looks like you don't have an account yet. If you want to register, please visit the nearest office. Our team will be happy to help you obtain your Digital ID quickly and easily!

Registration Successful!

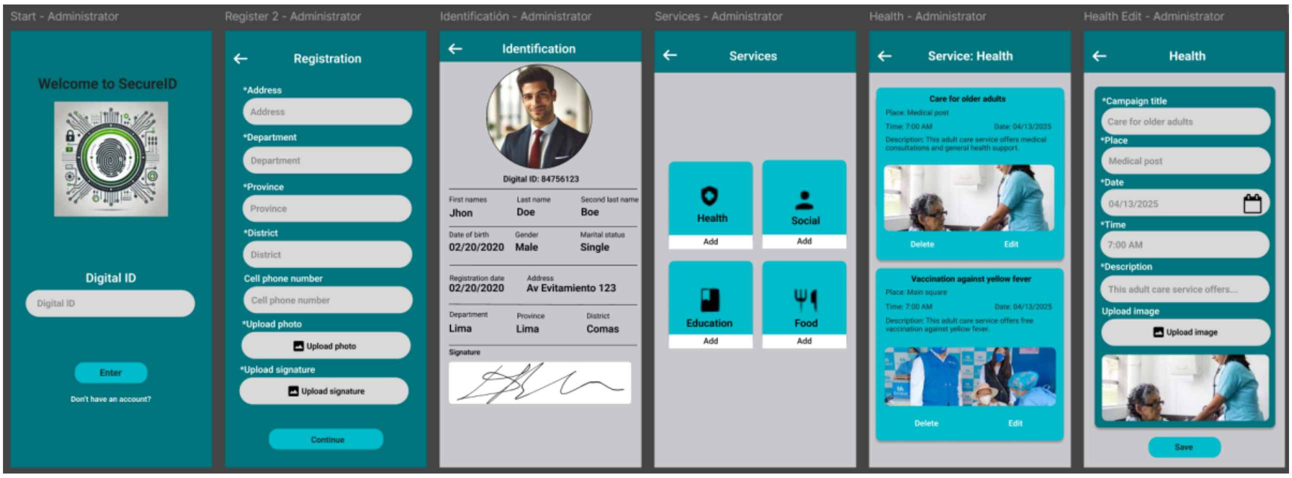
Digital ID created successfully.

Digital ID: 84756123

First names: Jhon
Last name: Doe
Second last name: Boe
Date of birth: 02/20/2020
Gender: Male
Marital status: Single
Registration date: 02/20/2020
Address: Av Evitamiento 123
Department: Lima
Province: Lima
District: Comas
Signature

XSL•FO
RenderX

Figure 6. Mockups: registration, identification, and dissemination of essential services.



Digital ID: 84756123

First names: Jhon
Last name: Doe
Second last name: Boe
Date of birth: 02/20/2020
Gender: Male
Marital status: Single
Registration date: 02/20/2020
Address: Av Evitamiento 123
Department: Lima
Province: Lima
District: Comas
Signature

Care for older adults
Place: Medical post
Time: 7:00 AM
Date: 04/13/2025
Description: This adult care service offers medical consultations and general health support.

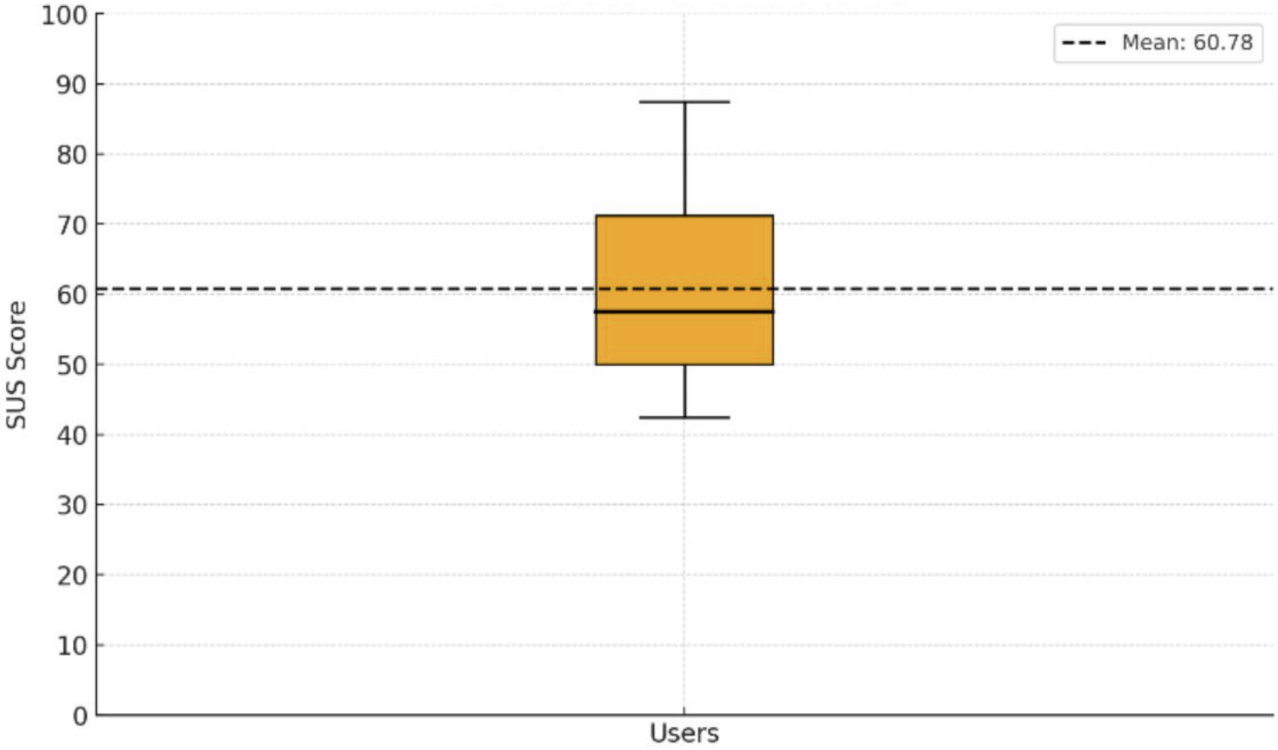
Vaccination against yellow fever
Place: Main square
Time: 7:00 AM
Date: 04/13/2025
Description: This adult care service offers free vaccination against yellow fever.

*Campaign title: Care for older adults
*Place: Medical post
*Date: 04/13/2025
*Time: 7:00 AM
*Description: This adult care service offers...

Testing Phase: Usability Evaluation

In this phase, the results obtained by applying the SUS to a sample of 16 older adults, representative users of the target population who interacted with the mobile app prototype, were analyzed. The mean SUS score was 60.78 (SD 13.68) on a scale from 0 to 100, where higher values indicate better perceived usability, as can be seen in Figure 7. This score is interpreted as an acceptable usability according to the standards established in the literature, which indicates that the designed prototype is perceived as easy to use, safe, and suitable for the target population.

Figure 7. Distribution of System Usability Scale (SUS) scores using a boxplot.



Discussion

Outline

In this study, we designed and evaluated the usability of a blockchain-based mobile app architecture intended to support secure digital ID and access to essential services for older adults in rural Peru. Usability testing with 16 older adults showed that the mean SUS score was 60.78 (SD 13.68), indicating acceptable usability. Older adults viewed the app as secure and mostly easy to use and specifically highlighted the intuitive interface, assisted registration process, and personalized notifications as positive features [25,27], although there were ongoing issues related to digital literacy, trust in data security, and accessibility for users with disabilities.

In terms of accessibility, the design targeted low digital literacy and disability-related barriers through large high-contrast fonts, simplified menus, intuitive iconography, and an assisted registration flow [19]. In our sample, most participants self-reported no visual or motor limitations; those who did generally required more time and occasional assistance during

XSL•FO
RenderX

formative sessions, which directly informed these adaptations [8]. The confirmatory evaluation will quantify effectiveness using prespecified metrics, including completion without assistance across core tasks, median task time, error rate, and SUS, stratified by limitation status.

The prototypes and C4 models developed demonstrate a clear and scalable architecture, since they are structured in modular layers that allow incorporating new functionalities without affecting the stability of the system [22]. This scalability is reflected in the use of the C4 model, which facilitates the expansion of the system both in terms of components and cloud services, adapting to different regions or user groups [28]. The solution presented is accessible and reliable, as evidenced by the results obtained in the usability tests with older adults [29].

Unlike the conceptual framework proposed by Tan et al [30], which focuses on a governance taxonomy for blockchain-based systems in the public sector at macro, meso, and micro levels, our proposal is based on a practical and applied approach, which directly implements such principles in a functional architecture with tested prototypes. This framework includes concrete technical decisions on decentralized authentication, user privacy, and accessibility. The designed app facilitates secure and efficient access to essential services, promoting digital inclusion and trust in the handling of personal information [11]. As a limitation, while the study initially focused on design and prototyping, the app is currently being evaluated in real-world settings with older adults from the target population to validate its effectiveness, usability, and adoption in practical environments.

Related Work

In recent years, academic studies have begun to examine how blockchain technology can reinforce governance structures, support the decentralization of public service delivery, enhance security for mobile apps, and improve the user experience in distributed digital environments. Below are four thematic categories that frame the most relevant work.

Application of Blockchain in Public Sector Governance Models

Tan et al [30] address the implementation of blockchain's potential to transform public services by improving transparency, efficiency, and security, key aspects for the design of blockchain-based mobile apps. Through a conceptual framework, the authors explore key governance decisions at three levels: micro, meso, and macro. These governance elements, illustrated in Figure 8, affect the design and implementation of blockchain-based systems in the public sector [31]. However, they identify limitations such as the lack of interoperable infrastructure and the need for effective governance models. This study contributes to these limitations by proposing a solution based on a more flexible approach to digital identity management and access to essential services, especially through mobile apps that enable better interaction with public services [30].

XSL•FO
RenderX

Figure 8. Key elements in the economic design of governance.

Specifically, the proposed architecture directly operationalizes Tan et al's [30] microlevel governance principle through user-centered control of digital credentials. For instance, the assisted registration system and blockchain-based authentication module enable users to manage their identity data without relying on centralized authorities. Likewise, the notification and verification features provide transparency and consent management, ensuring that users are aware of and approve any data exchanges. At the mesolevel, interoperability is supported by modular components that facilitate integration with local authority systems, while the macrolevel implications relate to potential scalability within national digital identity strategies. In this way, our design translates Tan et al's [30] theoretical governance taxonomy into practical app features that promote individual autonomy and trust in digital identity management.

For his part, Ibrahim [32] explores the impact of decentralization on improving public services, highlighting the need to optimize government efficiency and accountability by delegating authority to local governments. This approach, which highlights the importance of tailoring services to local needs, is essential when considering blockchain-based mobile apps for public services. However, Ibrahim notes that disparities in local resources and capabilities limit the effectiveness of decentralization. Our research complements this analysis by integrating blockchain-based technologies, providing a solution that improves security, accessibility, and efficiency in digital ID and access to essential services, overcoming local resource barriers through a decentralized and accessible infrastructure [32].

Innovations in Blockchain-Based Mobile Apps for Security and Education

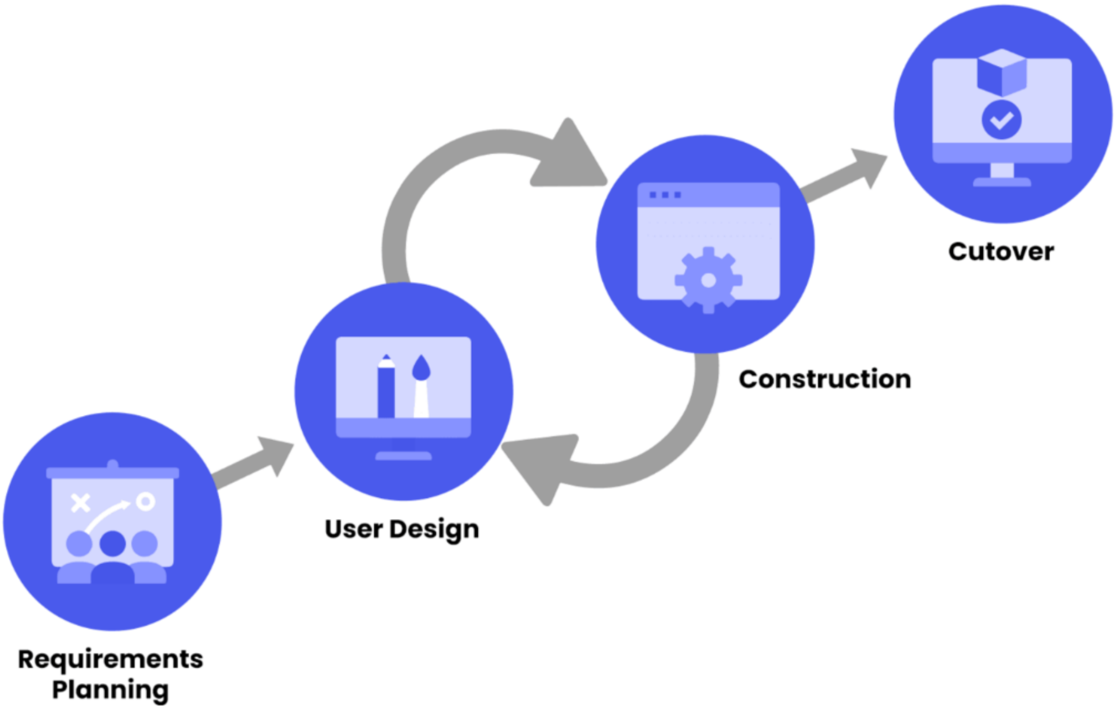
Rizky et al [33] propose a blockchain-based mobile app for decentralization of information management in the field of e-journals, addressing security and reliability issues in

XSL•FO
RenderX

centralized systems. Although their focus is on data security in academic platforms, their findings on decentralization and security are highly relevant to the design of mobile apps that manage digital ID in essential public services. Using a SWOT (strengths, weaknesses, opportunities, and threats) analysis and a waterfall development approach, the study implements blockchain to ensure a network resilient to external interference. However, scalability in larger environments remains a challenge. Our research advances this by proposing technical solutions that improve the efficiency and accessibility of essential services, overcoming these limitations by integrating blockchain for secure and decentralized digital ID.

Similarly, Asmawi et al [34] developed BlockScholar, a blockchain-based mobile educational app to facilitate the understanding of blockchain. Their research addresses the gap in interactive and accessible educational resources, highlighting the need for platforms that offer immersive and accessible learning on blockchain. They used the rapid application development model to create interactive and gamified content, as illustrated in Figure 9. Although the study made progress in accessibility and comprehension, it still faces challenges in adaptability to diverse educational contexts. Therefore, our research proposes solutions that enhance learning personalization and expand the coverage of blockchain-related topics.

Figure 9. Rapid application development model phases.



Security, Identity, and Digital Inclusion with Blockchain

Gumilar et al [35] explore the integration of financial technology in digital inclusion, with an emphasis on digital financial literacy as a catalyst for reducing economic and social disparities [35]. Through a systematic literature review, they used Scopus databases (2020-2024) to examine how the adoption of digital financial services optimizes inclusion. The results highlighted the importance of improving digital financial literacy but noted limitations in consumer use and protection. In line with these findings, our research extends the analysis by proposing innovative technological solutions and a more accessible design to improve adoption in underserved populations.

On the other hand, Musa et al [36] propose a blockchain-based approach to improve the security of data storage in Android mobile apps. They address the problem of vulnerability and unauthorized access to sensitive data by using blockchain to provide decentralized and secure storage. Their methodology includes the implementation of a 6-layer framework called BSADS (blockchain-based secure android data storage), which optimizes efficiency and security, as detailed in their proposed framework. Despite advances, scalability and costs remain limitations. This research contributes to overcoming these challenges by proposing lightweight node solutions and cost optimization techniques.

Identity and User Experience Management on Blockchain

Alanzi and Alkhatib's [37] study presents blockchain-based identity management solutions aimed at improving privacy and security in traditional centralized systems. Using blockchain technologies such as Ethereum and smart contracts, they address issues of third-party control, single point of failure, and vulnerability to data manipulation. However, limitations in scalability and weak authentication of some proposed systems are identified. In this context, our research extends these approaches, proposing improved decentralized infrastructure and optimizing the design of smart contracts to increase security and efficiency in digital identity management systems.

On the other hand, Jang and Han [38] developed a user experience framework for blockchain-based services, specifically focused on improving user interaction in contexts

XSL•FO
RenderX

such as finance and health care [38]. Through an analysis of active services, they identified both general and blockchain-specific functions, highlighting improved efficiency and security. However, their study faces limitations in general applicability due to a lack of standardization. Consequently, our research extends this framework, addressing the implementation of blockchain in digital identity and essential services, improving accessibility and user experience through advanced technological solutions [38]. Table 5 summarizes the key approaches and inputs underlying the development of a blockchain-based mobile app for digital ID and access to essential services.

Table . Fundamental approaches for blockchain-based mobile apps.

Approach	Contribution	Referenced works
Taxonomy of governance	• Rationale for the governance approach	• Tan et al [30]
Decentralization	• Decentralization analysis	• Ibrahim [32]
Patterns of mobile apps	• Mobile app design patterns	• Rizky et al [33] • Asmawi et al [34]
Inclusion, security, identity, and user experience	• Design and accessibility considerations • Blockchain security and authentication • Digital identity management • User experience and accessible design	• Gumilar et al [35] • Musa et al [36] • Alanzi and Alkhatib [37] • Jang and Han [38]

Conclusions

This study designed a functional architectural proposal for a blockchain-based mobile app that facilitates digital ID and access to essential services, with a focus on the inclusion and data security of older adults. The solution addresses usability, privacy, and decentralization issues, overcoming the limitations of traditional centralized and vulnerable systems. In addition to its technical and usability benefits, the proposed system entails ethical challenges related to data governance. If local authorities or institutions act as blockchain nodes, they could potentially exert undue control over users' identity data. To prevent this, future implementations should ensure transparent node management, community oversight, and independent auditing. These measures are key to promoting genuine digital inclusion while safeguarding autonomy and trust in rural contexts. Nevertheless, usability tests applied using the SUS revealed high acceptance and perceived ease of use by older adults, demonstrating that the proposal represents a solid basis for the comprehensive development of the app, which has been implemented in a functional prototype currently undergoing real-world testing with older adult participants in rural areas.

Acknowledgments

The authors would like to thank the Dirección de Investigación de la Universidad Peruana de Ciencias Aplicadas (UPC) for the support provided for the publication of this article.

Funding

This work received financial support from the Dirección de Investigación de la Universidad Peruana de Ciencias Aplicadas (UPC) exclusively for the article processing fee (APF). No additional funding was provided for the research activities.

Data Availability

The data collected in this study consist of usability evaluation responses (eg, survey data such as System Usability Scale scores and qualitative feedback) obtained from voluntary participants during prototype testing. No clinical, laboratory, or medical record data were collected. The dataset is not publicly available due to privacy and ethical considerations, but can be provided by the corresponding author upon reasonable request.

Authors' Contributions

WA-R and AFP-L contributed equally to this work. They led the study execution, including participant recruitment, data collection, usability evaluation, analysis of results, and drafting of the manuscript. JCM-A provided research supervision and leadership, contributed to the study planning and methodological guidance, ensured alignment with journal requirements, coordinated the submission and revision process, and performed critical review and editing of the manuscript.
All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
C4 model component diagram.
[PNG File, 134 KB - rehab_v13i1e79553_app1.png]

References

1. Cai S, Du J, Chen X, Li E, Chen Y. The relationship between e-health literacy and educational participation motivation among elderly individuals: the chained mediating effects of self-identity and social capital. Br J Hosp Med (Lond) 2024 Sep 30;85(9):1-13. [doi: 10.12968/hmed.2024.0261] [Medline: 39347685]
2. Farao A, Paparis G, Panda S, Panaousis E, Zarras A, Xenakis C. INCHAIN: a cyber insurance architecture with smart contracts and self-sovereign identity on top of blockchain. Int J Inf Secur 2024 Feb;23(1):347-371. [doi: 10.1007/s10207-023-00741-8]
3. Morte-Nadal T, Esteban-Navarro M. Recommendations for digital inclusion in the use of European digital public services. Humanit Soc Sci Commun 2025;12(1):273. [doi: 10.1057/s41599-025-04576-7]
4. Yi S, Yam ELY, Cheruvettolil K, et al. Perspectives of digital health innovations in low- and middle-income health care systems from South and Southeast Asia. J Med Internet Res 2024 Nov 25;26:e57612. [doi: 10.2196/57612] [Medline: 39586089]
5. Song W, Nokhbeh Zaeem R, Liau D, et al. Self-sovereign identity and user control for privacy-preserving contact tracing. Presented at: WI-IAT ’21; Dec 14-17, 2021; Melbourne, Australia p. 438-445 URL: https://dl.acm.org/doi/proceedings/10.1145/3486622 [doi: 10.1145/3486622.3493914]
6. George M, Chacko AM. Health passport: a blockchain-based PHR-integrated self-sovereign identity system. Front Blockchain 2023;6:1075083. [doi: 10.3389/fbloc.2023.1075083]
7. Ji H, Dong J, Pan W, Yu Y. Associations between digital literacy, health literacy, and digital health behaviors among rural residents: evidence from Zhejiang, China. Int J Equity Health 2024 Apr 9;23(1):68. [doi: 10.1186/s12939-024-02150-2] [Medline: 38594723]
8. Lu SY, Yoon S, Yee WQ, Heng Wen Ngiam N, Ng KYY, Low LL. Experiences of a community-based digital intervention among older people living in a low-income neighborhood: qualitative study. JMIR Aging 2024 Apr 25;7:e52292. [doi: 10.2196/52292] [Medline: 38662423]
9. Prieto-Egido I, Sanchez-Chaparro T, Urquijo-Reguera J. Impacts of information and communication technologies on the SDGs: the case of Mayu Telecomunicaciones in rural areas of Peru. Inf Technol Dev 2023 Jan 2;29(1):103-127. [doi: 10.1080/02681102.2022.2073581]
10. Li J, Liang N, Wu Y, Ma X, Liao Z. Effect of internet usage on the life satisfaction of the elderly spousal caregivers of disabled elderly. Sci Rep 2024;14(1):22294. [doi: 10.1038/s41598-024-73298-8]
11. Konopik J, Blunck D. Development of an evidence-based conceptual model of the health care sector under digital transformation: integrative review. J Med Internet Res 2023;25:e41512. [doi: 10.2196/41512]
12. Husain L, Greenhalgh T. Examining intersectionality and barriers to the uptake of video consultations among older adults from disadvantaged backgrounds with limited English proficiency: qualitative narrative interview study. J Med Internet Res 2025 Jan 6;27:e65690. [doi: 10.2196/65690] [Medline: 39761566]
13. Zuñe Chero L, Capuñay-Uceda OE, Idrogo Burga E, Capuñay-Uceda CE. Inequalities in learning performance caused by the digital divide during the health emergency. Presented at: 22nd LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI 2024; Jul 17-19, 2024 URL: https://laccei.org/LACCEI2024-CostaRica [doi: 10.18687/LACCEI2024.1.1.1042]
14. UPC Code of Ethics in Scientific Research of the Peruvian University of Applied Sciences (UPC). URL: https://investigacion.upc.edu.pe/assets/files/C%C3%B3digo-de-%C3%A9tica-en-la-investigacion-cientifica-de-la-UPC.pdf [accessed 2026-02-02]
15. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e29512. [doi: 10.2196/29512] [Medline: 35029535]
16. Sobrinho ADS, Gomes GDO, Bueno Júnior CR. Developing a multiprofessional mobile app to enhance health habits in older adults: user-centered approach. JMIR Form Res 2024 Apr 15;8(8):e54214. [doi: 10.2196/54214] [Medline: 38619865]
17. Essop H, Kekana R, Smuts H. Co-designing of a prototype mobile application for fetal radiation dose monitoring among pregnant radiographers using a design thinking approach. Health Informatics J 2024;30(3):14604582241284960. [doi: 10.1177/14604582241284960] [Medline: 39348214]
18. Chen E, Bishop J, Guge Cozon L, et al. Integrating human-centered design methods Into a health promotion project: supplemental nutrition assistance program education case study for intervention design. JMIR Form Res 2023 Apr 21;7:e37515. [doi: 10.2196/37515] [Medline: 37083485]
19. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023 Sep 21;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
20. Rashidi H, Rashidi Z. Software architecture tools: a classification and survey with recommendation for an organization. J Comput Secur 2023(10):61-81. [doi: 10.22108/jcs.2023.137862.1131]

XSL•FO
RenderX

21. Voorheis P, Zhao A, Kuluski K, et al. Integrating behavioral science and design thinking to develop mobile health interventions: systematic scoping review. JMIR Mhealth Uhealth 2022 Mar 16;10(3):e35799. [doi: 10.2196/35799] [Medline: 35293871]
22. Liang X, Zhao J, Chen Y, Bandara E, Shetty S. Architectural design of a blockchain-enabled, federated learning platform for algorithmic fairness in predictive health care: design science study. J Med Internet Res 2023 Oct 30;25:e46547. [doi: 10.2196/46547] [Medline: 37902833]
23. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry: Taylor & Francis; 1996:189-194. [doi: 10.1201/9781498710411]
24. Hill JR, Brown JC, Campbell NL, Holden RJ. Usability-in-place-remote usability testing methods for homebound older adults: rapid literature review. JMIR Form Res 2021 Nov 2;5(11):e26181. [doi: 10.2196/26181] [Medline: 34726604]
25. Sweeney M, Barton W, Nebeker C. Evaluating mobile apps targeting older adults: descriptive study. JMIR Form Res 2023 Apr 27;7:e37329. [doi: 10.2196/37329] [Medline: 37103995]
26. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults’ intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
27. Kim S, Park C, Park S, et al. Measuring digital health literacy in older adults: development and validation study. J Med Internet Res 2025 Feb 5;27:e65492. [doi: 10.2196/65492] [Medline: 39908081]
28. Liang X, Alam N, Sultana T, Bandara E, Shetty S. Designing a blockchain-empowered telehealth artifact for decentralized identity management and trustworthy communication: interdisciplinary approach. J Med Internet Res 2024 Sep 25;26:e46556. [doi: 10.2196/46556] [Medline: 39320943]
29. Koo JH, Park YH, Kang DR. Factors predicting older people’s acceptance of a personalized health care service app and the effect of chronic disease: cross-sectional questionnaire study. JMIR Aging 2023 Jun 21;6:e41429. [doi: 10.2196/41429] [Medline: 37342076]
30. Tan E, Mahula S, Crompvoets J. Blockchain governance in the public sector: a conceptual framework for public management. Gov Inf Q 2022 Jan;39(1):101625. [doi: 10.1016/j.giq.2021.101625]
31. Barrera C. A framework for blockchain governance design: the Prysm group wheel.: Medium; 2019. URL: https://medium.com/prysmeconomics/a-framework-for-blockchain-governance-design-the-prysm-group-wheel-703279c1b0dd [accessed 2025-06-01]
32. Ibrahim AHH. Decentralization and its impact on improving public services. Int J Soc Sci 2024;7(2):45-53. [doi: 10.21744/ijss.v7n2.2278]
33. Rizky A, Lutfiani N, Sri Mariyati W, Atika Sari A, Febrianto KR. Decentralization of information using blockchain technology on mobile apps e-journal. B-FronT 2022;1(2):1-10 [FREE Full text] [doi: 10.34306/bfront.v1i2.37]
34. Asmawi A, Saifulbahri EM, Mohd Ariffin NA. Development of BlockScholar as an educational mobile application on blockchain technology. ARASET 2023;34(1):15-23 [FREE Full text] [doi: 10.37934/araset.34.1.1523]
35. Gumilar DWA, Sangka KB, Totalia SA. Digital financial literacy and digital financial inclusion in the era of digital disruption: systematic literature review. FJMR 2024;3(5):1563-1576 [FREE Full text] [doi: 10.55927/fjmr.v3i5.9213]
36. Musa HS, Krichen M, Altun AA, Ammi M. Survey on blockchain-based data storage security for android mobile applications. Sensors (Basel) 2023 Oct 26;23(21):8749. [doi: 10.3390/s23218749] [Medline: 37960449]
37. Alanzi H, Alkhatib M. Towards improving privacy and security of identity management systems using blockchain technology: a systematic review. Appl Sci (Basel) 2022;12(23):12415. [doi: 10.3390/app122312415]
38. Jang H, Han SH. User experience framework for understanding user experience in blockchain services. Int J Hum Comput Stud 2022 Feb;158:102733. [doi: 10.1016/j.ijhcs.2021.102733]

Abbreviations

DT: design thinking
SSI: self-sovereign identity
SUS: System Usability Scale

Edited by S Munce; submitted 24.Jun.2025; peer-reviewed by C Wrigley, SB Guo; revised version received 20.Nov.2025; accepted 03.Dec.2025; published 02.Feb.2026.

Please cite as:
Arana-Ramos W, Pastrana-Leon AF, Morales-Arevalo JC
Blockchain-Based Mobile App for Digital Identification of Older Adults in Rural Peru: Design and Usability Evaluation Study
JMIR Rehabil Assist Technol 2026;13:e79553
URL: https://rehab.jmir.org/2026/1/e79553
doi:10.2196/79553

XSL•FO
RenderX

© Wilver Arana-Ramos, Aldo Francisco Pastrana-Leon, Juan Carlos Morales-Arevalo. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 2.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Social Justice Approach to Assistive Technology and Well-Being of People With Visual Disabilities in Low- and Middle-Income Countries: Qualitative Narrative Study

Luisa Maria Ortiz-Escobar[1,2*], MSc; Mario Andres Chavarria[3,4,5*], PhD; Samia Hurst-Majno[2*], MD, PhD, Prof Dr Med; Oscar Ivan Campo Salazar[3*], PhD; Celia Escobar-Hurtado[6], MSc; Michael Ashley Stein[7,8*], PhD; Minerva Rivas Velarde[2,5*], PhD, Prof Dr

[1]Grupo de Investigación en Ingeniería Biomédica, GBIO, Universidad Autónoma de Occidente, Santiago de Cali, Colombia
[2]Institute of Ethics, History, and Humanities, University of Geneva, Geneva, Switzerland
[3]Grupo de Investigaciones en Biomédica, Universidad Autónoma de Occidente, Santiago de Cali, Colombia
[4]EssentialTech Centre, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[5]School of Health Science, HES-SO Genève, Avenue de Champel 47 1206 Genève, Geneva, Switzerland
[6]Escuela de Rehabilitación Humana, Facultad de Salud, Universidad del Valle, Santiago de Cali, Colombia
[7]Harvard Law School Project on Disability, Harvard University, Cambridge, MA, United States
[8]Faculty of Law Centre for Human Rights, University of Pretoria, Pretoria, South Africa
[*]these authors contributed equally

Corresponding Author:
Minerva Rivas Velarde, PhD, Prof Dr
Institute of Ethics, History, and Humanities, University of Geneva, Geneva, Switzerland

Abstract

Background: The United Nations' third Sustainable Development Goal emphasizes ensuring healthy lives and promoting well-being (WB) for all, which requires effective assistive technology (AT) for persons with disabilities. In low- and middle-income countries (LMICs), however, AT remains largely inaccessible, and high abandonment rates indicate that many existing solutions fail to meet users' needs. To improve AT design and effectiveness, a deeper understanding of users' lived experiences and the ways AT influences WB is essential.

Objective: This study aimed to explore how technology creates opportunities or barriers in the daily lives of persons with visual disabilities in LMICs and how it affects their WB.

Methods: We conducted a qualitative narrative study guided by deductive qualitative analysis, using the capability approach (CA) and disadvantage theory (DT) as theoretical frameworks. Nineteen adults with visual disabilities from Cali, Colombia, participated in in-depth, semistructured interviews. A focus group (n=5) deepened the exploration of shared experiences. Data analysis followed three stages: (1) deductive coding using Nussbaum list of central capabilities and key CA constructs (functionings, conversion factors, and agency); (2) recoding through DT concepts (insecure functioning, corrosive disadvantages, and fertile functionings); and (3) inductive analysis to capture emergent sociocultural themes.

Results: AT shaped both opportunities and constraints in participants' lives. While functionings such as employment, mobility, and affiliation were highly valued, they often remained insecure due to systemic barriers. Corrosive disadvantages—such as unemployment, exclusion, and limited spatial autonomy—undermined multiple capabilities simultaneously. Conversely, fertile functionings such as equitable employment, adaptive sports, and access to well-designed AT supported agency and resilience. The inductive analysis revealed 3 interconnected themes: the aspiration to explore and expand movement, the desire to appear attractive, and the adoption of nonconfrontational strategies to maintain social harmony. These findings highlight how emotional, aesthetic, and cultural dimensions shape the experience and meaning of AT.

Conclusions: While AT research in LMICs often emphasizes availability, it rarely addresses how social norms, structural violence, and fear affect meaningful use. The combined CA and DT lens reveals that AT can either enable or constrain WB depending on how it aligns with users' lived contexts. Designing for fertile functionings—those that support agency, safety, and resilience—is essential. Participatory, context-sensitive design must prioritize not only functionality, but also aesthetic dignity, cultural relevance, and emotional security. Including the voices—and silences—of persons with disabilities in the Global South is crucial for transforming AT from a mere tool into a catalyst for real freedom and WB.

(JMIR Rehabil Assist Technol 2026;13:e72306) doi:10.2196/72306

XSL•FO
RenderX

KEYWORDS

assistive technology; persons with disabilities; visual disabilities; user-centered design; capabilities approach; disadvantage theory; LMIC; low- and middle-income countries

Introduction

Background

The incidence and prevalence of visual impairment worldwide have exceeded World Health Organization (WHO) projections for the second decade of the current century [1,2]. In low- and middle-income countries (LMICs), the rate of persons with severe visual impairments is approximately 2 to 3 times higher than in high-income countries [3].

Visual impairment significantly constrains individual autonomy and social participation, with far-reaching consequences for quality of life [4]. Assistive technology (AT) has emerged as a critical means of mitigating these effects by promoting independence and enabling greater inclusion.

Beyond its practical benefits, AT is increasingly recognized as instrumental in advancing the rights enshrined in the United Nations Convention on the Rights of Persons with Disabilities [5], as well as in supporting broader global commitments such as Sustainable Development Goal 3: Good Health and Well-Being [6].

Thus, AT contributes to overall well-being (WB) by enhancing autonomy, supporting participation in education, employment, and community life, and reinforcing psychosocial WB through increased agency and participation [7,8].

Nevertheless, despite its benefits, access to AT remains severely constrained—particularly in LMICs [9]. This lack of access is further exacerbated by high abandonment rates, which often result from inadequate alignment between technologies and the lived realities of users [10-12]. The literature increasingly points to the need for participatory, person-centered approaches that actively engage persons with disabilities in the design and delivery of AT solutions [13], ensuring that such interventions are both contextually relevant and sustainable.

Although AT development increasingly claims to adopt user-centered design [14], studies have identified a persistent gap between rhetoric and practice. Users are often consulted only to refine existing designs, rather than to inform fundamental decisions based on their needs and experiences [15,16].

To respond to this gap and advance a genuinely person-centered approach, we initiated a study to inform the development of AT for the navigation of persons with visual disabilities in LMICs. This paper presents the results of the first phase, a qualitative inquiry into users' lived experiences, aspirations, and interactions with technology.

While AT research in LMICs often focuses on availability, it neglects how structural and social contexts shape meaningful use. The capability approach (CA) addresses this by emphasizing individuals' real freedoms to achieve valued functionings, not just access to resources [17-19]. However, CA has been critiqued for undertheorizing how power, fear, and structural violence constrain agency [20]. Disadvantage theory (DT) complements CA by introducing concepts such as corrosive disadvantage, enabling deeper analysis of how AT may worsen exclusion under adverse conditions [21]. This combined CA and DT lens highlights how AT can both enable and limit WB in LMICs.

Approaches to Disability, WB, AT, and Justice

The Capabilities Approach

The CA evaluates human development through the lens of equity, emphasizing human flourishing as the ultimate goal. It focuses on the life individuals are able to lead, based on their "capabilities" (practical opportunities) and "functionings" (achievements) [22].

Building on this foundation, this approach recognizes the concept of agency, which is the ability to pursue valued goals. To achieve these goals, individuals need access to necessary resources and "conversion factors," which affect how resources are transformed into functionings [17].

This theoretical lens has gained traction in disability studies over the past 2 decades, where it emphasizes individual agency and enables a nuanced examination of resources and conversion factors. In this context, disability is framed in terms of the deprivation or restriction of capabilities and functionings [23-26].

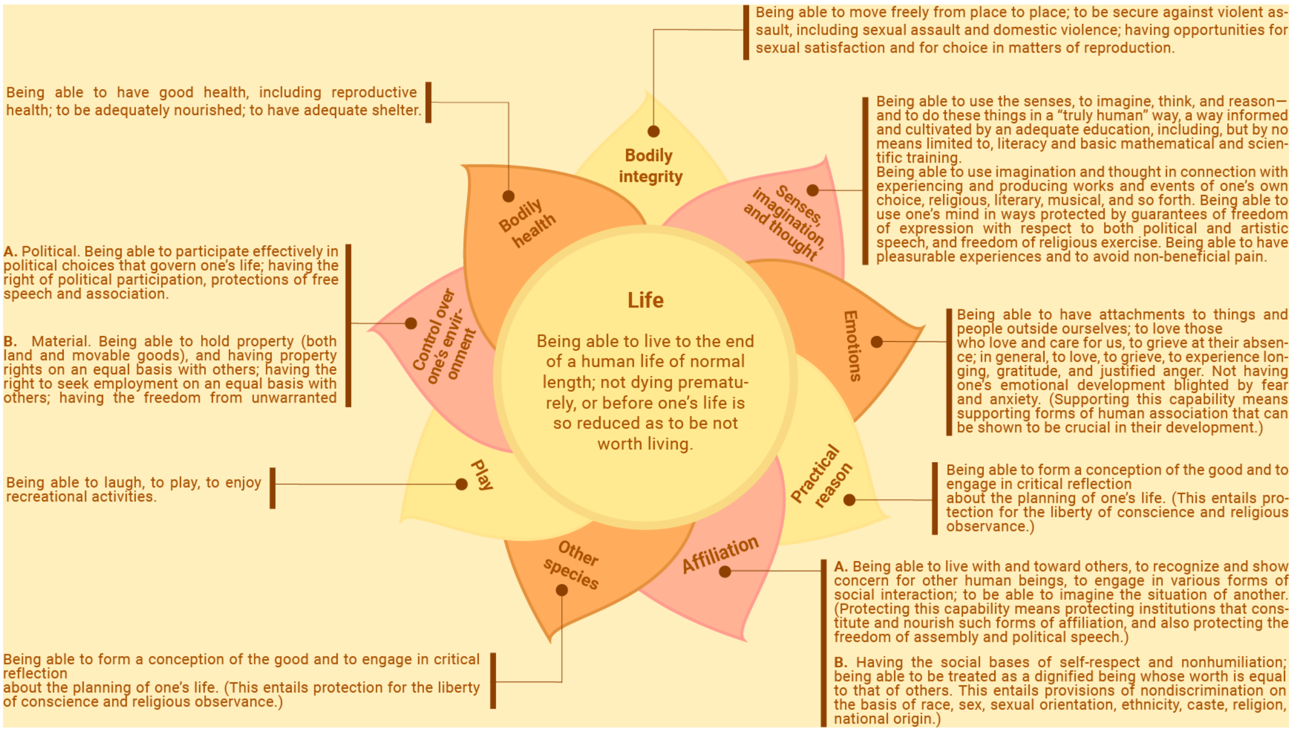
For this study, we chose the CA to analyze the experiences of persons with visual disabilities in a middle-income country, namely Colombia. This framework sheds light on what people value in their lives, guides researchers in identifying daily opportunities and barriers, and aids in evaluating how AT influences the pursuit of a "good life."

Recent literature situates the role of technology within this framework [27,28].

Nussbaum, who has made valuable contributions to the approach, proposed an open list of 10 capabilities, which she argues are central to the realization of self-determined, meaningful, and fulfilling lives [18,29]. This list (Figure 1) has been effectively applied in various contexts, including assessing dignity, monitoring human rights, and guiding social policy to reduce disadvantage [30].

XSL•FO
RenderX

Figure 1. Diagram of the 10 central capabilities proposed by Nussbaum and used as a framework for deductive analysis of participant narratives about technology and well-being.



Being able to move freely from place to place; to be secure against violent assault, including sexual assault and domestic violence; having opportunities for sexual satisfaction and for choice in matters of reproduction.

Being able to have good health, including reproductive health; to be adequately nourished; to have adequate shelter.

Being able to use the senses, to imagine, think, and reason—and to do these things in a "truly human" way, a way informed and cultivated by an adequate education, including, but by no means limited to, literacy and basic mathematical and scientific training.
Being able to use imagination and thought in connection with experiencing and producing works and events of one's own choice, religious, literary, musical, and so forth. Being able to use one's mind in ways protected by guarantees of freedom of expression with respect to both political and artistic speech, and freedom of religious exercise. Being able to have pleasurable experiences and to avoid non-beneficial pain.

A. Political. Being able to participate effectively in political choices that govern one's life; having the right of political participation, protections of free speech and association.

B. Material. Being able to hold property (both land and movable goods), and having property rights on an equal basis with others; having the right to seek employment on an equal basis with others; having the freedom from unwarranted

Being able to live to the end of a human life of normal length; not dying prematurely, or before one's life is so reduced as to be not worth living.

Being able to have attachments to things and people outside ourselves; to love those
who love and care for us, to grieve at their absence; in general, to love, to grieve, to experience longing, gratitude, and justified anger. Not having one's emotional development blighted by fear and anxiety. (Supporting this capability means supporting forms of human association that can be shown to be crucial in their development.)

Being able to laugh, to play, to enjoy recreational activities.

Being able to form a conception of the good and to engage in critical reflection
about the planning of one's life. (This entails protection for the liberty of conscience and religious observance.)

A. Being able to live with and toward others, to recognize and show concern for other human beings, to engage in various forms of social interaction; to be able to imagine the situation of another. (Protecting this capability means protecting institutions that constitute and nourish such forms of affiliation, and also protecting the freedom of assembly and political speech.)

B. Having the social bases of self-respect and nonhumiliation; being able to be treated as a dignified being whose worth is equal to that of others. This entails provisions of nondiscrimination on the basis of race, sex, sexual orientation, ethnicity, caste, religion, national origin.)

Being able to form a conception of the good and to engage in critical reflection
about the planning of one's life. (This entails protection for the liberty of conscience and religious observance.)

Disability research focused on understanding quality of life, living conditions, and WB in LMICs is scarce and predominantly quantitative, limiting insights into the lived experiences of disability and personal ideas of flourishing [31,32]. Research on AT within the CA for persons with disabilities is even rarer. This field could benefit from studies like this one, which aim to provide valuable information for developing AT to narrow the capabilities gap. Understanding the lived experience of disability and personal definitions of a good life can guide AT designers in making meaningful and desirable assistive devices (ADs).

Disadvantage Theory

Wolff and De-Shalit [21] developed a theory aimed at achieving equity and social justice by addressing disadvantage. Drawing from Sen CA and Nussbaum contributions, they define disadvantage as a "lack of genuine opportunities for secure functionings" [21].

These authors underscore the plural origins of disadvantages and the necessity of indexing them. They diverge from Sen and Nussbaum by using the terms genuine opportunities and secure functionings instead of capabilities and functionings. While Sen emphasizes freedom to achieve capabilities, Nussbaum defends a universal list of central capabilities, Wolff and De-Shalit focus on how disadvantage is experienced and addressed in real-world contexts. Their terminology reflects a concern with the stability of achieved functionings and the substantive nature of opportunity, aiming to make the framework more actionable for policy design.

They define a disadvantaged person as one whose functionings are involuntarily insecure. Another key concept is "corrosive disadvantage," where one disadvantage negatively impacts another functioning, making it insecure. To counteract this, the authors suggest identifying and addressing corrosive disadvantages while also seeking "fertile functionings," which positively impact and secure other functionings [21].

Persons with disabilities face multiple disadvantages, such as "reduced health, life span, opportunities for fulfilling work, social networks, control over their lives, leisure activities, and living standards" [21,33]. Given the likelihood of facing corrosive disadvantages, it is crucial to prioritize the capabilities and functionings of persons with disabilities to decluster their disadvantages.

WB, Visual Disability, and AT

Sen [34] addresses WB from an individual perspective, asserting that it should be evaluated based on the achievement of functionings that each person values for a good life. He states, "functionings are seen as central to the nature of WB, even though the sources of WB could easily be external to the person." This aligns with the current notion of subjective WB, "a person feeling and thinking his or her life is desirable regardless of how others see it" [35], which is used by organizations and countries to monitor compliance with Sustainable Development Goal 3 [36,37].

Most data on this subject come from high-income countries, showing that persons with disabilities generally have lower WB than those without disabilities, often linked to poverty, educational status, unemployment, and social exclusion [38-41]. A recent study including data from LMICs confirmed this trend, attributing lower WB in persons with disabilities to "differences in living conditions and life experiences in relation to household wealth, level of education, partnership status, and exposure to violence and discrimination in the previous year" [42].

XSL•FO
RenderX

In the context of visual disabilities, common challenges include social isolation, limited mobility, low education, unemployment, stigmatization, and injury risks [43].

ATs are seen as tools to overcome these limitations and improve WB. Despite their potential benefits, not all persons with visual disabilities adopt AT, particularly mobility aids, due to concerns about stigmatization and inadequate obstacle detection [44]. This reluctance indicates that AT does not always enhance WB, underscoring the need to understand user interactions with AT to develop solutions that are more desirable and supportive of users' WB.

Aims of This Study

We explored the role of technology in the lives of persons with visual disabilities in LMICs and aimed to answer the following questions: (1) How can AT create opportunities or barriers in daily life? (2) How does technology affect WB?

Context of Fieldwork

Lived experience, disability, capabilities, and functionings should be understood within their sociocultural context [32,45]. The cultural perception of disability significantly influences both coping strategies and the willingness to use AT. These factors cannot be separated from the historical and political contexts that have shaped the experiences of a particular group [46].

This study took place in Colombia, a country affected by internal armed conflict for nearly 6 decades. The "Comisión de la Verdad (Truth Commission)" reported 8,775,884 victims of human rights violations in the conflict [47].

In 2016, a peace agreement was signed between the Colombian government and the FARC-EP (Revolutionary Armed Forces of Colombia). The historical memory construction process, a process of reconstructing and interpreting the past to dignify victims, clarify the truth about human rights violations, and promote reconciliation in the context of Colombia's armed conflict, highlighted patriarchy, racism, and discrimination based on social class, religion, or ideology as conflict axes [48,49].

Although Colombia adopted the Convention on the Rights of Persons with Disabilities (CRPD) in 2011, disability has been scarcely addressed in the peace agreement and subsequent documents [50]. While this study does not focus on recognized conflict victims, it is important to situate the participants' reality within this sociopolitical context. Living in a region with complex social issues, partly due to the conflict and other factors, means that numerous disadvantaged population groups exist. Recognizing persons with disabilities as one of the "worst-off" groups for prioritizing attention is challenging, especially priorities after the peace agreement priorities include attending to direct victims of the conflict, which continues with other actors present.

The interviews were conducted in Cali, Colombia's third most populous city and economic center. Cali is noted for its multiethnic population, sports culture, and as a cultural hub, but also as one of the most violent cities in Colombia [51].

While ATs can potentially expand the capabilities of people with disabilities, analyzing the geographic, social, political, and economic context of potential users during the design phase is crucial. This consideration is particularly important since using ATs could increase security risks for persons with disabilities.

Methods

Study Design and Theoretical Framework

This study adopted a qualitative narrative approach framed within a deductive qualitative analysis strategy [52]. The research was grounded in the CA [18] and the DT [21] as its core theoretical frameworks, with a critical theory lens [53] guiding the inductive reflection. This hybrid strategy enabled the exploration of how AT supports or constrains human capabilities and WB among people with visual disabilities in low- and middle-income settings.

Participants and Data Collection

This study involved 19 voluntary participants. Eligibility criteria included being 18 years or older; residing in Cali, Colombia, or nearby municipalities; and being blind or visually impaired, without additional hearing, motor, or cognitive disabilities. Eligibility was verified using the Washington Group Questions on Disability Statistics [54]. Purposive and snowball sampling strategies were used. Semistructured interviews (60 - 105 minutes) captured life narratives and good life aspirations. A follow-up focus group (n=5) explored their relationship with technology in greater depth, addressing (1) commonly used technologies, (2) perceived utility, (3) opportunities created, (4) desired opportunities, (5) risks, and (6) features enhancing WB.

To ensure rigor, the study used triangulation through a focus group, member checking with the participants, peer debriefings, and a detailed audit trail. Trustworthiness was further supported by verbatim quotes in the findings, which preserved participant voice and supported transparency.

Data Analysis

Overview

Data were analyzed using deductive qualitative analysis guided by the CA. A preliminary deductive codebook was developed based on Nussbaum 10 central capabilities, supplemented with key CA constructs, including resources, conversion factors (personal, social, and structural), agency, and achieved functionings.

Once these were applied, a second layer of deductive analysis was conducted using DT, focusing on identifying insecure functionings, corrosive disadvantages, and fertile functionings.

In addition to the 2-phase deductive stage, an inductive thematic analysis was performed to identify emergent themes not captured by the theoretical framework. This inductive analysis was critically informed by critical theory, allowing interpretation of power asymmetries in technology design, structural exclusion, and user agency.

XSL•FO
RenderX

Step-by-Step Coding Procedure

Familiarization

The research team read all transcripts multiple times, documenting initial impressions and emotional, social, and technological cues.

Deductive Coding: Phase 1 (CA)

Using Nussbaum capabilities, data were categorized into capability domains and analyzed with subcodes for functionings, resources, and conversion factors.

Deductive Coding: Phase 2 (DT)

Narratives were recoded to assess the security and fertility of functionings and to identify corrosive disadvantages (eg, stigma and economic precarity).

Inductive Coding

Emergent codes such as desire to appear attractive, resistance to confrontation, and wish to explore movement were documented, grouped, and interpreted using critical theory.

Integration

Inductive and deductive codes were reconciled through iterative discussions, using analytic memos. For instance, fear of confrontation was related to insecure functionings under “Control over one’s environment,” while vanity versus disability was interpreted within “Affiliation.”

Synthesis

Final interpretation focused on how AT shaped each participant’s capability set, aspirations, and obstacles to achieving the good life.

Disagreements in coding were addressed through iterative team discussions until consensus was reached. A reflexivity journal was maintained by the lead researchers to document positionality, epistemic tensions, and interpretive decisions. All coding was conducted manually without the use of software for qualitative analysis. Transcripts were originally in Spanish, and coding was performed in Spanish by native speakers. English translations were collaboratively developed to preserve semantic and cultural accuracy.

Codebook Summary

Deductive and inductive codes are presented for illustrative purposes (Tables 1-3).

Table . Deductive codebook (capability approach).

Capability	Subcapability	Code	Definition	Conversion factor type	Example quote
Bodily integrity	Bodily sovereignty	Privacy respect	Situations of physical or decisional invasion	Social	“...they touch you… that’s not necessary. So, it’s like they don’t respect your bodily integrity...”
Bodily integrity	Mobility	Inaccessible transport	Lack of access to safe, navigable transport	Structural	“...the audio announcements don’t work... so how are we supposed to know if the bus is the one we’re waiting for?”
Affiliation	Dignity in service delivery	Discrimination	Being rejected or underestimated during services provision	Social	“...people assume I’m there to beg or ask for something.”
Control over one’s environment	Material control	Institutional barriers	Bureaucratic or policy-related access issues	Structural	“ATMs don’t talk... it’s both accessibility and safety.”

XSL•FO
RenderX

Table . Deductive codebook (disadvantage theory).

DT[a] code	Definition	Inclusion criteria	Example quote
Insecure functioning	A functioning being achieved but threatened	Describes fragile access to opportunities	“I only go to the bank with someone I trust. ATMs are not accessible and it’s unsafe.”
Corrosive disadvantage	A single disadvantage that creates multiple deprivations	Barriers that multiply exclusions (eg, inaccessible transport affects work and autonomy)	“If I can’t use public transport, I can’t go to training, or to work. Everything falls apart.”
Fertile functioning	A capability that enables many others	Mentions of tech or contexts that expand autonomy	“My cell phone is essential in my life. It helps me navigate, communicate with others, check when the bus is coming, know where to get off, and get any place.”

[a] DT: disadvantage theory.

Table . Integration of inductive themes with deductive frameworks (CA[a] and DT[b]).

Inductive theme	Definition	Example (quote)	Related capability (CA)	Associated concept (DT)	Interpretive notes
Wish to appear attractive	The desire to look good and express personal style, challenging stereotypes of dependency.	“I consider myself vain despite my disability... so, I would buy more clothes... I would dye my hair...."	Affiliation and bodily integrity	Insecure functioning or symbolic disadvantage	Challenges cultural narratives of disability as asexual or unattractive; signals agency in self-image.

[a] CA: capability approach.

[b] DT: disadvantage theory.

Reflexivity and Strategies for Rigor

Reflexivity was maintained via journals kept by both main researchers. These captured assumptions, affective responses, and positionality reflections. Disagreements in coding were addressed through iterative team discussions and memoing. Credibility was supported through member checking and triangulation using interview and focus group data. Saturation was considered achieved after 19 participants, as no new themes emerged during the final interviews and confirmatory focus group.

Researcher-Participant Relationships and Empowerment Strategies

Although this study did not follow a formal participatory design framework, several intentional strategies were used to promote ethical engagement and participant empowerment. Informed consent was provided in accessible formats, including Braille and audio, with participants choosing the version that best suited them. Potential interview locations were identified by the research team and evaluated in consultation with key participants, allowing for informed, context-sensitive decisions.

Participants were clearly informed of the study’s aims and the distinct roles of the research teams. Notably, the team responsible for data collection was not involved in AT design, helping to manage expectations and avoid instrumental relationships. Interviews were guided by an ethic of care, privileging participant-led narratives and emphasizing trust, autonomy, and open dialogue.

We also approached participation critically. While participatory methods are often promoted in inclusive design, they may reinforce power asymmetries if not carefully implemented. As Cooke and Kothari [55,56] warn, such approaches can result in a “tyranny of participation,” serving dominant interests rather than those of marginalized groups. To avoid this, our study prioritized epistemic justice by centering participants’ lived experiences as ends in themselves, not as inputs to a predefined design process. Member checking and a focus group enabled deeper reflection and participatory analysis, contributing to more equitable and meaningful knowledge production in line with the ethical commitments of the CA [17].

Ethical Considerations

The research protocol was approved by the Swiss Federal Institute of Technology Lausanne on research involving humans (056 - 2021 and 068 - 2022), the Human Ethics Committee of Universidad del Valle (008 - 022), Hospital Universitario del Valle (029 - 2022), and Instituto para Niños Ciegos y Sordos del Valle (CEI-2022 - 02), ensuring fully informed consent and participant anonymity. In accordance with ethical guidelines for research involving human participants, compensation was provided in a manner that did not constitute undue inducement. Participants received reimbursement for transportation costs and were offered food and beverages during data collection sessions. Additionally, they received a modest stipend of approximately 40,000 COP (equivalent to less than US $10) as compensation for their time and participation.

XSL•FO
RenderX

Results

Overview

This section presents the main findings of the study, starting with a general description of participant demographics (Table 4) to provide contextual background. The results are then organized into 2 main areas: the first presents the outcomes of the deductive analysis guided by the CA, structured under the subheadings “Notions of Well-Being and Autonomy,” “WB and the Good Life: Common Functionings Identified,” and “Lived Experiences through the Lens of Capabilities and Technology.” The second area includes the inductive themes that emerged beyond the predefined theoretical framework. Although the results of the secondary deductive analysis informed by DT—building upon the initial CA coding—are not included in this section, they are presented in the discussion, where they more effectively enrich the interpretation of structural constraints, genuine opportunities, and the security of valued functionings within participants’ lived realities.

Table . Participant demographics.

Participant characteristics	Overall
Age (years), mean (SD)	31.1 (10.9)
Sex, n (%)	
Female	7 (36.8)
Male	12 (63.2)
Marital status, n (%)	
Single	13 (68.4)
Common law	3 (15.8)
Married	3 (15.8)
Education, n (%)	
No formal education	1 (5.3)
Primary	1 (5.3)
Completed high school	6 (31.6)
College or university in process	5 (26.3)
Completed college or university	6 (31.6)
Employment status	
Employed	6 (31.6)
Unemployed	6 (31.6)
Self-employed	2 (10.6)
Student	5 (26.3)
Origin	
Urban	17 (89.4)
Rural	2 (10.6)
Self-reported ethnicity	
Afro-Colombian	2 (10.6)
None	17 (89.4)

Notions of WB

In exploring what constitutes WB for the participants, we inquired about their ideal life and contrasted this with their current experiences. It quickly became clear that disability significantly influences their conception of an ideal life. A couple of participants explicitly noted that their vision of an ideal life had to be “edited” due to their disability, such as stating “my ideal life now that I am blind” or recognizing that their disability constrains their aspirations.

Autonomy and WB

Participants identified autonomy as central to their WB, emphasizing the importance of freedom in decision-making and the ability to make meaningful choices. One participant articulated this need for autonomy as follows:

> *A more assertive life would be one in which people first see me as an equal, rather than focusing on my condition. It would be a life where I can make all my own decisions, where having a disability does not limit my opportunities.* [Male, 45 years]

The emphasis on autonomy extends to living without the constraints of disability-related prejudice and ensuring social acceptance. However, the frustration with societal limitations often leads to a pragmatic view of WB, focusing on achievable goals within their control and avoiding, in the process, being assisted by others. For example, a 29-year-old participant described her ideal life as

> *To have as much autonomy as possible—being confident to go out independently, having a stable job that meets my needs, maintaining good health, and practicing my sport. Total autonomy in my life is what I strive for.* [Participant, 29 years]

WB and the Good Life: Common Functionings Identified

Work is frequently cited as a critical element of WB. It is valued both for providing financial security and as a source of personal fulfillment. Achieving financial stability and having a fulfilling job are integral to their idea of a desirable life. As one participant who has several part-time jobs with limited hours a week, put it,

> *I have a job and I am grateful, but I would like to improve my working conditions to better provide for my family and ensure our own home.* [Female, 28 years]

Ownership of a home is another significant aspiration, often linked to job stability. Other desired functionings include pursuing education, maintaining good health, and excelling in sports. These aspirations reflect a broader understanding of WB that includes personal and professional fulfillment.

Moreover, family stability, emotional support, and social connections are also important components of WB for participants.

When one participant (male, 40 years) compared his current life with his ideals, he became aware that he had achieved much of what he considers a flourishing life. He reflected on his accomplishments with satisfaction while acknowledging areas for improvement. Despite this, he emphasized that his experience is not universal among persons with disabilities and that the limitations imposed by disability affect what can be achieved.

> *At this point, as a person with a visual disability, I believe I have achieved a great deal, despite the challenges. I see others who, unfortunately, lack employment, family, or social connections, or who haven’t been able to fully enjoy life. For me, job stability would make my life 100% ideal. I feel fulfilled because I’ve experienced love, marriage and the journey of being a parent. I have also traveled, engaged socially, and even gone bungee jumping.* [Male, 40 years]

Though the desire for family formation or partnership is less commonly mentioned by those who were single at the time of the interview.

Lived Experiences Through the Lens of Capabilities and Technology

In examining Nussbaum's central capabilities in the context of individuals with visual disabilities, the analysis reveals both opportunities and barriers that impact their daily lives.

Life

This capability is initially affected by the sociopolitical context of the country and specifically by the social and infrastructural conditions of the city. As one participant recounted,

> *We left Cauca, displaced by the violence. We also lived in Putumayo for a few years, but in 2009 they killed my brother, and we had to leave again—displaced by the violence.* [Female 21 years]

(Cauca and Putumayo are departments in southern Colombia that have been heavily affected by armed conflict.)

In some cases, participants explicitly linked their visual impairment to episodes of violence. One participant explained:

> *I’ve been living in Cali for four years, which is also how long I’ve had this condition (low vision). I was living in my village at the time, and I came here to study. It was around then that the incident happened — someone attacked me to steal my phone and other belongings.* [Male, 23 years]

Similarly, another participant shared:

> *I lost my sight when I was 20 years old, as a result of a firearm injury. After that—well, thank God—I recovered well after about a year. It was like a sabbatical year, where I didn’t really do anything, because at the time, I didn’t know how to get my life back on track.* [Male, 29 years]

Beyond the direct consequences of violence, participants also pointed to everyday urban dynamics that fail to prioritize the safety and dignity of people with visual disabilities. In particular, unsafe social behavior—such as the frequent disregard for traffic rules by drivers—emerged as a recurring concern. Falls, collisions, and traffic accidents were identified as major threats to personal safety and mobility. Several participants described incidents in which they had sustained serious injuries.

> *I fell into the gap and ended up on the road, between the station platform and the bus.* [Female, 31 years]

One participant told during the interview that her brother, who was also blind, was the fatal victim of a traffic accident 2 months earlier, when he was run over by a car while crossing the street.

Bodily Health

Participants have access to the Colombian health system and benefit from rehabilitation services that focus on orientation and mobility. Participants from rural areas only accessed these services once they arrived in Cali. Regarding access to basic AT, although there is a route to obtain it through the health system, a couple of participants reported that it is cumbersome. However, they say that canes are more readily available through the mayor’s office.

XSL•FO
RenderX

Financial instability has led to periods where participants' nutrition and living conditions have been compromised. As the following quotes illustrate:

> *When I was little and still living with my mom and siblings, we got evicted because we couldn't pay the rent anymore. We ended up moving into the first place we found, but it was full of rats—and I'm really scared of them.* [Female, 54 years]

> *There were hard times, times of real scarcity. Sometimes you had to really stretch things—like cutting back a lot on food—just to pay for basic services, because skipping those payments wasn't an option.* [Female, 21 years]

Bodily Integrity

Participants have received training that enables some degree of autonomy in navigating the city using a white cane. Nonetheless, the freedom to move around safely is frequently compromised by risks of physical injury and insecurity.

Technology plays an essential role in facilitating mobility for participants, with apps aiding navigation and transportation. However, while most participants own smartphones, many do not have consistent access to mobile data to ensure these apps function effectively.

Some participants limit navigation to familiar areas or rely on human assistance due to safety concerns and technological limitations. Experiences of mugging and other safety risks further hinder their mobility, independence, and willingness to use technology in public.

> *My phone was stolen at the station while I used a transportation app to check my bus arrival time for work. It was early, and I thought I was alone since I heard no one nearby.* [Male, 40 years]

Others are not unaware of the risks in public spaces—they just have a lower perception of the danger, or they choose to manage the risk rather than give up the technology altogether. Instead of avoiding it, they take precautions to reduce the chance of getting robbed.

> *I'm one of those people who says that blind folks don't get robbed — and yeah, people laugh when I say that, hahaha. Sure, I know plenty of blind people who've been robbed — but that's because they gave papaya (in Colombia, that means making yourself an easy target). You just can't give papaya! I've been in dangerous situations before, but thank God, I've always gotten out safely. I've never been robbed.* [Male, 39 years]

During the focus group, when participants discussed the disadvantages or risks of using technology, the first issue that came up was personal safety. As illustrated in the following quote:

> *The first thing that comes to mind is safety. For example, around here everyone talks about Lazarillo and all those apps—but I don't use any of them, because I don't take my phone out. We become very visible when we pull out our phones, hold them to our ear, or start using them. That makes me an easy target for someone to steal it. Given the context we live in, it just doesn't work for me—it actually makes me feel even more unsafe.* [Male, 37 years]

After sharing stories about people they knew who had been victims of phone theft, participants went on to describe the strategies they use to avoid becoming targets themselves. For example, the 2 female participants said they use headphones so they do not have to take their phones out of their pockets, and they hide the headphones with their hair.

Safe mobility is essential for achieving various functionings, but its impact extends beyond simply moving freely and avoiding aggression. One participant noted that communication is a major challenge for those unable to move independently. He added that autonomous mobility is crucial for meaningful interactions, including forming personal relationships, such as finding a partner.

> *Not walking alone makes others, even friends, perceive a person with a disability differently. They already have questions about your disability, which you must address when interacting with them. However, if they see you always relying on someone, it becomes much harder to connect with them.* [Male, 24 years]

Several interviewees noted that being a woman heightens their fear during independent travel, whether on public transport or walking in the streets. One participant shared,

> *Regardless of disability, ethnicity, or economic status, being a woman increases the risk of aggression, especially sexual violence. Being a woman in Cali, and in the areas I frequent, is terrifying. People often mock me with idiotic comments. Just the other day, someone said, 'Oh, you're so pretty, but what a pity about that stick (the cane).* [Female, 21 years]

Senses, Imagination, and Thought

With the exception of 2, all participants completed high school, and about half (n=10) pursued higher education. AT, such as screen readers and text-to-speech apps, are widely used. However, those who studied before advanced technologies faced significant barriers, such as manually converting documents to audio or advocating for accommodations. Despite these challenges, most described their education positively but noted limited career options due to inaccessible educational and professional pathways. As a 19-year-old participant explained,

> *I chose law by default. I picked it because I'm good at public speaking, but do I truly enjoy it? No. Our options are limited, even if they seem abundant. While barriers can be overcome, I'm not willing to exhaust myself doing so.* [Participant, 19 years]

Practical Reason

Participants' decision-making reflects strong priorities and resource management, aided by technology for accessing information and connecting with advocacy groups. However, uncertainty and anxiety about the future, especially among those

XSL•FO
RenderX

newly adjusting to disability, hinder long-term planning. As one 19-year-old male shared:

> *I don’t think about the future. It doesn’t exist—it’s just conjecture, you know? I’ve learned not to stress over it. For me, the future would mean a stable job, continued training, and… that’s it.* [Male, 19 years]

While technology plays a vital role in accessing information and social networks, accessibility issues remain a significant challenge.

Affiliation

People mostly connect with family members and others with visual impairments, but they find it hard to build relationships beyond those close groups. Strong support networks often form in educational or social contexts, such as public libraries, and through cultural or adapted sports activities. However, barriers persist, including discrimination, ableist remarks, and inaccessible public infrastructure, which often prioritizes property over safety.

> *Do you think people respect the few podotactile guides that exist? No, they park their cars on them. If you get too close, they say, "You're going to scratch my car," forcing you onto the road and into danger.* [Female, 54 years]

Participants also face employment challenges, frequently working in roles they are overqualified for or in part-time positions with insufficient pay. Visual impairment complicates social interactions, as one participant noted:

> *Sometimes relationships begin with a look, but without that, it’s already one point less, isn't it? And then there’s the hesitation people feel toward disabilities.* [Male, 45 years]

Accessing services adds to these difficulties. City buses equipped with loudspeakers frequently have these systems turned off, while shopping malls present physical accessibility barriers; additionally, salespeople often fail to recognize persons with disabilities as customers.

> *I’ve never gone shopping alone. I have no way to orient myself in a mall, and I don’t like arriving somewhere unfamiliar and having to wait for someone to help me. Salespeople usually don’t realize that a blind person with a cane can be a customer.* [Female, 31 years]

While technology, such as social networks and apps, has improved accessibility, significant issues remain, such as inaccessible ATMs, banking apps, and dating platforms such as Tinder (Match Group, Inc), where user photos lack proper descriptions.

Other Species

Participants express satisfaction with their ability to engage with nature through their other senses. Despite some limitations, they continue to find joy and fulfillment in their sensory experiences and interactions with the natural world.

Play

Available opportunities include participation in adapted sports and recreational activities supported by local initiatives. Participants enjoy activities such as listening to music, radio programs, and audio-described films. However, financial limitations restrict their access to a broader array of recreational options.

Control Over One’s Environment

Political involvement is marked by participation in advocacy activities and voting, though challenges such as discriminatory treatment and the need to navigate physical barriers persist.

Material control is constrained by the scarcity of full-time employment and inadequate financial stability. Issues with banking and consumer services, compounded by inaccessibility and discrimination, affect their ability to manage material resources effectively.

Themes Identified Through the Inductive Analysis

The inductive analysis of interviews and focus group discussions revealed 3 interrelated themes that reflect participants’ subjective experiences and priorities in everyday life. These themes include the aspiration to express and expand their movement potential, the importance placed on maintaining a positive and attractive personal appearance, and the strategies used to avoid confrontation and maintain social harmony.

Desire to Exploit Movement Potential

A prominent theme in the interviews was the frustration many participants felt regarding their limited mobility. They expressed this frustration over their inability to fully experience freedom of movement, navigate the city and its spaces because of infrastructural barriers, and socialize due to a lack of bodily experience. Additionally, they cited deficiencies in ADs as a contributing factor.

Participants frequently noted that the cane, while essential, places limitations on their freedom of movement and enforces a rigid posture, making their gait different from others.

> *The cane forces you to maintain a fixed posture, which limits your ability to move naturally. Unlike others who can easily turn and walk freely, we can't move the same way.* [Male, 47 years]

Many expressed a strong desire to engage in motor activities that are currently restricted for them, such as riding a bicycle, running, swimming, dancing, and skating. There was a clear enthusiasm for speed.

> *I long to run, but to run alone. I yearn to walk freely, without a cane, moving as I please. I would love to ride a bicycle.* [Male, 47 years]

In the realm of dance, beyond the desire for self-expression through movement and the joy of dancing itself, this activity serves as a crucial means of socialization—especially in this city, renowned as the salsa capital.

> *Well, I like to dance but I don't know how to do it well. I love dancing. ...I do it my own way, in a clumsy*

XSL•FO
RenderX

> *way, but I do it. ... Yet, I would love to master dancing.* [Female, 56 years]

In terms of technology, while some participants acknowledged the crucial role of the cane, they expressed a need for an AD that would enhance their mobility with greater security and fluidity, without compromising their sense of independence or human qualities. Additionally, they emphasized the importance of keeping at least one hand free, as the other one is currently occupied by the cane.

Although many express hope and enthusiasm for the potential of AT in enhancing mobility, others caution that the technology may not address the factors that limit their ability to navigate the city, particularly those related to human interaction. As one participant shared in the following quote:

> *Many variables are involved. The recklessness of drivers is one of them. I've heard that in other countries, people really respect traffic signs and regulations. If that were the case here, one could cross the street when the light is red, because you could trust that drivers would obey the rules. But that’s not the case here. Traffic lights are often ignored. Motorcycles also present a significant issue. Even when there’s a pedestrian crossing, not all vehicles line up properly. If they did, I would know the boundaries of the space I need to cross. Unfortunately, that doesn't happen. When I'm crossing at a traffic light, I may plan to pass on one side, but I find one car ahead of another, and then a motorcycle suddenly appears.* [Male, 40 years]

Desire to Look Attractive

The theme of looking attractive surfaced notably, particularly among female participants. One participant (female, 29 years) expressed frustration with online shopping due to insufficient garment descriptions. She also desired AD that would enable her to perform personal grooming tasks independently, such as painting her nails and plucking her eyebrows, activities she used to do on her own before acquiring a visual disability.

Another participant (female, 27 years) mentioned that if she had the financial means, she would purchase more clothes and enjoy activities such as hair dyeing and makeovers. Regarding mobility AT, she expressed a desire for a discreet or stylish device that could detect high obstacles and match her clothing. She confessed, somewhat apologetically, “I´m somewhat vain, despite my disability.”

A 21-year-old female participant noted her reluctance to use the cane when she had the option of human assistance, explaining that the cane did not complement her style.

> *I think it’s more like this: it’s not as if I say, 'Oh, that cane matches my style, how exciting, man!' No, it’s not like people come up and say, 'What a cool cane, it goes great with your shirt.* [Female, 21 years]

The desire to appear attractive also extends to social and romantic interactions. Participants revealed that concerns about attractiveness influenced their willingness to use certain AT. For example, one participant (male, 40 years) described how societal norms and personal shame affected his acceptance of using a cane.

> *Even though I knew how to get around with my cane, I always relied on my grandmother, who met me at the bus stop every day. One morning, I couldn’t let her know I’d be early, so I ended up waiting three hours instead of using my cane. That experience made me realize I needed to start using it. Honestly, the shame I felt was mostly about what the opposite sex might think of me. I guess those feelings are pretty common for young people, shaped by societal pressures and insecurities.* [Male, 40 years]

A female participant (57 years old) noted that her daughter’s initial embarrassment about using a cane was linked to teenage concerns about appearance:

> *She felt embarrassed to use the cane, especially because, as girls start to become aware of their attractiveness to boys, they may feel self-conscious. However, she eventually came to accept it.* [Female, 57 years]

Another female participant noted that her embarrassment about using the cane is heightened when she is around peers of her age and the opposite sex who do not have disabilities. She mentioned that she has rarely discussed this issue, as she recently recognized her feelings of shame and struggled to openly acknowledge them.

One male participant (45 years old) described his experience after trying a prototype of an AD that included glasses equipped with an obstacle-detection camera, expressing strong reservations.

> *I wouldn't even go out in public wearing those. First, they pose a danger—I could get robbed. Second, they look terrible. If you wore them to a party, do you really think anyone would find you attractive?* [Male, 45 years]

The Bid to Avoid Confrontation

The theme of avoiding confrontation appears in various contexts throughout the study. For instance, some participants avoid expressing their political views to prevent arguments or physical confrontation.

One participant noted that while he is open about his political convictions, he often faces derogatory remarks about his disability when others disagree with his views.

The desire to avoid confrontation also emerges when participants describe their experiences with ableism. Many report feeling frustrated by ableist treatment but choose to remain silent to avoid conflict. For example, a participant shared an experience:

> *They see you with a cane and automatically lock you in the world of the blind, of the <what a pity>, of <how do you manage to walk by yourself?>. Also, they think they have the right to ask questions and make silly expressions. One day I was with my cane, I sat in the Bus and a woman came up behind my ear*

XSL•FO RenderX

> *and said to me: <You have to pray a lot so that the spirit of blindness leaves your life>* [Female, 29 years]

Such incidents often deter individuals from using ADs, fearing further stigma.

Another participant recounted an incident where he felt threatened after unintentionally pushing another man while stumbling in a public bathroom. Since he was in a familiar location, he was not using his cane. To prevent a confrontation, he apologized and explained his condition.

This theme underscores the dual role of the cane. While it can signal vulnerability and deter aggression, it can also trigger stigmatization. Participants sometimes avoid using the cane to prevent negative reactions, which may impact their overall functionality and social participation.

Discussion

Principal Findings

The discussion begins by highlighting the centrality of autonomy in participants' conceptions of WB and its implications for AT design ("The Good Life and WB"). It then reorganizes functionings into 3 analytical categories: "Most Valued Functionings," "Insecure Functionings," and—based on their enabling or limiting effects—"Corrosive Disadvantages" and "Declustering Disadvantage." This is followed by a review of participants' experiences with specific technologies in comparison to existing literature ("AT and the Creation of Opportunities and Barriers in Participants' Daily Lives"). The final sections present emergent sociocultural themes ("Themes Identified Through the Inductive Analysis and AT") and integrate all theoretical strands into a situated analysis of agency, risk, and design ("AT Between Opportunity and Constraint").

Main Results

The Good Life and WB

This study aimed to examine how AT creates opportunities and barriers in participants' daily lives and influences their WB. We first explored the concept of an ideal life, which informed our understanding of participants' perceptions of a good life and WB.

This valued way of life includes, but is not limited to, WB [57]. WB, on the other hand, is assessed in terms of the achievement of functionings that people value and incorporate into their conception of a good life [34].

While interviewees widely valued agency, opinions on autonomy varied. A few participants preferred autonomy over independence, noting that dependence is a natural part of life. However, others viewed autonomy as minimizing reliance on others for daily activities.

This distinction is especially important when considering how AT is developed for persons with disabilities. While the desire for independence is a valid and often central driver in AT design, it can sometimes overshadow other dimensions of daily life. Bennett et al [58] argue for reframing AT through a lens of interdependence, emphasizing that technologies are embedded in relationships, routines, and environments—not isolated solutions. Similarly, Joskow et al [59], using the CA, show that the value of AT lies not only in enabling functional tasks but also in supporting participation, emotional WB, and meaningful choice. These perspectives highlight the need to design AT with a more holistic understanding of how persons with disabilities navigate and shape their worlds.

This becomes particularly evident when comparing the experiences of different users. One male participant, reflecting on the challenge of crossing at traffic lights, shared, "An AT device wouldn't work as well for that kind of situation, because it's still dangerous." Even with a guide dog, he noted, it's hard to manage the risks posed by reckless drivers. He also described the value of informal social interactions that occur in public space, "Sometimes I walk a bit on my own and someone will say, 'Hey, where are you headed?'...I've even met people that way—people who later became friends."

In contrast, a female participant expressed a clear preference for technological assistance over human help, "I would much rather have a device than a person helping me, because I'd feel like I'm working together with the device—and that makes me feel more autonomous...it's perfect!" She added, "Regardless of its size or how it looks, having a tool that helps us move around independently is incredibly important." Even if the device is "big, ugly, or heavy," she saw it as empowering—what mattered most was the sense of self-reliance it enabled.

Both perspectives offer valuable insights for AT design. Listening to this diversity of lived experience helps avoid defaulting to a one-size-fits-all model centered solely on individual autonomy. While preferences like those of the second participant—who expressed enthusiasm for any device that enhances independence regardless of size or appearance—may seem appealing to designers because they reduce complexity, relying solely on such responses can be misleading. If designers prioritize these simplified preferences without accounting for broader user expectations—such as comfort, discretion, aesthetics, or social dynamics—they risk developing yet another device that ultimately fails to meet the needs of a wider group of potential users. Truly inclusive design must engage with this heterogeneity, acknowledging not only aspirations for autonomy but also the embodied, emotional, and relational dimensions of navigating everyday life.

Most Valued Functionings

Among the functionings most valued by participants for a good life, material security—primarily through work as a source of income—stood out. Other key functions included affiliation, education, sports, health, and mobility. This aligns with Wolff and De-Shalit findings, which identify core categories of functioning, including life, bodily health, bodily integrity, affiliation, environmental control, and senses, imagination, and thought [21].

Insecure Valued Functionings

Although these functionings are highly valued by participants, they remain insecure for a significant proportion of them. Employment challenges are prevalent, reflecting findings from other studies that report low employability, poor working

XSL•FO
RenderX

conditions, lower wages, and job instability for people who are blind or have low vision [60-63].

Similarly, while participants emphasize the importance of family and social networks, they also highlight barriers to full inclusion and social participation, making affiliation another insecure functioning. These challenges are consistent with existing literature, which associates limited social engagement with reduced WB and identifies significant social stigma and restricted access to goods and services for persons with visual disabilities, in both public and private spaces [62,64-67].

Given the violence, physical barriers in urban areas, and societal attitudes, life and bodily integrity also emerge as insecure functionings. Empirical studies support these findings, citing falls, life-threatening injuries, and higher mortality risks among persons with visual disabilities [68,69].

Education, including higher education and vocational training, is generally seen as a secure functioning. Despite Colombia's high illiteracy rates among persons with disabilities, participants in this study appear to have attained higher levels of education than the national average for persons with disabilities [70,71].

Finally, although all participants have health insurance, some have faced precarious situations that question health's security as a functioning, particularly given the material insecurity caused by unemployment.

Corrosive Disadvantages

Overview

Efforts for social justice should prioritize groups experiencing multiple disadvantages. Participants highlighted unemployment, lack of affiliation, and mobility barriers as significant corrosive disadvantages.

Unemployment or underemployment threatens basic functionings such as housing, food, health, and life, while also limiting leisure, education, and recreation due to financial constraints. Studies show that income loss exacerbates social isolation, diminishes status, and is linked to reduced WB, limited access to healthcare and education, declines in health, and increased mortality risk [72-75].

Insecurity in affiliation similarly undermines other functionings. Participants often face disrespect and unequal treatment, which compromise job security, environmental control, and freedom of movement without risking safety and WB. Research links social exclusion with limited participation in economic, social, and cultural life, resulting in poverty, unemployment, inadequate health care, and poor housing for persons with disabilities [67,76-78].

Restricted mobility constitutes another corrosive disadvantage, as it directly undermines functionings such as access to essential services, physical activity, communication, WB, and social inclusion. Research highlights the importance of sensorimotor abilities in social interaction [79,80], with decreased mobility linked to isolation and lower satisfaction [81].

Declustering Disadvantage

Identifying corrosive disadvantages highlights the importance of fostering fertile functionings—those that support resilience and WB. Participation in adaptive sports is a clear instance, as research shows that each additional year in adaptive sport is associated with a 4% increase in employment likelihood among people with physical disabilities [82]. In our study, involvement in sports provided income, travel subsidies, and social connections. It supported capabilities such as affiliation, bodily integrity, and control over one's environment.

Employment also served as a fertile functioning when conditions were fair, remuneration equitable, and respect for dignity was present. Studies indicate that AT access in the workplace is strongly linked to employment retention and productivity among individuals with visual disabilities [83].

As pivotal conversion factors, AT tools such as screen readers and accessible mobile apps enabled participants to use digital platforms daily—facilitating email, social media, and work tasks [84]. As a result, these technologies not only support function, but also reinforce confidence and autonomy.

Although participants used diverse ATs, this paper focuses on representative examples. Further research should explore a broader array of technologies and how they foster capability and reduce disadvantage across different daily contexts.

Mobility technologies warrant particular attention. Studies show that interactive navigation tools can significantly enhance users' spatial confidence and independence [85]. Participants emphasized that AT, which enables seamless environmental interaction—not just orientation or obstacle avoidance—enhances their movement potential, autonomy, self-efficacy, and engagement with public space.

AT and the Creation of Opportunities and Barriers in Participants' Daily Lives

AT has opened multiple avenues for expanding capabilities among persons with visual disabilities, supporting valued functionings across educational, occupational, and social domains. Participants in this study described how information and communication technologies have enhanced their performance in academic tasks, enabled participation in the labor market, and expanded opportunities for leisure and connection. These tools facilitated informed decision-making, strengthened social networks, and increased personal visibility in both professional and interpersonal contexts [86].

Mobility technologies, including white canes and mobile navigation apps, were also perceived as crucial for promoting spatial autonomy. They allowed individuals to traverse their environments more independently, access public and private spaces, and engage in civic and social life. AT also enabled participants to perform essential tasks such as banking and online shopping, thereby enhancing their control over daily routines and improving their sense of independence.

However, participants also identified significant barriers to AT use—barriers that highlight the ambivalence of technology in their daily lives. As Moser [87] argues, technology does not simply empower—it also has the capacity to withdraw agency

XSL•FO
RenderX

by rendering difference hypervisible or by reinforcing stigma. This tension was evident in the accounts of participants who, while recognizing the utility of AT, expressed concern about its potential to attract unwanted attention, particularly in contexts marked by high levels of urban violence. In these settings, visibility becomes a liability, increasing the risk of theft or social targeting.

This ambivalence was especially pronounced in public spaces. While some participants preferred discreet technologies that do not mark them as disabled, others acknowledged that visible devices, such as white canes, could help signal their needs and elicit assistance. These findings resonate with Dos Santos et al, who found that aesthetic and symbolic dimensions strongly influence the acceptability and use of AT [88]. Similar insights have emerged in studies from the Global South, such as Barbareschi et al [89] work in Kenya, where ADs can either mitigate or exacerbate stigma depending on the sociocultural context.

Thus, the potential of AT to support WB is conditioned not only by its functional properties but by its alignment with users' social environments and identities. When well-designed and contextually appropriate, AT can foster fertile functionings—such as autonomy, affiliation, and environmental control. However, when AT is experienced as stigmatizing, risky, or misaligned with local realities, it can become a source of discomfort or even exclusion. This dynamic reinforces the importance of attending to not just technical efficiency, but also social safety, emotional meaning, and contextual sensitivity in design and policy.

Themes Identified Through the Inductive Analysis and AT

In Latin America, physical appearance holds significant value, rooted in colonial legacies that entrenched hierarchies by emphasizing differences between colonizers and the colonized. These hierarchies associated power with traits such as being European, patriarchal, White, wealthy, heterosexual, and able-bodied. Disability, conversely, was linked to negative characteristics such as unproductivity and rebellion, devaluing disabled bodies and dissociating them from beauty [90,91]. This detachment was reflected in interviews, where participants either avoided discussing physical appearance or addressed it obliquely. One participant remarked, "I am somewhat vain despite my disability."

The desire for attractiveness was particularly pronounced among women, aligning with research that connects beauty ideals to women's roles in Latin American cultures [92,93]. In Colombia and Brazil, beauty is often tied to socioeconomic status and professional opportunities [94,95]. While beauty remains culturally valued, ADs such as canes—symbolizing disability—are perceived as aesthetically unappealing and are sometimes rejected.

Aesthetics influence social interactions beyond economic implications. Participants expressed a desire to appear attractive to the opposite sex but seldom mentioned romantic aspirations, potentially reflecting aspirational frustrations. This tension between appearance and relational expectations emerged more clearly in the focus group discussions, which revealed the need for greater accessibility on dating platforms such as Tinder, highlighting participants' openness to using AT to facilitate social interactions and romantic relationships.

Limited mobility emerged as a recurring challenge, restricting participants' navigation skills and engagement with urban environments. Kukla's concept of spatial agency—the ability to autonomously use space—captures this dynamic. City navigation requires interaction with people and environments, relying on cues such as body language, eye contact, and street signs [96]. Existing mobility AT primarily addresses orientation and obstacle detection but lacks tools for effective environmental engagement. Participants' adaptive strategies highlight the need for improved AT design. Kukla also emphasizes that movement pace and aesthetic norms influence spatial navigation [96]. Participants underscored that AT should enhance autonomy while enabling natural interaction with tasks and people, without disrupting rhythms.

AT design must prioritize the user's holistic experience, accounting for the diverse environments they navigate [97]. This context extends beyond controlled settings, incorporating cultural meanings of disability, prior interactions, and broader social, historical, and political factors. The social realities of designers in the Global North often differ significantly from those of persons with disabilities or users in the Global South, as do their perceptions of aesthetics and its functional role in daily life.

AT Between Opportunity and Constraint: A Situated Analysis of Agency, Risk, and Design

In Colombia, defying dominant social norms often entails serious consequences, frequently enforced through violence. This violence is not limited to state institutions; nonstate actors—including armed groups and criminal organizations—have also enforced normative boundaries. Individuals who challenge socially constructed expectations—such as LGBTIQ+ (lesbian, gay, bisexual, transgender, intersex, and queer) persons, Indigenous peoples, Afro-Colombians, and social leaders—are disproportionately subjected to threats, forced displacement, and even assassination [98,99]. In cities such as Cali, such patterns have at times been reinforced by criminal governance structures that monitored daily life and punished perceived transgressions, exerting coercive control over entire neighborhoods [100]. In these contexts, expressing a nonnormative identity or dissent becomes a high-risk act, revealing how violence operates as a mechanism of civic regulation.

This coercive environment encourages conformity not through democratic consensus but through fear. Individuals internalize the threat of retaliation, adapting their appearance, behavior, and speech to avoid becoming targets [98]. Over time, this normalizes silence and compliance, particularly in communities with limited institutional protection. Violence, therefore, becomes a structuring force that shapes notions of "acceptable" citizenship and sustains systems of exclusion.

Within this broader context, ableism functions as a parallel yet distinct form of disciplinary control. In Colombia, persons with disabilities who defy normative expectations—by asserting

XSL•FO
RenderX

rights, demanding visibility, or rejecting passive roles—often face institutional neglect, attitudinal bias, and spatial exclusion. The Constitutional Court has described this as the product of "a segregating and exclusionary culture" that restricts life trajectories for persons with disabilities [101]. In conflict-affected areas, civic engagement can become a risk in itself, as some persons with disabilities avoid visibility out of fear of losing access to essential services [102]. Ableism reinforces conformity not through overt coercion, but through anticipated punishment—governing behavior, silencing dissent, and normalizing invisibility [103].

From the perspective of the CA, these exclusionary and coercive dynamics represent severe constraints on individuals' real freedoms to pursue lives they have reason to value. Structural violence, institutional neglect, and everyday surveillance undermine central capabilities such as bodily integrity, affiliation, and control over one's environment [18]. These forces not only restrict access to resources but also impair agency itself, fostering anticipatory compliance and silencing non-normative expression.

While the CA emphasizes agency and meaningful choice, it has been critiqued for undertheorizing the sociopolitical and ideological conditions that suppress agency—particularly in contexts marked by systemic domination and fear [20,104]. Although Sen acknowledges power structures through the notion of conversion factors and has addressed issues such as identity-based violence and gender inequality, he provides only a partial theorization of how social and institutional forces constrain choice [20]. In response, critical theorists emphasize that agency is always situated—shaped by hegemonic ideologies, historical inequalities, and structural constraints [20]. These insights enrich the CA by highlighting how oppression is not only material but also discursive and normative.

AT becomes a key analytical entry point in this discussion. Within the CA, AT functions as both a resource and a conversion factor—potentially expanding real freedoms by enabling valued functionings such as mobility, communication, and social participation [9,18]. Yet in contexts shaped by criminal governance, stigma, or weak institutional presence, AT's capacity to enhance capabilities may be compromised. Fear of surveillance, ridicule, or exclusion can deter individuals from using AT, effectively transforming it into a corrosive disadvantage—a concept from DT that refers to resources that worsen users' conditions in hostile environments [21].

In cities such as Cali, where high levels of urban violence intersect with limited protection for vulnerable groups, participants reported avoiding certain technologies out of fear for their personal safety. While navigation apps and mobile-based tools might enhance autonomy in principle, in practice, they often increase visibility and perceived vulnerability in public space. As shown in the findings, such risks lead many to forgo potentially beneficial technologies, not because of a lack of interest or understanding, but because the act of using them carries a social and physical cost that undermines the very freedoms they are meant to support.

This reflects a broader pattern in which structural violence becomes embedded in everyday decision-making. Technology adoption is shaped not only by access and design, but also by how individuals negotiate exposure, safety, and autonomy in hostile or unstable environments. In this sense, the very tools intended to promote inclusion may reinforce exclusion when they are not aligned with the social conditions and lived experiences of their users.

Moreover, AT design and policy often reflect biomedical or market-driven logics, failing to account for cultural preferences, affective experiences, and social risks. Research shows that when AT is perceived as overly medicalized or aesthetically stigmatizing, it is more likely to be abandoned—despite its functional benefits [105,106]. In Latin American contexts, aesthetic appeal and discretion are not merely preferences but protective strategies in public spaces marked by insecurity. Orellano-Colón et al [105] found that Hispanic users preferred AT devices that did not draw attention, valuing discretion and aesthetics alongside functionality. In Colombia, where individuals must navigate public visibility with caution, design decisions can significantly impact the acceptability and sustainability of AT use.

Emotional responses—such as shame, pride, or resignation—often remain unspoken but can be inferred from narrative silences, reflecting adaptive preferences. Rather than ignore these silences, designers should interpret them as meaningful indicators of unmet needs—shaped by structural disadvantage—and take them into account in participatory design.

Given that the aim of this study was to understand the role of technology in the lives of people with visual disabilities in LMICs, and how AT shapes opportunities, barriers, and WB, the findings underscore both the promise and limits of AT. While AT cannot resolve the structural inequalities that underpin marginalization, it can be designed to mitigate their effects. This requires that designers pay careful attention to users' cultural values, local forms of risk management, and context-specific aspirations. Culturally sensitive design that values emotional WB and aesthetic dignity is more likely to foster meaningful adoption and sustained use.

Applying the DT alongside the CA allows for a more nuanced understanding of how AT impacts people's lives. It invites researchers and designers to consider not only what a technology enables, but what it may jeopardize, particularly when social or environmental conditions are hostile. As the findings show, AT that is context-blind risks becoming ineffective or even harmful. Design processes must therefore attend to fertile functionings—those that support broader forms of agency and resilience—while actively avoiding the reproduction of corrosive disadvantages.

Limitations

Although the study used rigorous qualitative methods and achieved thematic saturation, the sample size (N=19) and composition—skewed toward younger and more educated urban residents—may limit the diversity of perspectives captured. Future research should include larger and more demographically

XSL•FO
RenderX

diverse samples, especially individuals from rural areas, older adults, and persons with visual disabilities with intersecting marginalized identities, to further explore how AT and structural disadvantage interact in varied sociocultural contexts.

While the study aimed to include participants from both urban and rural areas, the final sample was predominantly urban. This may have limited the representation of geographically diverse experiences.

Additionally, the snowball sampling strategy, while effective in reaching a specific population, may have introduced bias by concentrating perspectives within existing social networks.

Certain themes, such as romantic relationships, were underexplored due to study design limitations. However, the focus group offered deeper insights, showing that group discussions encourage openness on sensitive topics.

Even if this study does not cover all 10 of Sen's capabilities, it supports the framework’s relevance in analyzing AT users’ needs. Future research should further explore the specific characteristics of AT and how they align with users’ daily realities.

Conclusions

This study sheds light on how AT is woven into the everyday lives of people with visual disabilities in Colombia—not just as tools, but as companions, barriers, enablers, and sometimes sources of tension or silence. In a context marked by inequality, insecurity, and deep social stigma, AT does not operate in a vacuum. Whether it supports or hinders WB depends on how it fits into people’s lived realities, relationships, environments, and desires.

By combining the CA with DT and a critical lens, we were able to see not only the functionings participants valued—such as education, work, and independence—but also the emotional and social costs of trying to pursue them. Technologies that might promise autonomy in one context could increase vulnerability in another. A navigation app or white cane, for instance, might enhance freedom, but also attract unwanted attention in unsafe public spaces. For many, using AT meant balancing aspirations with risk.

Participants told stories of resilience, but also frustration. They spoke of wanting to be seen as attractive, of yearning for movement without fear, of choosing silence over confrontation to stay safe. Others left things unsaid—particularly around relationships and intimacy—pointing to desires that are often overlooked, even by well-intentioned designers and researchers.

These findings remind us that the voices of people with disabilities in the Global South need to be at the center of technology design—not just in consultations or interviews, but as full collaborators in research and development. The choices they make, the things they value, and even what they choose not to say are powerful forms of knowledge. Listening deeply to these experiences helps us understand what truly makes a technology meaningful.

To create technologies that genuinely support WB, we must look beyond technical fixes. We must ask, does this technology help someone live the life they have reason to value? Does it support their dignity, their safety, their hopes? These questions cannot be answered by engineers or policymakers alone. They require interdisciplinary collaboration—between designers, social scientists, humanists, and most importantly, people with disabilities themselves.

Ultimately, this study shows that AT holds enormous promise—but only if it is rooted in empathy, shaped by real-world complexity, and open to the diverse ways in which people navigate, resist, and imagine their worlds.

Funding

MR was supported by the Ambizione grant (PZOOP1_186035), funded by the Swiss National Science Foundation and by the Latin American Leading House of the University of St Gallen. LMO-E and MAC were supported by the Fondation Gelbert (0751-2020).

Conflicts of Interest

None declared.

References

1. Blindness and vision impairment. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment [accessed 2026-03-31]
2. World report on vision. : World Health Organization; 2019 URL: https://www.who.int/publications/i/item/world-report-on-vision [accessed 2026-03-31]
3. Rizzo JR, Beheshti M, Hudson TE, et al. The global crisis of visual impairment: an emerging global health priority requiring urgent action. Disabil Rehabil Assist Technol 2023 Apr;18(3):240-245. [doi: 10.1080/17483107.2020.1854876] [Medline: 33332166]
4. Brunes A, B Hansen M, Heir T. Loneliness among adults with visual impairment: prevalence, associated factors, and relationship to life satisfaction. Health Qual Life Outcomes 2019 Feb 1;17(1):24. [doi: 10.1186/s12955-019-1096-y] [Medline: 30709406]

XSL•FO
RenderX

5. Smith EM, Huff S, Wescott H, et al. Assistive technologies are central to the realization of the Convention on the Rights of Persons with Disabilities. Disabil Rehabil Assist Technol 2024 Feb;19(2):486-491. [doi: 10.1080/17483107.2022.2099987] [Medline: 35900971]
6. Disability and development report: realizing the sustainable development goals by, for and with persons with disabilities. : Department of Economic and Social Affairs, United Nations; 2019 URL: https://www.un.org/development/desa/disabilities/wp-content/uploads/sites/15/2019/10/UN-flagship-report-on-disability-and-development.pdf [accessed 2026-03-31]
7. Tebbutt E, Brodmann R, Borg J, MacLachlan M, Khasnabis C, Horvath R. Assistive products and the sustainable development goals (SDGs). Global Health 2016 Nov 29;12(1):79. [doi: 10.1186/s12992-016-0220-6] [Medline: 27899117]
8. Jang HY, Chung D, Oh E, Son Hong GR. Experiences of individuals with severe disabilities using assistive devices: a qualitative study. Disabil Health J 2025 Jul;18(3):101833. [doi: 10.1016/j.dhjo.2025.101833] [Medline: 40194902]
9. Global report on assistive technology. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240049451 [accessed 2026-03-31]
10. Barbet I, Hartmann L, Deville D, Ferreira MS. Design of an assessment tool for implementing assistive technology (AT) reuse programs in France. Resour Conserv Recycl Adv 2022 Nov;15:200094. [doi: 10.1016/j.rcradv.2022.200094]
11. Grott R. On technology abandonment or discontinuance. In: Assistive Technology Service Delivery: Academic Press; 2019:211-216. [doi: 10.1016/B978-0-12-812979-1.00015-1]
12. Sugawara AT, Ramos VD, Alfieri FM, Battistella LR. Abandonment of assistive products: assessing abandonment levels and factors that impact on it. Disabil Rehabil Assist Technol 2018 Oct;13(7):716-723. [doi: 10.1080/17483107.2018.1425748] [Medline: 29334475]
13. Menich N. Each person as an end? The users' choices in the service delivery process for assistive technology in Hungary. Societies (Basel) 2022;12(5):130. [doi: 10.3390/soc12050130]
14. Real S, Araujo A. Navigation systems for the blind and visually impaired: past work, challenges, and open problems. Sensors (Basel) 2019 Aug 2;19(15):3404. [doi: 10.3390/s19153404] [Medline: 31382536]
15. Blackburn SJ, Cudd PA. A discussion of systematic user requirements gathering from a population who require assistive technology. Technol Disabil 2012;24(3):193-204. [doi: 10.3233/TAD-2012-0352]
16. Ortiz-Escobar LM, Chavarria MA, Schönenberger K, et al. Assessing the implementation of user-centred design standards on assistive technology for persons with visual impairments: a systematic review. Front Rehabil Sci 2023;4:1238158. [doi: 10.3389/fresc.2023.1238158] [Medline: 37744430]
17. Sen A. Development as Freedom: Alfred A. Knopf; 1999. URL: https://raggeduniversity.co.uk/wp-content/uploads/2025/01/1_x_senDevelopmentasFreedom-_compressed.pdf [accessed 2026-03-31]
18. Nussbaum MC. Creating Capabilities: The Human Development Approach: Harvard; 2011. [doi: 10.4159/harvard.9780674061200]
19. Drèze J, Sen A. India: Development and Participation: Oxford University Press; 2002. URL: https://archive.org/details/indiadevelopment0000drez_u8n8/page/n7/mode/2up [accessed 2024-03-24]
20. Robeyns I. The capability approach: a theoretical survey. J Hum Dev 2005 Mar;6(1):93-117. [doi: 10.1080/146498805200034266]
21. Wolff J, De-Shalit A. Disadvantage: Oxford University Press; 2007.
22. Sen A. Inequality Reexamined: Oxford University Press; 1995. URL: https://www.jstor.org/stable/2940919 [accessed 2022-12-10]
23. Terzi L. Beyond the dilemma of difference: the capability approach to disability and special educational needs. J Philos Educ 2005 Aug;39(3):443-459 [FREE Full text] [doi: 10.1111/j.1467-9752.2005.00447.x]
24. Mitra S. The capability approach and disability. J Disabil Policy Stud 2006 Mar;16(4):236-247. [doi: 10.1177/10442073060160040501]
25. Trani JF, Bakhshi P, Bellanca N, Biggeri M, Marchetta F. Disabilities through the capability approach lens: implications for public policies. ALTER 2011;5(3):143-157. [doi: 10.1016/j.alter.2011.04.001]
26. Mitra S. Disability, Health and Human Development: Palgrave Pivot New York; 2018. URL: http://link.springer.com/10.1057/978-1-137-53638-9 [accessed 2026-03-05] [doi: 10.1057/978-1-137-53638-9]
27. Haenssgen MJ, Ariana P. The place of technology in the capability approach. Oxf Dev Stud 2018 Jan 2;46(1):98-112. [doi: 10.1080/13600818.2017.1325456]
28. Ryan M. The ethics of dietary apps: technology, health, and the capability approach. Technol Soc 2022 Feb;68:101873. [doi: 10.1016/j.techsoc.2022.101873]
29. Nussbaum MC. Women and Human Development: The Capabilities Approach: Cambridge University Press; 2000. [doi: 10.1017/CBO9780511841286]
30. Reed R. Dignity in transgender lives: a capabilities approach. J Human Dev Capabil 2020 Jan 2;21(1):36-48. [doi: 10.1080/19452829.2019.1661982]
31. Kuper H, Davey C, Banks LM, Shakespeare T. Trials and Tribulations of Collecting Evidence on Effectiveness in Disability-Inclusive Development: A Narrative Review. Sustainability 2018;12(18):7823 [FREE Full text] [doi: 10.3390/su12187823]

XSL•FO
RenderX

32. Shakespeare T. Nasty brutish and short? on the predicament of disability and embodiment. In: Bickenbach J, Felder F, Schmitz B, editors. Disability and the Good Human Life: Cambridge University Press; 2013:93-112. [doi: 10.1017/CBO9781139225632.005]
33. Riddle CA. Disability and disadvantage in the capabilities approach. In: Oxford Handbook of Philosophy and Disability: Oxford University Press:2020. [doi: 10.1093/oxfordhb/9780190622879.013.19]
34. Sen A. Capability and well-being. In: The Quality of Life: Oxford University Press; 1993, Vol. 30:270-293. [doi: 10.1093/0198287976.003.0003]
35. Das KV, Jones-Harrell C, Fan Y, Ramaswami A, Orlove B, Botchwey N. Understanding subjective well-being: perspectives from psychology and public health. Public Health Rev 2020 Nov 19;41(1):25. [doi: 10.1186/s40985-020-00142-5] [Medline: 33292677]
36. OECD. Health at a Glance 2013: OECD Indicators: OECD Publishing; 2013. [doi: 10.1787/health_glance-2013-en]
37. Diener E, Heintzelman SJ, Kushlev K, et al. Findings all psychologists should know from the new science on subjective well-being. Can Psychol 2017;58(2):87-104. [doi: 10.1037/cap0000063]
38. Lucas RE. Long-term disability is associated with lasting changes in subjective well-being: evidence from two nationally representative longitudinal studies. J Pers Soc Psychol 2007 Apr;92(4):717-730. [doi: 10.1037/0022-3514.92.4.717] [Medline: 17469954]
39. Emerson E, Llewellyn G, Honey A, Kariuki M. Lower well-being of young Australian adults with self-reported disability reflects their poorer living conditions rather than health issues. Aust N Z J Public Health 2012 Apr;36(2):176-182. [doi: 10.1111/j.1753-6405.2011.00810.x] [Medline: 22487354]
40. Office for National Statistics. Understanding well-being inequalities: who has the poorest personal well-being. 2018 URL: https://www.ons.gov.uk/releases/thecharacteristicsandcircumstancesofthosewhoreportlowestpersonalwellbeingratings [accessed 2026-03-31]
41. Emerson E, Fortune N, Aitken Z, Hatton C, Stancliffe R, Llewellyn G. The wellbeing of working-age adults with and without disability in the UK: associations with age, gender, ethnicity, partnership status, educational attainment and employment status. Disabil Health J 2020 Jul;13(3):100889. [doi: 10.1016/j.dhjo.2020.100889] [Medline: 32046927]
42. Emerson E, Llewellyn G. The wellbeing of women and men with and without disabilities: evidence from cross-sectional national surveys in 27 low- and middle-income countries. Qual Life Res 2023 Feb;32(2):357-371. [doi: 10.1007/s11136-022-03268-y] [Medline: 36273048]
43. Gothwal VK, Kanchustambam J, Kodavati K, Subramanian A. Young adults with vision impairment in India: loneliness and social networks. Ophthalmic Physiol Opt 2024 Jul;44(5):808-818. [doi: 10.1111/opo.13317] [Medline: 38619126]
44. Zapata MA. Disability affirmation predicts employment among adults with visual impairment and blindness. Rehabil Couns Bull 2022 Jan;65(2):120-128. [doi: 10.1177/0034355220957107]
45. Halvorsen R, Hvinden B, Brown JB, Biggeri M, Tøssebro J. In: Waldschmidt A, editor. Understanding the Lived Experiences of Persons with Disabilities in Nine Countries: Active Citizenship and Disability in Europe: Routledge; 2017. [doi: 10.4324/9781315623924]
46. Ripat J, Woodgate R. The intersection of culture, disability and assistive technology. Disabil Rehabil Assist Technol 2011;6(2):87-96. [doi: 10.3109/17483107.2010.507859] [Medline: 20698763]
47. Ganem Maloof K. Comisión para el esclarecimiento de la verdad. hay futuro si hay verdad: informe final, convocatoria a la paz grande [there is a future if there is truth: final report. a call for a great peace]. : Comisión para el Esclarecimiento de la Verdad; 2022 URL: https://www.comisiondelaverdad.co/ [accessed 2022-03-22]
48. Building memory. National Center for Historical Memory. 2025 May 20. URL: https://centrodememoriahistorica.gov.co/construccion-de-la-memoria-historica/ [accessed 2026-03-05]
49. Yaffe L. Conflicto armado en Colombia: análisis de las causas económicas, sociales e institucionales de la oposición violenta: [Armed conflict in Colombia: analysis of the economic, social, and institutional causes of violent opposition]. RevCS 2011(8):187-208. [doi: 10.18046/recs.i8.1133]
50. Velarde MR, Lord JE, Stein MA, Shakespeare T. Disarmament, demobilization, and reintegration in Colombia: lost human rights opportunities for ex-combatants with disabilities. J Hum Rights 2022 Jan 1;21(1):18-35. [doi: 10.1080/14754835.2021.1969648]
51. Martínez L, Prada S, Estrada D. Homicides, public goods, and population health in the context of high urban violence rates in Cali, Colombia. J Urban Health 2018 Jun;95(3):391-400. [doi: 10.1007/s11524-017-0215-5] [Medline: 29204844]
52. Fife ST, Gossner JD. Deductive qualitative analysis: evaluating, expanding, and refining theory. Int J Qual Methods 2024 Jan;23:16094069241244856. [doi: 10.1177/16094069241244856]
53. Kincheloe JL, McLaren P. Rethinking critical theory and qualitative research. In: Denzin NK, Lincoln YS, editors. The Landscape of Qualitative Research, 2nd edition: Sage; 2005:303-342. [doi: 10.1007/978-94-6091-397-6_23]
54. The Washington Group Short Set on Functioning (WG-SS). Washington Group on Disability Statistics. 2006. URL: https://www.washingtongroup-disability.com/question-sets/wg-short-set-on-functioning-wg-ss/ [accessed 2022-11-20]
55. Cooke B, Kothari U. Participation: The New Tyranny: Zed Books; 2001. [doi: 10.1016/S0738-0593(02)00022-6]
56. Kothari U. In: Cooke B, Kothari U, editors. Power, Knowledge and Social Control in Participatory Development: Zed Books; 2001:139-152.

XSL•FO
RenderX

57. Restrepo-Ochoa DA. La salud y la vida buena: aportes del enfoque de las capacidades de Amartya Sen para el razonamiento ético en salud pública [Health and the good life: contributions of Amartya Sen's capability approach to ethical reasoning in public health]. Cad Saúde Pública 2013 Dec;29(12):2371-2382. [doi: 10.1590/0102-311X00069913]
58. Bennett CL, Brady E, Branham SM. Interdependence as a frame for assistive technology research and design. Presented at: Proceedings of the 20th ACM SIGACCESS Conference on Computers and Accessibility; Oct 8, 2018. [doi: 10.1145/3234695.3236348]
59. Joskow R, Patel D, Landre A, et al. Understanding the impact of assistive technology on users' lives in England: a capability approach. Bioengineering (Basel) 2025 Jul 9;12(7):750. [doi: 10.3390/bioengineering12070750] [Medline: 40722442]
60. McDonnall MC, Sui Z. Employment and unemployment rates of people who are blind or visually impaired: estimates from multiple sources. J Vis Impair Blind 2019 Nov;113(6):481-492. [doi: 10.1177/0145482X19887620]
61. Llorente-Barroso C, Mañas-Viniegra L, Sierra-Sánchez J, García-García F. Disability and employability in the audio-visual sector: the (dis)connection between corporate social sustainability goals and the employment experiences of people with disabilities. EPI 2023;32(6). [doi: 10.3145/epi.2023.nov.04]
62. Tsatsou P. Is digital inclusion fighting disability stigma? Opportunities, barriers, and recommendations. Disabil Soc 2021 May 28;36(5):702-729. [doi: 10.1080/09687599.2020.1749563]
63. Viñarás-Abad M, Sánchez-Valle M, Vázquez-Barrio T. The situation of people with disabilities in the Communication sector in Spain: labor, professional and academic aspects. EPI 2021;30(2):e300202. [doi: 10.3145/epi.2021.mar.02]
64. Tshuma C, Ntombela N, Mabvurira V. Challenges and coping strategies of visually impaired adults in Zeerust, South Africa. J Soc Sci Humanit 2021;18(5):53-65 [FREE Full text]
65. Lamoureux EL, Hassell JB, Keeffe JE. The determinants of participation in activities of daily living in people with impaired vision. Am J Ophthalmol 2004 Feb;137(2):265-270. [doi: 10.1016/j.ajo.2003.08.003] [Medline: 14962415]
66. Jackson SE, Hackett RA, Pardhan S, Smith L, Steptoe A. Association of perceived discrimination with emotional well-being in older adults with visual impairment. JAMA Ophthalmol 2019 Jul 1;137(7):825-832. [doi: 10.1001/jamaophthalmol.2019.1230] [Medline: 31145413]
67. Tobias EI, Mukhopadhyay S. Disability and social exclusion: experiences of individuals with visual impairments in the Oshikoto and Oshana regions of Namibia. Psychol Dev Soc 2017;29(1):22-43. [doi: 10.1177/0971333616689203]
68. Saur R, Hansen MB, Jansen A, Heir T. Visually impaired individuals, safety perceptions and traumatic events: a qualitative study of hazards, reactions and coping. Disabil Rehabil 2017 Apr;39(7):691-696. [doi: 10.3109/09638288.2016.1161836] [Medline: 27027571]
69. Ehrlich JR, Ramke J, Macleod D, et al. Association between vision impairment and mortality: a systematic review and meta-analysis. Lancet Glob Health 2021 Apr;9(4):e418-e430. [doi: 10.1016/S2214-109X(20)30549-0] [Medline: 33607015]
70. International Disability Alliance. Our right to education: a compilation of evidence gathered by opds on progress towards SDG 4 and CRPD article 24. 2021 URL: https://www.internationaldisabilityalliance.org/sites/default/files/sdg4_crpd24_compilation_april_2021_final.pdf [accessed 2024-10-29]
71. UNESCO Institute for Statistics (UIS). Education and disability: analysis of data from 49 countries. 2018 URL: https://unesdoc.unesco.org/ark:/48223/pf0000262805 [accessed 2025-04-07]
72. Jahoda M. Employment and Unemployment: Cambridge University Press; 1982. URL: https://www.cambridge.org/us/universitypress/subjects/psychology/social-psychology/employment-and-unemployment-social-psychological-analysis?format=PB&isbn=9780521285865 [accessed 2024-02-03]
73. Kuklek NM, Cséplő M, Pozsonyi E, Pusztafalvi H. Raising employment and quality of life among people with disadvantages - results of a Hungarian project. BMC Public Health 2021 Sep 23;21(1):1729. [doi: 10.1186/s12889-021-11763-z] [Medline: 34556054]
74. Korpi T. Accumulating disadvantage. Longitudinal analyses of unemployment and physical health in representative samples of the Swedish population. Eur Sociol Rev 2001 Sep 1;17(3):255-273. [doi: 10.1093/esr/17.3.255]
75. Dănăcică DE. Unemployment of highly educated disabled individuals in Romania. Ann Constantin Brancusi Univ Targu Jiu Econ Ser 2023;1:45-53 [FREE Full text]
76. Bates * P, Davis FA. Social capital, social inclusion and services for people with learning disabilities. Disabil Soc 2004 May;19(3):195-207. [doi: 10.1080/0968759042000204202]
77. Pal GC. Disability, intersectionality and deprivation. Psychol Dev Soc J 2011 Sep;23(2):159-176. [doi: 10.1177/097133361102300202]
78. Grech S, Soldatic K. Disability and colonialism: (dis)encounters and anxious intersectionalities. Soc Ident 2015 Jan 2;21(1):1-5. [doi: 10.1080/13504630.2014.995394]
79. Spaulding S. Introduction to debates on embodied social cognition. Phenom Cogn Sci 2012 Dec;11(4):431-448. [doi: 10.1007/s11097-012-9275-x]
80. Wagner-Ferrari A, Kuschel Ruiz C. Understanding the embodied nature of social cognition through visuospatial perspective-taking. LíMITE (Arica) 2022;17:61-76. [doi: 10.4067/S0718-50652022000100204]
81. Schölvinck AFM, Pittens C, Broerse JEW. The research priorities of people with visual impairments in the Netherlands. J Vis Impair Blind 2017 May;111(3):201-217. [doi: 10.1177/0145482X1711100302]

XSL•FO
RenderX

82. Lastuka A, Cottingham M. The effect of adaptive sports on employment among people with disabilities. Disabil Rehabil 2016 Apr;38(8):742-748. [doi: 10.3109/09638288.2015.1059497] [Medline: 26114627]
83. Akkiraju K, Rao P, Srinivas KV, et al. Assistive technology use and employment outcomes among users in a low-income urban setting in Bangalore, India. ACM Trans Access Comput 2012;5(3):1-19. [doi: 10.1145/2160673.2160711]
84. McDonnall MC, Boydstun J, Steverson A. Is one enough? Screen reader use among employed people who are blind or have low vision in the U.S. Disabil Rehabil Assist Technol 2025 Nov;20(8):2725-2735. [doi: 10.1080/17483107.2025.2528858] [Medline: 40614067]
85. Yao F, Zhou W, Hu H. A review of vision-based assistive systems for visually impaired people: technologies, applications, and future directions. arXiv. Preprint posted online on May 20, 2025 URL: https://arxiv.org/abs/2505.14298 [accessed 2026-03-05]
86. Montenegro-Rueda M, Fernández-Batanero JM, Fernández-Cerero J. Impact of ICT on university students with visual impairment. Br J Spec Educ 2023;50(1):28-48. [doi: 10.1111/1467-8578.12433]
87. Moser I. Disability and the promises of technology: technology, subjectivity and embodiment within an order of the normal. Inf Commun Soc 2006 Jun;9(3):373-395. [doi: 10.1080/13691180600751348]
88. Dos Santos ADP, Ferrari ALM, Medola FO, Sandnes FE. Aesthetics and the perceived stigma of assistive technology for visual impairment. Disabil Rehabil Assist Technol 2022 Feb;17(2):152-158. [doi: 10.1080/17483107.2020.176830810.1080/17483107.2020.1768308] [Medline: 32501732]
89. Barbareschi G, Carew MT, Johnson EA, Kopi N, Holloway C. “When They See a Wheelchair, They’ve Not Even Seen Me”-factors shaping the experience of disability stigma and discrimination in Kenya. Int J Environ Res Public Health 2021 Apr 17;18(8):4272. [doi: 10.3390/ijerph18084272] [Medline: 33920601]
90. Grech S. Decolonising Eurocentric disability studies: why colonialism matters in the disability and global South debate. Soc Ident 2015 Jan 2;21(1):6-21. [doi: 10.1080/13504630.2014.995347]
91. Fanon F. The Wretched of the Earth: Penguin twentieth-century classics; 1963. URL: https://www.penguin.co.uk/books/57385/the-wretched-of-the-earth-by-frantz-fanon-trans-constance-farrington-preface-by-jean-paul-sartre/9780141186542 [accessed 2024-10-02]
92. Gruber E, Kalkbrenner MT, Hitter TL. A complex conceptualization of beauty in Latinx women: a mixed methods study. Body Image 2022 Jun;41:432-442. [doi: 10.1016/j.bodyim.2022.04.008] [Medline: 35533521]
93. Adames HY, Chavez-Dueñas NY. Cultural Foundations and Interventions in Latino/a Mental Health: History, Theory and within Group Differences: Routledge; 2016. [doi: 10.4324/9781315724058]
94. Jarrin A. Towards a biopolitics of beauty: eugenics, aesthetic hierarchies and plastic surgery in Brazil. J Lat Am Cult Stud 2015 Oct 2;24(4):535-552. [doi: 10.1080/13569325.2015.1091296]
95. Figueroa MGM, Moore MR. Beauty, race and feminist theory in Latin America and the Caribbean. Fem Theor 2013 Aug;14(2):131-136. [doi: 10.1177/1464700113483233]
96. Kukla QR. City Living: How Urban Dwellers and Urban Spaces Make One Another: Oxford University Press; 2021. [doi: 10.1093/oso/9780190855369.001.0001]
97. Chavarria MA, Rivas Velarde M. When innovation is perpetrating inequities: the pivotal role of contextualisation in the development of assistive technology for low- and middle-income countries (LMICs). Bioethica Forum 2025;17(1). [doi: 10.24894/BF.2024.17003]
98. Cuerpos excluidos, rostros de impunidad. informe de violencia hacia personas LGBT en Colombia, 2015. Colombia Diversa, Caribe Afirmativo, Santamaría Fundación. 2015. URL: https://colombiadiversa.org/ddhh-lgbt/Informe-Violencia-LGBT-Colombia-DDHH-2015.pdf [accessed 2026-03-31]
99. Mosquera Vera C. Gender and sexuality in the Colombian armed conflict: from patriarchal oppression to feminist resistance. Rocznik Filozoficzno-Społeczny Civitas Hominibus 2023(18):73-84. [doi: 10.25312/2391-5145.18/2023_06cmv]
100. Ordóñez-Valverde J. From gangs to bands: transformations of gang violence in slums in Cali. Soc Econ 2017(32):107-126 [FREE Full text]
101. Corte Constitucional de Colombia. Judgment C-671 of 2014. Bogotá, Colombia: Constitutional Court of Colombia.: Corte Constitucional; 2014. URL: https://www.alcaldiabogota.gov.co/sisjur//normas/Norma1.jsp?i=63728 [accessed 2024-03-05]
102. Rodríguez Caicedo N, León-Giraldo S, González-Uribe C, Bernal O, Gallego-Pantoja L, Carrasquilla G. Access to health services during the Colombian armed conflict: a challenge for the population with disabilities in the department of Meta. BMC Health Serv Res 2023 Jun 13;23(1):628. [doi: 10.1186/s12913-023-09472-x] [Medline: 37312099]
103. Díaz Velázquez E. Citizenship, identity and social exclusion of persons with disability. Polit Sci 2010(1):115-135 [FREE Full text]
104. Zheng Y, Stahl BC. Technology, capabilities and critical perspectives: what can critical theory contribute to Sen’s capability approach? Ethics Inf Technol 2011 Jun;13(2):69-80. [doi: 10.1007/s10676-011-9264-8]
105. Orellano-Colón EM, Mann WC, Rivero M, et al. Hispanic older adults’ perceptions of personal, contextual and technology-related barriers for using assistive technology devices. Occup Ther Int 2016;23(2):136-145. [doi: 10.1002/oti.1415]
106. MacLachlan M, Banes D, Bell D, et al. Assistive technology policy: a position paper from the first global research, innovation, and education on assistive technology (GREAT) summit. Disabil Rehabil Assist Technol 2018 Jul;13(5):454-466. [doi: 10.1080/17483107.2018.1468496] [Medline: 29790393]

XSL•FO
RenderX

Abbreviations

AD: assistive device
AT: assistive technology
CA: capability approach
CRPD: Convention on the Rights of Persons with Disabilities
DT: disadvantage theory
FARC-EP: Revolutionary Armed Forces of Colombia
LGBTIQ+: lesbian, gay, bisexual, transgender, intersex, and queer
LMIC: low- and middle-income country
WB: well-being
WHO: World Health Organization

Edited by S Munce, U Gopaul; submitted 10.Feb.2025; peer-reviewed by R Wu, T Hodson; revised version received 20.Jul.2025; accepted 08.Oct.2025; published 08.Apr.2026.

Please cite as:
Ortiz-Escobar LM, Chavarria MA, Hurst-Majno S, Campo Salazar OI, Escobar-Hurtado C, Stein MA, Rivas Velarde M
A Social Justice Approach to Assistive Technology and Well-Being of People With Visual Disabilities in Low- and Middle-Income Countries: Qualitative Narrative Study
JMIR Rehabil Assist Technol 2026;13:e72306
URL: https://rehab.jmir.org/2026/1/e72306
doi:10.2196/72306

© Luisa Maria Ortiz-Escobar, Mario Andres Chavarria, Samia Hurst-Majno, Oscar Ivan Campo Salazar, Celia Escobar-Hurtado, Michael Ashley Stein, Minerva Rivas Velarde. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Therapeutic Use of Virtual Reality for Patients With Fibromyalgia and Chronic Neck Pain: Randomized Controlled Trial

Edurne Úbeda-D’Ocasar[1,2*], PhD, Prof Dr; Yaiza Moreno-Crespo[1,2*], PT; Eduardo Cimadevilla-Fernández-Pola[1,2*], PhD, Prof Dr; Juan Hernández-Lougedo[1,2*], PhD, Prof Dr; Álvaro Navas-Mosqueda[1,2*], PT; Mario Caballero-Corella[1,2*], PT; Noemí Mayoral-Gonzalo[1,2*], PT; Blanca Pedauyé-Rueda[1,2*], RDN; María Jesús Fernández-Aceñero[3*], Dr med, PhD; Juan Pablo Hervás-Pérez[4*], PhD, Prof Dr; Cristina Ojedo-Martín[1,2*], PT

[1]Physiotherapy and Health Research Group (FYSA), Faculty of Health Sciences – HM Hospitals, Camilo José Cela University, Villanueva de la Cañada, Madrid, Spain

[2]Instituto de Investigación Sanitaria HM Hospitales, Madrid, Spain

[3]Department of Legal Medicine, Psychiatry and Pathology, Facultad de Medicina, Universidad Complutense de Madrid, Fundación de Investigación Clínico San Carlos (IDiSCC), Madrid, Spain

[4]Department of Chemistry in Pharmaceutical Sciences, Analytical Chemistry Unit, Faculty of Pharmacy, Complutense University of Madrid, Plaza Ramón y Cajal s/n, Ciudad Universitaria, Madrid, Spain

[*]all authors contributed equally

Corresponding Author:
Juan Pablo Hervás-Pérez, PhD, Prof Dr
Department of Chemistry in Pharmaceutical Sciences, Analytical Chemistry Unit, Faculty of Pharmacy, Complutense University of Madrid, Plaza Ramón y Cajal s/n, Ciudad Universitaria, Madrid, Spain

Abstract

Background: Fibromyalgia (FM) causes widespread pain, fatigue, and cognitive abnormalities. Cervical pain is a common and debilitating symptom.

Objective: This study aims to evaluate the effectiveness of virtual reality (VR) as a treatment for chronic cervical pain experienced by patients with FM.

Methods: A single-blind randomized clinical trial was conducted. A total of 56 women were randomly assigned to 3 groups: G1 (VR+cervical mobility exercises), G2 (cervical mobility exercises), and the control group. Therapy was administered twice a week for 4 weeks. Variables such as disease impact, quality of life, kinesiophobia, pain, range of motion, fatigue, and treatment adherence were measured.

Results: The mean age of the participants was 54.26 (SD 7.7) years. Participants were overweight, with a mean BMI of 28.7 (SD 7.8). The mean visual analog scale value was 6.72 (SD 1.8). The baseline values for age, BMI, visual analog scale, algometric measures, and functional capacity (measured using the Timed Up and Go test, cervical rotation, and lateral displacement) were similar across the 3 groups. Following the intervention therapy, the control group did not exhibit notable improvement (mean 3.5, SD 1.4; differences of mean values –0.46, 95% CI –1.1 to 0.2; *P*=.15), particularly in pain perception, while both therapy groups did show improvements (G1: mean 3.8, SD 1.1; differences of mean values 1.2, 95% CI 0.78-1.54; *P*<.001; G2: mean 2.8, SD 1.8; differences of mean values 1.2, 95% CI 0.66-1.7; *P*<.001). Both intervention groups improved significantly compared to control postintervention in FM impact (CG vs G1: differences of mean values 9.31, 95% CI 14.7-3.8; *P*<.001; CG vs G2: differences of mean values 8.4, 95% CI 13.84-3.06; *P*<.001), central sensitization (CG vs G2: differences of mean values 7.53, 95% CI 12.12-2.95; *P*<.001), and cervical disability (CG vs G2: differences of mean values 6.44, 95% CI 9.93-2.94; *P*<.001). However, at 1 month, only G1 maintained superior improvements across all measures, including a reduction in kinesiophobia (G2: differences of mean values 6.2, 95% CI 4.7-9.8; *P*<.001), indicating a more sustained effect of the combined approach.

Conclusions: The combination of VR with cervical mobility and strengthening exercises produced superior and sustained improvements in women with FM compared to exercise alone or control. Significant benefits were observed in disease impact, central sensitization, cervical disability, and kinesiophobia, with effects maintained at 1 month only in the VR group. These findings support VR as an effective adjunct to enhance symptom management and treatment adherence in FM.

Trial Registration: ClinicalTrials.gov NCT05933941; https://clinicaltrials.gov/study/NCT05933941

(JMIR Rehabil Assist Technol 2026;13:e81158) doi:10.2196/81158

KEYWORDS

chronic neck pain; exercise; fibromyalgia; virtual reality; randomized controlled trial

XSL•FO
RenderX

Introduction

Fibromyalgia (FM) is a disease that causes widespread chronic pain and intense fatigue along with many other symptoms [1-3]. Historically, there was some skepticism among health care professionals regarding the existence of this disease and its classification as a mental disorder. Such skepticism may have arisen from concerns about the optimal use of resources in providing care for individuals with FM [3,4]. It is estimated that this disease affects between 2% and 4% of the Spanish population, making it the most common cause of diffuse chronic musculoskeletal pain. Notably, the prevalence is much higher in women, with a ratio of 8:1 [5-7]. Its incidence is increasing because of the improvements in diagnostic criteria and advances in research that are bringing this pathology to light [7,8]. Given its high prevalence, this disease represents a significant public health concern [9,10].

Currently, there are no specific diagnosis tests available for confirming an FM diagnosis. Historically, an FM diagnosis was based on the presence of widespread pain in at least 11 of 18 designated tender points for a duration of 3 months. However, this criterion proved insufficient to capture the complexity of the disorder. Consequently, more comprehensive diagnostic frameworks have been developed, including a range of associated symptoms such as fatigue, joint stiffness, cervical and lumbar pain, muscle weakness, depressive symptoms, gastrointestinal issues, cognitive impairments, and balance alterations. Collectively, these manifestations significantly compromise the quality of life of individuals affected by FM. This pain can become disabling during flare-ups and even become chronic, often being a cause of recurrent absences from work [11].

Unlike pain associated with a sedentary lifestyle or poor posture, neck pain in FM results from a complex interplay between physiological, neurological, and psychological factors. Chronic pain often involves the cervical spine, leading to acute and prolonged pain episodes, along with muscle tension in the suboccipital, trapezius, and elevator scapulae muscles. It is thought that several factors may contribute to this chronic pain, including central pain sensitization, nervous system dysregulation, sleep disturbances, musculoskeletal changes, and psychological and emotional factors [11-13].

In recent years, new therapeutic approaches have been explored to improve symptomatology and thus patients' quality of life. One of the main challenges in treating patients with chronic pain, particularly those with FM, is their low adherence to physical therapy exercises. This may be due to factors such as severe pain and fatigue, a fear of worsening symptoms, a lack of understanding about exercise benefits, inadequate support and supervision, and frustration with slow improvement. To address these issues, it would be beneficial to adopt a comprehensive approach that involves educating patients regarding the benefits of exercise, customizing programs according to individual needs, providing ongoing support via health care professionals, managing pain and fatigue during exercise, setting realistic expectations, and recognizing achievements [14-16]. It may be beneficial to consider a treatment plan that incorporates aerobic activities, strength training, and stretching exercises [17-19]. However, this type of therapy may be limited by kinesiophobia. One possible solution to this issue could be the use of virtual reality (VR) technology. Exercise interventions delivered through VR have shown superior efficacy in alleviating the core symptoms of FM, including pain, fatigue, and stiffness, while also promoting improvements in balance and postural control [20]. The affordability of portable VR devices, combined with their sustained effectiveness as a nonpharmacological intervention for chronic pain management, positions VR as a promising tool with potential future applications as an analgesic modality [21]. Numerous randomized controlled trials have demonstrated that VR constitutes an alternative and accessible therapeutic approach for pain management [22].

Immersive VR is an innovative and interactive technology that generates 3D scenarios, enabling patients to experience highly realistic simulations while interacting with the virtual environment using their own hands. VR has been researched and applied in various contexts for pain management, demonstrating considerable potential as a therapeutic tool. The effectiveness of this therapeutic approach has been demonstrated for patients with chronic pain, as it addresses several factors, such as fatigue, kinesiophobia, and range of motion (ROM) [23]. This therapy can be adapted to suit the specific requirements of each patient, providing the option to adjust the level of difficulty from low to intermediate or high. Additionally, it offers a variety of game modes that could be beneficial in targeting different clinical conditions, such as low back pain, neck pain, and balance disorders [24,25].

The therapeutic mechanisms and components of VR are distraction, activity management and behavioral activation, skills-based cognitive behavioral therapy, relaxation training and biofeedback, positive emotion induction, neuromodulation, physical rehabilitation, and reduction of kinesiophobia [8,23,26].

VR can be categorized into 2 main types: immersive and nonimmersive. Immersive VR involves the use of head-mounted displays or VR goggles that fully cover the visual field and may include motion sensors for the hands or feet. In contrast, nonimmersive VR is accessed through conventional screens or computers, without any device that isolates the user from the external environment [27] .

Chronic pain has been extensively investigated over the past decades, and research efforts persist in identifying novel strategies to mitigate its impact and enhance patients' quality of life. VR in both modalities has demonstrated efficacy in reducing chronic musculoskeletal pain. These findings support the integration of VR as a therapeutic tool in clinical populations affected by conditions associated with this type of pain [28].

Patients may perceive amplified motion in the virtual world, depending on the configuration of the headset used. This could result in small neck flexion in real life being rewarded with greater apparent motion in VR, which might increase kinesiophobia [26]. Studies have indicated that for patients with chronic low-back pain, kinesiophobia may be addressed through VR interventions [20,29,30]. VR helps to distract an individual from painful stimuli while exercising in a way that improves

XSL•FO
RenderX

their perception of pain [31]. In addition, there have been indications that improvements in fatigue, sleep quality, and ROM may also be achieved [9].

In this study, we aimed to evaluate the effectiveness of VR as a therapeutic option for neck pain in patients with FM. We hypothesized that VR could decrease cervical pain and kinesiophobia and increase ROM, facilitating improvements in quality of life and adherence to treatment.

Methods

Trial Design

A single-blind, randomized experimental clinical trial was conducted among women diagnosed with FM recruited through the Association of Fibromyalgia, “AFINSFACRO,” in Móstoles, Spain. Following the selection of patients who met the inclusion criteria and the collection of informed consent forms and information sheets, according to ethical standards, the patients were randomly assigned to 1 of the 3 groups. The first group performed 20 minutes of exercise plus 10 minutes of VR (G1), the second group performed 30 minutes of exercise (G2), and the control group (CG) did not perform any exercise or VR treatment.

This study was registered at ClinicalTrials.gov (NCT05933941) and approved by the Clinical Research Committee of the Hospital Clínico San Carlos (Madrid, Spain; ID: VRTCNPPFM-13/07/2023). The research was conducted from April 1, 2024, to January 30, 2025. All participants gave informed consent, and the general scope of this study was explained to them via a participant information sheet. This clinical trial was designed and reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) guidelines, with specific adherence to the CONSORT-Harms (Consolidated Standards of Reporting Trials Harms Extension) 2022 statement (Checklist 1) to ensure transparent and comprehensive reporting of adverse events and safety outcomes.

Participants

The QuestionPro survey software (QuestionPro, Inc) was used to calculate the required sample size, considering a 95% CI and a 5% α error. The sample size calculation was based on the study by Gulsen et al [8]. A total of 56 women were enrolled, of whom 2 withdrew for personal reasons unrelated to the study. The inclusion criteria comprised female patients aged 20 to 65 years with a confirmed diagnosis of FM and chronic cervical pain lasting more than 3 months. The exclusion criteria included vertigo, claustrophobia, epilepsy, pregnancy, or refusal to provide informed consent.

Type of Sampling

A consecutive, nonprobabilistic convenience sampling strategy was applied to participants who met the inclusion criteria.

Variables and Outcomes

Sociodemographic variables included age, height, weight, BMI, marital status, pain medication use, employment status, smoking history, and comorbid conditions such as restless legs syndrome, chronic fatigue syndrome, temporomandibular joint dysfunction, migraines, irritable bowel syndrome, multiple chemical sensitivity, anxiety, and depression. All these sociodemographic variables were collected using a self-administered survey.

The following tools were used in this study:

1. Pain: we used a visual analog scale (VAS) to assess self-reported pain intensity, ensuring maximum reproducibility among different observers [32]. Furthermore, we used an analog pressure FPK 20 algometer (Wagner algometer, Force Dial FDK 20; Wagner Instruments) to evaluate pain among individuals diagnosed with FM. This instrument allows the assessment of a patient’s central sensitivity to pain. Measurements focused on tender points in the occipital and upper trapezius regions bilaterally [33-35].
2. Subjective intensity of effort was assessed using the Borg Category-Ratio Scale (CR-10), which quantifies the subjective intensity of effort experienced during physical exercise or functional testing. Participants were asked to rate their exertion immediately after each test or exercise session on a 0 to 10 scale, where 0 indicates “no exertion at all” and 10 represents “maximal exertion” [36].
3. Neck Disability Index (NDI): the validated Spanish version of this questionnaire was used to assess pain and neck-related disability [37].
4. Fear of movement: the Tampa Scale for Kinesiophobia (TSK) was used to assess the fear of movement [38].
5. Exercise adherence: we used the Exercise Adherence Rating Scale, which is a validated questionnaire with 2 sections, one assessing exercise performance and another evaluating frequency, motivation, and consistency [39].
6. Impact of FM: the Fibromyalgia Impact Questionnaire (FIQ) was used to assess the impact of FM on health-related quality of life [40-42].
7. Quality of life: we used EQ-5D (EuroQol 5-Dimensions) questionnaire, which is a tool that allows the evaluation of a patient’s overall quality of life in primary care settings [43-45].
8. Symptoms of central sensitization: we used the validated Spanish version of the Central Sensitization Inventory (CSI). This questionnaire assesses symptoms using a scale of 0 (“never”) to 4 (“always”). The total score ranges from 0 to 100. A score above 40 indicates the presence of central sensitization [46].
9. ROM: cervical flexion, extension, lateral flexion, and rotation were assessed using a goniometer [47]. This instrument is used to evaluate one’s degree of joint mobility, thereby facilitating the determination of an individual’s restrictions.
10. Time Up Go test: the test consisted of measuring the time it took participants to get up from a chair with a height of 46 cm, walking 3 meters, turning around a cone, and sitting down again. This test was performed to assess physical performance, gait, and dynamic balance [48,49].

Procedure

The intervention was conducted over a period of 1 month. G1 and G2 participated in 2 sessions per week. G1 performed 20 minutes of cervical mobility and strengthening exercises

XSL•FO
RenderX

(Multimedia Appendix 1), followed by 10 minutes of immersive VR therapy. G2 completed 30 minutes of cervical mobility and strengthening exercises. The CG did not perform any cervical mobility or strengthening exercises nor participate in immersive VR therapy. The immersive VR therapy was delivered using Meta Quest 2000 headsets, using the game "Interkosmos 2000." In the game, participants assumed the role of a spacecraft pilot navigating through a series of rings while avoiding meteoroids by performing neck movements according to the game's instructions. The difficulty level was adjusted as follows: the first 3 sessions were carried out in "easy" mode, allowing a maximum cervical mobility of 30°. The subsequent 3 sessions were conducted in "medium" mode, allowing up to 60° of cervical mobility. The final 2 sessions were performed in "hard" mode, which not only increased the spacecraft's speed but also enabled a full cervical range of motion. All groups underwent assessments of cervical ROM including flexion, extension, right and left lateral inclination, and right and left rotation. In addition, pain intensity in the upper trapezius (bilaterally) and in the right and left occipital regions was quantified using an algometer. Evaluations were conducted at 4 time points: baseline (preintervention), immediately after the intervention, 15 days postintervention, and 1 month postintervention. All follow-up assessments were performed at the association's premises. After each session, participants from G1 and G2 reported their perceived levels of cervical pain and fatigue. All outcome measures were collected by a research assistant blinded to the study objectives and group allocations.

Evaluations were conducted at 4 time points: baseline (preintervention), immediately postintervention, 15 days postintervention, and 1-month postintervention. All follow-up assessments were conducted at the association's headquarters. After each session, participants in G1 and G2 reported perceived levels of cervical pain and fatigue.

All outcome measures were collected by a research assistant who was blinded to both the study's objectives and the participants' group assignments.

Statistical Analysis

Statistical analyses were performed using SPSS 29.0 (IBM Corp). The normality of quantitative variables was assessed with the Kolmogorov-Smirnov test. Depending on distribution, data were described using means and SDs or medians, IQRs, and ranges. Qualitative variables were expressed as percentages and absolute values. Baseline comparisons between the groups were conducted using chi-squared tests or ANOVA with Bonferroni post hoc analysis. Temporal changes in outcomes were analyzed with paired tests, and intergroup differences were analyzed using ANOVA, Bonferroni post hoc, or Kruskal-Wallis tests, as appropriate. Correlations between quantitative variables were assessed using the Pearson or Spearman tests, and associations between qualitative variables were assessed using chi-squared tests. Statistical significance was set at $P<.05$.

Ethical Considerations

This study received approval from the institutional review board of Hospital Clínico San Carlos (23‑458 EC) and was conducted in accordance with the Spanish legislation, including Law 41/2002 on patient autonomy and Organic Law 3/2018 on data protection and digital rights. These laws prohibit the processing of sensitive personal data such as racial or ethnic origin, political or religious beliefs, biometric identifiers, health information, and sexual orientation. The study also adhered to the ethical principles outlined in the World Medical Association's Declaration of Helsinki 2014 [50]. All procedures conducted in this study complied with applicable ethical standards. Informed consent was obtained from all participants prior to their inclusion in the study. Participant privacy and confidentiality were strictly protected, and all data were handled and stored in accordance with relevant data protection regulations. No personally identifiable information was collected or disclosed. Participants did not receive any financial or material compensation for their participation.

Results

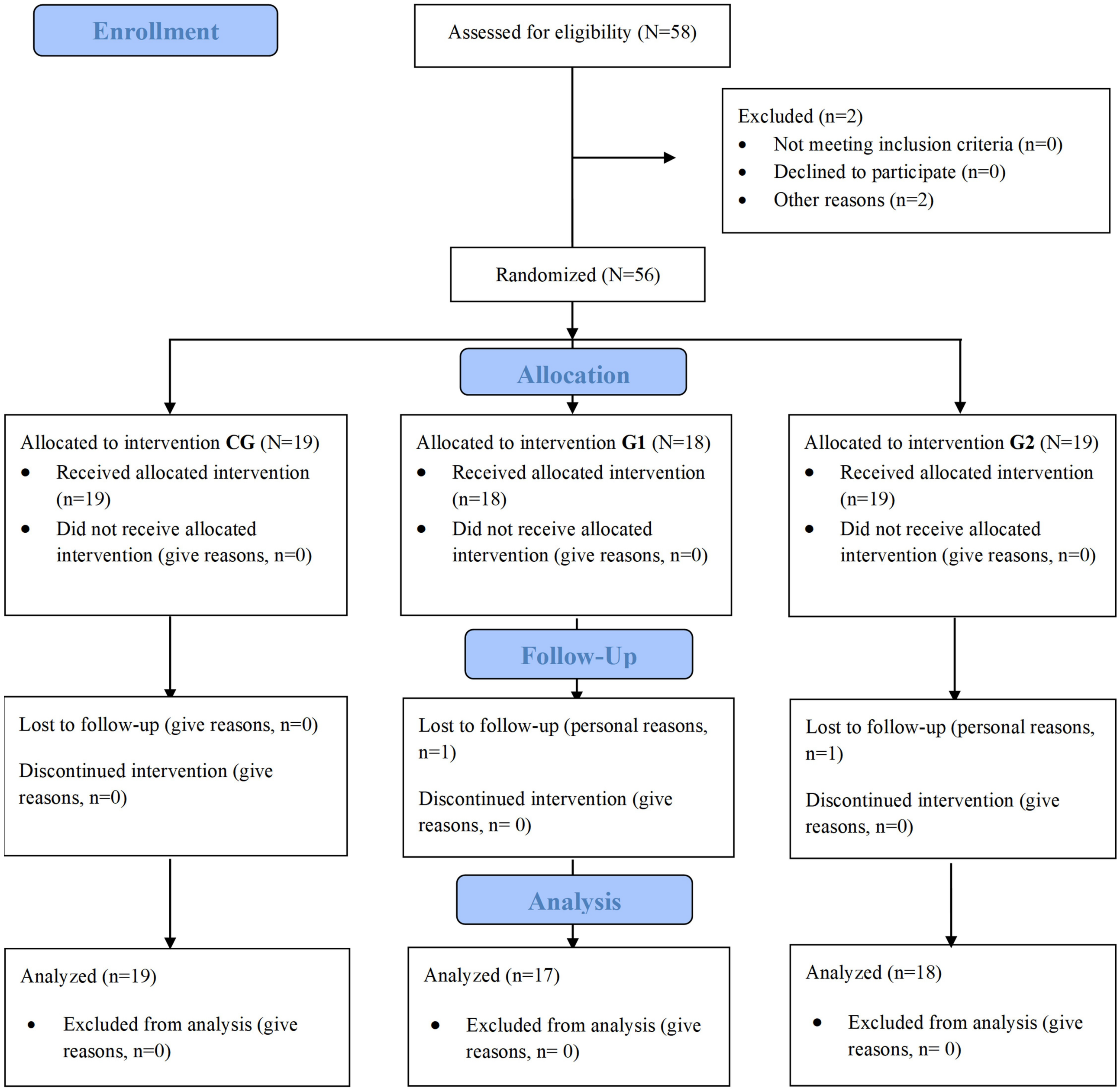
Participant and Baseline Characteristics

The final sample included 54 women with a mean age of 54.26 (SD 7.7) years. Participants were distributed among the groups by computer-generated random number sequence (Figure 1).

Among the participants, 33 (61%) presented chronic fatigue syndrome, 37 (68%) had temporomandibular joint dysfunction, and 16 (33%) reported restless legs syndrome. Regarding the psychological comorbidities, 33 (61%) participants presented anxiety and 13 (24%) depression. Baseline comparisons revealed no statistically significant differences among groups in age, BMI, or other demographic variables. However, subjective measures showed group-level differences in several baseline scales, including FIQ, EQ-5D, CSI, NDI, and Borg scale. Almost all variables at baseline start from similar values, although G1 has a value of 1 point lower on the VAS scale, indicating that they have less pain at the outset. However, in variables such as algometry, the averages reflect similar values compared to the other 2 groups (Multimedia Appendix 2).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. CG: control group.



Postintervention Outcomes

Table 1 shows the differences in the mean values measured before therapy and immediately after its end. We found significant differences in all groups that received intervention in FIQ, CSI, NDI, right and left trapezius, occipital algometer, and in all movements in which ROM was evaluated. For the right and left trapezius and occipital muscles, algometric measurements were obtained, and ROM was assessed across all evaluated movements. About the treatment adherence variable, G1 obtained an average of 94 points out of 100, compared to G2, whose post-treatment average was 74.72. The summary of the comparison between baseline measures and measures immediately after intervention (intragroup comparisons). Shown are the differences in the mean values between both time points (Multimedia Appendix 3).

XSL•FO
RenderX

Table . Intragroup (G1, G2, and CG[a]) analysis of variables measured using questionnaires between baseline and immediately after intervention.

Value and intervention group	Differences of mean values (95% CI)	*P* value
FIQ[b] (1-100)		
Whole series	5.4 (3.2 to 7.5)	<.001
G1	9 (4.6 to 13.3)	<.001
G2	8.1 (5.5 to 10.6)	<.001
CG	0.36 (−3.2 to 2.51)	.80
EQ-5D[c] (–0.59 to 1)		
Whole series	4.1 (2.6 to 5.6)	<.001
G1	3.1 (0.6 to 5.6)	.01
G2	7.7 (5.6 to 9.8)	<.001
CG	−1.7 (−4.4 to 1)	.20
TSK[d] (17-37)		
Whole series	3.2 (0.8 to 5.6)	.01
G1	1.2 (−5.2 to 7.6)	.70
G2	6.2 (4.7 to 9.8)	<.001
CG	−2.04 (−5.9 to 1.86)	.30
CSI[e] (0-100)		
Whole series	4.2 (2.4 to 6)	<.001
G1	6 (3.5 to 8.5)	<.001
G2	7.2 (4.7 to 9.8)	<.001
CG	0.32 (−3.5 to 2.9)	.80
NDI[f] (0-50)		
Whole series	4.5 (3.1 to 5.9)	<.001
G1	5.35 (3.2 to 7.5)	<.001
G2	7.4 (5.1 to 9.7)	<.001
CG	−0.95 (−2.9 to 1.05)	.33

[a] CG: control group.
[b] FIQ: Fibromyalgia Impact Questionnaire.
[c] EQ-5D: EuroQol 5-Dimensions.
[d] TSK: Tampa Scale for Kinesiophobia.
[e] CSI: Central Sensitization Inventory.
[f] NDI: Neck Disability Index Questionnaire.

Follow-Up at 1 Month

Summary of the comparison between baseline and 1-month post-intervention scores for FIQ, EQ-5D, TSK, CSI, and NDI (intragroup comparisons). Shown are the differences in the mean values between both time points (Table 2). The differences in the measurement times of the other variables can be observed in Multimedia Appendix 4.

XSL•FO
RenderX

Table . Intragroup (G1, G2, and CG[a]) analysis of variables measured using questionnaires between baseline and 1 month after intervention.

Value and intervention group	Differences of mean values (95% CI)	*P* value
FIQ[b]		
Whole series	4.4 (2.5 to 76.3)	<.001
G1	10.71 (6.6 to 14.8)	<.001
G2	2.71 (0.55 to 4.87)	.02
CG	0.33 (–2.08 to 1.42)	.70
EQ-5D[c]		
Whole series	4.7 (2.6 to 5.7)	<.001
G1	7.7 (5.46 to 9.97)	<.001
G2	4.17 (1.7 to 6.63)	.002
CG	1.07 (–3.87 to 1.73)	.43
TSK[d]		
Whole series	3.3 (1.4 to 5.1)	.001
G1	8.41 (5.25 to 11.58)	<.001
G2	2.1 (0.06 to 4.13)	.04
CG	0.2 (–2.87 to 3.28)	.89
CSI[e]		
Whole series	3.9 (2.18 to 5.8)	<.001
G1	9.78 (7.35 to 12.22)	<.001
G2	2.94 (1.25 to 4.64)	.02
CG	0.26 (–2.76 to 3.3)	.85
NDI[f]		
Whole series	4.03 (2.5 to 6.31)	<.001
G1	10.65 (8.35 to 12.94)	<.001
G2	4.22 (2.39 to 6.06)	<.001
CG	2.05 (–0.33 to 4.43)	.08

[a] CG: control group.

[b] FIQ: Fibromyalgia Impact Questionnaire.

[c] EQ-5D: EuroQol 5-Dimensions.

[d] TSK: Tampa Scale for Kinesiophobia.

[e] CSI: Central Sensitization Inventory.

[f] NDI: Neck Disability Index Questionnaire.

After confirming the effect of both interventions, we compared the intervention groups to see whether VR had a significant influence on the effect of the therapy. Table 3 shows that the ANOVA between groups was significant for FIQ, CSI, and NDI at both postintervention and 1-month follow-up (*P*<.001). Immediately after treatment, both intervention groups (G1 and G2) improved significantly compared with the CG, with no differences between G1 and G2. At 1 month, G1 maintained superior outcomes across all measures, showing significantly better scores than both GC and G2 (*P*<.001), while G2 no longer differed from GC for FIQ and CSI. For NDI, all pairwise comparisons were significant (*P*<.001), confirming that both interventions reduce neck disability, with the VR-enhanced program providing the greatest and most sustained benefit.

Table . Summary of the comparison between the groups for the change in FIQ[a], CSI[b], and NDI[c], both immediately after intervention and 1 month after.

Variable and intervention group	*P* value for ANOVA	Mean difference between time points (95% CI)	*P* value
FIQ			
Immediate	<.001		
CG[d] vs G1		9.31 (14.7 to 3.8)	<.001
CG vs G2		8.4 (13.84 to 3.06)	.001
G2 vs G1		0.85 (6.39 to −4.68)	>.99
1 month	<.001		
CG vs G1		10.38 (5.8 to 14.9)	<.001
CG vs G2		2.38 (–2.12 to 6.9)	.59
G2 vs G1		7.99 (3.35 to 12.63)	<.001
CSI			
Immediate	<.001		
CG vs G1		6.31 (10.96 to 1.66)	.04
CG vs G2		7.53 (12.12 to 2.95)	<.001
G2 vs G1		–1.22 (3.4 to −5.93)	.51
1 month	<.001		
CG vs G1		10.04 (14.16 to 5.93)	<.001
CG vs G2		3.20 (7.26 to −0.84)	.16
G2 vs G1		6.84 (11.01 to 2.67)	.001
NDI			
Immediate	<.001		
CG vs G1		4.40 (7.95 to 0.85)	.01
CG vs G2		6.44 (9.93 to 2.94)	<.001
G2 vs G1		–2.03 (1.55 to −5.62)	.50
1 month	<.001		
CG vs G1		12.69 (16.33 to 9.06)	<.001
CG vs G2		6.27 (9.85 to 2.69)	<.001
G2 vs G1		6.42 (10.10 to 2.74)	<.001

[a] FIQ: Fibromyalgia Impact Questionnaire.

[b] CSI: Central Sensitization Inventory.

[c] NDI: Neck Disability Index Questionnaire.

[d] CG: control group.

Discussion

Principal Findings

This study evaluated the effectiveness of VR as an adjunctive therapy for cervical pain in women with FM. The main findings revealed that combining VR with mobility and strengthening exercises produced significant improvements in pain perception, ROM, and functional performance compared with physical therapy alone. These benefits were also sustained at 1-month follow-up, emphasizing the potential of immersive VR interventions to enhance therapeutic outcomes in individuals with FM.

In the intergroup analysis, the G1 consistently outperformed both G2 and CG in those outcomes most closely related to global disease impact and NDI. Specifically, G1 showed significantly greater and more sustained improvements in FIQ and CSI scores than both CG and G2, whereas improvements in G2 tended to converge toward control values at 1-month follow-up. Similarly, NDI scores decreased in both intervention groups immediately after treatment, but at follow-up, a clear gradient GC>G2>G1 was observed, indicating the largest and most persistent reduction in neck-related disability in the combined intervention. These effects were paralleled by a recurrent pattern of superior gains in cervical ROM and higher-pressure pain thresholds in the trapezius and occipital in G1, while the Time Up Go test

XSL•FO
RenderX

improved in both intervention groups compared with CG and TSK decreased more markedly in G1 at 1 month. By contrast, the VAS and Borg scale showed only limited discrimination between the groups, suggesting that the added value of VR was more evident in multidimensional impact and sensitization-related and functional domains than in isolated pain ratings.

This intergroup pattern is consistent with previous evidence in FM and chronic neck pain, indicating that VR, particularly when used as an adjunct to active exercise, preferentially enhances global impact, disability, and pain-modulation outcomes. In patients with FM, a study reported that fully immersive VR combined with aerobic and pilates training produced greater improvements in pain, kinesiophobia, fatigue, physical activity levels, and mental quality of life than exercise alone, while both groups improved in FM impact, supporting an adjunctive role of immersive VR in comprehensive rehabilitation programs [8]. Other studies showed that, in individuals with chronic neck pain, VR-based cervical training led to larger gains in pressure pain thresholds at the upper cervical levels and greater reductions in functional limitation compared with motor control exercises, despite the absence of between-group differences in pain intensity and quality of life, which parallels our finding of more robust intergroup differences in pressure pain thresholds, ROM, and disability than in VAS score [51]. Furthermore, a randomized crossover trial in women with FM found that VR increased cold pain thresholds and tolerance in both FM patients and pain-free controls, whereas effects on pain intensity were limited, reinforcing the notion that VR may primarily modulate pain processing rather than consistently decreasing reported pain intensity [52]. Taken together, these converging results suggest that combining immersive VR with targeted cervical exercise may primarily potentiate mechanisms related to central sensitization, motor control, and the global impact of FM rather than merely amplifying short-term analgesic effects.

VR has emerged as a promising nonpharmacological intervention for the symptom management of FM. Its immersive and interactive features provide cognitive distraction, emotional engagement, and motor stimulation, which together enhance adherence and reduce pain intensity. Previous studies have demonstrated that VR can improve mood, motivation, and functional capacity when applied alongside traditional rehabilitation or psychological therapies. This approach not only benefits patients with FM but may also be applicable to other chronic pain conditions [49,50,53].

The use of VR has been demonstrated to significantly influence pain relief, motor function, and joint mobility among patients with a range of chronic pathologies [53]. The primary applications of VR in health care include the management of pain and anxiety as well as the enhancement of patient motivation [47,54,55].

While it is not possible to make direct comparisons between this study and research conducted by other authors, the described implementation of an 8-session intervention resulted in statistically significant outcomes for the variables under investigation. This suggests VR is effective when combined with active exercise therapy.

The comparative analysis concerning VR and physical exercise indicated that there was minimal divergence in the measured variables. In a separate study [56], 44 patients underwent an intervention for 4 weeks, with 2 sessions per week. The intervention was divided into 2 groups: G1 (whose members performed only cervical mobility exercises) and G2 (whose members conducted a treatment based solely on VR exercises). The findings indicated that there were no statistically significant differences between the 2 groups across the different variables, including pain, ROM, neck disability, incapacitating pain, and anxiety. However, in this study, the combination of VR with cervical mobility exercises yielded a significantly better outcome. Furthermore, these improvements were maintained over time (1 month) to a greater degree in the group subjected to both therapies combined when compared to the group that performed only mobility exercises. It is important to highlight that the study with which this investigation was compared implemented its intervention in a population without FM. Consequently, the observed discrepancy in outcomes could be due to this factor. A similar study was conducted [51], with an equivalent number of sessions implemented. The outcomes observed in this research were analogous to those obtained in our research, as the investigators combined active therapy and VR exercises for patients with FM [55]. However, the FM cohort in that report comprised only 20 patients. In addition, the active component focused on aerobic training and Pilates instead of the cervical exercises used in our research.

Another study assessed the efficacy of VR for patients with chronic neck pain versus a CG that performed motor control exercises. In this case, the study spanned a period of 6 weeks, comprising 3 weekly sessions. Their findings revealed that the utilization of VR was advantageous when measuring pain due to pressure with an algometer at C1-C2 and C5-C6, and it was also beneficial when considering functional limitations. However, no significant differences were observed between the groups with respect to pain intensity, muscle performance, or quality of life (36-Item Short Form Health Survey). It is evident that the results of this study deviate from those previously observed. However, it should be noted that the mentioned study did not include the population with FM, and no group underwent combined VR and active therapy [51].

Following a thorough review of the scientific literature, we found no studies addressing the use of VR in the cervical region using this protocol. Other studies have applied VR in FM use, games, and global exercises, focusing on observing aspects such as anaerobic capacity, balance, or fatigue [7,54,55]. In this study, the focus was pain in the cervical region and other more global variables such as fatigue, VAS, and FIQ. Therefore, our study offers a novel contribution to this field of research by addressing this gap in the existing literature. Given that patients with FM experience cervical pain, headaches, and tender points in the occipital regions, it is imperative to reduce the associated symptoms.

The limitations of this study lie in the fact that the results are not representative of a heterogeneous population with FM, since the cohort of participants in this research consisted solely of women. Due to the challenging nature of the disease, it was not possible to consider medication used by participants, which

XSL•FO
RenderX

may have influenced the results. Finally, the study did not include a group that used only VR, so we cannot conclude that the results are attributable to the VR component or its combination with exercise.

This approach not only benefits patients with FM but may also be applicable to other chronic pain conditions [49,50,53].

Conclusions

Virtual reality, when added to cervical mobility and strengthening exercises, produced greater and more sustained improvements in disease impact, neck disability, central sensitization, kinesiophobia, and functional outcomes in women with FM than exercise alone or no intervention. The improvements were maintained for 1 month in the G1 group in the variables evaluated. These findings support VR as an effective adjunct to active therapy, with potential benefits for symptom reduction and warrant further research on long-term effects, cost-effectiveness, and applicability to broader chronic pain populations.

Funding

This research received external funding for the FIBYSAR project of Camilo Jose Cela University.

Data Availability

The data presented in this study are available upon reasonable request from the corresponding author.

Authors' Contributions

Conceptualization: AN-M, CO-M, EÚ-D, YM-C
Data curation: BP-R, EC-F-P, EÚ-D, JPH-P
Formal analysis: JH-L, JPH-P, MJF-A, NM-G
Investigation: BP-R, EC-F-P, EÚ-D, JH-L, JPH-P, NM-G, YM-C
Methodology: EC-F-P, JH-L, JPH-P, MC-C
Resources: CO-M, EC-F-P, EÚ-D, JPH-P
Software: AN-M, JH-L, YM-C
Supervision: CO-M, EÚ-D, JPH-P
Validation: EÚ-D, EC-F-P, JPH-P, MC-C, NM-G
Visualization: EÚ-D, JPH-P, MC-C, YM-C
Writing – original draft: AN-M, CO-M, EÚ-D, NM-G, YM-C
Writing – review & editing: BP-R, EC-F-P, EÚ-D, JH-L, JPH-P, MJF-A

Conflicts of Interest

None declared.

Multimedia Appendix 1
Recommended exercises.
[DOCX File, 15 KB - rehab_v13i1e81158_app1.docx]

Multimedia Appendix 2
Baseline measurements of pain and functional capacity (mean and SD).
[DOCX File, 23 KB - rehab_v13i1e81158_app2.docx]

Multimedia Appendix 3
Intragroup analysis of baseline and immediately after intervention.
[DOCX File, 20 KB - rehab_v13i1e81158_app3.docx]

Multimedia Appendix 4
Intragroup analysis of baseline and 1 month after intervention.
[DOCX File, 26 KB - rehab_v13i1e81158_app4.docx]

Checklist 1
CONSORT-Harms-2022 checklist.
[PDF File, 208 KB - rehab_v13i1e81158_app5.pdf]

References

XSL•FO
RenderX

1. Richard JY, Hurley RA, Taber KH. Fibromyalgia: Centralized pain processing and neuroimaging. J Neuropsychiatry Clin Neurosci 2019;31(3):A6-187. [doi: 10.1176/appi.neuropsych.19050107] [Medline: 31322995]
2. Häuser W, Fitzcharles MA. Facts and myths pertaining to fibromyalgia. Dialogues Clin Neurosci 2018 Mar;20(1):53-62. [doi: 10.31887/DCNS.2018.20.1/whauser] [Medline: 29946212]
3. Solà JF. El síndrome de fatiga crónica [Article in Spanish]. Medicina Integral 2002;40(2):56-63 [FREE Full text]
4. Sumpton JE, Moulin DE. Fibromyalgia. Handb Clin Neurol 2014;119:513-527. [doi: 10.1016/B978-0-7020-4086-3.00033-3] [Medline: 24365316]
5. Ruschak I, Montesó-Curto P, Rosselló L, Aguilar Martín C, Sánchez-Montesó L, Toussaint L. Fibromyalgia syndrome pain in men and women: a scoping review. Healthcare (Basel) 2023 Jan 11;11(2):223. [doi: 10.3390/healthcare11020223] [Medline: 36673591]
6. Liang D, Guo X, Zhang J, Jiang R. Pain characteristics of patients with fibromyalgia: a comparison between gender and different emotional states. Pain Physician 2024 Jan 20;27(1):E109-E118. [doi: 10.36076/ppj.2024.27.E109] [Medline: 38285043]
7. Cabo-Meseguer A, Cerdá-Olmedo G, Trillo-Mata JL. Fibromyalgia: prevalence, epidemiologic profiles and economic costs. Med Clin 2017 Nov;149(10):441-448. [doi: 10.1016/j.medcli.2017.06.008]
8. Gulsen C PT, MSc, Soke F PT, PhD, Eldemir K PT, MSc, et al. Effect of fully immersive virtual reality treatment combined with exercise in fibromyalgia patients: a randomized controlled trial. Assist Technol 2022 May 4;34(3):256-263. [doi: 10.1080/10400435.2020.1772900] [Medline: 32543290]
9. Villanueva VL, Valía JC, Cerdá G, et al. Fibromialgia: diagnóstico y tratamiento. El estado de la cuestión [Article in Spanish]. Rev Soc Esp Dolor 2004;11(7):430-443 [FREE Full text]
10. Borchers AT, Gershwin ME. Fibromyalgia: A critical and comprehensive review. Clinic Rev Allerg Immunol 2015 Oct;49(2):100-151. [doi: 10.1007/s12016-015-8509-4]
11. Wang JJ, Tam KW, Hsiao HY, Liou TH, Rau CL, Hsu TH. Effect of resistance exercises on function and pain in fibromyalgia. Am J Phys Med Rehabil 2024;103(4):275-283. [doi: 10.1097/PHM.0000000000002318]
12. Giesecke T, Gracely RH, Grant MAB, et al. Evidence of augmented central pain processing in idiopathic chronic low back pain. Arthritis Rheum 2004 Feb;50(2):613-623. [doi: 10.1002/art.20063]
13. Yunus MB. The prevalence of fibromyalgia in other chronic pain conditions. Pain Res Treat 2012;2012:584573. [doi: 10.1155/2012/584573] [Medline: 22191024]
14. Prikhodkina M, Melnikov S. Factors that influence medication adherence in women with fibromyalgia: a path analysis. J Clin Nurs 2024 Oct;33(10):3943-3953. [doi: 10.1111/jocn.17044] [Medline: 38284436]
15. Desai R, Marlow N. Effect of non-adherence on healthcare resource utilization in fibromyalgia patients: a retrospective cohort study. Value Health 2018 May;21:S255. [doi: 10.1016/j.jval.2018.04.1776]
16. Dobkin PL, Sita A, Sewitch MJ. Predictors of adherence to treatment in women with fibromyalgia. Clin J Pain 2006;22(3):286-294. [doi: 10.1097/01.ajp.0000173016.87612.4b] [Medline: 16514330]
17. Iglesias AR, Soria Ayuda RE, Martinez AB, Sánchez AJ, Villaroya Bielsa E, Callejero Guillen AJ. La efectividad de un programa de ejercicio terapéutico en pacientes con fibromialgia [Web page in Spanish]. Revista Sanitaria de Investigacion. URL: https://revistasanitariadeinvestigacion.com/la-efectividad-de-un-programa-de-ejercicio-terapeutico-en-pacientes-con-fibromialgia/ [accessed 2025-12-30]
18. Winkelmann A, Häuser W, Friedel E, et al. Physiotherapy and physical therapies for fibromyalgia syndrome. Systematic review, meta-analysis and guideline [Title in German]. Schmerz 2012 Jun;26(3):276-286. [doi: 10.1007/s00482-012-1171-3] [Medline: 22760460]
19. Sarac AJ, Gur A. Complementary and alternative medical therapies in fibromyalgia. Curr Pharm Des 2006;12(1):47-57. [doi: 10.2174/138161206775193262] [Medline: 16454724]
20. Collado-Mateo D, Dominguez-Muñoz FJ, Adsuar JC, Merellano-Navarro E, Gusi N. Exergames for women with fibromyalgia: a randomised controlled trial to evaluate the effects on mobility skills, balance and fear of falling. PeerJ 2017;5:e3211. [doi: 10.7717/peerj.3211] [Medline: 28439471]
21. Ahmadpour N, Randall H, Choksi H, Gao A, Vaughan C, Poronnik P. Virtual reality interventions for acute and chronic pain management. Int J Biochem Cell Biol 2019 Sep;114(September):105568. [doi: 10.1016/j.biocel.2019.105568] [Medline: 31306747]
22. Huang Q, Lin J, Han R, Peng C, Huang A. Using virtual reality exposure therapy in pain management: a systematic review and meta-analysis of randomized controlled trials. Value Health 2022 Feb;25(2):288-301. [doi: 10.1016/j.jval.2021.04.1285] [Medline: 35094802]
23. Cortés-Pérez I, Zagalaz-Anula N, Ibancos-Losada MDR, Nieto-Escámez FA, Obrero-Gaitán E, Osuna-Pérez MC. Virtual reality-based therapy reduces the disabling impact of fibromyalgia syndrome in women: systematic review with meta-analysis of randomized controlled trials. J Pers Med 2021 Nov 9;11(11):1167. [doi: 10.3390/jpm11111167] [Medline: 34834518]
24. Carvalho MSD, Carvalho LC, Menezes FDS, Frazin A, Gomes EDC, Iunes DH. Effects of exergames in women with fibromyalgia: a randomized controlled study. Games Health J 2020 Oct 1;9(5):358-367. [doi: 10.1089/g4h.2019.0108]
25. Gómez-Pérez L, López-Martínez AE, Ruiz-Párraga GT. Psychometric properties of the Spanish version of the Tampa Scale for Kinesiophobia (TSK). J Pain 2011 Apr;12(4):425-435. [doi: 10.1016/j.jpain.2010.08.004] [Medline: 20926355]

XSL•FO
RenderX

26. Sato K, Fukumori S, Matsusaki T, et al. Nonimmersive virtual reality mirror visual feedback therapy and its application for the treatment of complex regional pain syndrome: an open-label pilot study. Pain Med 2010 Apr;11(4):622-629. [doi: 10.1111/j.1526-4637.2010.00819.x] [Medline: 20202141]
27. Freeman D, Reeve S, Robinson A, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017 Oct;47(14):2393-2400. [doi: 10.1017/S003329171700040X] [Medline: 28325167]
28. Lo HHM, Zhu M, Zou Z, et al. Immersive and nonimmersive virtual reality-assisted active training in chronic musculoskeletal pain: systematic review and meta-analysis. J Med Internet Res 2024 Aug 19;26:e48787. [doi: 10.2196/48787] [Medline: 39159449]
29. Harvie DS, Broecker M, Smith RT, Meulders A, Madden VJ, Moseley GL. Bogus visual feedback alters onset of movement-evoked pain in people with neck pain. Psychol Sci 2015 Apr;26(4):385-392. [doi: 10.1177/0956797614563339] [Medline: 25691362]
30. Trost Z, Zielke M, Guck A, et al. The promise and challenge of virtual gaming technologies for chronic pain: the case of graded exposure for low back pain. Pain Manag 2015;5(3):197-206. [doi: 10.2217/pmt.15.6] [Medline: 25971643]
31. Sarig Bahat H, Takasaki H, Chen X, Bet-Or Y, Treleaven J. Cervical kinematic training with and without interactive VR training for chronic neck pain—a randomized clinical trial. Man Ther 2015 Feb;20(1):68-78. [doi: 10.1016/j.math.2014.06.008] [Medline: 25066503]
32. Ubillos-Landa S, García-Otero R, Puente-Martínez A. Validation of an instrument for measuring chronic pain in nursing homes. An Sist Sanit Navar 2019 Apr 25;42(1):19-30. [doi: 10.23938/ASSN.0390] [Medline: 30706908]
33. Úbeda-D’Ocasar E, Valera-Calero JA, Hervás-Pérez JP, Caballero-Corella M, Ojedo-Martín C, Gallego-Sendarrubias GM. Pain intensity and sensory perception of tender points in female patients with fibromyalgia: a pilot study. Int J Environ Res Public Health 2021 Feb 4;18(4):1461. [doi: 10.3390/ijerph18041461] [Medline: 33557288]
34. Ayala Pastorino S, Varaldi Azcoytia G, Illescas Calegari L, Castroman Espasandín Espasandin PJ. Test cuantitativo sensorial con algometría de presión: experiencia preliminar en un servicio de anestesiología [Article in Spanish]. Rev Soc Esp Dolor 2022;28(6):311-318. [doi: 10.20986/resed.2022.3947/2021]
35. Ge HY, Nie H, Madeleine P, Danneskiold-Samsøe B, Graven-Nielsen T, Arendt-Nielsen L. Contribution of the local and referred pain from active myofascial trigger points in fibromyalgia syndrome. Pain 2009 Dec 15;147(1-3):233-240. [doi: 10.1016/j.pain.2009.09.019] [Medline: 19819074]
36. Walthall H, Jenkinson C, Boulton M. Living with breathlessness in chronic heart failure: a qualitative study. J Clin Nurs 2017 Jul;26(13-14):2036-2044. [doi: 10.1111/jocn.13615] [Medline: 27731919]
37. Jovicic MD, Konstantinovic LM, Grgurevic AD, et al. Validation of the Neck Disability Index in Serbian patients with cervical radiculopathy. J Manipulative Physiol Ther 2018;41(6):496-502. [doi: 10.1016/j.jmpt.2017.10.018] [Medline: 30107938]
38. Roelofs J, Sluiter JK, Frings-Dresen MHW, et al. Fear of movement and (re)injury in chronic musculoskeletal pain: evidence for an invariant two-factor model of the Tampa Scale for Kinesiophobia across pain diagnoses and Dutch, Swedish, and Canadian samples. Pain 2007 Sep;131(1-2):181-190. [doi: 10.1016/j.pain.2007.01.008] [Medline: 17317011]
39. Newman-Beinart NA, Norton S, Dowling D, et al. The development and initial psychometric evaluation of a measure assessing adherence to prescribed exercise: the Exercise Adherence Rating Scale (EARS). Physiotherapy 2017 Jun;103(2):180-185. [doi: 10.1016/j.physio.2016.11.001] [Medline: 27913064]
40. Gauffin J, Hankama T, Kautiainen H, Arkela-Kautiainen M, Hannonen P, Haanpää M. Validation of a Finnish version of the Fibromyalgia Impact Questionnaire (Finn-FIQ). Scand J Pain 2012 Jan 1;3(1):15-20. [doi: 10.1016/j.sjpain.2011.10.004] [Medline: 29913759]
41. Monterdea S, Salvata I, Montulla S, Fernández-Ballart J. Validación de la versión española del Fibromyalgia Impact Questionnaire [Article in Spanish]. Rev Esp Reumatol 2004;31:507-513 [FREE Full text]
42. Esteve-Vives J, Rivera Redondo J, Isabel Salvat Salvat M, de Gracia Blanco M, de Miquel CA. Propuesta de una versión de consenso del Fibromyalgia Impact Questionnaire (FIQ) para la población española. Reumatol Clín 2007 Jan;3(1):21-24. [doi: 10.1016/S1699-258X(07)73594-5] [Medline: 21794391]
43. Herdman M, Badia X, Berra S. El EuroQol-5D: una alternativa sencilla para la medición de la calidad de vida relacionada con la salud en atención primaria. Aten Primaria 2001;28(6):425-429. [doi: 10.1016/S0212-6567(01)70406-4]
44. Wolfe F, Michaud K, Li T, Katz RS. EQ-5D and SF-36 quality of life measures in systemic lupus erythematosus: comparisons with rheumatoid arthritis, noninflammatory rheumatic disorders, and fibromyalgia. J Rheumatol 2010 Feb;37(2):296-304. [doi: 10.3899/jrheum.090778] [Medline: 20032098]
45. Carralero García P, Rosa Hoyos Miranda F, Deblas Sandoval Á, López García M. Calidad del sueño según el Pittsburgh Sleep Quality Index en una muestra de pacientes recibiendo cuidados paliativos. Med Paliat 2013;20(2):44-48. [doi: 10.1016/j.medipa.2012.05.005]
46. Salaffi F, Farah S, Mariani C, Sarzi-Puttini P, Di Carlo M. Validity of the Central Sensitization Inventory compared with traditional measures of disease severity in fibromyalgia. Pain Pract 2022 Nov;22(8):702-710. [doi: 10.1111/papr.13162] [Medline: 36097821]
47. Prushansky T, Dvir Z. La prueba de la movilidad cervical: metodología e implicaciones clínicas [Article in Spanish]. Osteopat Cient 2008;3:108-114. [doi: 10.1016/S1886-9297(08)75759-X]

XSL•FO
RenderX

48. Collado-Mateo D, Domínguez-Muñoz FJ, Adsuar JC, Merellano-Navarro E, Olivares PR, Gusi N. Reliability of the timed up and go test in fibromyalgia. Rehabil Nurs 2018;43(1):35-39. [doi: 10.1002/rnj.307] [Medline: 27781288]
49. Beaudart C, Rolland Y, Cruz-Jentoft AJ, et al. Assessment of muscle function and physical performance in daily clinical practice: a position paper endorsed by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO). Calcif Tissue Int 2019;105(1):1-14. [doi: 10.1007/s00223-019-00545-w]
50. Declaración de Helsinki de la AMM [Web page in Spanish]. Asociación Médica Mundial. 2024. URL: https://www.wma.net/es/policies-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/ [accessed 2026-01-05]
51. Cetin H, Kose N, Oge HK. Virtual reality and motor control exercises to treat chronic neck pain: a randomized controlled trial. Musculoskelet Sci Pract 2022 Dec;62(August):102636. [doi: 10.1016/j.msksp.2022.102636] [Medline: 35952621]
52. Christensen SWM, Almsborg H, Vain TS, Vaegter HB. The effect of virtual reality on cold pain sensitivity in patients with fibromyalgia and pain-free individuals: a randomized crossover study. Games Health J 2023 Aug;12(4):295-301. [doi: 10.1089/g4h.2022.0138] [Medline: 36454199]
53. Garcia-Palacios A, Herrero R, Vizcaíno Y, et al. Integrating virtual reality with activity management for the treatment of fibromyalgia: acceptability and preliminary efficacy. Clin J Pain 2015 Jun;31(6):564-572. [doi: 10.1097/AJP.0000000000000196] [Medline: 25551475]
54. Darnall BD, Krishnamurthy P, Tsuei J, Minor JD. Self-administered skills-based virtual reality intervention for chronic pain: randomized controlled pilot study. JMIR Form Res 2020 Jul 7;4(7):e17293. [doi: 10.2196/17293] [Medline: 32374272]
55. Herrero R, García-Palacios A, Castilla D, Molinari G, Botella C. Virtual reality for the induction of positive emotions in the treatment of fibromyalgia: a pilot study over acceptability, satisfaction, and the effect of virtual reality on mood. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):379-384. [doi: 10.1089/cyber.2014.0052] [Medline: 24892201]
56. Tejera DM, Beltran-Alacreu H, Cano-de-la-Cuerda R, et al. Effects of virtual reality versus exercise on pain, functional, somatosensory and psychosocial outcomes in patients with non-specific chronic neck pain: a randomized clinical trial. Int J Environ Res Public Health 2020 Aug 16;17(16):5950. [doi: 10.3390/ijerph17165950] [Medline: 32824394]

Abbreviations

CG: control group
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-Harms: Consolidated Standards of Reporting Trials—Harms Extension
CSI: Central Sensitization Inventory
EQ-5D: EuroQol Dimensions
FIQ: Fibromyalgia Impact Questionnaire
FM: fibromyalgia
NDI: Neck Disability Index Questionnaire
ROM: range of motion
TSK: Tampa Scale for Kinesiophobia
VAS: visual analog scale
VR: virtual reality

Edited by S Munce; submitted 23.Jul.2025; peer-reviewed by D Parmak, GGD Alencar; revised version received 10.Dec.2025; accepted 19.Dec.2025; published 23.Jan.2026.

Please cite as:
Úbeda-D'Ocasar E, Moreno-Crespo Y, Cimadevilla-Fernández-Pola E, Hernández-Lougedo J, Navas-Mosqueda Á, Caballero-Corella M, Mayoral-Gonzalo N, Pedauyé-Rueda B, Fernández-Aceñero MJ, Hervás-Pérez JP, Ojedo-Martín C
Therapeutic Use of Virtual Reality for Patients With Fibromyalgia and Chronic Neck Pain: Randomized Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e81158
URL: https://rehab.jmir.org/2026/1/e81158
doi:10.2196/81158

© Edurne Úbeda-D'Ocasar, Yaiza Moreno-Crespo, Eduardo Cimadevilla-Fernández-Pola, Juan Hernández-Lougedo, Álvaro Navas-Mosqueda, Mario Caballero-Corella, Noemí Mayoral-Gonzalo, Blanca Pedauyé-Rueda, María Jesús Fernández-Aceñero, Juan Pablo Hervás-Pérez, Cristina Ojedo-Martín. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of High-Intensity Laser Therapy Combined With Plantar Fascia Stretching Exercises in the Treatment of Plantar Fasciitis: Randomized, Double-Blind, Sham-Controlled Trial

Nantaporn Jitpimolmard, MSc, MD; Phonlawat Ouemphancharoen, MD; Preeda Arayawichanon, MD

Department of Rehabilitation Medicine, Faculty of Medicine, Khon Kaen University, 123 Mittraphap Road, Mueang, Khon Kaen, Thailand

Corresponding Author:
Nantaporn Jitpimolmard, MSc, MD
Department of Rehabilitation Medicine, Faculty of Medicine, Khon Kaen University, 123 Mittraphap Road, Mueang, Khon Kaen, Thailand

Abstract

Background: Plantar fasciitis causes heel pain and functional limitations; conservative treatment typically includes plantar fascia and calf stretching. High-intensity laser therapy (HILT) offers deeper photobiomodulation and potential tissue-healing benefits, but robust evidence of added clinical benefit remains limited.

Objective: This study aims to evaluate the efficacy of HILT combined with plantar fascia stretching exercises compared with a sham control for the treatment of plantar fasciitis.

Methods: This study was designed as a randomized, double-blind, sham-controlled trial conducted at the outpatient clinic of a university hospital. Participants were randomly allocated into 2 groups: the HILT group and the sham treatment group. Both groups received 9 treatment sessions over 3 weeks. The HILT group received active laser therapy, while the sham group received identical treatment without laser emission. The HILT used a wavelength of 1064 nm in continuous mode, with a power output of 12 W applied for 250 seconds, delivering an energy density of 120 J/cm² applied to a 25-cm² area for a total energy of 3000 J. In addition to the assigned interventions, all participants performed a standardized self-stretching exercise program targeting the plantar fascia and Achilles tendon throughout the study period. The primary outcome was pain intensity measured using a visual analog scale. Secondary outcomes included ultrasonographic measurement of plantar fascia thickness (PFT) and the subjective foot and ankle ability measure (FAAM) score, recorded before and immediately after the intervention.

Results: A total of 34 patients diagnosed with unilateral plantar fasciitis were enrolled in this study. Based on intragroup comparison, both groups demonstrated statistically significant improvements in all outcomes compared with baseline (*P*<.001). However, no significant differences were found between the 2 groups across all outcomes. The mean difference in pain reduction, measured by the visual analog scale, was –35.3 (95% CI –45.3 to –25.0) mm in the HILT group and –30.4 (95% CI –46.3 to –14.4) mm in the sham group (–5.0 mm, 95% CI –14.3 to 4.3; *P*=.59). Similarly, reductions in PFT and improvements in FAAM scores showed no significant differences between groups (mean difference –0.02 mm, 95% CI –0.2 to 0.1; *P*=.90 and 5.6 points, 95% CI –1.1 to 12.4; *P*=.40, respectively).

Conclusions: There was no additional clinical effectiveness of HILT on pain reduction, decreased PFT, or increased FAAM scores compared with sham laser when combined with standard stretching exercises for plantar fasciitis and the Achilles tendon.

Trial Registration: Thai Clinical Trials Registry TCTR20200724001; https://www.thaiclinicaltrials.org/show/TCTR20200724001

(JMIR Rehabil Assist Technol 2026;13:e77419) doi:10.2196/77419

KEYWORDS

plantar fasciitis; high-intensity laser therapy; lasers; pain measurement; ultrasonography; fascia; foot; ankle

Introduction

Plantar fasciitis is a common and often disabling condition characterized by pain at the plantar aspect of the heel, which can substantially limit daily activities [1]. Risk factors include increasing age, elevated body mass index, prolonged standing, and foot posture abnormalities that increase stress on the plantar fascia [2]. First-line management is conservative and heterogeneous, encompassing manual therapy, night splints, taping, orthoses, and exercise programs; among these, stretching of the plantar fascia and associated calf musculature is widely recommended and frequently used in clinical practice [3].

For this study, passive stretching was chosen as the standardized backbone intervention. Mechanistically, stretching directly targets gastrocnemius–soleus tightness and plantar fascia strain, two important contributors to symptom persistence, and has been shown in clinical studies to reduce pain and improve function in the short to medium term [4]. Practically, passive stretching is easily standardized, low cost, self-administered, well tolerated, and associated with good patient adherence compared with more resource-intensive or operator-dependent modalities (eg, manual therapy or supervised exercise programs). Using a single, well-established conservative treatment as the common background therapy therefore reduces between-participant variability and more clearly isolates any incremental effect of the adjunctive intervention under study.

Adjunctive physical modalities—including extracorporeal shockwave therapy, ultrasound, and low-intensity laser therapy—are commonly used to enhance recovery [5]. Laser therapy delivers photons that may stimulate cellular processes involved in tissue repair, including increased adenosine triphosphate production, enhanced collagen synthesis, and improved microcirculation [6-8]. Systematic reviews indicate that low-intensity laser therapy can reduce pain in plantar fasciitis [5,9-12]. In theory, high-intensity laser therapy (HILT) has greater tissue penetration and may produce stronger collagen stimulation, greater blood flow effects, and more pronounced anti-inflammatory responses than low-intensity devices [13,14]. However, studies comparing HILT with low-intensity laser or sham treatment have reported conflicting results [15,16], and recent meta-analyses have not comprehensively evaluated HILT for plantar fasciitis.

Prior HILT studies have often been limited by methodological weaknesses, inconsistent control treatments, and a scarcity of objective outcomes such as ultrasound-measured plantar fascia thickness (PFT) or functionally relevant measures. To address these gaps, this randomized, sham-controlled trial evaluated the efficacy of HILT added to a standardized passive stretching program compared with sham laser plus the same conservative regimen. We hypothesized that adjunctive HILT would produce greater improvements in pain, functional capacity, and PFT compared with conservative care using sham laser.

Methods

Study Design

This study was a prospective, randomized, double-blinded (participant and assessor), sham-controlled trial conducted in an outpatient rehabilitation setting at Srinagarind Hospital, Khon Kaen University.

Ethical Considerations

The study protocol was approved by the Khon Kaen University Ethics Committee (HE621318). Written informed consent was obtained from all participants prior to enrollment. Participant privacy and confidentiality were strictly maintained. All data were deidentified prior to analysis, and personal identifiers were removed. The dataset was stored securely and was accessible only to authorized members of the research team. Participants received compensation of 900 for their participation. Based on an exchange rate of US $1 to 31, this is approximately US $29 per participant. Compensation was provided regardless of study completion.

The study adhered to the CONSORT (Consolidated Standards of Reporting Trials) guidelines.

Participants

A total of 36 participants diagnosed with unilateral plantar fasciitis were recruited between January 2020 and September 2020.

Inclusion Criteria

Participants were eligible for inclusion if they met all of the following criteria:

- Adults aged 18 years or older
- Foot pain persisting for a minimum of 6 weeks
- Diagnosis of plantar fasciitis confirmed by a rehabilitation medicine physician based on clinical history, characterized by pain primarily felt during the first steps after waking, which worsens with prolonged weight-bearing activities throughout the day [17]
- Tenderness at the medial calcaneal tubercle upon palpation [17]

Exclusion Criteria

Participants were excluded from the study if they met any of the following criteria:

- History of foot surgery
- Chronic inflammatory arthritis (eg, rheumatoid arthritis, gout, systemic lupus erythematosus)
- Conditions affecting sensory perception (eg, diabetes mellitus with peripheral neuropathy)
- History of corticosteroid injection to the plantar fascia within the preceding 6 months
- Neuropathic pain in the foot, or evidence of intrinsic foot muscle weakness
- Pregnancy

Baseline Assessment

At baseline, the following data were collected: age, gender, body mass index, side with the most severe pain, and Foot Posture Index (FPI-6) score.

Randomization and Blinding

Participants were randomly assigned to either the HILT group or the sham laser (control) group using a permuted block randomization method with a block size of 4. The randomization sequence was computer generated and concealed within sequentially numbered, opaque, sealed envelopes. Envelopes were opened by the treating therapist immediately prior to the first treatment session. Participants were blinded to their treatment allocation. The assessor performing outcome measurements was also blinded to treatment allocation.

Intervention

Overview

All participants received 9 treatment sessions administered every 3 days. Both groups received a standardized home exercise program consisting of plantar fascia and Achilles tendon stretching exercises. Participants were instructed to perform the following stretch while seated or supine: with a towel or strap around the forefoot, the knee was extended and the forefoot pulled toward the body, maintaining dorsiflexion for 30 seconds. The exercise was performed for 10 repetitions per session, once

XSL•FO
RenderX

daily. Adherence to the home exercise program was monitored using patient logs.

HILT Group

HILT was administered by a trained medical doctor using the BTL-6000 high-intensity laser device (BTL Industries) operating at a wavelength of 1064 nm. The treatment parameters were as follows: continuous mode, power output of 12 W applied for 250 seconds, energy density of 120 J/cm² delivered over a 25-cm² area, resulting in a total energy of 3000 J [15,16]. The participant was positioned prone with the foot in a neutral position. The laser was applied using a 10-mm pen applicator and swept over the calcaneal insertion and along the medial border of the plantar fascia.

Sham Laser Group

The sham treatment mimicked the HILT procedure, including the visual and auditory cues of the laser device, but without actual laser energy delivery. The same treatment protocol as the HILT group was followed in terms of positioning, applicator movement, and device sounds.

Medication and High Plantar Fascial Loading Activity

Participants were instructed to maintain their usual pain medication regimen, which was limited to 500 mg of paracetamol, and not to exceed 3000 mg per day. They were also advised to strictly refrain from strenuous exercises that involve high plantar fascial loading. Crucially, the use of anti-inflammatory drugs was strictly prohibited during the study per the protocol. The number of pain medication pills taken was recorded throughout the study period.

Outcome Measures

Outcomes were assessed at baseline and immediately postintervention by a blinded assessor. Pain intensity was measured using a 100-mm visual analog scale (VAS). Participants were asked to recall the pain experienced during the first 3 steps taken in the morning. Scores were assessed immediately preintervention and again via telephone 1 day after the final session to capture pain during participants' first steps in the morning following the treatment period.

PFT was measured by ultrasound using a LOGIQ E Premium system (GE Healthcare) with a 5 MHz linear probe at the calcaneal insertion. Participants were positioned prone with feet in a neutral position, and the feet were placed at the end of the examination bed. The ultrasound probe was placed perpendicular to the plantar fascia on the plantar surface, and the thickness of the thickest portion was measured [18]. The reported value is the mean of 3 measurements.

Foot and ankle function was assessed using the activities of daily living subscale of the Thai version of the foot and ankle ability measure (FAAM). The FAAM subjective form evaluates foot and ankle function for daily activities and sports; however, only the activities of daily living subscale was used in this study. The reliability of the FAAM was examined by Arunakul et al [19], who reported strong internal consistency.

Data Analysis

All analyses were conducted using Python (version 3.10), executed in Google Colaboratory. Continuous variables are presented as mean (SD) or as median (IQR) for nonnormally distributed data, while categorical variables are presented as frequency (percentage). Normality of continuous data was assessed using the Shapiro-Wilk test.

The primary analysis followed the intention-to-treat (ITT) principle. Missing outcome data (5%) were addressed using multiple imputation by chained equations with 20 imputations. The imputation model included baseline outcome values, treatment group, and relevant covariates. Each imputed dataset was analyzed separately, and the results were pooled using Rubin's rules.

Between-group comparisons of continuous posttreatment outcomes (primary and secondary) were conducted using independent *t* tests when distributional assumptions were met. For nonnormally distributed outcomes, between-group comparisons used the Mann-Whitney *U* test on observed data, with additional sensitivity analyses performed on the imputed datasets. Within-group changes were evaluated using paired *t* tests or the Wilcoxon signed rank test, as appropriate.

A per-protocol analysis including participants who completed at least 7 treatment sessions was conducted as a sensitivity analysis. Effect sizes (Cohen *d*) and 95% CIs were reported for parametric comparisons. All statistical tests were 2-tailed, and a *P* value <.05 was considered statistically significant.

Sample Size Calculation

Based on prior studies [15], a sample size of 16 participants per group was required to detect a mean difference of 15 (SD 15) mm on the VAS, with 80% power and a 2-sided α of .05. To allow for an anticipated 10% dropout rate, 36 participants (18 per group) were recruited.

Withdrawal and Discontinuation Criteria

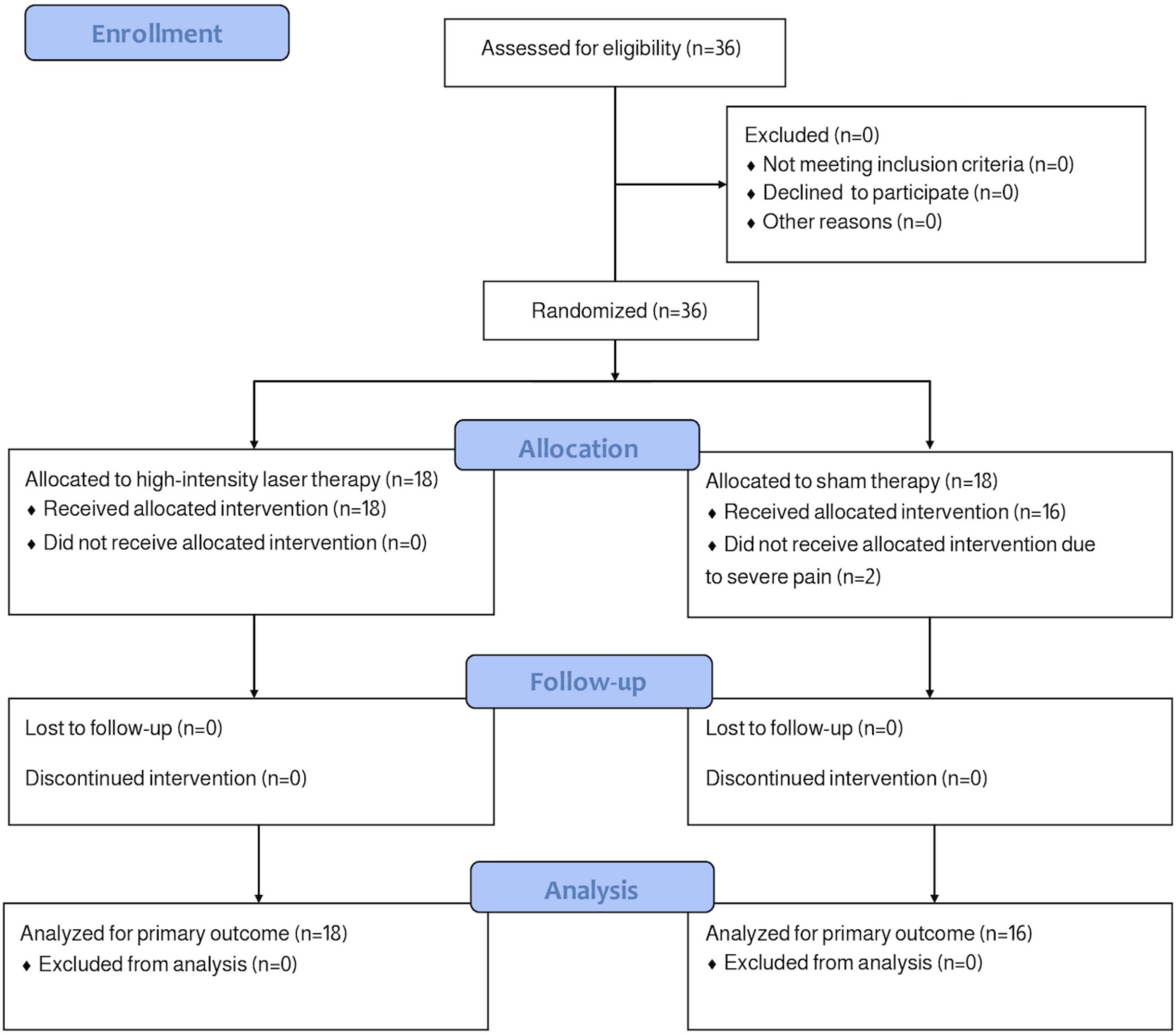
Participants could withdraw from the study at any time without affecting their care. Discontinuation criteria included withdrawal of consent or adverse effects (eg, burns, eye irritation).

Results

A total of 36 participants were enrolled, with 34 participants completing the study. Two participants withdrew due to severe pain and opted for alternative treatments (Figure 1).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram outlining the identification, enrollment, and allocation of participants to the 2 groups: high-intensity laser therapy group and sham group.



The HILT group included 18 participants (n=2 men, n=16 women; mean age 46, SD 11.4 years), while the sham laser group included 16 participants (n=1 man, n=15 women; mean age 49, SD 13.0 years). The mean body mass index was 23.3 (SD 2.9) kg/m² in the HILT group and 26.0 (SD 2.8) kg/m² in the sham laser group (*P*=.20). The FPI-6 mean scores were –0.4 (SD 1.1) in the HILT group and –0.6 (SD 1.1) in the sham group, indicating a slightly supinated foot posture [20]. Moreover, no patients in our study had markedly supinated or pronated feet. There was no significant difference in the duration of foot pain, with a mean of 14.7 (SD 4.2) weeks in the HILT group and 14.8 (SD 4.4) weeks in the sham group (Table 1).

XSL•FO
RenderX

Table . Demographic characteristics of the 2 study groups.

Characteristics	High-intensity laser therapy group (n=18)	Sham group (n=16)	*P* value
Age (y), mean (SD)	46 (11.4)	49 (13.0)	.52
Gender, n (%)			.61
Male	2 (11.1)	1 (6.3)	
Female	16 (88.9)	15 (93.7)	
FPI-6[a] (points), mean (SD)	–0.4 (1.1)	–0.6 (1.1)	.51
BMI (kg/m²), mean (SD)	23.3 (2.9)	26.0 (2.8)	.21
Duration of foot pain (wk), mean (SD)	14.7 (4.2)	14.8 (4.4)	>.99

[a]FPI-6: Foot Posture Index.

The ITT analysis (n=36; pooled across 20 imputations) showed the following posttreatment results. For pain (VAS), the mean posttreatment score was 23.0 (SD 17.5) mm in the HILT group versus 26.8 (SD 20.7) mm in the sham group. The between-group mean difference (laser and sham groups) in posttreatment VAS was –5.0 mm (95% CI –14.3 to 4.3; *P*=.59). Cohen *d* for the between-group difference at posttreatment (pooled SD) was small (d≈–0.20).

Mean PFT decreased from 4.8 (SD 1.4) mm to 4.1 (SD 1.2) mm (mean change –0.7 mm, 95% CI –1.0 to –0.5; *P*<.001) in the HILT group. For PFT, the median decreased from 5.3 (IQR 4.5 - 5.4) mm to 4.5 (IQR 3.7 - 4.7) mm (median change –0.7 mm, 95% CI –1.0 to −0.5; *P*<.001) in the sham group. The between-group difference in change was –0.02 mm (95% CI –0.2 to 0.1; *P*=.90). Because values in the sham group were nonnormally distributed, a nonparametric test (Mann-Whitney *U* test) was used for between-group comparisons of this outcome, and Cohen *d* was not reported for this comparison.

For FAAM, the mean posttreatment scores were 82.3 (SD 12.8) in the HILT group and 82.8 (SD 17.7) in the sham group. Within-group improvements were significant in both groups. The mean change was 18.7 (95% CI 10.4 to 27.1; *P*<.001) in the HILT group and 14.5 (95% CI 3.2 to 25.8; *P*=.02) in the sham group. The between-group mean difference in change was 5.6 (95% CI –1.1 to 12.4; *P*=.40). Cohen *d* for the between-group posttreatment comparison was very small (d≈0.03; Table 2).

Table . Comparison of pain intensity measured by the 100-mm VAS, plantar fascial thickness, and FAAM score.

Outcomes	High-intensity laser therapy group (n=18)				Sham group (n=16)				High-intensity laser therapy versus sham	
	Before, mean (SD)	After, mean (SD)	Mean difference (95% CI)	*P* value	Before, mean (SD)	After, mean (SD)	Mean difference (95% CI)	*P* value	Between-group mean difference (95% CI)	*P* value
VAS[a] (mm)	58.3 (19.1)	23.0 (17.5)	–35.3[b] (–45.3 to –25.0)	<.001	57.1 (22.2)	26.8 (20.7)	–30.4[b] (–46.3 to –14.4)	.001	–5.0 (–14.3 to 4.3)[c]	.59
Plantar fascia thickness (mm)	4.8 (1.4)	4.1 (1.2)	–0.7[b] (–1.0 to –0.5)	<.001	5.3 (4.5 to 5.4)[d]	4.5 (3.7 to 4.7)[d]	–0.7[e] (–1.0 to –0.5)	<.001	–0.02 (–0.2 to 0.1)[c]	.90
FAAM[f] (point)	63.6 (13.6)	82.3 (12.8)	18.7[b] (10.4 to 27.1)	<.001	68.3 (23.8)	82.8 (17.7)	14.5[b] (3.2 to 25.8)	.02	5.62 (–1.1 to 12.4)[c]	.40

[a]VAS: visual analog scale.

[b]Paired *t* test was used.

[c]Independent *t* test was used.

[d]Median (IQR).

[e]Wilcoxon signed rank test was used.

[f]FAAM: foot and ankle ability measure.

The per-protocol analysis (n=34 completers) produced results consistent with the ITT analysis; between-group differences in change scores remained small and nonsignificant for VAS, PFT, and FAAM, supporting the robustness of the findings. Among participants with plantar fasciitis who had a history of strenuous weight-bearing exercise, 1 participant (a runner) in the HILT

XSL•FO
RenderX

group was advised to refrain from such activity during the treatment period. Furthermore, 2 participants (1 from each group) reported taking paracetamol at a dosage of 2000 mg per day for 2 days. No adverse events (eg, increased pain, skin burns) were reported in the HILT group. Two participants in the sham group withdrew owing to pain intolerance. Adherence to the exercise protocol did not differ significantly between groups: there were a mean 15.1 (SD 3.8) sessions in the HILT group versus 17.5 (SD 3.7) in the sham group (*P*=.06).

Discussion

This randomized controlled trial assessed the effectiveness of HILT combined with conventional therapy (stretching exercises) compared with sham laser plus conventional therapy in patients with plantar fasciitis. A key strength of our methodology, which distinguishes it from many prior trials, was the inclusion of objective structural outcomes (ultrasound) and validated quality-of-life assessment (FAAM), coupled with a rigorous sham control. Both treatment arms showed significant improvements in pain scores, a reduction in PFT, and an increase in foot and ankle function following treatment. However, HILT did not demonstrate superiority over the sham intervention.

This primary finding—that HILT offered no clear added benefit—aligns with several previous investigations. For instance, Tkocz et al [21], studying chronic heel pain (including plantar fasciitis and calcaneal spurs) using a HILT device at a lower power (7 W) and high total energy, similarly found no clear advantage of HILT over sham combined with ultrasound. Furthermore, Naruseviciute and Kubilius [16] also reported no superiority of HILT versus low-intensity laser therapy regarding either pain or fascial thickness. Conversely, Akkurt et al [22] observed improved outcomes when HILT was combined with insoles; however, the absence of a true sham laser control in that study makes it impossible to definitively exclude a significant placebo contribution.

The reasons for this consistent lack of incremental benefit are likely multifactorial. High-intensity laser protocols vary widely across trials concerning power, energy per session, mode, number of sessions, and application technique [15,16,21,22]. This substantial heterogeneity may partly explain discrepant results in the literature, as these parameters dictate whether the tissue effects are primarily photobiomodulatory or thermal. Nevertheless, the majority of trials to date do not show a consistent, clinically meaningful incremental effect of HILT over standard conservative care. Future research must prioritize robust dose-finding studies, provide complete dosimetry and application details, and include mechanistic measures to determine whether any specific parameter set reliably improves outcomes in plantar fasciitis before routine adoption.

Beyond methodological variations, anatomical considerations may temper the potential benefit of HILT. The plantar fascia lies only a few millimeters below the skin, varying by location and individual subcutaneous thickness. Consequently, the deeper penetration advantage often ascribed to HILT may be less relevant for plantar fasciitis compared with deeper targets (eg, knee osteoarthritis), where reaching intra-articular structures is more pertinent [23]. This superficial anatomical location of the target may inherently limit the potential incremental benefit of HILT in this disease.

Given that all 3 primary outcomes improved significantly in both study arms, it strongly suggests that the standardized stretching program was the major driver of clinical benefit. Plantar fascia and calf stretching target the gastrocnemius–soleus complex and the fascia because tightness increases tensile strain on the arch [4]. Mechanistically, regular stretching likely contributed to the reduction in pain and functional gains by increasing ankle dorsiflexion, lowering passive tension transmitted to the fascial insertion, and improving tissue viscoelasticity.

The consistent improvements observed in our objective and patient-centric measures validate the effectiveness of this conservative approach. The reduction in PFT, quantified objectively via ultrasound, confirms that the therapy successfully induced a positive structural change in the affected tissue, moving beyond mere symptomatic relief. PFT is considered the anatomical hallmark of chronic plantar fasciitis (reflecting chronic microtearing and disorganized collagen repair) [18]. Similarly, the significant increase in scores on the FAAM further substantiates the clinical relevance of the treatment. Since the FAAM is a validated patient-reported outcome measure focusing specifically on function and participation in daily living and sports activities [19], the documented improvements indicate that the structural and pain changes successfully translated directly into better functional capacity for participants.

An alternative explanation for the null between-group difference lies in statistics. The study was powered to detect a 15-mm VAS difference, yet the observed between-group difference was statistically small (only –5.0 mm). This small effect size combined with substantial variability reduced the power to find a statistically significant difference. Additionally, because many participants improved substantially with stretching (and some used short courses of acetaminophen), a ceiling effect is possible, leaving little room for additional measurable improvement attributable solely to HILT. Reporting responder rates, change-score distributions, and conducting larger trials or trials enrolling participants with higher baseline pain would help determine whether a clinically meaningful subgroup benefit exists.

Most participants in our sample had near-normal alignment with slight supination, and baseline alignment was balanced between the 2 groups, so imbalance in foot posture is unlikely to account for the null between-group findings. Nonetheless, because plantar fascia loading differs across foot types and many patients with plantar fasciopathy present with pes planus or overpronation [2], the sample's alignment profile may limit external validity. Future trials should objectively quantify foot posture (eg, FPI-6, navicular drop, calcaneal eversion, or dynamic pressure measures), consider stratified randomization or prespecified subgroup analyses by alignment, and, where appropriate, tailor background conservative care to alignment to test whether foot posture modifies the effect of adjunctive treatments such as HILT.

In terms of safety and adherence, a small number of participants used brief acetaminophen therapy, and 1 HILT participant was

XSL•FO
RenderX

advised to avoid strenuous running during treatment. This activity modification is an independent potential confounder. Importantly, however, no adverse events were reported among HILT recipients, supporting short-term tolerability in our sample. Although 2 participants in the sham group withdrew because of pain intolerance, both ITT and per-protocol sensitivity analyses produced similar results, supporting the robustness of our conclusions.

Finally, this study assessed outcomes only immediately postintervention, meaning the durability of effects remains unknown. Furthermore, the lack of standardized HILT protocols for plantar fasciitis complicates cross-trial comparison. Future research must prioritize longer follow-up, dose-finding studies, standardized reporting of dosimetry, and adequately powered trials (including stratified or targeted enrollment) to identify the specific patient subgroups most likely to benefit from HILT.

This study provides high-quality evidence that HILT does not offer additional benefits over standard stretching exercises in the treatment of plantar fasciitis. Given the statistically similar outcomes between the HILT and sham laser groups, our findings reinforce the effectiveness of conservative treatment approaches as the primary management strategy for plantar fasciitis.

Acknowledgments

The authors would like to express their sincere gratitude for the financial support provided, which made this study possible.

Funding

This research was supported by the Research and Innovation Division, Faculty of Medicine, Khon Kaen University, Thailand.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH (V 1.6.1) checklist.
[PDF File, 1162 KB - rehab_v13i1e77419_app1.pdf]

References

1. Scher DL, Belmont PJ, Bear R, Mountcastle SB, Orr JD, Owens BD. The incidence of plantar fasciitis in the United States military. J Bone Joint Surg Am 2009 Dec;91(12):2867-2872. [doi: 10.2106/JBJS.I.00257] [Medline: 19952249]
2. Johannsen FE, Rydahl JP, Jacobsen AS, Brahe CCH, Magnusson PS. Foot posture and ankle dorsiflexion as risk factors for developing Achilles tendinopathy and plantar fasciitis: a case-control study. Foot Ankle Int 2024 Dec;45(12):1380-1389. [doi: 10.1177/10711007241281289] [Medline: 39422991]
3. Koc TA, Bise CG, Neville C, Carreira D, Martin RL, McDonough CM. Heel pain - plantar fasciitis: revision 2023. J Orthop Sports Phys Ther 2023 Dec;53(12):CPG1-CPG39. [doi: 10.2519/jospt.2023.0303] [Medline: 38037331]
4. DiGiovanni BF, Nawoczenski DA, Lintal ME, et al. Tissue-specific plantar fascia-stretching exercise enhances outcomes in patients with chronic heel pain. a prospective, randomized study. J Bone Joint Surg Am 2003 Jul;85(7):1270-1277. [doi: 10.2106/00004623-200307000-00013] [Medline: 12851352]
5. Guimarães JDS, Arcanjo FL, Leporace G, et al. Effects of therapeutic interventions on pain due to plantar fasciitis: a systematic review and meta-analysis. Clin Rehabil 2023 Jun;37(6):727-746. [doi: 10.1177/02692155221143865] [Medline: 36571559]
6. Kujawa J, Zavodnik L, Zavodnik I, Buko V, Lapshyna A, Bryszewska M. Effect of low-intensity (3.75-25 J/cm2) near-infrared (810 nm) laser radiation on red blood cell ATPase activities and membrane structure. J Clin Laser Med Surg 2004 Apr;22(2):111-117. [doi: 10.1089/104454704774076163] [Medline: 15165385]
7. De Oliveira MF, Johnson DS, Demchak T, Tomazoni SS, Leal-junior EC. Low-intensity LASER and LED (photobiomodulation therapy) for pain control of the most common musculoskeletal conditions. Eur J Phys Rehabil Med 2021;58(2):282-289. [doi: 10.23736/S1973-9087.21.07236-1]
8. de Oliveira VLC, Silva JA Jr, Serra AJ, et al. Photobiomodulation therapy in the modulation of inflammatory mediators and bradykinin receptors in an experimental model of acute osteoarthritis. Lasers Med Sci 2017 Jan;32(1):87-94. [doi: 10.1007/s10103-016-2089-2] [Medline: 27726041]
9. Dos Santos SA, Sampaio LM, Caires JR, et al. Parameters and effects of photobiomodulation in plantar fasciitis: a meta-analysis and systematic review. Photobiomodul Photomed Laser Surg 2019 Jun;37(6):327-335. [doi: 10.1089/photob.2018.4588] [Medline: 31107161]
10. Guimarães JS, Arcanjo FL, Leporace G, et al. Effect of low-level laser therapy on pain and disability in patients with plantar fasciitis: a systematic review and meta-analysis. Musculoskelet Sci Pract 2022 Feb;57:102478. [doi: 10.1016/j.msksp.2021.102478] [Medline: 34847470]

XSL•FO
RenderX

11. Naterstad IF, Joensen J, Bjordal JM, Couppé C, Lopes-Martins RAB, Stausholm MB. Efficacy of low-level laser therapy in patients with lower extremity tendinopathy or plantar fasciitis: systematic review and meta-analysis of randomised controlled trials. BMJ Open 2022 Sep 28;12(9):e059479. [doi: 10.1136/bmjopen-2021-059479] [Medline: 36171024]
12. Wang W, Jiang W, Tang C, Zhang X, Xiang J. Clinical efficacy of low-level laser therapy in plantar fasciitis: a systematic review and meta-analysis. Medicine (Baltimore) 2019 Jan;98(3):e14088. [doi: 10.1097/MD.0000000000014088] [Medline: 30653125]
13. Kaub L, Schmitz C. Comparison of the penetration depth of 905 nm and 1064 nm laser light in surface layers of biological tissue ex vivo. Biomedicines 2023 May 4;11(5):1355. [doi: 10.3390/biomedicines11051355] [Medline: 37239026]
14. Astri SW, Murdhana N, Nusdwinuringtyas N, Kekalih A, Sunarjo P, Soewito F. The comparison of the low-level laser therapy and high intensity laser therapy on pain and functional ability in knee osteoarthritis. J Indon Med Assoc 2022;72(6):275-283. [doi: 10.47830/jinma-vol.72.6-2022-826]
15. Ordahan B, Karahan AY, Kaydok E. The effect of high-intensity versus low-level laser therapy in the management of plantar fasciitis: a randomized clinical trial. Lasers Med Sci 2018 Aug;33(6):1363-1369. [doi: 10.1007/s10103-018-2497-6] [Medline: 29627888]
16. Naruseviciute D, Kubilius R. The effect of high-intensity versus low-level laser therapy in the management of plantar fasciitis: randomized participant blind controlled trial. Clin Rehabil 2020 Aug;34(8):1072-1082. [doi: 10.1177/0269215520929073] [Medline: 32513018]
17. Goff JD, Crawford R. Diagnosis and treatment of plantar fasciitis. Am Fam Physician 2021 Sep 15;84(6) [FREE Full text]
18. Manske RC, Wolfe C, Page P, Voight M. Diagnostic musculoskeletal ultrasound in the evaluation of the plantar fascia. Int J Sports Phys Ther 2025;20(7):1091-1096. [doi: 10.26603/001c.141177] [Medline: 40620402]
19. Arunakul M, Arunakul P, Suesiritumrong C, Angthong C, Chernchujit B. Validity and reliability of thai version of the foot and ankle ability measure (FAAM) subjective form. J Med Assoc Thai 2015 Jun;98(6):561-567. [Medline: 26219160]
20. Kyung MG, Cho YJ, Lee JH, Shin MS, Park JH, Lee DY. Reliability and radiographic correlation of the foot posture index-6: a multi-rater analysis in symptomatic and asymptomatic individuals. Diagnostics (Basel) 2025 May 12;15(10):1214. [doi: 10.3390/diagnostics15101214] [Medline: 40428206]
21. Tkocz P, Matusz T, Kosowski Ł, et al. A randomised-controlled clinical study examining the effect of high-intensity laser therapy (HILT) on the management of painful calcaneal spur with plantar fasciitis. J Clin Med 2021 Oct 23;10(21):4891. [doi: 10.3390/jcm10214891] [Medline: 34768411]
22. Akkurt F, Akkurt HE, Yılmaz H, Olgun Y, Sen Z. Efficacy of high-intensity laser therapy and silicone insole in plantar fasciitis. Int J Phys Med Rehabil 2018;06(05). [doi: 10.4172/2329-9096.1000484]
23. Wyszyńska J, Bal-Bocheńska M. Efficacy of high-intensity laser therapy in treating knee osteoarthritis: a first systematic review. Photomed Laser Surg 2018 Jul;36(7):343-353. [doi: 10.1089/pho.2017.4425] [Medline: 29688827]

Abbreviations

FAAM: foot and ankle ability measure
FPI-6: Foot Posture Index
HILT: high-intensity laser therapy
ITT: intention-to-treat
PFT: plantar fascia thickness
VAS: visual analog scale

Edited by K Kokorelias; submitted 13.May.2025; peer-reviewed by P Areeudomwong, P Sakulsriprasert, SP Madani, W Namwong; revised version received 22.Dec.2025; accepted 22.Jan.2026; published 20.Feb.2026.

Please cite as:
Jitpimolmard N, Ouemphancharoen P, Arayawichanon P
Efficacy of High-Intensity Laser Therapy Combined With Plantar Fascia Stretching Exercises in the Treatment of Plantar Fasciitis: Randomized, Double-Blind, Sham-Controlled Trial
JMIR Rehabil Assist Technol 2026;13:e77419
URL: https://rehab.jmir.org/2026/1/e77419
doi:10.2196/77419

© Nantaporn Jitpimolmard, Phonlawat Ouemphancharoen, Preeda Arayawichanon. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology,

XSL•FO
RenderX

is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Natural Language Prompts With AI Models for Low-Cost Assistive Software Design: Exploratory Comparative Evaluation

Francesc Antoni Bañuls-Lapuerta[1*], PhD; Vicent Marti-Miralles[1,2*], MSc; Rómulo Jacobo Gónzalez-García[1*], PhD; Gabriel Martínez-Rico[1*], Prof Dr

[1]Campus Capacitas, Valencia Catholic University Saint Vincent Martyr, Carrer de Joaquin Navarro, 37, Burjassot, Spain
[2]Doctorate School, Valencia Catholic University Saint Vincent Martyr, Valencia, Spain
[*]all authors contributed equally

Corresponding Author:
Vicent Marti-Miralles, MSc
Campus Capacitas, Valencia Catholic University Saint Vincent Martyr, Carrer de Joaquin Navarro, 37, Burjassot, Spain

Abstract

Background: This study investigates the capacity of 7 artificial intelligence (AI) models, 5 free and 2 paid, to generate functional software for designing low-cost, personalized assistive products.

Objective: The objective was to determine which models are most effective, accessible, and consistent in supporting nontechnical professionals in developing inclusive digital solutions and to assess the capabilities of commercially available and easy-to-access AI models to generate code from natural language interactions in the shape of a nontechnical assistive technology design process.

Methods: Each AI model was prompted using natural language, without any technical input, to create a Python program that converts an arcade gamepad into an adapted mouse-like controller. Sixteen progressively complex functions were requested through standardized prompts, delivered without additional feedback or correction. Model performance was evaluated based on the number of successfully implemented functions and the average number of prompts required.

Results: Paid models demonstrated markedly superior performance. Gemini Pro (Google) successfully implemented 14 of 16 requested functions with an average of 1.25 (SD 0.45) prompts, while ChatGPT Plus (GPT-5) achieved 11 functions with an average of 1.31 (SD 0.48) prompts. In contrast, free models produced between 0 and 4 functional outcomes, with DeepSeek and Gemini Free ranking the highest within their category. The enhanced outcomes of paid models were linked to improved contextual understanding, greater tolerance for natural language, and reduced conversational drift.

Conclusions: Paid AI models, particularly Gemini Pro and ChatGPT Plus, exhibit strong potential as tools for bridging the gap between health or education professionals and software development. They enable the creation of affordable, user-centered assistive technology without requiring advanced programming skills. Nevertheless, human oversight and foundational literacy in prompt design remain crucial to guarantee functionality, reliability, and ethical use.

(JMIR Rehabil Assist Technol 2026;13:e86786) doi:10.2196/86786

KEYWORDS

artificial intelligence; assistive products; digital accessibility; Gemini; ChatGPT; inclusive software

Introduction

Background

Artificial intelligence (AI) can be defined as the ability of nonhuman systems, machines, or software to simulate functions inherent to human intelligence, such as perceiving, reasoning, learning, planning, and anticipating future situations [1]. The recent advancements in the field have made clear that generative AI will open new scenarios in which human-machine collaboration will enable faster and more adaptive software solutions to meet the needs of people with disabilities [2]. In these contexts, the professional programmer assumes a supervisory role, while users without technical experience can create functional prototypes through intuitive AI-powered interfaces. It is thought that, by 2030, integrated assistants, such as the so-called HyperAssistant, will be capable of accompanying individuals throughout all stages of development, bridging the gap between software conception and implementation [3].

AI is transforming the way software is conceived and produced by allowing nontechnical users to participate actively in programming. Traditionally, this process was restricted to those mastering specific languages and methodologies; however, generative models in the shape of conversational AI systems have reduced the gap between natural language and coding, fostering digital inclusion and democratizing access to digital creation [4,5]. Generative assistants increase productivity and allow users with limited skills to build functional prototypes by

XSL•FO
RenderX

using natural language–oriented programming, where users express requirements in everyday language and AI translates them into executable code, bridging human thought and computational structures. This AI integration in development environments additionally enhances efficiency through real-time suggestions and code autocompletion, while warning against technological overreliance [6-8].

Design and Characteristics of Effective Prompts for Interaction With AI

The dialogue between a user and an AI system depends on how instructions or prompts, as they are known in the AI environment, are formulated. Prompts act as mediators between human intention and automatic language generation, shaping the trajectory of interaction, structuring intent, and bridging human goals with machine logic [9]. The first key element to achieving functional prompts is goal definition. Effective communication with conversational systems requires explicit purposes that limit and contain the semantic context. A prompt without a defined end, such as "tell me about economics," produces generic responses, whereas "summarize in 200 words the effects of remote work on business productivity" provides focus and precision [10].

The second principle is linguistic specificity, referring to the fact that grammatically well-structured prompts enhance coherence. Since AI operates through probabilistic predictions, lexical clarity, thematic delimitation, and details about format or audience reduce ambiguity [11-13]. The internal structure should ideally include 5 essential elements: topic, style, tone, format, and context. It should also contain action verbs (such as analyze, compare, or synthesize) to guide specific cognitive-like operations [14,15]. This level of precision is fundamental, as prompt design directly influences the reasoning processes activated by the model [16].

Using these key principles and unifying different prompt explanations from the literature, we could divide prompts into 6 specific and differentiated types with their own goals (Table 1).

Table . Types of prompts used.

Prompt type	Goal	Example
Zero-shot or decision	Execute a task or produce a direct response without prior examples [17].	Explain the concept of neuroplasticity in simple terms.
Few-shot or iterative	Guide generation using prior examples [12,18].	Example 1: write a scientific abstract. Example 2: create a new one about artificial intelligence.
Chain of thought	Promote logical processes or explanations step-by-step [19].	Explain, step by step, how you arrive at the conclusion about the benefits of automation.
Sequential	Develop an idea or process step by step [11].	First, define the concept, then provide an example, and finally conclude.
Argumentative	Request a reasoned, well-supported position [11].	Argue why automation can be beneficial.
Verification	Review one's own output and detect errors or biases [20].	Review your previous answer and indicate possible inaccuracies.

Communicative Limitations of Nonexpert Users in Interactions With AI

Several studies agree that interactions between users and AI systems present communicative limitations stemming from the spontaneous and imprecise use of natural language. Users without technical training often formulate ambiguous prompts, making it difficult for the system to interpret communicative intent [21]. Additionally, the absence of nonverbal cues and the limited adaptability of artificial discourse exacerbate these differences, revealing that the effectiveness of interaction depends on the precision and structure of the language used [22].

From a qualitative perspective, users tend to treat chatbots as human interlocutors, reproducing communicative patterns based on empathy and reciprocity [23]. Users tend to apply Grice conversational maxims when evaluating virtual assistants. These maxims include quantity (providing the right amount of information), quality (saying only what is true and verifiable), manner (expressing oneself clearly and orderly, avoiding ambiguity), and relevance (keeping responses focused on the topic at hand) [24]. Complementarily, anthropomorphism, or trying to treat AI like a human, seems to induce a sort of relational dissonance that clearly shows a contradiction as users understand AI as nonhuman yet interact with it as if it were a human [25,26].

Generative models do not interpret implicit intentions but rather textual correlations; therefore, effectiveness depends on the specificity of the prompt [12,14]. Many users are unaware of this mechanism and rely on trial and error [27,28]. Furthermore, the absence of mental models about AI processing creates a cognitive gap between intention and response, reinforced by interfaces that simulate naturalness while not processing information cognitively naturally [29,30].

In summary, most users lack the metalinguistic skills required to design effective prompts, and bridging this gap demands a form of communicative literacy oriented toward AI.

AI in Education: Barriers, Assistive Products, and Opportunities for Digital Inclusion

At this point, it is essential to analyze how AI can be integrated into educational and disability support contexts, expanding

XSL•FO
RenderX

opportunities for student participation and learning. This link between technology and inclusion requires examining the barriers that still limit its effective use in classrooms. The analysis of possibilities for students with disabilities within the school setting must address the obstacles that hinder their access and active participation, emphasizing the importance of identifying factors that prevent the functional use of educational materials [31]. Students with disabilities face multiple barriers—including negative attitudes, lack of appropriate resources, and insufficient technological accessibility—which restrict their access to inclusive and equitable education [32].

Since the COVID-19 pandemic, information and communication technology has played a central role both in education and society [33]. However, the lack of technological accessibility remains a major barrier, as current learning materials heavily depend on digital tools. This not only affects curricular access but also the acquisition of digital skills essential for adulthood. Studies and hiring companies, such as iHire, highlight that 76% of people seek employment online, illustrating the importance of digital literacy for future labor inclusion [34,35].

To address these barriers, assistive products (APs) are key tools for ensuring the educational and social participation of students with disabilities. According to ISO 9999:2022, APs are products that optimize functioning and reduce disability, serving as intermediaries between personal abilities and environmental demands. The previous version of this standard (ISO 9999:2016) specified 3 key functions: facilitating participation, supporting or substituting body functions, and preventing limitations or restrictions in participation. APs associated with information and communication technology include both hardware and software, and the ISO 9999:2023 standard emphasizes their mixed nature, explicitly stating “including software” and recognizing their educational value within category e130 of the International Classification of Functioning [36-38].

Society and public administrations are responsible for ensuring that children with disabilities have access to the APs necessary for their personal and social development [39-41]. However, the World Health Organization warns that high costs and technical complexity limit access, particularly in educational contexts [42]. Although technological advances and globalization have partially reduced prices, technical barriers remain in specific contexts [43].

While many education and health professionals are familiar with physical APs, they often lack expertise in software-based solutions, which usually require technical support from IT specialists. This dependency increases costs and delays implementation. In this context, AI represents a major opportunity: its ability to generate and modify code accessibly allows the creation or adaptation of low-cost technological products, narrowing the gap between technical design and real student needs [44]. Although AI-generated software may not reach the technical refinement of professionally supervised development, it offers an effective, economical, and flexible alternative to enhance accessibility and educational participation for students with disabilities, particularly in low-resource or time-constrained contexts.

Methods

Overview

The software development process was carried out iteratively with the assistance of 8 different AI models, 6 of which were free and 2 were paid (Table 2). These models were selected to generate results applicable to contexts where cost may be a limiting factor. The chatbots were asked to program a software solution to convert an arcade fighting-style controller into a computer control device similar to an adapted mouse. Communication with the AIs was conducted using predetermined prompts prior to testing.

The objective was not only to obtain a functional program but also to significantly reduce the feedback provided to each AI and subsequently analyze each model’s ability to solve the problem as efficiently as possible, assessing whether it could produce a complete and functional solution. This approach enabled the comparison of performance, accuracy, and coherence across different platforms within the same experimental context, providing objective data on their real effectiveness in supporting software development through natural language instructions.

Table . AIs used, models, and price.

Company	AI[a]	Model	Price
OpenAI	ChatGPT (Free)	GPT-4.1 mini	N/A[b]
OpenAI	ChatGPT (Pro)	GPT-5	€23 (US $26.36)/month
Google	Gemini (Free)	Gemini 2.5 Pro	N/A
Google	Gemini (Paid)	Gemini 2.5 Pro	€21.99 (US $25.20)/month
Anthropic	Claude (Free)	Sonnet 4.5	N/A
DeepSeek	DeepSeek	DeepSeek V3.2	N/A
Microsoft	Copilot	o3 mini	N/A

[a] AI: artificial intelligence.

[b] N/A: not applicable.

To ensure experimental validity and avoid prior learning or contamination from conversational memory, each evaluation was conducted in entirely new conversations and, in the case of free plans, through newly created accounts. This strategy is

XSL•FO RenderX

grounded in empirical evidence showing that conversational history can induce bias or interference between tasks, thereby affecting model responses [45]. Similarly, literature on data contamination in language models indicates that any prior exposure to the evaluated information alters the reliability of results [46,47]. Even slight reformulations or lingering contextual information can modify the direction and quality of the generated output [48]. Therefore, initiating each session in a clean environment constitutes a methodologically necessary measure to control carryover bias, prevent contextual leakage, and ensure independence between trials, thereby preserving the external comparability of performance and accuracy across the evaluated models.

Ethical Considerations

The Research Ethics Committee of Valencia Catholic University Saint Vincent Martyr approved the study (UCV/2023-2024/010).

Test Protocol

The first step in initiating the coding process was to define the functions that could be beneficial and link each of them to the prompts that would be sent to the different AIs (Table 3). The definition of these prompts and the design of the conversation were based on two key concepts:

1. Iterative execution: the initial prompt is sent; the generated code is tested for functionality; and if it works successfully, the next prompt is sent with the intention of iteratively adding new features to the designed program.
2. Error handling: no feedback is provided, only the message "it doesn't work, fix it." This approach is based on the understanding that a person without technical qualifications might not comprehend the underlying problems in the code and would likely just ask the AI to fix the errors encountered. Such a person might be able to inform the AI that, for example, the joystick is not working but not describe more complex interaction failures or interpret Python error messages.

Table . Function desired and chosen prompt.

Function	Prompt
1. Move mouse with joystick	I have no knowledge of programming. I have a generic fighting style Game Pad that has a single joystick and several buttons. I need you to code a program in Python that allows the joystick on this device to generate the movements of the computer mouse.
2. Single right and left click	I need you to code left and click buttons like the mouse has.
3. Allow button reassignment	I need you to code in an app that allows the program to reassign keys in the gamepad.
4. Double left click from a single button	I need you to add an extra button that performs a double click in single click.
5. Toggle hold for left click	I need you to add an extra button that holds left click when pressed and releases when pressed again.
6. Scroll page up and down	I need you to add two extra buttons for scrolling up and down.
7. Increase and decrease volume	I need you to add extra buttons for volume up and down.
8. Assign a program to a button	I need you to add an extra button that opens Google Chrome when pressed.
9. Perform an automatic click after 3 seconds of mouse inactivity	Make it so that when the mouse automatically clicks once when standing still for 3 seconds.
10. Allow double-click with a configurable delay	I need you to add a delay that allows two clicks made within 3 seconds to register as a double click.
11. Type the user's name with a button	I need you to add a button that writes the name Pelayo.
12. Pause and unpause	I need you to add a button that toggles between pause and play.
13. Ignore movement during the first second of joystick input	I need you to add a delay that ignores the first seconds of movement in any direction and then starts moving.
14. Automatically send a help email	I need you to add an extra button that automatically sends an email saying "Help" to francesc.banuls@ucv.es.
15. Generate an app that allows modifying sensitivity and delay time for automatic clicking	I need you to code in an app that allows me to change sensitivity of the joystick and how long the mouse waits after being still to automatically click.
16. Create a program and installer	I need you to compile the program made into an app by coding installer for this app.

XSL•FO
RenderX

Based on this premise, and to reduce inconsistencies, it was considered that modern AIs increasingly incorporate more integrated self-feedback mechanisms, independent from human input, and that models should be capable of performing such functions [49]. All testing was conducted by the authors, in the context of an autonomous reference center for disability in Valencia, Spain, during 4 days, from October 10, 2025, to October 14, 2025. Subsequently, all testing was conducted with the approval of the university's ethics committee with code UCV/2023-2024/010.

Additionally, all testing was conducted using a Kubii USB Arcade Controller reference ODARCADE, connected via USB to a Windows 11 computer using Python 3.13 and having installed the libraries pygame and pyautogui.

The selection and design of the prompt constitute a key methodological element in this study, as its structure determines the comparative validity among the different AI models analyzed. Its uniformity follows the principle of instructional consistency, emphasizing that coherence in the formulation of commands is essential to ensure comparability between conversational systems [10]. Complementarily, the structure of the prompt conditions the type of cognitive processing activated by the model; therefore, keeping it constant allows the isolation of the intrinsic effect of each architecture [11]. Minimal wording differences can change both performance and the relative ranking of models, reinforcing the need to use a standardized prompt. Accordingly, the prompts were designed based on 3 theoretical principles [12,28,50].

The first is defining a nontechnical user role, representative of professionals without programming training. Contextualizing the interlocutor's identity allows for adjusting response complexity, and that assigning an explicit role guides the model's generation toward a semantically coherent and functional framework. This role also demands that all models are accessed from their web-found standard versions without adjusting models or using specific more technical or specially designed models for programming. This can be seen in the use of Copilot instead of specific instances coded into GitHub or other programming platforms [10,11].

The second is presenting a sequence of actions formulated in natural language and logical order. Procedural prompts activate step-by-step reasoning that promotes coherent and structured results, facilitating the translation of human descriptions into computational processes (as the de facto interaction mode with chatbots), although such prompts may carry a higher risk of drift [11].

The third is detailing functional requirements, such as cursor control or command execution. Precision in parameters or specific conditions reduces ambiguity and improves consistency of the output, making the prompt a cognitive interface between human language and automated execution [12,28].

These 3 concepts serve as guiding principles to mediate between the natural language of a nontechnical professional and the AIs. To more faithfully emulate the real process of software design by a person without advanced technical knowledge, 2 conditions were established to determine when to terminate the programming attempts.

The first one is that 3 consecutive nonfunctional generations either fail to execute, only partially implement, or lose communication with the gamepad. It is understood that if 3 consecutive iterations cannot fix inherited errors without feedback, it is unlikely that subsequent ones will do so.

Prompts are considered successful when they can correctly implement the full capabilities described in Table 1 for each prompt without compromising existing functionality. They are considered partially successful when they either maintain previous capabilities but only partially implement new functions or when they correctly implement new functions but lose previous ones. Finally, prompts are considered unsuccessful when they either fail to execute or lose communication with the gamepad. Partially successful prompts are marked as such to better understand the performance of different models but are functionally considered unsuccessful as they did not achieve the desired function. As such, 3 unsuccessful or partially successful prompts (with 2 tries each) stop interactions with a model.

The other way in which interactions with models are stopped is when reaching the daily prompt limit in free plans, to better reflect that the prompt limit is one of the most significant constraints of these models.

Goals

The experimental design was developed with the explicit understanding that current AI systems are inherently nondeterministic and that the replication of results will be inconsistent [51]. Consequently, the study objectives reflect this fact:

- General goal 1: compare the effectiveness of the most widely used AI solutions as software programmers for creating personalized assistive products using nontechnical natural language.
    - Specific goal 1: describe the relationship between cost and effectiveness of the models used.
    - Specific goal 2: analyze the ease of use, consistency, and accuracy of each model.

Results

Overview

The 8 AI solutions initially proposed were used to progressively design robust code solutions, incorporating functions of increasing complexity inspired by APs with similar objectives.

The results can be divided according to the cost of the alternatives, distinguishing between those that are free of charge and those requiring a subscription. The latter provided more consistent results and better adaptation to the project's functional requirements, demonstrating a greater understanding of the requested functions. However, it is important to note that these results should be interpreted within the framework of the specific characteristics of the study and that even under identical

XSL•FO
RenderX

conditions, different outcomes could be obtained due to the nondeterministic nature of these technologies.

Paid Alternatives

The subscription-based alternatives delivered consistent results, generating code that can generally be considered successful in meeting the required functions. This category includes ChatGPT Pro, based on the GPT-5 model, and Gemini Pro, which uses Google's advanced Gemini Pro-2.5 model.

Regarding Gemini Pro, the AI solution developed by Google proved to be the most robust option. It successfully implemented 14 out of the 16 required functions within a single application, without any rollback of previously implemented features during the design process and generated functional code with an average of 1.25 (SD 0.45) prompts. Gemini was even partially able to anticipate the implementation of future functions in earlier items: item 2 was successfully integrated within item 1; item 6 was partially included in item 3 by adding a scrolling mode, although without the 2 specific keys intended for that function; and item 15 was partially addressed in item 3 through a joystick sensitivity slider. The only functions not achieved were items 14 and 16. Item 14 could be considered partially achieved, as the AI managed to open the email app and compose a message addressed to the correct recipient but was unable to send it. All partially achieved items are considered not achieved when counting successful items implemented. Regarding item 16, Gemini could not autonomously generate an installable app, and the alternative solutions proposed in Python were also nonfunctional.

One thing to note is that both items (14 and 16) did not work across the board for security reasons, as AIs are not capable of generating programs that have such deep access to your operative system. Seeing how the AIs approached a task they knew beforehand was impossible to fully be able to perform is interesting and shows the AIs' problem-solving skill navigating natural language generated prompts, which may be difficult or impossible to fulfill.

Complementarily, OpenAI's paid version (ChatGPT Pro using GPT-5) also demonstrated solid performance, successfully implementing 11 out of the 16 required functions and producing functional code with an average of 1.31 (SD 0.48) prompts. GPT-5 showed the same limitation as Gemini Pro with respect to sending the email in item 14 and was likewise unable to compile an executable file or provide stable Python-based compilation solutions. Additionally, OpenAI's AI had difficulties implementing joystick controls in item 1, which caused the device to move the cursor only downward and to the right. This error persisted in subsequent iterations, affecting later results such as items 13 and 15, which were only partially achieved due to the lack of precise joystick control. ChatGPT also displayed inconsistency, producing multiple files for different functions instead of integrating them into a single app and rolling back previously implemented features without justification or explicit request.

Free Alternatives

The free alternatives, in contrast to the subscription-based ones, produced notably less consistent results. In some cases, this was due to the limitations of the models themselves, which were unable to meet the functional requirements set for the task. In other cases, restrictions on the number of interactions significantly limited development, constituting a key distinction. It cannot be asserted that greater time investment would have yielded results equivalent to the paid versions, since the possibility of submitting additional prompts was not available. Therefore, in this study, the maximum number of interactions allowed by each model in its free version was established as a methodological limitation. Conversely, the models that failed to generate functional solutions can, for the time being, be considered inoperative for the intended function.

The free AI with the best performance was DeepSeek. This model generated the first 3 items in the initial prompt and even implemented item 2 without it being explicitly requested in item 1. However, the AI was unable to technically complete items 4, 5, and 6, implementing item 5 through a bind function that only allowed the assignment of a single key and caused the program to crash when attempting to add more combinations. This model generated functional items with an average of 1.88 (SD 0.34) prompts, achieving 4 out of the 16 required functions.

The second-best performance can be attributed to GPT Free. The free version of GPT successfully produced a functional prototype on the first attempt for the first 3 prompts but maintained the inconsistency observed in its paid counterpart. From the fourth prompt onward, it began producing nonfunctional code, leading to discontinuation after the seventh attempt. On average, it generated functional prototypes with 1.81 prompts and achieved 3 out of the 16 intended functions.

After GPT, in terms of performance, is Gemini Pro Free. Although it theoretically uses the same model as the paid version, its results differed appreciably. The free version produced correct and functional outputs up to the permitted prompt limit, forcing the process to stop prematurely. It successfully implemented the first 2 items and partially the third, generating a program that ran natively in the browser. Subsequently, time-based restrictions prevented continuation. On average, it produced functional prototypes with 1.88 prompts and achieved 2 out of the 16 complete functions.

The second-to-last worst-performing model is Claude. The Sonnet 4.5 model managed to implement only the first of the planned functions. After failing to execute functions 3, 4, and 5, its use was discontinued. Claude generated functional prototypes with an average of 1.94 (SD 0.25) prompts, achieving 1 out of the 16 proposed functions.

Finally, the individually worst-performing model is Copilot. Using its Deep Think model, Microsoft's AI failed to generate any executable Python files within the first 3 prompts, leading to the termination of testing after the fourth attempt. It did not produce any functional prototypes, with an average of 2 prompts and a total of 0 out of 16 functions achieved.

All models and a color-coded list of the items successfully implemented by each, along with the number of prompts required, are presented in Table 4. Additionally, prompts achieved and average prompts per model are presented in Figure 1.

XSL•FO
RenderX

Table . Functions achieved and prompts needed.

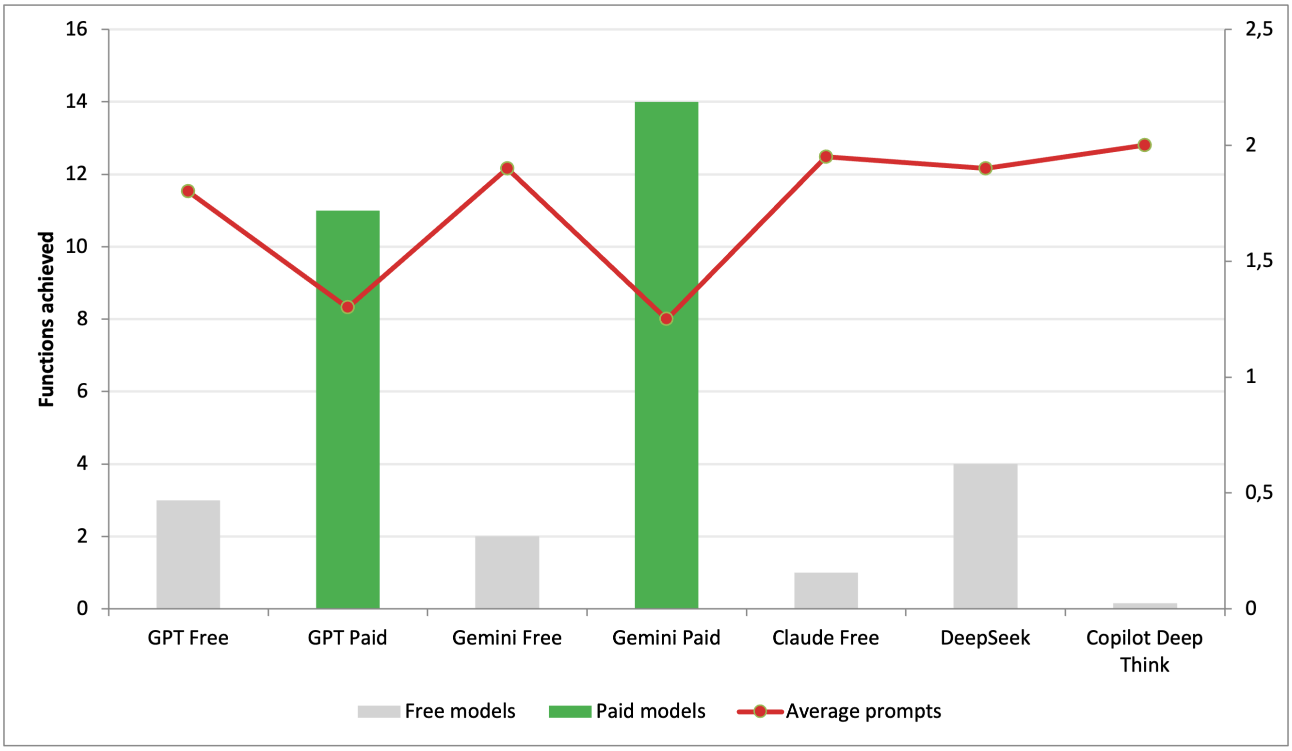
Required function	GPT Free	GPT Paid	Gemini Free	Gemini Paid	Claude Free	DeepSeek	Copilot Deep Think
1	✓[a]	×[b]	✓	✓	✓	✓	×
2	✓	✓	×	✓	×	✓[c]	×
3	✓	✓	✓[c]	×	×	✓	×
4	×	✓	×	✓	×	×	×
5	×	✓	×	✓	×	✓	×
6	×	✓	×	✓[c]	×	×	×
7	×	✓	×	✓	×	×	×
8	×	✓	×	✓	×	×	×
9	×	✓	×	✓	×	×	×
10	×	✓	×	×	×	×	×
11	×	✓	×	✓	×	×	×
12	×	✓	×	✓	×	×	×
13	×	×	×	✓	×	×	×
14	×	×	×	×	×	×	×
15	×	×	×	✓[c]	×	×	×
16	×	×	×	×	×	×	×

[a] Items achieved.

[b] Items not achieved or partially achieved. Partially achieved items are counted as unsuccessful.

[c] Functions already implemented or partially already implemented in previous interactions without direct prompts for their inclusion.

Figure 1. Prompts achieved and average prompts per model.



XSL•FO
RenderX

Discussion

Principal Findings

The results obtained show that the paid AI alternatives produced robust code, achieving 14 out of 16 functions with Gemini Pro and 11 out of 16 functions with ChatGPT Plus, respectively. The free-to-use alternatives, on the other hand, managed to integrate a variable, though consistently lower, number of functions ranging from none to 4 successfully implemented. The discussion of the results obtained can therefore be divided into 3 key areas (Figure 1).

Superiority of Paid AIs

The use of paid AI systems offers several advantages, among which the most significant are access to more advanced models and the elimination or substantial expansion of the number of prompts allowed within a given period. These advantages naturally make paid AI solutions more reliable for performing complex tasks. The main improvements can be summarized in 3 key points.

Paid versions typically operate on newer, more advanced architectures with higher parameter capacity and improved optimization strategies and training based on larger and higher-quality datasets. In literature, larger and better-trained models consistently achieve superior results in both code generation and natural language understanding tasks. For instance, GPT-4 performs substantially better in programming tasks than earlier versions [52].

Due to these technical enhancements, paid models generally demonstrate a stronger grasp of natural language and communicative context, which increases the likelihood of successful interactions in this specific case scenario. They are also better at maintaining coherence across sequential instructions, retaining previously defined functions, and being better at avoiding contradictions and hallucinations. This allows them to integrate multiple features coherently and progressively with fewer rollbacks or loss of content [53-55].

The removal or significant reduction of restrictions on the number of prompts permitted per time interval is particularly relevant when considered alongside the highlighted advantages. Older free models often have weaker memory capabilities and greater difficulty minimizing inconsistencies, rollbacks, and errors across different interaction sessions [56]. Increasing the interaction limits from, for example, 5 per session in the free version of Gemini to 100, or from 10 every 5 hours in GPT-4.1 mini to 160 every 3 hours in GPT-5, not only eases time constraints but also reduces errors and can substantially improve code quality [57,58].

In summary, the differences in model capacity, contextual understanding, memory, and reasoning largely explain why paid AIs may achieve significantly more functional results than their free counterparts. This outcome is expected, as paying for a service naturally entails the expectation of better performance. However, another relevant point of discussion lies in comparing paid AIs with one another and questioning why Google's AI seems to have outperformed GPT in this specific scenario, despite the latter's longer development history and its recent update to version 5.

Differences Between Gemini Pro and GPT-5

Although it is not possible to access the internal technical details of Gemini Pro, or of any major AI model, the existing literature published by the companies behind these models appears to support the results obtained in this specific case scenario, pointing to 3 key aspects that may explain why Gemini stood out as the most robust alternative for generating code from natural language given the restrictions and tasks asked.

On the one hand, Gemini may have better alignment between intention and execution. According to the Google Gemini team, the training of Gemini Pro placed greater emphasis on the correspondence between high-level instructions and structured code generation, giving it an advantage in translating complex requirements into functional code. The model achieved scores above 74% on specialized programming benchmarks [59].

This may also lead to much greater stability and coherence in long sessions. In tasks involving multiple functions (such as the 16 required in this study), the ability to maintain coherence across many iterations is crucial. If Gemini Pro manages internal session states more effectively, it becomes less prone to "forget" previously implemented functions or to inconsistently rename elements [60].

On the other hand, and especially relevant to the study, prompt optimization may be more tolerant of natural language. A distinct training approach regarding the model's handling of natural language results in greater adaptability to intricate instructions, improved contextual understanding, and enhanced problem-solving abilities. This makes Gemini Pro more resilient to inaccuracies or ambiguities in prompts written in nontechnical natural language, whereas ChatGPT Plus may tend to require more carefully formulated instructions to produce reliable outcomes [59,60].

Limitations and Strengths of Free AIs

As previously mentioned, free AI systems may present significant limitations related to the use of simpler and more restricted models, limited numbers of prompts, and, in some cases, the restriction to process multimodal information. These constraints make free models more challenging to use or less adaptable to users' needs. The standardized subscription fee of US $20, as implemented by companies such as OpenAI and Google, remains possibly too expensive for low-resource contexts in many parts of the world.

In recent months, both companies have begun introducing country-based pricing, adjusting subscription costs to local purchasing power. OpenAI (2025) implemented multicurrency billing to reduce conversion costs, and in India, the Plus plan is now offered in rupees, alongside a lower-cost alternative named ChatGPT Go [61]. This strategy echoes *The Economist*'s Big Mac Index, which illustrates how multinational companies adjust prices according to economic context. AI can enhance accessibility in education, but its impact is constrained by the economic barriers preventing access to paid tools [62]. However, if AI design fails to account for income inequalities within a

XSL•FO
RenderX

single country, despite global cost-adjustment strategies, it may reinforce digital exclusion. Thus, exploring the use and functionality of free versions and comparing them with paid alternatives as a potential source of social inequality becomes a social necessity to ensure equitable access to knowledge and accessibility [63,64].

In our test, among the free options, Gemini Free managed to implement 2 functions with relative stability during the first iteration, approaching a third before reaching the interaction limit. Considering the performance of its paid counterpart under fewer restrictions, it is likely that free versions could offer solid, no-cost solutions, albeit requiring a significant time investment, provided the model seems to be able to maintain internal coherence across sessions [59].

In contrast, GPT Free and DeepSeek achieved a greater number of items but lost track of the ongoing programming process, producing unstable code that would require the user to identify errors and explain them to the chatbot. While this may be possible, it is difficult to fully assess how much feedback a nontechnical user may be able to provide, and as such, it is difficult to give certain conclusions about these models. Professionals with intermediate technical knowledge could use these 2 models effectively to develop software solutions, but nontechnical users may be unable to [65,66].

Finally, Claude and Copilot, in their free versions, showed no real ability to generate operational code. In the case of Claude, this is understandable, since its website explicitly states that only the Pro version provides such capabilities [67]. The second model, Copilot, simply appears unable to generate functional, plug-and-play code. This is also due to the several factors, such as copilot being used in the web version (not optimized to do such tasks) without any specific model or implementation chosen [68]. In both case scenarios, using Claude Pro with programming capabilities and a version of Copilot optimized for programming may generate more competent results.

These findings may point to the fact that not all free versions are equivalent: some models are sufficiently robust for basic tasks, while others are less capable with the task chosen, the prompt limit, the version used, and many other factors contributing to their success.

One key limitation to have in mind is that only 1 run has been carried out with each model to mimic a real case scenario of a nontechnical user using these models. Given the nondeterministic nature of these models, more different runs around the same task may produce different results.

Observations and Practical Implications

The observations, limitations, and implications for practice derived from this study can be summarized in 10 key points that integrate the main insights gained throughout the research. Four main points stand out as the implications of the results obtained.

Some AI models, such as ChatGPT (both 4.1 mini and 5), may produce inconsistent results, hallucinating or forgetting its progress. This can be seen by these models including renaming functions without apparent reason, changing project names, generating new programs instead of adding features to the existing one, or rolling back previously correct functions. This suggests that OpenAI models can generate the intended products but require a technical skill level beyond that of nonspecialist professionals, significantly complicating the process of obtaining functional prototypes for nontechnical users in this specific case scenario [66,69].

This behavior is sometimes categorized as AI “drift” and means that, during extended sessions, models may change behavior, forget previously defined functions, or misinterpret prompts that are conceptually equivalent. Maintaining a clear “narrative thread” and consistent prompt formulation is therefore essential. Natural language alone may not suffice, and the risk of accumulating errors that render the code nonfunctional remains high [70,71].

In this specific test, Gemini showed greater consistency interpreting natural-language instructions and required fewer technical reformulations or corrections. It also showed a lower risk of drift. In contexts where users represent the target population, these traits may reduce the barrier to entry and increase the likelihood of producing functional prototypes. The Pro version is notably more efficient and reliable, but the free version can achieve similar outcomes with sufficient time investment if the time between tries does not induce AI drift [72].

While AIs can generate code that is operational and useful, it must be remembered that its efficiency (execution time, memory usage, and design simplicity) is generally suboptimal. For this reason, professional supervision, or ideally full implementation, by a qualified software engineer remains preferable. Studies, such as EffiBench, have shown that AI-generated code tends to be, on average, less efficient than optimized human solutions [73].

The results obtained in this domain (low-cost AP adaptations with 16 functions with no feedback and in a single run) are constrained to the models and prompts used and to the absence of user feedback, maybe reflecting the condition of a nontechnical user. Other models or differently designed prompts might yield different outcomes. The practical implications are as follows:

1. For real-world AP adaptations: Collaboration with software professionals should always be prioritized when budgets allow. If it is not feasible, both paid models tested may be the most effective choice, as they perform better in programming tasks than other paid alternatives. When access to paid tools is possible, they remain the most robust option. However, free models such as Gemini Free, GPT Free, or even DeepSeek can provide functional results in exchange for greater time investment.
2. For low-budget or educational contexts: Free versions, particularly Gemini Free or GPT Free, can serve as valuable tools for initial prototyping, though additional iterations, revisions, and acceptance of functional limitations will be necessary [74].
3. Effective prompt design and problem segmentation: Breaking the task into smaller subfunctions and crafting

XSL•FO
RenderX

clear, stepwise prompts improves the success rate, especially in free models.

4. Continuous monitoring and automated testing: Implementing automated unit tests (eg, by prompting the AI to generate its own test cases) may help detect errors or inconsistencies in the generated code.
5. Caution with long sessions or evolving prompts: Avoid overly long or evolving prompts within a single session and restart the interaction when the model shows signs of drift to preserve consistency.

Conclusions

In conclusion, the comparative analysis of different AI models applied to the generation of functional software for designing personalized APs using natural language reveals clear differences in performance, usability, and consistency. Paid models, such as Gemini Pro and ChatGPT Plus, demonstrated greater efficiency and reliability, achieving 14 and 11 out of 16 required functions, respectively, while free alternatives ranged between 0 and 4. These results confirm that the technical advancements of newer and more sophisticated models, such as expanded context windows, improved multimodal communication, enhanced natural language comprehension, reduced drift, and fewer prompt limitations, directly influence the functional quality of the generated code.

Regarding the relationship between cost and effectiveness, Gemini Pro may stand out as the most balanced option, in this specific case scenario offering higher precision and coherence at a cost equivalent to other paid alternatives. For low-resource contexts, DeepSeek emerges as a viable free alternative that, despite its limitations, can produce acceptable results given sufficient time and technical supervision, while the free versions of Gemini can deliver solid outputs with adequate time investment.

In terms of ease of use and consistency, Gemini demonstrated greater tolerance for nontechnical natural language and more stable performance in long sessions, making it particularly suitable for users without programming knowledge, an essential aspect when the goal is to promote inclusive and low-cost design of APs.

Overall, the findings suggest that AI can serve as an effective bridge between health care or educational professionals and programming, enabling the creation of personalized assistive solutions without the need for advanced software development expertise. Nevertheless, the results also highlight the critical need for collaboration with qualified IT professionals, given the importance of human technical oversight, careful prompt design, and continuous testing to prevent drift and ensure the generation of stable software alternatives.

Future research could focus on exploring how AI models handle problem-solving from visual inputs (eg, images sent to chatbots), their capacity for self-generated prompt design, and the evaluation of emerging or alternative models.

Funding

This study was carried out within the framework of the research project PID2022-142309OB-100 funded by the Ministry of Science and Innovation, State Research Agency and in collaboration with the Capacitas Campus of the Universidad Católica de Valéncia San Vicente Mártir (UCV).

Conflicts of Interest

None declared.

References

1. Xu Y, Liu X, Cao X, et al. Artificial intelligence: a powerful paradigm for scientific research. Innovation (Camb) 2021 Oct 28;2(4):100179. [doi: 10.1016/j.xinn.2021.100179] [Medline: 34877560]
2. Sauvola J, Tarkoma S, Klemettinen M, Riekki J, Doermann D. Future of software development with generative AI. Autom Softw Eng 2024 May;31(1):26. [doi: 10.1007/s10515-024-00426-z]
3. Qiu K, Puccinelli N, Ciniselli M, Di Grazia L. From today's code to tomorrow's symphony: the AI transformation of developer's routine by 2030. ACM Trans Softw Eng Methodol 2025 Jun 30;34(5):1-17. [doi: 10.1145/3709353]
4. Alenezi M, Akour M. AI-driven innovations in software engineering: a review of current practices and future directions. Appl Sci (Basel) 2025;15(3):1344. [doi: 10.3390/app15031344]
5. Yang EW, Waldrup B, Velazquez-Villarreal E. Conversational AI agent for precision oncology: AI-HOPE-WNT integrates clinical and genomic data to investigate WNT pathway dysregulation in colorectal cancer. Front Artif Intell 2025;8:1624797. [doi: 10.3389/frai.2025.1624797] [Medline: 40860720]
6. Yu L. Paradigm shift on coding productivity using GenAI. Presented at: Proceedings of the 29th International Conference on Evaluation and Assessment in Software Engineering (EASE '25); Jun 17-20, 2025. [doi: 10.1145/3756681.3757081]
7. Sergeyuk A, Titov S, Izadi M. In-IDE human-AI experience in the era of large language models; a literature review. Presented at: IDE '24: Proceedings of the 1st ACM/IEEE Workshop on Integrated Development Environments; Apr 20, 2024. [doi: 10.1145/3643796.3648463]
8. Beheshti A. Natural language-oriented programming (NLOP): towards democratizing software creation. Presented at: 2024 IEEE International Conference on Software Services Engineering (SSE); Jul 7-13, 2024. [doi: 10.1109/SSE62657.2024.00047]

XSL•FO
RenderX

9. Dalsgaard P. Thinking through prompting: cognitive mediation in human–AI interaction. 2025 Presented at: Proceedings of the European Conference on Cognitive Ergonomics (ECCE ’25); Oct 7-10, 2025. [doi: 10.1145/3746175.3747192]
10. McTear M, Callejas Z, Griol D. The Conversational Interface Talking to Smart Devices: Springer International Publishing; 2016. [doi: 10.1007/978-3-319-32967-3]
11. Morales-Chan M. Explorando el potencial de chat GPT: una clasificación de prompts efectivos para la enseñanza [Report in Spanish]. : Universidad Galileo; 2023 URL: https://biblioteca.galileo.edu/tesario/handle/123456789/1348 [accessed 2026-03-07]
12. Brown TB, Mann B, Ryder N, et al. Language models are few-shot learners. Presented at: Proceedings of the 34rd International Conference on Neural Information Processing Systems (NeurIPS 2020); Dec 6-12, 2020 URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf [accessed 2026-03-07]
13. Serban I, Sordoni A, Lowe R, et al. A hierarchical latent variable encoder-decoder model for generating dialogues. 2017 Presented at: Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI 2017); Feb 4-9, 2017. [doi: 10.1609/aaai.v31i1.10983]
14. Dathathri S, Madotto A, Lan Z, Fung P, Neubig G. Plug and play language models: a simple approach to controlled text generation. Presented at: Findings of the Association for Computational Linguistics: EMNLP 2021; Nov 7-11, 2021. [doi: 10.18653/v1/2021.findings-emnlp.334]
15. Korzynski P, Mazurek G, Krzypkowska P, Kurasinski A. Artificial intelligence prompt engineering as a new digital competence: analysis of generative AI technologies such as ChatGPT. Entrepren Bus Econ Rev 2023;11(3):25-37. [doi: 10.15678/EBER.2023.110302]
16. Kumar H, Musabirov I, Shi J, et al. Exploring the design of prompts for applying GPT-3 based chatbots: a mental wellbeing case study on Mechanical Turk. arXiv. Preprint posted online on Sep 22, 2022. [doi: 10.48550/arXiv.2209.11344]
17. Zhou K, Zhang J, Liu X, Sun M. A systematic survey of prompt engineering in large language models: techniques and applications. arXiv. Preprint posted online on Feb 5, 2024. [doi: 10.48550/arXiv.2402.07927]
18. Fagbohun O, Harrison RM, Dereventsov A. An empirical categorization of prompting techniques for large language models: a practitioner’s guide. J Artif Intell Mach Learn Data Sci ;1(4):1-11. [doi: 10.51219/JAIMLD/Oluwole-Fagbohun/15]
19. Wei J, Wang X, Schuurmans D, et al. Chain-of-thought prompting elicits reasoning in large language models. Presented at: Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS 2022); Nov 28 to Dec 9, 2022 URL: https://openreview.net/pdf?id=_VjQlMeSB_J [accessed 2026-03-07]
20. Madaan A, Tandom N, Gupta P. Self-refine: iterative refinement with self-feedback. Presented at: Advances in Neural Information Processing Systems 36; Dec 10-16, 2023 URL: https://proceedings.neurips.cc/paper_files/paper/2023/hash/91edff07232fb1b55a505a9e9f6c0ff3-Abstract-Conference.html [accessed 2026-03-18]
21. Anam RK. Prompt engineering and the effectiveness of large language models in enhancing human productivity. arXiv. Preprint posted online on May 10, 2025. [doi: 10.48550/arXiv.2507.18638]
22. Xu Y, Thomas T, Yu CL, Pan EZ. What makes children perceive or not perceive minds in generative AI? Comput Hum Behav Artif Hum 2025 May;4:100135. [doi: 10.1016/j.chbah.2025.100135]
23. Brandtzaeg PB, Skjuve M, Følstad A. My AI friend: how users of a social chatbot understand their human–AI friendship. Hum Commun Res 2022 Jun 29;48(3):404-429. [doi: 10.1093/hcr/hqac008]
24. Panfili L, Duman S, Nave A, Ridgeway KP, Eversole N, Sarikaya R. Human-AI interactions through a Gricean lens. Proc Ling Soc Amer ;6(1):288. [doi: 10.3765/plsa.v6i1.4971]
25. Zheng Q, Tang Y, Liu Y, Liu W, Huang Y. UX research on conversational human-AI interaction: a literature review of the ACM digital library. 2022 Apr 29 Presented at: CHI ’22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; Apr 29 to May 5, 2022. [doi: 10.1145/3491102.3501855]
26. Gülay E, Picco E, Glerean E, Coupette C. Relational dissonance in human–AI interactions: the case of knowledge work. arXiv. Preprint posted online on Sep 19, 2025. [doi: 10.48550/arXiv.2509.15836]
27. Kim Y, Lee J, Kim S, Park J, Kim J. Understanding users’ dissatisfaction with ChatGPT responses: types, resolving tactics, and the effect of knowledge level. Presented at: Proceedings of the 29th International Conference on Intelligent User Interfaces (IUI 2024); Mar 18-21, 2024. [doi: 10.1145/3640543.3645148]
28. Hernández Caralt M, Sekulić I, Carevic F, et al. Stupid robot, I want to speak to a human!” User frustration detection in task-oriented dialog systems. 2025 Presented at: Proceedings of the 31st International Conference on Computational Linguistics: Industry Track; Jan 19-24, 2025 URL: https://aclanthology.org/2025.coling-industry.23/ [accessed 2026-03-07]
29. Eiband M, Schneider H, Bilandzic M, Fazekas-Con J, Haug M, Hussmann H. Bringing transparency design into practice. Presented at: Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI 2018); Mar 7-11, 2018 URL: https://dl.acm.org/doi/proceedings/10.1145/3172944 [doi: 10.1145/3172944.3172961]
30. Degachi C, Freire SK, Niforatos E, Kortuem G. Understanding mental models of generative conversational search and the effect of interface transparency. arXiv. Preprint posted online on Jun 4, 2025. [doi: 10.48550/arXiv.2506.03807]
31. Jardinez MJ, Natividad LR. The advantages and challenges of inclusive education: striving for equity in the classroom. Shanlax Int J Educ 2024;12(2):57-65. [doi: 10.34293/education.v12i2.7182]
32. Bani Odeh K, Lach LM. Barriers to, and facilitators of, education for children with disabilities worldwide: a descriptive review. Front Public Health 2023;11:1294849. [doi: 10.3389/fpubh.2023.1294849] [Medline: 38292375]

XSL•FO
RenderX

33. Pérez Echeverría MP, Cabellos B, Pozo JI. The use of ICT in classrooms: the effect of the pandemic. Educ Inf Technol 2025 Jul;30(10):14069-14093. [doi: 10.1007/s10639-024-13124-w]
34. Alanazi AS, Benlaria H. Understanding the landscape: assistive technology and work challenges for people with disabilities in Saudi Arabia. Humanit Soc Sci Commun 2024;11(1):1608. [doi: 10.1057/s41599-024-04023-z]
35. The state of online recruiting 2024. iHire. URL: https://www.ihire.com/resourcecenter/employer/pages/the-state-of-online-recruiting-2024 [accessed 2025-10-22]
36. ISO 9999:2022—assistive products for persons with disability: classification and terminology. : International Organization for Standardization (ISO); 2022 URL: https://cdn.standards.iteh.ai/samples/72464/3f3608ed0bff4545bd53c02373f8cdbb/ISO-9999-2022.pdf [accessed 2026-03-07]
37. ISO 9999:2016—assistive products for persons with disability: classification and terminology. : International Organization for Standardization (ISO); 2016 URL: https://cdn.standards.iteh.ai/samples/60547/4694b28419324e45810538491357903c/ISO-9999-2016.pdf [accessed 2026-03-07]
38. International Classification of Functioning, Disability and Health (ICF). : World Health Organization (WHO); 2001 URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2026-03-07]
39. Tomás V. La discapacidad como elemento de discriminación positiva. In: Los Derechos de Los Niños, Responsabilidad de Todos [Book in Spanish]: Universidad de Murcia; 2007:213-218 URL: https://dialnet.unirioja.es/servlet/articulo?codigo=2306106 [accessed 2026-03-07]
40. Vila-Merino ES, Rascón-Gómez T, Calderón-Almendros I. Discapacidad, estigma y sufrimiento en las escuelas. Narrativas emergentes por el derecho a la educación inclusiva [Article in Spanish]. Educ XX1 2024;27(1):353-371. [doi: 10.5944/educxx1.36753]
41. de Beco G. The right to inclusive education according to Article 24 of the UN Convention on the Rights of Persons with Disabilities: background, requirements and (remaining) questions. Neth Q Hum Rights 2014 Sep;32(3):263-287. [doi: 10.1177/016934411403200304]
42. Assistive technology. World Health Organization (WHO). 2024. URL: https://www.who.int/news-room/fact-sheets/detail/assistive-technology [accessed 2025-10-22]
43. Fteiha M, Al-Rashaida M, Elsori D, Khalil A, Al Bustami G. Obstacles for using assistive technology in centres of special needs in the UAE. Disabil Rehabil Assist Technol 2024 Nov;19(8):2934-2944. [doi: 10.1080/17483107.2024.2323698] [Medline: 38436086]
44. Fernández-Batanero JM, Montenegro-Rueda M, Fernández-Cerero J, García-Martínez I. Assistive technology for the inclusion of students with disabilities: a systematic review. Education Tech Research Dev 2022;70(5):1911-1930. [doi: 10.1007/s11423-022-10127-7]
45. Gupta A, Sheth I, Raina V, Gales M, Fritz M. LLM task interference: an initial study on the impact of task-switch in conversational history. Presented at: Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing; Nov 12-16, 2024. [doi: 10.18653/v1/2024.emnlp-main.811]
46. Golchin S, Surdeanu M. Time travel in LLMs: tracing data contamination in large language models. arXiv. Preprint posted online on Aug 16, 2023. [doi: 10.48550/arXiv.2308.08493]
47. Cheng Y, Chang Y, Wu Y. A survey on data contamination for large language models. arXiv. Preprint posted online on Feb 20, 2025. [doi: 10.48550/arXiv.2502.14425]
48. Yang S, Chiang WL, Zheng L, Gonzalez JE, Stoica I. Rethinking benchmark and contamination for language models with rephrased samples. arXiv. Preprint posted online on Nov 8, 2023. [doi: 10.48550/arXiv.2311.04850]
49. Chen Z, Cox D, Gan C, et al. Principle-driven self-alignment of language models from scratch with minimal human supervision. Presented at: 37th Conference on Neural Information Processing Systems (NeurIPS 2023); Dec 10-16, 2023. [doi: 10.52202/075280-0115]
50. Mizrahi M, Kaplan G, Malkin D, Dror R, Shahaf D, Stanovsky G. State of what art? A call for multi-prompt LLM evaluation. Trans Assoc Comput Linguist 2024 Aug;12:933-949. [doi: 10.1162/tacl_a_00681]
51. Ouyang S, Zhang JM, Harman M, Wang M. An empirical study of the non-determinism of ChatGPT in code generation. ACM Trans Softw Eng Methodol 2023;34(2):1-28. [doi: 10.1145/3697010]
52. Moussiades L, Zografos G, Papakostas G. GPT-4 vs. GPT-3.5 as coding assistants. Research Square. Preprint posted online on Feb 7, 2024. [doi: 10.21203/rs.3.rs-3920214/v1]
53. Shuvo UA, Dip SA, Vaskar NR, Al Islam ABMA. Assessing ChatGPT's code generation capabilities with short vs long context programming problems. 2024 Presented at: Proceedings of the 11th International Conference on Networking, Systems, and Security (NSysS 2024); Dec 19-21, 2024. [doi: 10.1145/3704522.3704535]
54. OpenAI. GPT-4 technical report. arXiv. Preprint posted online on Mar 15, 2023. [doi: 10.48550/arXiv.2303.08774]
55. Hou Y, Zhan Z, Zhang R. Benchmarking GPT-5 for biomedical natural language processing. arXiv. Preprint posted online on Aug 28, 2025. [doi: 10.48550/arXiv.2509.04462]
56. Wang C, Sun JV. Unable to forget: proactive lnterference reveals working memory limits in LLMs beyond context length. arXiv. Preprint posted online on Jul 9, 2025. [doi: 10.48550/arXiv.2506.08184]

XSL•FO
RenderX

57. GPT-5.3 and GPT-5.4 in ChatGPT. OpenAI Help Center. 2025. URL: https://help.openai.com/en/articles/11909943-gpt-5-in-chatgpt [accessed 2025-10-20]
58. Gemini Apps limits & upgrades for Google AI subscribers. Gemini Apps Help. 2025. URL: https://support.google.com/gemini/answer/16275805?hl=en [accessed 2025-10-15]
59. Anil R, Borgeaud S, Alayrac JB, et al. Gemini: a family of highly capable multimodal models. : Google DeepMind; 2023 URL: https://storage.googleapis.com/deepmind-media/gemini/gemini_1_report.pdf [accessed 2026-03-07]
60. Alsajri A, Salman HA, Steiti A. Generative models in natural language processing: a comparative study of ChatGPT and Gemini. Babylonian J Artif Intell 2024;2024:134-145. [doi: 10.58496/BJAI/2024/015]
61. OpenAI goes local in India, ChatGPT now available in INR: check ChatGPT Plus, Pro prices here. India Today. 2025. URL: https://www.indiatoday.in/technology/news/story/openai-goes-local-in-india-chatgpt-now-available-in-inr-check-chatgpt-plus-pro-prices-here-2770840-2025-08-13 [accessed 2025-10-22]
62. Melo-López VA, Basantes-Andrade A, Gudiño-Mejía CB, Hernández-Martínez E. The impact of artificial intelligence on inclusive education: a systematic review. Educ Sci 2025;15(5):539. [doi: 10.3390/educsci15050539]
63. Umucu E. Artificial intelligence and health equity for people with disabilities: an integrated framework for disability-inclusive AI design. Inquiry 2025;62:469580251365472. [doi: 10.1177/00469580251365472] [Medline: 40847466]
64. Daepp MIG, Counts S. The emerging generative artificial intelligence divide in the United States. Presented at: Proceedings of the Nineteenth International AAAI Conference on Web and Social Media (ICWSM 2025); Jun 23-26, 2025. [doi: 10.1609/icwsm.v19i1.35825]
65. Laban P, Hayashi H, Zhou Y, Neville J. LLMs get lost in multi-turn conversation. arXiv. Preprint posted online on May 9, 2025. [doi: 10.48550/arXiv.2505.06120]
66. Liu Y, Le-Cong T, Widyasari R, et al. Refining ChatGPT-generated code: characterizing and mitigating code quality issues. ACM Trans Softw Eng Methodol 2024 Jun 30;33(5):1-26. [doi: 10.1145/3643674]
67. Claude AI. 2025. URL: https://claude.ai/upgrade [accessed 2025-10-21]
68. Zhang B, Liang P, Zhou X, Ahmad A, Waseem M. Practices and challenges of using GitHub Copilot: an empirical study. Presented at: The 35th International Conference on Software Engineering and Knowledge Engineering (SEKE 2023); Jul 10-12, 2023. [doi: 10.18293/SEKE2023-077]
69. Bartsch H, Jorgensen O, Rosati D, Hoelscher-Obermaier J, Pfau J. Self-consistency of large language models under ambiguity. Presented at: Proceedings of the 6th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP; Dec 7, 2023. [doi: 10.18653/v1/2023.blackboxnlp-1.7]
70. Donge V, Rossi RA, Lai VD, Yoon DS, Hakkani-Tür D, Bui T. Drift no more? Context equilibria in multi-turn LLM interactions. arXiv. Preprint posted online on Oct 9, 2025. [doi: 10.48550/arXiv.2510.07777]
71. Luo Y, Yang Z, Meng F, Li Y, Zhou J, Zhang Y. An empirical study of catastrophic forgetting in large language models during continual fine-tuning. IEEE Trans Audio Speech Lang Process 2023;33:3786-3776. [doi: 10.1109/TASLPRO.2025.3606231]
72. Comanici G, Bieber E, Schaeckerman M, et al. Gemini 2.5: pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. : Google DeepMind; 2025 URL: https://storage.googleapis.com/deepmind-media/gemini/gemini_v2_5_report.pdf [accessed 2026-03-07]
73. Huang D, Qing Y, Shang W, Cui H, Zhang JM. Effibench: benchmarking the efficiency of automatically generated code. Presented at: Proceedings of the 38th International Conference on Neural Information Processing Systems (NeurIPS 2024); Dec 9-17, 2024 URL: https://papers.nips.cc/paper_files/paper/2024/file/15807b6e09d691fe5e96cdecde6d7b80-Paper-Datasets_and_Benchmarks_Track.pdf [accessed 2026-03-07]
74. Rakotonirina NC, Hamdy M, Campos JA, et al. From tools to teammates: evaluating LLMs in multi-session coding interactions. Presented at: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); Jul 27 to Aug 1, 2025. [doi: 10.18653/v1/2025.acl-long.964]

Abbreviations

AI: artificial intelligence
AP: assistive product

XSL•FO
RenderX

Edited by S Munce; submitted 30.Oct.2025; peer-reviewed by A Elnashar, J Liu, M Al-Agil; accepted 19.Feb.2026; published 24.Mar.2026.

Please cite as:
Bañuls-Lapuerta FA, Marti-Miralles V, Gónzalez-García RJ, Martínez-Rico G
Using Natural Language Prompts With AI Models for Low-Cost Assistive Software Design: Exploratory Comparative Evaluation
JMIR Rehabil Assist Technol 2026;13:e86786
URL: https://rehab.jmir.org/2026/1/e86786
doi:10.2196/86786

© Francesc Antoni Bañuls-Lapuerta, Vicent Marti-Miralles, Rómulo Jacobo Gónzalez-García, Gabriel Martínez-Rico. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient-Perceived Factors Influencing Physical Activity Sensor Use in Stroke Prevention and Rehabilitation: Systematic Review of Qualitative Studies Using Thematic Synthesis

Paul Harris[1], MClinRes; Ingrid Maine[1,2], BAppSc (OT), MClinRes

[1]Department of Medical Education, The University of Melbourne, Grattan Street, Parkville, Melbourne, Victoria, Australia

[2]Division of Chronic and Complex Care, Western Health, St Albans, Melbourne, Victoria, Australia

Corresponding Author:
Paul Harris, MClinRes
Department of Medical Education, The University of Melbourne, Grattan Street, Parkville, Melbourne, Victoria, Australia

Abstract

Background: A robust correlation exists between physical activity (PA) and stroke risk reduction, and wearable PA sensors have emerged as promising adjuncts for rehabilitation and risk self-management. However, evidence regarding their long-term efficacy in facilitating sustained behavioral change remains equivocal.

Objective: This study aimed to explore the barriers and facilitators influencing the effective use of PA sensors in older adults at risk of primary or recurrent stroke.

Methods: A systematic search and thematic synthesis of qualitative research was conducted, focusing on PA sensor use among older adults at risk of primary or recurrent stroke. Given the emerging qualitative evidence base, inclusion was extended to proxy populations with analogous cardiovascular risk and functional profiles. Data were analyzed using line-by-line coding of primary text to generate descriptive themes and synthesize overarching analytical constructs.

Results: A search of 6 bibliographic databases (January 2010 to December 2023) identified 18 eligible studies. This systematic review and thematic synthesis revealed predominantly technological (user experience and device attributes) and psychological (motivation) barriers. Key facilitators were psychological (feedback and motivation), technological (user experience), and social/environmental supports. Higher-level analysis revealed a critical interrelationship between effective user engagement and optimally assistive device characteristics.

Conclusions: This review revealed a synergistic user-device interaction driving sustained PA in older adults at risk of primary or recurrent stroke. Future interventions developed in collaboration with patients and informed by the factors identified by this study will improve participation in rehabilitation and functional outcomes in this population.

International Registered Report Identifier (IRRID): RR2-10.1371/journal.pone.0301983

(JMIR Rehabil Assist Technol 2026;13:e86915) doi:10.2196/86915

KEYWORDS

sensors; stroke; physical activity; rehabilitation; qualitative research

Introduction

The terms “physical activity” (PA) and “exercise,” although often used interchangeably, are distinct concepts. PA broadly refers to any bodily movement resulting in energy expenditure, while exercise is more precisely defined as a subcomponent of PA that is “planned, structured, and... done to improve or maintain one or more components of physical fitness” [1]. Within this context, physical fitness encompasses the attributes and abilities an individual possesses or acquires to perform PA or exercise. Physical inactivity, or sedentary behavior, is a significant global risk factor for mortality associated with numerous adverse health outcomes including stroke [2].

As with other cardiovascular diseases, the pathogenesis of stroke is influenced by a range of risk factors, broadly categorized into nonmodifiable, medically modifiable, and lifestyle-modifiable attributes [2]. Nonmodifiable factors include inherent characteristics such as age and genetic predisposition. Medically modifiable factors are amenable to pharmacological or surgical interventions. Lifestyle-modifiable (behavioral) factors are those that can be altered through changes to daily habits and practices. Hypertension, for example, the most significant modifiable risk factor for stroke, can be effectively managed through both medical and lifestyle interventions, including pharmacotherapy, PA, dietary changes, and weight management. Despite the established efficacy of these interventions, one study reported that fewer than 50% of patients had their stroke or cardiovascular disease risk factors appropriately assessed,

XSL•FO
RenderX

treated, or managed [3]. Similarly, a recent audit of stroke rehabilitation services in Australia revealed that a comparable proportion of patients received information on self-management strategies or support programs [3]. PA is a critical, accessible, and cost-effective means of mitigating underlying risk factors, such as hypertension, contributing to the primary prevention of initial stroke events and the secondary prevention of recurrent stroke [4,5].

A growing body of evidence suggests that increased PA confers substantial health benefits across all age groups [6]. Notably, reduced sedentary time demonstrably lowers health risks and mortality, even when accounting for activity levels and established risk factors. Regular PA also mitigates the progression of hypertension and has been linked to improved weight management, reduced adiposity, and a decreased incidence of type 2 diabetes mellitus—all major risk factors for stroke. While increased PA significantly improves walking speed and balance, the evidence regarding the impact on poststroke mortality, dependence, and long-term disability remains insufficient or unclear [7]. A survey conducted in the United States revealed that a significant majority (71%) of older adults, 65 years and older, engage in self-monitoring for health indicators such as weight, diet, or exercise [8]. However, despite the prevalence of digital health tools, only a small minority (12%) utilize PA sensors and companion mobile apps. The predominant method for tracking health measures in this group was pen-and-paper [8]. The limited adoption of technology for health monitoring among this demographic may be attributable to a combination of objective factors such as PA sensor accuracy and subjective considerations such as perceived usability, which may influence engagement levels among older adults.

The need for increasingly sophisticated and persuasive strategies to motivate sustained healthy behavior remains a central challenge for clinicians designing interventions and engineers developing wearable PA sensors. Remarkably, few qualitative studies have thoroughly investigated the use of PA sensors in older adults, specifically exploring reasons for nonadoption and discontinuation. A deeper understanding of these factors is essential for the effective design and implementation of PA sensor-based interventions for stroke prevention and rehabilitation [9]. With that in mind, the current study aimed to explore the barriers and facilitators influencing the effective use of PA sensors in older adults at risk of primary or recurrent stroke.

Methods

Methodological Framework

This systematic review and thematic synthesis was conducted and reported according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Checklist 1) and the JBI (Joanna Briggs Institute) framework for evidence synthesis. Study quality was evaluated using JBI critical appraisal tools. Review methods were established a priori, as detailed in the published protocol [10]. The inclusion of proxy groups, such as individuals with hypertension or age-related motor deficits, is justified by their functional and psychological alignment with stroke survivors. These cohorts experience near-identical ergonomic barriers, such as the manual dexterity required for small device interfaces, and share psychological burdens, including diminished self-efficacy due to device failure, for example. By prioritizing the phenomenology of functional experience over narrow neurological diagnoses, this study applies conceptual transferability to enhance thematic depth where disease-specific evidence is still emerging [11-13]. The primary objective of the study was to identify and synthesize patient-perceived factors acting as either barriers or facilitators, influencing the use of wearable PA sensors in the context of stroke prevention and rehabilitation for older adults.

Inclusion and Exclusion Criteria

Peer-reviewed articles published in English between January 1, 2010, and December 31, 2023, that had a primary focus on wearable PA sensor use were included in the analysis. Studies were required to have a qualitative design or a significant qualitative component. The target population included older stroke survivors and populations with analogous clinical profiles. Eligible participants were characterized by shared cardiovascular risks (eg, atrial fibrillation, hypertension), metabolic conditions (eg, type 2 diabetes, dyslipidemia), and lifestyle-related factors. Inclusion was further defined by stroke-related functional deficits, specifically, impaired physical mobility (gait and balance), cognitive decline (executive dysfunction), and reduced independence in activities of daily living [14]. Study participants were stratified into 3 age cohorts to reflect distinct clinical and life-stage profiles. The 50- to 64-year age group ("Young-Old") was included to capture the critical transition period where stroke risk doubles, and this allowed for an analysis of whether PA sensor–based interventions performed differently in individuals with higher neuroplasticity and higher work/life stress [2,15]. The 65- to 84-year group ("Old-Old") represents the standard geriatric stroke demographic. Finally, those aged 85+ years ("Oldest-Old") were analyzed separately to ensure the unique frailty and comorbid complexities of the oldest-old did not skew the results of the younger cohorts.

Search Strategy

The methodology for this synthesis followed the prepublished protocol [10]. A high-sensitivity search strategy was conducted across MEDLINE, EMBASE, CINAHL, Cochrane, PsycINFO, and Scopus using MeSH and keywords related to wearable PA sensors and exercise. To maximize recall within an emerging qualitative evidence base, population-specific filters (eg, "stroke") were omitted; instead, the search was broadened to include all studies of populations with analogous risk and functional profiles. These results were supplemented by manual bibliography scans of included studies. This approach ensured a comprehensive capture of qualitative data despite the paucity of stroke-specific literature.

Screening and Data Extraction

Following the prepublished protocol, publication titles and abstracts were independently screened by two reviewers (IM and PH), with a pilot test conducted to ensure interrater reliability. Disagreements were resolved through consensus or third-party adjudication. Full-text screening and deduplication

XSL•FO
RenderX

were managed via Covidence. A standardized form was used to extract study characteristics (eg, demographics, research design, intervention specifics) and participant data, which were subsequently managed in NVivo. Methodological quality was evaluated using CASP (Critical Appraisal Skills Program) and JBI checklists [16,17]. Quality assessment focused on congruity of perspective, research questions or objectives, methodology, bias, participant voice, ethical standards, analysis, and conclusions. Papers were also evaluated for analytical richness, reflecting qualitative depth and interpretation with respect to the study aims.

Data Analysis

This systematic review and thematic synthesis was conducted in 3 stages, summarized below, and discussed in more detail in the study protocol [10].

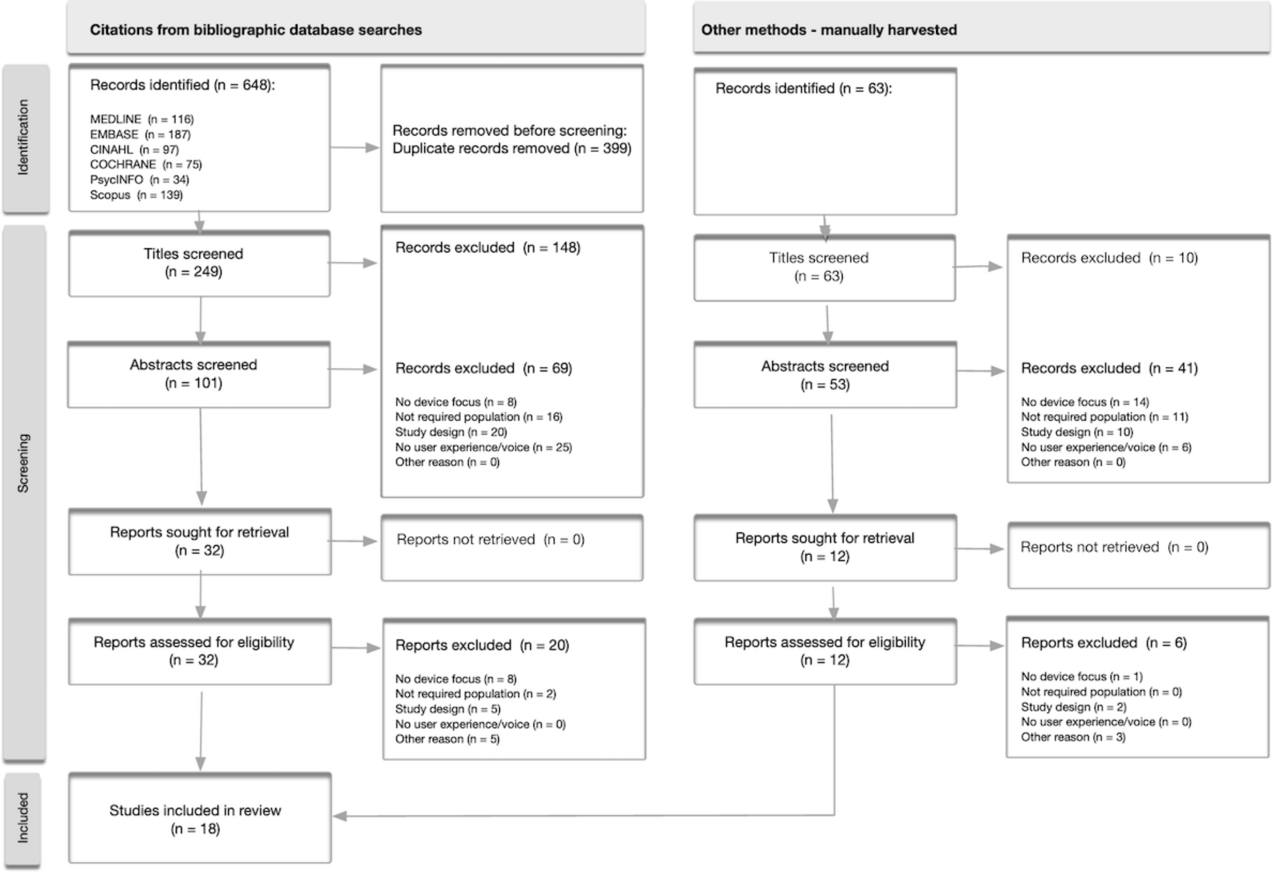
- *Stage 1. Inductive coding of included studies*: Reviewers independently coded the included studies, identifying salient sections and discussing coded sections as a team. A set of a priori codes based on the Person-Environment-Occupation-Performance and Technology Acceptance Model provided a provisional structure, with themes differentiated according to valence under binary (barrier and facilitator) top-level nodes [18,19].
- *Stage 2. Summarize under descriptive themes*: Related codes were aggregated into broader "descriptive themes" that summarized concepts across studies without going beyond the source texts. These higher-level descriptive themes were generated, discussed, and refined iteratively. Coded region frequencies (coding densities) and referenced study counts were analyzed to identify major and minor themes. The use of density metrics is well established in qualitative research, somewhat like matrix-coding queries, and is indicative of theme recurrence for included papers only [20,21].
- *Stage 3. Generate analytical constructs*: Descriptive themes were analyzed in relation to the research questions and objectives. Links between descriptive themes were mapped to generate analytical constructs and visualized diagrammatically to develop a coherent understanding from the evidence base.

Results

Study Selection and Characteristics

A systematic search of 6 electronic bibliographic databases yielded 648 records. An additional 63 records were identified through manual harvesting from other methods (ie, screening citation lists of included studies). After duplicate removal and screening, 44 studies were sought for full-text retrieval. Finally, 18 studies met the inclusion criteria and were included in this systematic review and thematic synthesis. The selection process and search results are summarized in the PRISMA flow diagram (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for systematic reviews.



XSL•FO
RenderX

To categorize heterogeneous participant data across the 18 studies listed in Table 1, a standardized coding system was applied:

- Population (N): [PRX] proxy; [STR] stroke
- Age: [YNG] 50-64; [OLD] 65-84; [ELD] ≥85 years
- Risk factors: [LFS] lifestyle; [MET] metabolic; [VAS] vascular
- Functional deficits: [ADL] activities of daily living; [COG] cognitive; [MOT] motor

XSL•FO
RenderX

Table . Characteristics of included studies.

Studies	Year	Location	Design	Duration	Pop (n)	Age group	Risk profile	Functional profile	Device(s)
Batsis et al [22]	2016	USA[a]	QLT[b]	1 mo	PRX[c] (8)	[OLD[d]]	[LFS[e], MET[f], VAS[g]]	[ADL[h], MOT[i]]	Fitbit
Nguyen et al [23]	2017	AUS[j]	QLT	18 wk	PRX (14)	[YNG[k], OLD]	[LFS, MET, VAS]	[ADL, MOT]	Various
Schlomann [24]	2017	DEU[l]	QLT	1 y	PRX (6)	[YNG, OLD]	[LFS, MET, VAS]	[MOT]	ViFit connect
Hamilton et al [25]	2018	AUS	QLT	6 mo	STR[m] (15)	[YNG, OLD, ELD[n]]	[LFS, MET, VAS]	[ADL, COG[o], MOT]	Fitbit
Johansson et al [26]	2018	SWE[p]	SYS[q], MIX[r]	NR[s]	STR (~43) PRX (54)	[YNG, OLD, ELD]	NR	[ADL, COG, MOT]	Various
Farina et al [27]	2019	GBR[t]	QLT	1 mo	STR (1) PRX (28)	[OLD, ELD]	[LFS, VAS]	[ADL, COG, MOT]	GENEActiv
Kononova et al [28]	2019	USA	QLT	6 wk	PRX (48)	[OLD, ELD]	[LFS, MET, VAS]	[ADL, MOT]	Various
Western et al [29]	2019	GBR	QLT	1 wk	PRX (31)	[YNG, OLD]	[LFS, MET, VAS]	[MOT]	SenseWear Pro
Whelan et al [30]	2019	GBR	MIX	6 wk	PRX (45)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Fitbit, Freestyle
Støve and Larson [31]	2019	DNK[u]	QLT	3 wk	STR (20) PRX (5)	[YNG, OLD]	[LFS, VAS]	[ADL, COG, MOT]	Garmin
Ummels et al [32]	2019	NLD[v]	QLT	2 - 8 wk	PRX (30)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Tracker
Östlind et al [33]	2022	SWE	QLT	12 wk	PRX (18)	[YNG, OLD]	[LFS, MET, VAS]	[MOT]	Fitbit
Gualtieri et al [34]	2016	USA	QLT	14 wk	PRX (10)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Withings Pulse
Mercer et al [35]	2016	CAN[w]	MIX	3+ d	STR (1) PRX (33)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Various
Randriambelonoro et al [36]	2017	CHE[x]	QLT	7 mo	PRX (18)	[YNG, OLD]	[LFS, MET, VAS]	[ADL, MOT]	Fitbit
Ehn et al [37]	2018	SWE	QLT	9 d	PRX (8)	[OLD, ELD]	[VAS, MET, LFS]	[MOT, ADL]	Activite, Jawbone
Takemoto et al [38]	2018	USA	QLT	3 wk	PRX (15)	[YNG, OLD, ELD]	[VAS, MET, LFS]	[MOT, ADL]	Various
Brickwood et al [39]	2020	AUS	MIX	1 y	PRX (20)	[OLD]	[VAS, MET, LFS]	[MOT, ADL]	Jawbone

[a] USA: United States.

[b] QLT: qualitative.

[c] PRX: proxy.

[d] OLD: Old-Old (65–84 y).

[e] LFS: lifestyle.

[f] MET: metabolic.

[g] VAS: vascular.

[h] ADL: activities of daily living.

[i] MOT: motor.

[j] AUS: Australia.

[k] YNG: Young-Old (50–64 y).

[l] DEU: Germany.

[m] STR: stroke survivor.

[n] ELD: Oldest-Old (≥85 y).

[o] COG: cognitive.

[p] SWE: Sweden.

[q] SYS: systematic review.

[r] MIX: mixed methods.

[s] NR: not reported.

[t] GBR: Great Britain.

[u] DNK: Denmark.

[v] NLD: Netherlands.

[w] CAN: Canada.

[x] CHE: Switzerland.

Of the included studies, 1 was a systematic review that included quantitative and qualitative studies, 3 employed mixed methods, and 14 were qualitative. Studies were conducted in North America (n=5), Australia (n=3), Great Britain (n=3), and Europe (n=7). A detailed breakdown of the quality and risk of bias assessment for each study is provided in Table S1 in Multimedia Appendix 1. The characteristics of the included studies with country designations according to International Organization for Standardization 3-letter standards (eg, DEU: Germany and CHE: Switzerland) are summarized in Table 1.

Approximately 78% (n=14) of the studies included individuals with metabolic risks like obesity, prediabetes, or type 2 diabetes. Participants with vascular risks, such as hypertension, heart disease, and high cholesterol, were even more common, appearing in 83% (n=15) of the studies. All 18 studies included participants with lifestyle-related risks, specifically high sedentary behavior and poor dietary choices. Functional deficits further defined these populations: 89% (n=16) of studies reported on participants with motor or physical limitations such as joint pain from osteoarthritis, reduced walking speed, or the use of mobility aids like walkers. Participants with cognitive deficits (eg, memory loss or forgetfulness, impaired task comprehension, and execution) were represented in 4 (22%) studies, and participants with impaired activities of daily living affecting self-care and mobility were represented in 6 (33%) studies.

Specific stroke-related representation was documented in 5 studies. Johansson et al [26] systematically reviewed 24 studies of stroke survivors, while Støve and Larsen [31] included patients with stroke (80%) in active inpatient rehabilitation with hemiparesis. Similarly, Hamilton et al [25] reported a sample where 60% had a neurological primary diagnosis, including stroke, alongside significant motor deficits. Finally, Mercer et al [35] included a participant with stroke-related gait asymmetry (shuffling) that impacted activity sensor accuracy, and Farina et al [27] included a participant with stroke-related vascular dementia but no acute motor symptoms.

Synthesis of Findings

The thematic coding frame in Table 2 was applied to all included studies and reviewed for consistency of interpretation. This process, known as “axial coding,” ensured coded themes remained connected to and understood within the context of the source studies [13].

Table . Low-level theme and subtheme nodes.

Theme	Subtheme
Built environment and technology	Acceptance, accuracy, attitudes, community, customizable, damage, familiarity, functionality, home, influences, privacy, reliability, safety, usability, usefulness, wearability
Cognitive	Attention, awareness, comprehension, language, memory, planning
Natural environment	Safety, terrain, weather
Neurobehavioral	Balance, gait, motor, sensory, vision
Physiological	Endurance, fatigue, flexibility, injury, mobility, strength
Psychological	Accountability, adherence, attainment, autonomy, competence, feedback, goals, habits, mood, motivation, preconceptions, privacy, self-awareness, self-efficacy, trust, well-being
Social support	Caregiver, information, peer, professional, training
Societal and economic	Attitudes, commitments, cost, cultural, norms, strata
Spiritual	Fulfillment, meaningfulness, values

XSL•FO
RenderX

Coding densities informed the generation of the higher-level descriptive (aggregate) themes and subthemes listed in Table 3. Density cutoff points were used to classify major (n≥30), minor (n=10 - 29), and negligible (n<10) themes. Finally, descriptive themes were analyzed with respect to the research questions and objectives of the current study.

Table . Descriptive themes and subthemes.

Theme	Subtheme
Technological	User experience, device attributes, security, other
Psychological	Self-identity, motivation, affective
Neurological	Neurophysiological, cognitive
Support	Social, environment

Densities were also used to visualize descriptive barrier and facilitator themes and subthemes. For readability, descriptive themes in this section have been indicated by quotation marks, and subthemes have been italicized.

Major Barriers

Major barrier themes were largely "technological," more specifically related to user *experience* (24.5% of barrier themes) and *device attributes* (20.1%), and "psychological," related to *motivation* (15.6%). *User experience* themes summarized in Multimedia Appendix 1 (Table S2) included older adults perceiving themselves as not "tech-savvy" but able to use basic functions such as step count. The perceived obtrusiveness of the PA sensor, both aesthetically and functionally, was a recurrent barrier theme. Devices were perceived to be designed for younger people and not older adults due to form factor, usability, and functionality. A lack of customization and the perceived irrelevance of some metrics hindered sustained engagement. Underlying *device attribute* themes referred to metric accuracy, the ability to register low-intensity activities such as standing, slow walking (strolling), or gardening, and to correctly record other activity types like yoga or strength training. The inability of PA sensors to correctly differentiate between similar activities such as sitting versus standing or to accurately track specific movements relevant to stroke recovery (eg, differentiated movements in hemiparesis) were seen as serious limitations. Users wanted more granular and relevant data (eg, time spent in moderate to vigorous PA) and exercise categorization (eg, fat-burning, cardio, peak); cadence and stride length for running and walking; stairs climbed as measures of vertical movement; swim metrics including laps, stroke type, and count; and strength training metrics such as repetition and set counts. *Motivation* themes referred to the "honeymoon period" or novelty effect of PA sensors quickly diminishing. Users lost interest once they felt they fully understood their activity patterns or had achieved their desired goals. Buzzing reminders were seen as annoying, especially when activity change was not possible, such as while driving. Goals that were not established in partnership with a therapist were often perceived as unrealistic, presenting a significant barrier to adherence. PA sensor unreliability or an inability to accurately measure certain activities (eg, nonlinear occupations and nonrepetitive movements) was demotivating. Users expressed a desire for real-time feedback that helped them to understand the impact of their PA level, for example, on blood sugar, rather than just step counts.

Major Facilitators

Major facilitator themes were "psychological" related to *motivation* (31.2% of facilitator themes), "technological" related to *user experience* (24.4%), and "support" related to both *social* and *environmental* support (21.4%). Underlying *motivation* themes summarized in Multimedia Appendix 1 (Table S3) included PA sensors that significantly increased self-awareness of PA levels. Achieving goals leads to feelings of inspiration, satisfaction, accountability, and a desire for additional effort. Long-term use was perceived to be strongly linked to intrinsic motivation, driven by personal enjoyment, feelings of accomplishment, and perceived benefits. Features like virtual incentives, nudges, reminders, and other forms of positive encouragement could act as effective extrinsic motivators, especially initially. PA sensors that effectively incorporated behavior change techniques such as self-monitoring, goal setting, and feedback promoted self-efficacy and PA. Individualized and tailored goals, and personalized feedback, for example, based on health status and functional capacity, were seen as critical for sustained motivation. Findings suggested that engagement with PA sensors was often nonlinear, with some users temporarily disengaging and later resuming use rather than abandoning the device entirely. Recognizing these flexible usage patterns is essential for developing interventions that foster long-term adherence. Underlying *user experience* themes included preferences for devices with simple, intuitive interfaces that required minimal user input. Users preferred PA sensors that were comfortable and unobtrusive, which did not interfere with daily activities and routines. Users expressed interest in tracking a more diverse range of activities such as swimming and cycling. The ability to customize goals, feedback type (vibration or visual), frequency of prompts, and data display was critical for effective motivation and sustained use. *Social* support themes related to therapists introducing and setting up the PA sensor, providing clear usage instructions, and interpreting data. This included personalized advice, goal setting, and addressing emotional responses to feedback. Social interaction and group engagement, both face-to-face and online via social networks, were regarded as powerful motivators. Embedding PA sensors into the clinical care process, with therapists actively discussing data and translating it into actionable insights, enhanced the perceived value and utility of the PA sensor. *Self-identity* themes referred to PA sensors significantly enhancing self-awareness of activity levels. Users were often surprised by their baseline level. Objective data helped users clarify real versus perceived activity. Tracking

XSL•FO
RenderX

progress and achieving goals, especially with immediate feedback, directly contributed to increased self-efficacy and confidence in the user's ability to be active and manage their health. For some, using the tracker shifted their self-perception, moving them from a state of pre-contemplation (unmotivated) to maintenance of a healthier lifestyle.

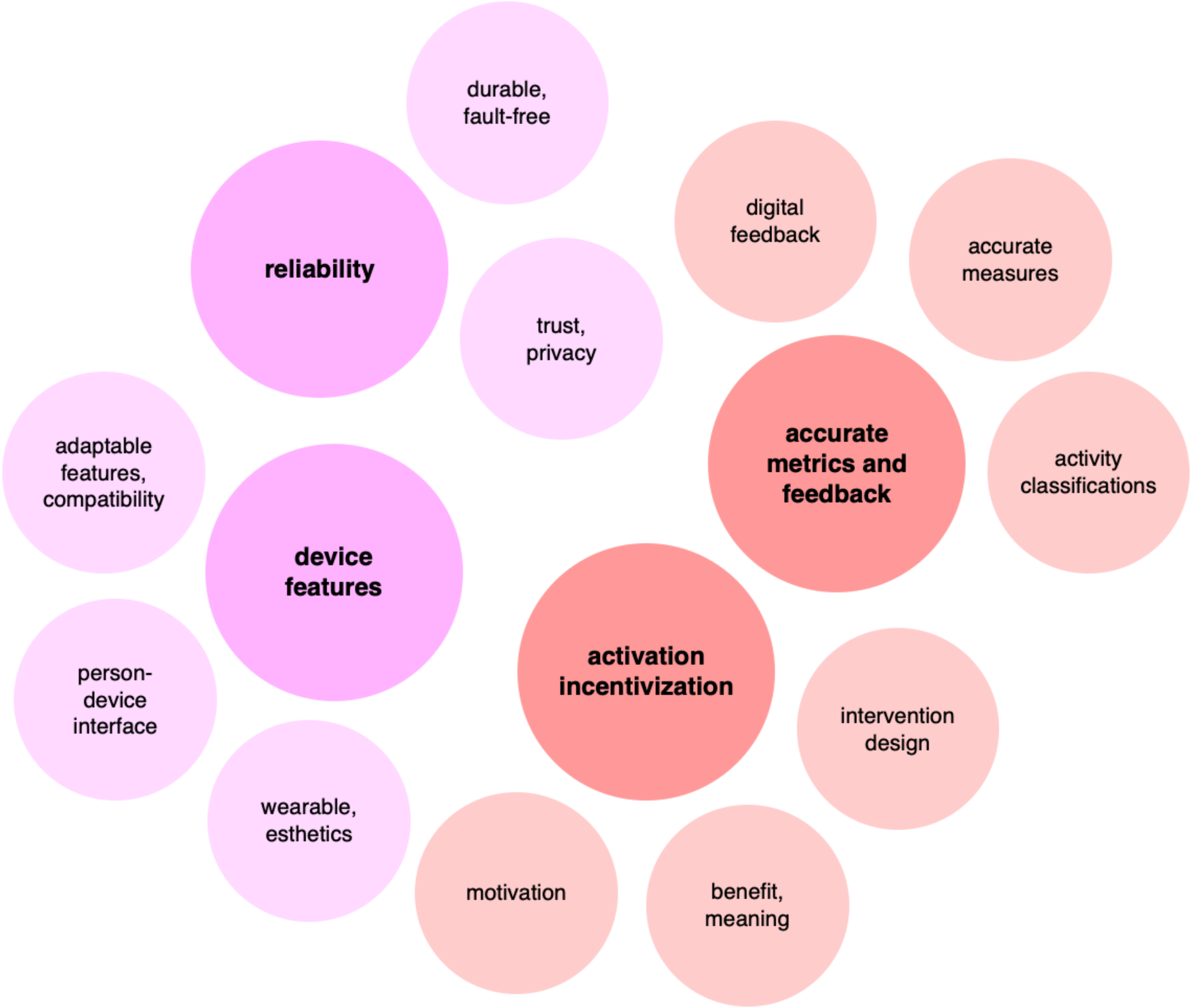
Visualized together in Figure 2, major barriers clustered mainly around *user experience* and *device attribute* aggregates, while facilitators were clustered around *user experience* and *motivation*. Drilling down to the underlying (low-level) themes, it was apparent that major *user experience* barriers were related to *wearability*, *usability*, and *familiarity* in descending order, and major *device attribute* barriers were *accuracy*, *reliability*, and *familiarity* in descending order. For facilitators, which comprised 60% of themes, most were clustered around *motivation*, *user experience*, and *support*. Underlying *motivation*, *feedback* (*motivation* more specifically), and *goals* in descending order were coded most frequently. Major *user experience* descriptive facilitators were *usability*, *wearability*, and *functionality*. Social and environmental *support* facilitators included most notably *professional*, that is, therapist and *caregiver* assistance.

Figure 2. Major barrier/facilitator descriptive and substrate themes. Dev. attr.: device attributes; Neurophys.: neurophysiological

The initial analysis revealed major barriers clustered around device themes: subjective *accuracy*, *wearability*, and *usability* (Figure 3). In this context, *accuracy* referred to the perceived precision of the PA sensor; *wearability* to characteristics that influenced the user's ability and willingness to wear the PA sensor consistently for extended periods (eg, size, fit, esthetics); and *usability* referred to the ease with which users interacted with the PA sensor and companion app.

XSL•FO
RenderX

Figure 3. Major barrier clustering by descriptive theme.



Predominant facilitators were "psychological," that is, *feedback*, *motivation*, *self-awareness*; device-specific; or related to professional support from the physician or therapist (Figure 4).

XSL•FO
RenderX

Figure 4. Major facilitator clustering by descriptive theme.

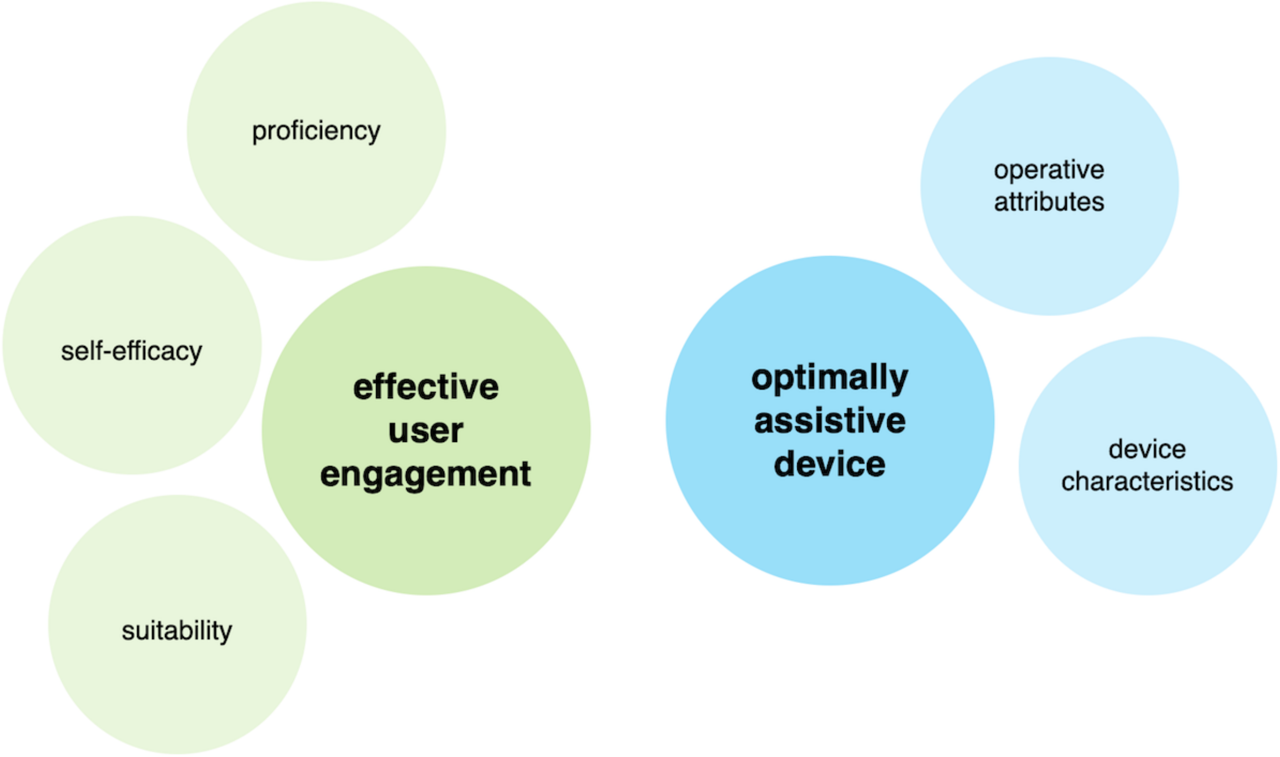
Building on the descriptive themes, higher-level analytical or theory-driven constructs were developed. Two distinct analytical constructs emerged, one person-centric, emphasizing factors related to user experience and engagement, and the other essentially device-centric, focused on aspects attributable to the PA sensor itself. Together, these suggested a synergistic interrelationship between effective user engagement and optimally assistive device factors, highlighting their combined contribution to long-term behavior change (Table 4 and Figure 5).

XSL•FO
RenderX

Table . Analytical constructs, descriptive themes, and subthemes.

Analytical construct and descriptive themes	Subthemes
Effective user engagement	
Proficiency	• Requisite skill(s) and knowledge: *degree of complexity; training and information* • Independent operation: *acceptability; ease of use*
Self-efficacy	• Supportive environment: *cost, commitments, natural environment; professional support; family, peers, friends, carers*
Suitability	• Appropriate for population: *cognitive, sensorimotor ability/impairment; cultural congruence; older adults, chronic conditions* • Thoughts and preconceptions: *values, attitudes; expectations; self-awareness*
Optimally assistive device	
Operative attributes	• Activation, incentivization: *motivation; intervention design; perceived benefit and meaningfulness* • Accurate metrics, feedback: *accurate measures; activity type/classification; useful and appropriate feedback*
Device characteristics	• Reliability: *durable, fault-free; trust, privacy* • Device features: *aesthetics and wearability; person-device interface; adaptable features, compatibility*

Figure 5. High-level analytical constructs.



Effective user engagement, as a “person-centric” analytical construct, represented the extent to which users meaningfully and effectively interacted with PA sensors. Contributing to this overarching construct were 3 key descriptive themes: *Proficiency* referred to the level of skill and technical knowledge required by the user and their capacity to use the PA sensor independently. Inadequate training, lack of technical skills, inability to resolve usage issues independent of external support, and poor instructions were major technical barrier themes evident in 9 of the studies included in the current analysis [25,26,28,30,32,34-37]. When users felt confident in their ability to operate a PA sensor correctly, they were more likely to engage with it. *Self-efficacy* referred to the individual’s belief in their capacity or readiness to initiate PA to achieve their goals. This encompassed the ability to integrate PA within personal and broader social/cultural contexts to establish or

XSL•FO
RenderX

participate in enabling social networks, exercise buddies, peers, and family, and to become better informed about their health, the benefits of regular activity, and maintaining long-term self-motivation. An interaction between motivation and adherence was also evident, suggesting an interplay between the analytical themes discussed in this section, that is, how PA sensors might facilitate a sense of self-efficacy or self-empowerment that in turn promoted increased activity. A lack of self-efficacy or belief in one's ability to effectively engage in PA and achieve goals was identified as a significant theme in 8 of the studies [22,25,28-30,33,34,39]. Human support from peers, family, coaches, and clinicians was seen as an essential adjunct that aided increased PA together with PA sensor usage; the absence of these supports was evident in 4 studies as major barriers [24,25,32,36]. *Suitability* referred to the appropriateness and acceptability of the PA sensor within the target population. This reflected how well the PA sensor aligned with the physical, cognitive, and cultural needs of the user.

Optimally assistive device, as a “device-centric” analytical construct, referred to the inherent characteristics of the PA sensor and its capacity to effectively promote increased PA. Contributing to this were 2 primary descriptive themes: *Operative attributes* that referred to the functional performance of the PA sensor, accurately capturing and conveying relevant activity data to the user. This included precision, for example, accurate step counts, correct activity classification, and the clarity, relevance, and appropriateness of feedback provided. *Device characteristics*, on the other hand, referred to the design elements of the PA sensor. This included user-friendliness and ease of use of the PA sensor, functional capacity, that is, support for required or expected functionality, timeliness and appropriateness of information provided, and integration with other devices such as the companion app. Perceived inaccuracies and PA sensor defects were major barriers evident in 8 of the studies [23,27,29,33,36,37,40,41]. Similarly, themes that referred to a perceived lack of PA sensor validation, or characteristics that did not appear to have been designed for the abilities of older persons outlined earlier (eg, form factor, usability, functionality), highlighted broader device characteristic issues. Wearability themes reflected how easily people were able to use the PA sensor in everyday life situations and adapt usage to daily routines.

Discussion

Principal Findings

This study was primarily undertaken to address noted inconsistencies in the long-term efficacy of device-based PA interventions, particularly in older adults at risk of primary or recurrent stroke. Through a review of the qualitative literature, nuanced human and device-related factors influencing sustained engagement in this specific population were explored through a qualitative lens. By synthesizing lived experiences, insights emerged that explained why these inconsistencies occurred, reflecting individual and contextual factors at play. This qualitative depth is crucial for bridging the implementation gap between technological potential and real-world impact. These findings extend the understanding of existing models by demonstrating how device characteristics must be inherently assistive and person-centric to promote sustained engagement, especially in more vulnerable populations.

The preliminary synthesis was underpinned by a binary schema, categorizing overarching descriptive themes under top-level barrier and facilitator nodes. This implicitly assumed influencing factors were mutually exclusive. However, emergent themes within the included studies suggested a more nuanced reality. A similar observation might be made regarding the conceptual dichotomy between human and device factors. Indeed, one systematic review highlighted an interrelationship between these factors, reporting improved adherence in studies with longer monitoring durations [26]. The authors of the review posited that this could be attributed to increased user confidence over time. Consistent with this finding, studies included in the current analysis similarly indicated an interplay between motivation, technical integrity (eg, absence of defects, data accuracy), and adherence factors, suggesting that these elements are not isolated but rather interact synergistically in complex ways to influence outcomes. Studies included in the current analysis consistently indicated an interplay between motivation and adherence to PA.

Numerous theoretical frameworks have attempted to elucidate behavior change mechanisms through constructs such as self-efficacy (perceived confidence), decisional balance (weighing pros and cons), and processes of change (covert and overt activities facilitating progression) [42]. In the context of PA, self-efficacy specifically refers to an individual's confidence in identifying and achieving planned PA goals. This confidence is reinforced through mastery experiences, vicarious learning (ie, observing others' success), persuasion, and performance feedback [40]. While some behavioral goals are characterized by defined end states, others, such as increased PA, necessitate ongoing motivation for sustained maintenance. It has been suggested that individuals are more inclined to adopt healthy behaviors when confronted with a substantial threat and when their actions are perceived as effectively reducing the likelihood or severity of that threat [43]. This applies to patients with stroke concerned about secondary events, as well as individuals at risk of a primary stroke. However, the absence of universally agreed-upon construct measures complicates efforts to isolate which specific intervention components successfully drive behavior change and which do not.

The effectiveness of health care interventions and PA sensors is often complicated by social dynamics, as individuals' interactions can lead to unforeseen outcomes that alter the intended effects of the intervention [44]. For researchers, deciphering the complex evolution and adaptation of behaviors stemming from these interactions is a formidable task. This complexity is arguably amplified in PA sensor–based interventions, given the intricate interplay of human and device factors. The current study revealed prominent device-specific themes, with technical barriers featuring more significantly than psychological barriers related to motivation, support, adherence, or physiological limitations such as sensorimotor impairment.

While PA sensor–based devices offer objective measures of PA, potentially reducing biases inherent in self-report

XSL•FO
RenderX

instruments, themes related to perceived accuracy and trustworthiness were nonetheless evident in this synthesis [28]. The unregulated nature of the consumer market and clinometric quality of consumer-grade PA sensors warrant consideration in this context. A systematic review assessing consumer-grade PA sensors in older adults (mean age 70.2, SD 4.8 y) reported high accuracy for average daily step counts compared to research-grade reference PA sensors [41]. However, the reliability of daily step counts varied considerably across PA sensors, with gait, device placement, and walking speed influencing reliability. Several of the consumer-grade PA sensors contributing to the current thematic synthesis were also evaluated in the aforementioned review. While these findings offer some reassurance regarding accuracy, they simultaneously raise questions about public perceptions of consumer PA sensor accuracy and critical intervention design factors such as training and support. The level of support required for effective PA sensor use also emerged as a recurring major theme in the current study. The degree of intuitiveness and interactivity of the PA sensors may be contributing factors.

The synergistic factors identified in the current study suggest that future interventions should prioritize co-design methodologies, actively involving older patients with stroke in the development and refinement of PA sensors and health care interventions based on these technologies. This would ensure that PA sensors are not only functionally more robust but also intuitively usable, comfortable, and psychologically reinforcing for this specific cohort. For clinicians, these findings underscore the importance of tailored support, including training and goal setting, while addressing individual user concerns regarding PA sensor accuracy and usability when employing these devices in rehabilitation or preventive health care interventions. Future research should focus on developing and rigorously testing co-designed PA sensor–based interventions, employing mixed methods approaches to quantitatively validate the impact of the qualitative insights identified here on sustained levels of PA. Specific areas for investigation might include the optimal frequency and type of professional support, the most effective feedback mechanisms, and ergonomic design features critical in the context of stroke prevention and rehabilitation.

While the hypothesized advantages of PA sensors have garnered considerable scholarly attention, a notable gap exists in the literature regarding their application and, more critically, the lived experience of their use by stroke survivors and older adults at risk of stroke. The current study partially addresses that gap by synthesizing existing qualitative research concerning the utilization of PA sensors within this specific demographic. Through a rigorous qualitative synthesis, both major thematic elements and less predominant, yet illuminating, experiential aspects associated with PA sensor use were identified. The insights gleaned from these findings hold significant implications for the development of more efficacious interventions aimed at mitigating sedentary behavior and fostering increased PA in older adults at risk of primary or recurrent stroke.

Limitations

Despite a comprehensive search, only 2 included studies specifically addressed older stroke survivors: a qualitative study of 15 survivors and a systematic review incorporating one qualitative stroke study and 5 qualitative proxy population studies [25,26]. While limited stroke-specific data might traditionally constrain the generalizability of these findings, the methodology employed prioritized *conceptual transferability* over statistical representativeness [13]. By extending inclusion to proxy populations with analogous risk and functional profiles, this synthesis identified higher-level analytical constructs that are specifically relevant to older adults at risk of primary or recurrent stroke. These overarching themes provide a robust explanatory framework for PA sensor–based interventions that is theoretically relevant to the stroke experience, even as empirical qualitative data for this population are emerging. However, it is acknowledged that the inherent subjectivity of qualitative interpretation means this synthesis may represent one of several possible frameworks. Furthermore, the lack of detailed technical specifications across the diverse range of included PA sensors precluded an analysis of device-specific effects, such as wearability or discrete usability features. Future research should aim to validate these conceptual constructs through larger, stroke-specific cohorts and more granular investigations into device-specific factors.

Conclusion

By pooling existing qualitative research, this study contributes a novel and deeper understanding of the benefits and limitations of wearable PA sensors through the subjective experiential lens of older adults at risk of primary or recurrent stroke. This systematic review and thematic synthesis identified a synergistic interplay between human and device-specific factors that influence engagement and sustained increased PA. However, the precise nature of this relationship and its impact on objective PA levels would require verification through higher levels of evidence. These findings provide a framework for more effective PA sensor–based interventions in stroke prevention and rehabilitation, offering valuable insights for technologists designing PA sensors and therapists developing PA sensor–based health interventions for older stroke cohorts.

Acknowledgments

This study was based on a clinical research thesis written by PH under the supervision of A/Prof Anita Horvath from the Department of Medical Education, University of Melbourne. The coding structure and review process were jointly developed by IM and PH based on the thematic synthesis methodology. Dr Phillis Lau from the Department of General Practice, University of Melbourne, reviewed the methodological rigor of both the thesis and published protocol. Review comments for a preliminary analysis were also provided by Prof Kathleen Gray from the Centre for Digital Transformation of Health, University of Melbourne. All authors have read and approved the final version of the manuscript.

XSL•FO
RenderX

Funding

The authors received no specific funding for this work.

Data Availability

The detailed search strategy, selection process documentation, PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)/JBI (Joanna Briggs Institute) checklists, and Data Management Plan are openly available on Figshare [45].

Authors' Contributions

PH conceived and designed the study, conducted the search and data curation, and performed the primary project administration. PH and IM collaborated on the methodology and formal analysis. PH drafted the original manuscript. IM and PH contributed to the critical review and editing of the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Quality and risk of bias assessments for included studies; barrier themes, coded region frequencies (density), and source study counts; facilitator themes, coded region frequencies (density), and study references; and methodological data.
[DOCX File, 279 KB - rehab_v13i1e86915_app1.docx]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 85 KB - rehab_v13i1e86915_app2.pdf]

References

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131. [Medline: 3920711]
2. Boehme AK, Esenwa C, Elkind MSV. Stroke risk factors, genetics, and prevention. Circ Res 2017 Feb 3;120(3):472-495. [doi: 10.1161/CIRCRESAHA.116.308398] [Medline: 28154098]
3. Cooper C, Gross A, Brinkman C, et al. The impact of wearable motion sensing technology on physical activity in older adults. Exp Gerontol 2018 Oct 2;112:9-19. [doi: 10.1016/j.exger.2018.08.002] [Medline: 30103025]
4. Kelley GA, Kelley KS. Leisure time physical activity reduces the risk for stroke in adults: a reanalysis of a meta-analysis using the inverse-heterogeneity model. Stroke Res Treat 2019;2019:8264502. [doi: 10.1155/2019/8264502] [Medline: 31275539]
5. Lee CD, Folsom AR, Blair SN. Physical activity and stroke risk: a meta-analysis. Stroke 2003 Oct;34(10):2475-2481. [doi: 10.1161/01.STR.0000091843.02517.9D] [Medline: 14500932]
6. Chodzko-Zajko WJ, Proctor DN, Fiatarone Singh MA, et al. Exercise and physical activity for older adults. Med Sci Sports Exerc 2009;41(7):1510-1530. [doi: 10.1249/MSS.0b013e3181a0c95c]
7. Saunders DH, Sanderson M, Hayes S, et al. Physical fitness training for stroke patients. Cochrane Database Syst Rev 2016 Mar 24;3(3):CD003316. [doi: 10.1002/14651858.CD003316.pub6] [Medline: 27010219]
8. Fox S, Duggan M. Tracking for health. : Pew Research Center; 2013 Jan 28 URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/media/Files/Reports/2013/PIP_TrackingforHealth-with-appendix.pdf? [accessed 2026-02-21]
9. Zhang MW, Chew PY, Yeo LL, Ho RC. The untapped potential of smartphone sensors for stroke rehabilitation and after-care. Technol Health Care 2016;24(1):139-143. [doi: 10.3233/THC-151099] [Medline: 26484884]
10. Harris PT, Maine I. Patient-perceived factors influencing physical activity sensor use in stroke prevention and rehabilitation: a thematic synthesis protocol. PLOS ONE 2024;19(4):e0301983. [doi: 10.1371/journal.pone.0301983] [Medline: 38687706]
11. Booth A, Papaioannou D, Sutton A. Systematic Approaches to a Successful Literature Review: SAGE Publications; 2012. URL: https://www.researchgate.net/profile/Andrew-Booth-2/publication/235930866_Systematic_Approaches_to_a_Successful_Literature_Review/links/5da06c7f45851553ff8705fa/Systematic-Approaches-to-a-Successful-Literature-Review.pdf [accessed 2026-02-21]
12. Sandelowski M. “To be of use”: enhancing the utility of qualitative research. Nurs Outlook 1997;45(3):125-132. [doi: 10.1016/s0029-6554(97)90043-9] [Medline: 9210160]
13. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
14. O’Donnell MJ, Chin SL, Rangarajan S, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. Lancet 2016 Aug 20;388(10046):761-775. [doi: 10.1016/S0140-6736(16)30506-2] [Medline: 27431356]

XSL•FO
RenderX

15. Lundebjerg NE, Trucil DE, Hammond EC, Applegate WB. When it comes to older adults, language matters: Journal of the American Geriatrics Society adopts modified American Medical Association style. J Am Geriatr Soc 2017 Jul;65(7):1386-1388. [doi: 10.1111/jgs.14941] [Medline: 28568284]
16. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Healthc 2015 Sep;13(3):132-140. [doi: 10.1097/XEB.0000000000000055] [Medline: 26360830]
17. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc 2015 Sep;13(3):179-187. [doi: 10.1097/XEB.0000000000000062] [Medline: 26262565]
18. Baum C, Christiansen C. The person-environment-occupation-performance (PEOP) model. In: Christiansen CH, Baum CM, Bass JD, editors. Occupational Therapy: Performance, Participation, and Well-Being, 4th edition: SLACK Incorporated; 2015:49-56.
19. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
20. Bassett B. Computer-based analysis of qualitative data: NVIVO. In: Mills AJ, Durepos G, Weibe E, editors. Encyclopedia of Case Study Research: Sage Publications; 2010:193-194. [doi: 10.4135/9781412957397.n71]
21. Dulmage BO, Akintilo L, Welty LJ, Davis MM, Colavincenzo M, Xu S. A qualitative, cross-sectional study of positive and negative comments of residency programs across 9 medical and surgical specialties. Am J Med 2018 Sep;131(9):1130-1134. [doi: 10.1016/j.amjmed.2018.05.019] [Medline: 29908767]
22. Batsis JA, Naslund JA, Gill LE, Masutani RK, Agarwal N, Bartels SJ. Use of a wearable activity device in rural older obese adults: a pilot study. Gerontol Geriatr Med 2016;2:2333721416678076. [doi: 10.1177/2333721416678076] [Medline: 28138502]
23. Nguyen NH, Hadgraft NT, Moore MM, et al. A qualitative evaluation of breast cancer survivors' acceptance of and preferences for consumer wearable technology activity trackers. Support Care Cancer 2017 Nov;25(11):3375-3384. [doi: 10.1007/s00520-017-3756-y] [Medline: 28540402]
24. Schlomann A. A case study on older adults' long-term use of an activity tracker. Gerontechnology 2017 Aug 9;16(2):115-124. [doi: 10.4017/gt.2017.16.2.007.00]
25. Hamilton C, McCluskey A, Hassett L, Killington M, Lovarini M. Patient and therapist experiences of using affordable feedback-based technology in rehabilitation: a qualitative study nested in a randomized controlled trial. Clin Rehabil 2018 Sep;32(9):1258-1270. [doi: 10.1177/0269215518771820] [Medline: 29696990]
26. Johansson D, Malmgren K, Alt Murphy M. Wearable sensors for clinical applications in epilepsy, Parkinson's disease, and stroke: a mixed-methods systematic review. J Neurol 2018 Aug;265(8):1740-1752. [doi: 10.1007/s00415-018-8786-y] [Medline: 29427026]
27. Farina N, Sherlock G, Thomas S, Lowry RG, Banerjee S. Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia. Int J Geriat Psychiatry 2019 Apr;34(4):617-624. [doi: 10.1002/gps.5064]
28. Kononova A, Li L, Kamp K, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. JMIR Mhealth Uhealth 2019 Apr 5;7(4):e9832. [doi: 10.2196/mhealth.9832] [Medline: 30950807]
29. Western MJ, Thompson D, Peacock OJ, Stathi A. The impact of multidimensional physical activity feedback on healthcare practitioners and patients. BJGP Open 2019 Apr;3(1):bjgpopen18X101628. [doi: 10.3399/bjgpopen18X101628] [Medline: 31049409]
30. Whelan ME, Orme MW, Kingsnorth AP, Sherar LB, Denton FL, Esliger DW. Examining the use of glucose and physical activity self-monitoring technologies in individuals at moderate to high risk of developing type 2 diabetes: randomized trial. JMIR Mhealth Uhealth 2019 Oct 28;7(10):e14195. [doi: 10.2196/14195] [Medline: 31661077]
31. Støve MP, Larsen BT. Self-monitoring – usability evaluation of heart rate monitoring using wearable devices in patients with acquired brain injury. Eur J Physiother 2020 Nov 1;22(6):364-372. [doi: 10.1080/21679169.2019.1628300]
32. Ummels D, Beekman E, Moser A, Braun SM, Beurskens AJ. Patients' experiences with commercially available activity trackers embedded in physiotherapy treatment: a qualitative study. Disabil Rehabil 2020 Nov;42(23):3284-3292. [doi: 10.1080/09638288.2019.1590470] [Medline: 30973026]
33. Östlind E, Ekvall Hansson E, Eek F, Stigmar K. Experiences of activity monitoring and perceptions of digital support among working individuals with hip and knee osteoarthritis - a focus group study. BMC Public Health 2022 Aug 30;22(1):1641. [doi: 10.1186/s12889-022-14065-0] [Medline: 36042425]
34. Gualtieri L, Rosenbluth S, Phillips J. Can a free wearable activity tracker change behavior? The impact of trackers on adults in a physician-led wellness group. JMIR Res Protoc 2016 Nov 30;5(4):e237. [doi: 10.2196/resprot.6534] [Medline: 27903490]
35. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7. [doi: 10.2196/mhealth.4225] [Medline: 26818775]
36. Randriambelonoro M, Chen Y, Geissbuhler A, Pu P. Exploring physical activity monitoring devices for diabetic and obese patients. Presented at: UbiComp/ISWC'15 Adjunct: Adjunct Proceedings of the 2015 ACM International Joint Conference

XSL•FO
RenderX

on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers; Sep 7-11, 2015 URL: http://dl.acm.org/citation.cfm?doid=2800835 [accessed 2026-02-21]
37. Ehn M, Eriksson LC, Åkerberg N, Johansson AC. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 1;6(2):e34. [doi: 10.2196/mhealth.8345] [Medline: 29391342]
38. Takemoto M, Lewars B, Hurst S, et al. Participants' perceptions on the use of wearable devices to reduce sitting time: qualitative analysis. JMIR Mhealth Uhealth 2018 Mar 31;6(3):e73. [doi: 10.2196/mhealth.7857] [Medline: 29599105]
39. Brickwood KJ, Williams AD, Watson G, O'Brien J. Older adults' experiences of using a wearable activity tracker with health professional feedback over a 12-month randomised controlled trial. Digit Health 2020;6:2055207620921678. [doi: 10.1177/2055207620921678] [Medline: 32426152]
40. Jones F, Riazi A. Self-efficacy and self-management after stroke: a systematic review. Disabil Rehabil 2011;33(10):797-810. [doi: 10.3109/09638288.2010.511415] [Medline: 20795919]
41. Straiton N, Alharbi M, Bauman A, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]
42. Prochaska JM. The transtheoretical model applied to the community and the workplace. J Health Psychol 2007 Jan;12(1):198-200. [doi: 10.1177/1359105307071754] [Medline: 17158853]
43. Jones CJ, Smith H, Llewellyn C. Evaluating the effectiveness of health belief model interventions in improving adherence: a systematic review. Health Psychol Rev 2014;8(3):253-269. [doi: 10.1080/17437199.2013.802623] [Medline: 25053213]
44. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. BMJ 2001 Sep 15;323(7313):625-628. [doi: 10.1136/bmj.323.7313.625] [Medline: 11557716]
45. Harris P, Maine I. Patient-perceived factors influencing physical activity sensor use in stroke prevention and rehabilitation. Figshare. URL: https://doi.org/10.6084/m9.figshare.c.7992074 [accessed 2026-02-26]

Abbreviations

CASP: Critical Appraisal Skills Program
JBI: Joanna Briggs Institute
PA: physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by S Munce; submitted 31.Oct.2025; peer-reviewed by A Al-Muqarm, H Tariq; revised version received 09.Feb.2026; accepted 10.Feb.2026; published 12.Mar.2026.

Please cite as:
Harris P, Maine I
Patient-Perceived Factors Influencing Physical Activity Sensor Use in Stroke Prevention and Rehabilitation: Systematic Review of Qualitative Studies Using Thematic Synthesis
JMIR Rehabil Assist Technol 2026;13:e86915
URL: https://rehab.jmir.org/2026/1/e86915
doi:10.2196/86915

© Paul Harris, Ingrid Maine. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Technology-Driven Group Exercise Program Implementation in an Underserved Community: Multimethod Retrospective Evaluation Study

Whitney N Neal[1*], PhD; Laurie A Malone[2*], PhD, MPH

[1]Department of Health Behavior, School of Public Health, University of Alabama at Birmingham, 1665 University Boulevard, Birmingham, United States

[2]Department of Occupational Therapy, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States

[*]all authors contributed equally

Corresponding Author:
Whitney N Neal, PhD
Department of Health Behavior, School of Public Health, University of Alabama at Birmingham, 1665 University Boulevard, Birmingham, United States

Abstract

Background: Physical inactivity among people with disabilities and older adults is a persistent public health concern, particularly those from racial minority groups living in underserved communities where structural barriers limit access to exercise opportunities. Technology-driven exercise programs offer scalable solutions, but the contextual factors that influence their uptake, fidelity, and sustainability remain underexplored.

Objective: This study applied the Consolidated Framework for Implementation Research (CFIR) to identify barriers, facilitators, and lessons learned during the implementation of 2 technology-driven group exercise programs—synchronous online fitness and active virtual reality gaming—delivered to predominantly African American older adults with disabilities and/or chronic health conditions in an underserved community.

Methods: Using a multimethod design, we conducted semistructured interviews and focus groups with program participants, staff, and administrators (n=15) and administered implementation surveys to program deliverers (n=9). Qualitative data were analyzed deductively using CFIR domains and constructs, while survey data provided descriptive context.

Results: Both programs achieved high acceptability and appropriateness scores, with feasibility rated slightly lower for virtual reality. Themes mapped to 4 CFIR domains (innovation, outer setting, inner setting, and individuals). Qualitative analysis identified adaptability, affordability, engaging design, and community support as key facilitators, while barriers included technology and space constraints, resistance to new technology, and sustainability challenges.

Conclusions: Technology-driven exercise programs can expand access to physical activity in underserved communities when they are adaptable, affordable, and socially engaging. Addressing multilevel barriers, including resource limitations, technology hesitancy, and long-term sustainability, is critical to scaling and sustaining these interventions.

(JMIR Rehabil Assist Technol 2026;13:e79598) doi:10.2196/79598

KEYWORDS

physical activity; underserved communities; group fitness intervention; program implementation; exercise

Introduction

Despite the known health benefits of being physically active, people with disabilities report high levels of physical inactivity. Reports worldwide indicate that approximately 54% to 91% of adults without disabilities meet physical activity recommendations, compared with only 21% to 60% of people with disabilities [1]. In the United States, 1 in 4 adults is living with a disability [2], and sedentary people with disabilities are 3 times more likely to report primary health conditions such as heart disease, stroke, diabetes, or cancer compared to active people with disabilities [3]. Furthermore, physical inactivity exacerbates loneliness and social isolation, which are also associated with adverse health outcomes, including heart disease and stroke [4]. Of concern is that individuals with disabilities are not only physically inactive but often experience higher levels of social isolation and loneliness than persons without disabilities [5]. Together, these disparities underscore the urgent need for effective, inclusive strategies to promote physical activity among people with disabilities.

Barriers to physical activity for people with disabilities span multiple levels of the socioecological model, including individual (eg, lack of knowledge, physical limitations, pain,

XSL•FO
RenderX

and fatigue), interpersonal (eg, lack of social support), institutional or organizational (eg, lack of accessible fitness facilities and exercise equipment), community (eg, poorly designed outdoor spaces and street connectivity), and systemic (eg, limited or no transportation) factors [1,6-10]. These multifactorial barriers limit opportunities for people with disabilities to engage in effective and sustainable physical activity and are further compounded in underserved communities—those characterized by racial and ethnic health disparities, high poverty rates, low insurance coverage, and limited health care and physical activity resources—underscoring the need to address physical activity barriers within the socioecological framework [11,12].

Technology-driven physical activity [13] and health promotion programming can help eliminate many of these barriers by delivering accessible, scalable interventions that foster social engagement and a sense of community [14-17]. Recent systematic reviews and meta-analyses demonstrate that mobile and wearable technologies, telehealth-delivered exercise, online fitness programs, and interactive gaming platforms can increase physical activity levels and improve body composition and physical function, with consistent benefits shown across home-based and community settings [18-20]. Telehealth programs have shown improvements in physical functioning even among individuals with frailty, cognitive challenges, or mobility limitations, highlighting the value of remote delivery for populations facing barriers to traditional exercise access [21]. Adherence to technology-based exercise is generally favorable when programs are accessible, engaging, and designed to meet participant needs [22]. Interactive and game-based technologies further extend reach and appeal: virtual reality (VR) exergame programs have been shown to improve balance, mobility, and functional outcomes in older adults, including those living in long-term care, while enhancing enjoyment and motivation to be active [23-25]. Work by Malone et al [15,26] demonstrates that adults with mobility impairments can engage with adapted exergaming systems—such as modified Wii Fit balance boards and gaming mats—to achieve greater energy expenditure and report high enjoyment, suggesting these technologies can reduce barriers and support participation in meaningful movement. Additionally, online membership–based fitness programs have shown feasibility and preliminary effectiveness for increasing physical activity in people with mobility impairments [27]. Collectively, this body of evidence supports the use of technology-driven physical activity interventions as accessible, engaging, and adaptable options for promoting physical activity and supporting health in populations traditionally underserved by conventional exercise models.

Although technology-mediated exercise shows promise, evidence of efficacy and feasibility alone is insufficient to ensure successful deployment in resource-limited settings. To translate and sustain such programs in underserved settings, it is essential to examine multilevel determinants of implementation. Implementation science offers a structured approach to understanding and addressing contextual factors that influence the uptake, fidelity, and sustainability of health interventions [28]. The Consolidated Framework for Implementation Research (CFIR) provides a comprehensive, validated taxonomy for assessing contextual factors that influence implementation success across 5 domains: outer setting, inner setting, innovation, implementation process, and individuals [29,30]. Using a framework like the CFIR ensures that barriers and facilitators are identified and addressed in a structured, reproducible manner, strengthening both the science and practice of health promotion program delivery in underserved settings. The CFIR has been used to identify implementation barriers and facilitators of physical activity interventions for older adults and people with disabilities [31-33], and in a systematic review of community-based physical activity interventions [34]. However, research on technology-driven exercise interventions in underserved communities remains limited.

To address this gap, this study used the CFIR to identify facilitators, barriers, and lessons learned during the implementation of 2 technology-driven group exercise programs (synchronous online fitness and VR exergaming) delivered to older adults (aged >60 y) and individuals with disabilities and/or chronic health conditions living in an underserved community. Our primary aims were to (1) characterize the CFIR constructs that most strongly influenced each program's implementation outcomes and (2) compare implementation team member perspectives (program staff, administrators, and participants) on technology-specific and context-specific challenges and enablers. By leveraging a validated implementation framework, this work examines how technological and contextual factors shape program delivery in underserved communities and informs strategies to translate and sustain effective technology-based exercise in traditionally under-resourced settings.

Methods

Study Design and Setting

Overview

This study followed a multimethod research design, whereby the core of the research was qualitative, utilizing in-depth interviews and focus groups, and was supplemented by a minor quantitative component in the form of a brief survey provided to program staff and administrators. We chose both methods not for the purpose of triangulation or formal mixing of qualitative and quantitative paradigms, but rather for complementarity and expansion [35], allowing the survey data to offer a broader, descriptive overview of factors associated with the Lakeshore Online Fitness (LOF) and VR program implementation and the rich qualitative data to provide deep context.

This research was a retrospective evaluation of 2 technology-based interventions deployed between July 2022 and October 2022 at a neighborhood recreation center located in an underserved community (population <5800; ≈50% Black, ≈15% Hispanic or Latino; ≥25% living below the federal poverty line; 1 in 4 uninsured; 1 in 4 with a disability). The synchronous online fitness program was deployed in partnership with a local nonprofit organization that provides physical activity, recreation, and sport programs for people with physical disabilities and chronic health conditions (Lakeshore Foundation). For the VR program component, content- and age-appropriate (ie, limited violence or foul language) active video games were identified by research staff. The LOF and Get Active with VR programs were delivered on-site at the municipal recreation facility building. Program recipients were current members of the recreation center's senior programs who met the following eligibility criteria: (1) aged 18 years or older, (2) self-reported a physical disability or health condition, (3) aged 60 years or older regardless of health condition, and (4) were not enrolled in any other physical activity program. All successfully screened participants provided informed consent prior to enrollment. Sessions for both programs were held twice a week, lasted 60 minutes, and included a warm-up, main activity, and cooldown. Participants completed 6 weeks of each program (LOF or VR) in a randomized crossover design using alternating assignments to either the LOF (Figure 1) or VR program (Figure 2).

Figure 1. Participants engaging in the "Lakeshore Online Fitness" program.

XSL•FO
RenderX

Figure 2. Participants engaging in the "Get Active with Virtual Reality" program.

Lakeshore Online Fitness

Online classes were delivered via Zoom (Zoom Communications, Inc) from Lakeshore to the municipal recreation facility [27,36]. Exercise dose was based on the predetermined length of online classes provided by Lakeshore, and frequency was based on scheduling needs at the recreation center. An instructor from Lakeshore taught all classes, and a community facilitator (staff or volunteer) and research team member oversaw the on-site group activity.

Get Active With Virtual Reality

Oculus Quest 2 headsets were used for the VR program. A community facilitator and research team member oversaw VR classes. For each session, program recipients would choose to play a predetermined game individually or competitively against other participants. Exercise dose and frequency were matched to those of the LOF program.

Study Participants

This research study used convenience sampling techniques to identify factors related to program implementation. Our sample size was determined by the availability of respondents and met minimum guidelines for thematic analysis [37,38]. We gathered input from individuals who (1) participated in the LOF and/or VR programs and (2) played a role in program implementation and/or delivery. Invitations for interviews were extended to program recipients who completed the LOF and VR programs (n=10). Invitations for focus groups were extended to 2 LOF instructors, 2 Lakeshore Foundation administrators, 1 recreation center administrator, 1 recreation center program staff member, and 2 community volunteers (n=9). The recreation center staff member's primary role was to help with the physical setup for LOF, assist participants during classes, and take down program equipment. Both community volunteers were recruited through the University of Alabama at Birmingham Center for the Study of Community Health to act as community liaisons, assist with classes, and identify potential new locations for implementation of the programs. These volunteers also helped with the physical setup for the VR program, assisted participants during classes, and took down program equipment.

Ethical Considerations

The University of Alabama at Birmingham Institutional Review Board approved this study (IRB-300008762), and all program participants provided written informed consent before taking part, which included agreement to participate in the study and permission for their images to appear in photos used in the study. Program participants received US $15 for completing the interviews. Data were stored on encrypted, password-protected servers accessible only to the research team.

Data Collection

Quantitative

Administrators, program staff, and volunteers involved in implementing the LOF and VR programs were invited to complete an online survey assessing accessibility, appropriateness, and feasibility of the two programs in October 2023. The Acceptability of Intervention Measure (AIM), Intervention Appropriateness Measure (IAM), and Feasibility of Intervention Measure (FIM) surveys were completed by all

XSL•FO
RenderX

9 relevant staff members and administrators. The AIM assesses the “perception that a given treatment, service, practice, or innovation is aggregable, palatable, or satisfactory,” the IAM evaluates the “perceived fit of the innovation to address a particular issue or problem,” while the FIM measures the ease at which a “new treatment or innovation can be successfully used or carried out within a given agency or setting.” [39] Each measure consisted of 4 thematic questions, with 5 responses ranging from completely disagree to completely agree on a 5-point Likert scale [40]. Outcomes for the LOF and VR programs were assessed separately. Scores for each scale range from 4 to 20 and were classified as low (<12), medium (12-15), and high (>15) [40].

Qualitative

Program participants took part in one-on-one, semistructured interviews at the recreation center in January 2023 after completing the intervention. A CFIR-informed interview guide, developed in collaboration with the study principal investigator (LAM) and interviewer (WNN), was used to gauge perceptions of the most important factors influencing implementation of the 2 programs for older adults and people with disabilities in underserved communities. Interview guide questions focused on the individual, innovation, and inner setting domains of the CFIR (Multimedia Appendix 1) [29].

A separate CFIR-informed focus group guide was developed to capture staff members’ opinions about barriers, facilitators, and lessons learned from implementation of the programs (Multimedia Appendix 2). Focus group questions covered the innovation, outer setting, inner setting, and implementation process domains of the CFIR [29]. Two semistructured, in-depth focus groups and 1 interview were carried out. The first focus group (n=3) included the recreation center program staff member and two community volunteers who helped set up and deliver the LOF and VR programs. The second focus group (n=3) comprised LOF fitness instructors who delivered the LOF program and 1 Lakeshore Foundation senior staff member. The focus group with recreation center staff and volunteers was conducted in person at the community center where participants completed the programs, while the focus group with LOF program deliverers and senior staff was carried out in person at the Lakeshore Foundation in November 2023. One Lakeshore Foundation senior staff member was unable to attend the focus group session, so they completed a follow-up, one-on-one, semistructured interview online in December 2023. Interviews and focus groups were completed with 79% of program participants and relevant staff and volunteers.

Data Analysis and Rigor

Survey data and participant characteristics are summarized using descriptive statistics. Means and SDs are reported for the continuous variables (ie, age and AIM, IAM, and FIM surveys), while categorical variables (ie, race and sex) are presented as frequencies and percentages.

All qualitative data were coded and organized using NVivo 14 (Lumivero). In the first stage of analysis, the Zoom recordings were transcribed verbatim using a professional transcription service (Landmark Associates, Inc), with time stamps at the beginning of each response and speakers labeled as interviewer or interviewee to ensure clarity and anonymity of the content. The first author became familiar with the data by listening to the audio recordings and repeatedly reading the transcripts. The data were then analyzed using a deductive content analysis approach, following the 3-phase guide outlined by Elo and Kyngäs [41], which includes preparation, organization, and reporting. During the preparation phase, it was decided that sentences, phrases, and themes from the transcripts would serve as the units of analysis to capture both detailed content and broader thematic patterns. Both manifest (explicit, surface-level content) and latent (underlying meanings) content were coded, allowing for a more nuanced understanding of factors influencing implementation. A deductive approach was chosen to align the identified themes with the CFIR domains and constructs, facilitating a theory-driven analysis of implementation factors.

In phase 2 of the qualitative analysis, a structured categorization matrix based on the 5 CFIR domains and 39 constructs was developed to guide organization and interpretation of themes. The first author then reviewed the data for content and coding based on correspondence with the predetermined domains and constructs. In phase 3, results were synthesized by identifying the most frequently reported themes. To ensure accurate alignment with CFIR constructs, the first and second authors reviewed and refined the synthesized results, confirming that each theme was appropriately categorized and all relevant data were considered. Integration occurred during the interpretation phase, where the survey results were used to provide supplementary context and illustration for the core qualitative themes. In this analysis, the main findings remain rooted in the qualitative evidence and are organized by facilitators and barriers, specific CFIR domains, and subthemes that correspond to the CFIR constructs [29].

To ensure quality and rigor, the evaluative markers coherence, transparency, and substantive contribution were selected from Smith and Caddick’s [42] ongoing list and can be found in Table 1. At the time of data collection, the first author (WNN) was a predoctoral student in health behavior with formal training in qualitative research methods. With nearly a decade of experience working with individuals with chronic health conditions and no prior involvement in the study’s program delivery portion, she brought an outsider perspective. The second author (LAM) has been involved in several qualitative studies related to physical activity and disability and served as a “critical friend” throughout the process. This role involved reviewing the first author’s interpretations of the data and providing feedback on the accuracy and consistency of the coding and categorization process. Rather than reach an agreement, this collaborative, reflective approach was intended to generate nuanced and multifaceted interpretations of meaning.

XSL•FO
RenderX

Table . Evaluative markers chosen for quality and rigor.

Marker	Definition	Application
Coherence	The way that different parts of the interpretation create a complete and meaningful picture.	Demonstrated through the alignment of the research questions, methodology, and validated theoretical framework to guide the evaluation.
Transparency	Honesty about the research process.	Maintenance of an audit trail through detailed notes and critical friend throughout analytical process.
Substantive contribution	The extent to which a study makes a meaningful impact on our understanding of what is being researched theoretically, conceptually, practically, or methodologically.	Offers a theoretical understanding of factors influencing the implementation of technology-driven community programs.

Results

Implementation Surveys

All administrators, program staff, and volunteers involved in implementing the LOF and VR programs (n=9, 100%) completed the surveys. Results from the AIM, IAM, and FIM surveys are provided in Table 2. Overall, both programs attained high ratings across all 3 domains. However, one-third of respondents (n=3, 33.3%) rated the VR program's feasibility as medium (FIM score <15) rather than high.

Table . Acceptability, appropriateness, and feasibility measure results.

	LOF[a] program, mean (SD)	VR[b] program, mean (SD)
AIM[c]	16.89 (1.62)	17.56 (2.74)
IAM[d]	16.33 (1.58)	16.56 (2.70)
FIM[e]	16.78 (1.64)	16.22 (2.86)

[a] LOF: Lakeshore Online Fitness

[b] VR: virtual reality

[c] AIM: Acceptability of Intervention Measure

[d] IAM: Intervention Appropriateness Measure

[e] FIM: Feasibility of Intervention Measure

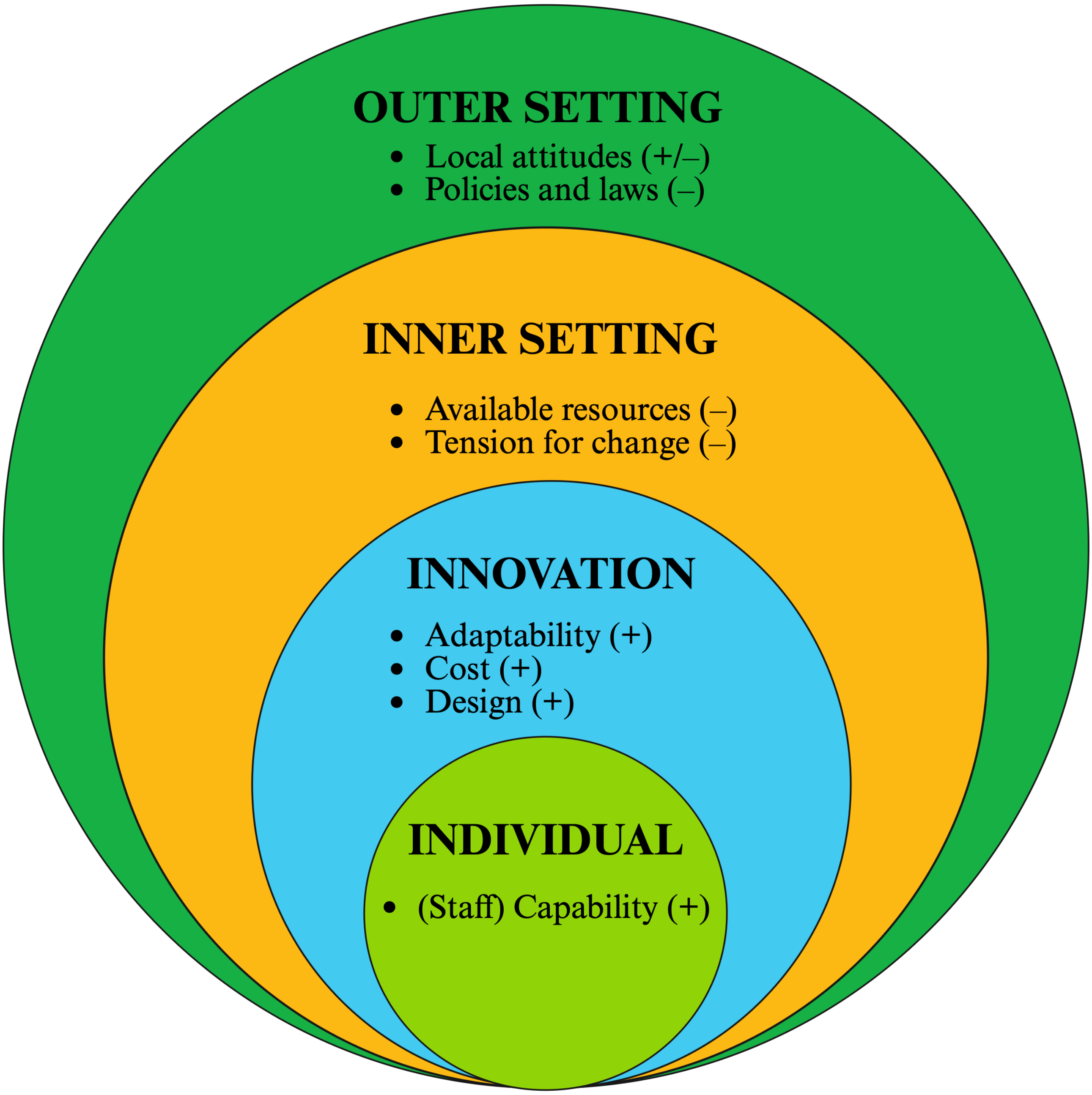
Interview and Focus Group Results

A total of 15 individuals took part in the interviews (mean 41.3 min, SD 16.1 min; range 24 - 63 min) or focus groups (mean 49.3 min, SD 7.1 min), including 8 of 10 program participants, 1 recreation center staff member, 2 community volunteers, 2 fitness instructors, and 2 administrators [1]. Interview participants (n=8) were primarily Black or African American (n=7, 87.5%) and female (n=5, 62.5%) participants with a mean age of 65.0 (SD 14.0; range 44 - 86) years. The majority of interview participants had hypertension (n=6, 75%) and arthritis (n=4, 50%), with 37.5% having both concurrently. Focus group participants were primarily White (5 /7, 71.4%) and male (4/7, 57.1%) individuals; age was not collected. Program participants completed 78.3% and 85.4% of LOF and VR sessions, respectively.

Qualitative data analysis identified 4 overarching domains and 8 constructs that represented key facilitators and barriers to implementing the LOF and VR programs. The results mapped to CFIR domains and constructs are depicted in Figure 3.

XSL•FO RenderX

Figure 3. Implementation barriers and facilitators coded into Consolidated Framework for Implementation Research (CFIR) domains and constructs. "–" indicates barriers, "+" indicates facilitators. Created in BioRender [43].



Key Facilitators to LOF or VR Implementation

Six subthemes based on 5 CFIR constructs were identified as facilitators to program implementation. These constructs belonged to 3 CFIR domains ("innovation," "outer setting," and "individuals") and are described in more detail under the "Innovation," "Outer Setting," and "Individuals" subsections below.

Innovation

Overview

The innovation domain comprises features of an intervention that might influence implementation. Key facilitators related to program characteristics centered on the adaptability, cost, design quality, and packaging constructs. Implementation team members praised both the LOF and VR programs for their flexibility, allowing for modifications to accommodate diverse physical abilities. The low cost of equipment and free access at the community center made both programs accessible and affordable, while the socially interactive and engaging design enhanced motivation and enjoyment for program recipients.

Flexible LOF Exercises (Adaptability)

All program recipients and deliverers noted the flexibility of the LOF program for participants with diverse physical abilities, such as low mobility, dexterity limitations, low vision, or stroke-related impairments. Participants with dexterity limitations mentioned the ease of modifying the LOF program components to facilitate engagement, "Once I could pick up the weights, when the resistance ball isn't—none of it bothered me, it was just putting it in my hand" (program recipient 04). Instructors adjusted exercises to participants' needs without full

XSL•FO
RenderX

information on their abilities but noted that experience was essential for making these on-the-fly adaptations.

VR Program Technology Considerations (Adaptability)

Recreation center staff and volunteers emphasized the importance of the VR program's adaptability to diverse settings, such as standing versus seated sessions to support group activity and ensure participant safety:

> *I think if they were standing up, they would likely free roam. They're in their virtual world, and they can't see where they're going or what they're doing. When you start moving, you just movin'. They might be bumping into each other. I think sitting was the perfect thing to stay in the circle and not bump each other. I think that was good.* [Community volunteer 01]

Although some participants initially struggled with dexterity, this issue was resolved quickly by the research staff, who modified the VR equipment for one-handed use.

Program Access and Affordability (Cost)

Program deliverers felt that the LOF program's low-cost exercise equipment, such as balls, bands, and lightweight dumbbells, would make the program relatively simple to replicate in other locations:

> *In theory, the way this works is we're bringing a lot of the expense to them. We're bringing the TV. We're trying to cut away as much of the red tape as we can. What we really need is for them to say, yes, here's a room. Here's somebody to turn on the TV. Here's somebody who can hover in there to help with stuff, but you guys do the hard part of buying the TV and having the expertise to teach the class. It's just [a] partnership.* [Administrator 01]

Program recipients appreciated that both programs were free and available at the community center, citing affordability and ease of access as key facilitators of engagement. They acknowledged that many older adults and people with chronic health conditions in their community lacked transportation to attend paid physical activity programs, but that, "…they would be happy to do exercise like that, long as they don't have to leave the building'' (program recipient 09).

Socially Interactive and Engaging Design (Design)

This subtheme pertains to participant perceptions of the quality of the materials associated with the LOF and VR programs. Program recipients valued the socially supported, group-based nature of the LOF program, which enhanced motivation. "We did those stretch band things. We did the balls and the dumbbells…I liked it. I liked being on the equipment" (program recipient 02). They found the competitive and interactive elements of the VR program particularly motivating, as these elements allowed them to compete against others or themselves, "We had a competition going together shooting games and stuff like that, bowling, and stuff like that" (program recipient 03). These program components made physical activity more enjoyable and immersive:

> *I like the [VR] headset because you get the feeling of really being in there. That's the part I really like. The sceneries are so beautiful. You really did, you felt like you had actually stepped into a space age, or a tennis court or baseball diamond.* [Program recipient 08]

Outer Setting: Social and Community Engagement (Local Attitudes)

This subtheme reflects the degree to which sociocultural values (eg, shared responsibility in helping recipients) and beliefs (eg, convictions about the worthiness of recipients) encourage the outer setting (eg, the city the program was implemented in) to support implementation and/or delivery of the intervention." [30,31] Those involved in program implementation and delivery felt that, "Having a community champion like the assistant mayor" (administrator 01) was crucial to the program's success, as their advocacy facilitated connections with community leaders and strengthened local engagement. Positive reactions from program recipients further affirmed the programs' value and energized program deliverers. These groups noted that this enthusiasm made the program more enjoyable for the staff and was a key indicator of its success:

> *Just seeing them enjoying made me feel good, because they [were] enjoying something. It's not like they [were] forced to do it or "Okay, we're in this program here at the rec, so this is part of it. We have to do it." "No, this is a program that I enjoy. I'm glad when it start[s]. I hate when it end[s]. We just want to have fun." They all got along. That was a good thing.* [Community volunteer 01]

> *...it was fun to coach, because you could see them move and you could see 'em doing things and it seemed like they really enjoyed it and then when it ended, they would clap and say thank you, and all that and give you that reassurance.* [Fitness instructor 01]

Individuals: Staff Training and Role Clarity (Capability)

Effective implementation required clear role distribution and adequate staffing to assist participants, manage technology, and troubleshoot issues onsite. Program deliverers felt that staff with strong interpersonal and problem-solving skills and at least some knowledge of exercise were beneficial to LOF and VR program success:

> *So, ideally, it would be someone with fitness experience…[but] minimally, it has to be, in my opinion, someone who is consistent, dependable, and someone that we could train.* [Administrator 02]

> *It was a very easy setup and once the class was going, you were hands off. If someone didn't really seem like they knew what they were doing, you would try and mimic what the instructor was doing. I don't really feel like there was too much training needed for that portion of it.* [Recreation center staff 01]

Key Barriers to LOF or VR Implementation

There were 4 barriers identified from the interviews and focus groups, with subthemes divided evenly between the inner and outer setting domains.

XSL•FO
RenderX

Inner Setting

Overview

Barriers to LOF and VR program implementation within the inner setting domain (eg, the organization in which the intervention is being implemented) [30,31] centered on the available resources and tension for change constructs. The LOF and VR programs at the recreation center faced technical and spatial challenges, and program recipients expressed dissatisfaction with the limited accessible physical activity options in their area, highlighting the need for sustainable community resources to address underlying infrastructure deficits.

Technology and Space Challenges (Available Resources)

This subtheme reflects “the degree to which resources are available to implement and deliver the intervention” [30,31] Addressing technical and spatial barriers, including technology setup and space constraints, was essential to successful implementation of the LOF and VR programs at the recreation center. Stable Wi-Fi, reliable equipment, and adequate space were critical to overcoming these challenges and ensuring successful program delivery. Program users reported minimal technical issues:

> *The only [LOF program] technical issue I think we had—well, then it wasn't an issue—was that we could not hear what she was sayin', but we could see what she was doin', so we still continued because we could see her doin' the exercises.* [Program recipient 05]
>
> *One day the [VR program] computer wasn’t acting right for some. … [But] we never had to miss a day because of the technical difficulties.* [Program recipient 08]

However, program providers felt that LOF program delivery was complicated by spatial constraints at the recreation center:

> *I think the biggest complaint with the [LOF] class was we just didn’t have enough space… Sometimes we had 10 people, and so we had all the chairs. It was like they couldn’t fully put out their arms, or [they’d] be running into each other and stuff like that. … I feel like just the size of the room was the biggest issue.* [Recreation center staff 01]
>
> *You could see 'em moving and everything, but you weren't able to read people’s eyes, you saw gross movements, but you didn't see the finer movements.* [Fitness instructor 01]

LOF instructors noted that participant monitoring could be enhanced with larger screens and/or additional cameras, “…so you can see people a little bit closer, ’cause we were at a distance, and you really couldn’t see their faces” (fitness instructor 01). Instructors also emphasized the importance of standardized and simplified technological setups on their end to minimize the burden of adjusting to new rooms, equipment, and lighting conditions.

Need for Infrastructure and Resource Support (Tension for Change)

This subtheme aligns with the CFIR construct “tension for change,” where implementation team member recognition of significant gaps creates momentum for implementation efforts [30,31]. Program recipients expressed a strong desire for more accessible physical activity options in their area, noting that the absence of such resources created what they perceived as an urgent community need: “There aren’t no other [free physical activity opportunities] here...People, they go out down there to the community store and buy ’em a can of beer or a pack of cigarettes” (program recipient 04). Participants highlighted both the immediate value of the LOF and VR programs, “Well, me and my other two siblings, we got high blood pressure, so we need somethin’ to motivate us” (program recipient 02), and emphasized the need for additional community resources to support these programs long-term:

> *...Cause we don’t have nothin’ there. I needed it. ...Those people need it. They really need somethin’ like that…* [Program recipient 09]

The emphasis on sustainability concerns further reflected program recipients’ insight that temporary interventions, while beneficial, would not address the fundamental infrastructure deficits underlying their community’s physical activity challenges.

Outer Setting

Key barriers to LOF and VR program implementation within the outer setting domain centered on the local attitudes and policies and laws constructs. Resistance to new technology and sustainability challenges were identified as barriers to implementing and sustaining the LOF and VR programs.

Resistance to New Technology and Approaches (Local Attitudes)

Program recipients, administrators, recreation center staff, and volunteers all mentioned resistance to adopting new technology and program approaches. Some program recipients expressed concerns that others in the community might resist the VR program due to discomfort with wearing headsets, influenced by cultural and personal preferences: “...I don’t think they would like it ’cause just puttin’ stuff on their face…” (program recipient 09). Similarly, recreation center administration initially doubted participants’ willingness to engage with the LOF and VR programs, reflecting skepticism about the programs’ acceptance and feasibility:

> *When we said, well, how about this group take the class, [senior staff] at the rec center were like, y’all are out of your minds. They are never going to participate in this* [Administrator 01]

Sustainability Challenges (Policies and Laws)

This subtheme reflects the “degree to which legislation, regulations, professional group guidelines and recommendations, or accreditation standards support implementation and/or delivery of an intervention” [30,31]. One administrator felt that relying solely on grant funding made long-term sustainability of the LOF program uncertain, especially without a financial

XSL•FO
RenderX

model for continued implementation, “So how do we deliver it when it’s not through a grant, and how do we try to align as closely to the way that we would onboard members now, but on the online, which is different” (administrator 02). This administrator also noted liability and membership issues, including the need for waivers and consistent liability protocols, as barriers to continuing the LOF program at the community center, “...Because we can’t deliver any product without some kinda safeguard set up...and I don’t believe we ever figured that out” (administrator 02). Similarly, program deliverers expressed concerns about the VR program’s long-term feasibility due to equipment maintenance, cultural unfamiliarity, and staff burden—factors that align with the relatively lower feasibility scores reported in the FIM survey.

Discussion

Key Findings and Contributions

This study applied the CFIR to explore barriers and facilitators influencing implementation of two technology-driven group exercise programs—an online platform (LOF) and VR gaming—delivered to predominantly African American older adults living in an underserved community. Through semistructured interviews, focus groups, and implementation surveys, we identified factors across 4 CFIR domains (innovation, outer setting, inner setting, and individuals) that collectively influenced the programs’ uptake, delivery, and sustainability. Our findings highlight that successful implementation was shaped by a combination of program adaptability, affordability, social engagement, and community support, while sustainability and resource constraints emerged as persistent barriers.

A key contribution of this study lies in demonstrating how technology-enabled interventions can create structured exercise opportunities in community centers that previously lacked formal fitness programming. Unlike settings where technology supplements existing exercise classes, the recreation center in this study had no structured physical activity programs prior to implementation. Technology served as the catalyst for introducing accessible, engaging, and adaptable exercise options, addressing a critical gap in underserved communities where traditional programs are often unavailable due to resource limitations. This approach aligns with growing evidence that technology-based interventions can overcome environmental and logistical barriers by delivering scalable, customizable programs that promote physical activity and social engagement among older adults and individuals with disabilities [18-20]. By leveraging online platforms and VR gaming, the programs provided not only exercise opportunities but also novel experiences that enhanced enjoyment and motivation—factors shown to improve adherence in populations facing mobility and transportation challenges [15,23,26].

Facilitators of Implementation

Adaptability emerged as a key facilitator across both programs. The ability to adjust exercises and technology to accommodate diverse physical abilities allowed for broad participation, even among individuals with mobility limitations and dexterity challenges. These findings align with previous research highlighting the importance of customizable interventions that support engagement among populations with heterogeneous physical limitations [44,45]. Affordability and ease of access were also critical enablers, extending beyond equipment costs to include transportation and membership fees, which are common deterrents among people with disabilities [1]. By eliminating these costs, the programs created equitable access to structured exercise opportunities, reinforcing prior evidence that cost-sensitive solutions are vital for program sustainability in resource-constrained settings [46].

The socially interactive and gamified design of both programs further enhanced motivation and enjoyment. LOF’s group-based format fostered camaraderie, while the VR program’s competitive elements introduced novelty and fun, supporting the value of incorporating social and gamified components into technology-based group physical activity programs for older adults and people with disabilities to sustain engagement [47]. Future implementation of technology-driven exercise interventions should leverage these design principles to strengthen both physical and psychosocial outcomes in this population.

Finally, community support amplified program success. Advocacy from community champions (eg, city leaders), enthusiastic endorsement from community volunteers and recreation center administrators, and visible program recipient enthusiasm strengthened buy-in among staff and volunteers, creating a sense of shared ownership. This is consistent with previous research emphasizing the role of community advocacy in successful implementation of health promotion programs in underserved areas [45] and suggests that early identification of community champions should be integral to implementation planning.

Barriers to Implementation

Despite these strengths, several barriers impeded implementation. Technology and space limitations were the most immediate obstacles. Unstable Wi-Fi, suboptimal camera setups, and inadequate physical space complicated LOF program delivery and monitoring, while the VR sessions required careful configuration to ensure safety, highlighting the need for robust infrastructure and standardized protocols [45,46,48]. Resistance to new technology reflected cultural and personal discomfort, particularly with VR headsets, and skepticism among staff about program feasibility. While program recipients spoke highly of the VR program, initial hesitancy underscores the importance of tailored program orientation and gradual exposure to technology to build confidence and acceptance [45,49,50].

Additionally, sustainability challenges stemming from reliance on grant funding and unresolved liability protocols raised concerns about program continuity. Administrations cited liability and membership protocols as unresolved issues for the LOF program, while survey responses from program deliverers reflected reservations about the VR program’s feasibility. This underscores the need for alternative funding strategies (eg, partnerships with local health departments and sliding-scale fees), infrastructure development, and policy alignment to support long-term program continuation. Without such planning,

XSL•FO
RenderX

even well-received programs risk discontinuation once initial funding ends.

Implications

Our findings suggest that technology-driven exercise programs can reduce health disparities in underserved communities when they are adaptable, affordable, socially engaging, and supported by community champions. However, implementation efforts must anticipate and address resource constraints, technology resistance, and sustainability challenges. Multilevel implementation strategies, including investment in digital infrastructure, community advocacy, targeted orientation to reduce technology hesitancy, and sustainable financing models, are essential to scale and sustain such interventions. Future research should explore strategies for scaling these programs, such as partnerships with local organizations, integration into existing community services, or development of cost-sharing models. Programs should engage community champions from inception, build standardized technical protocols, and test these programs across diverse geographic and demographic contexts.

Strengths and Limitations

Study strengths include its use of a validated implementation framework (CFIR), integration of multiple implementation team member perspectives, and use of both qualitative and quantitative methods to identify barriers and facilitators to program implementation. However, the current study has limitations that warrant consideration. Our sample, though achieving 79% participation among relevant implementation team members, was relatively small and drawn from a single, predominantly African American community, which may limit generalizability. Additionally, racial distribution differed between data collection formats: interview participants were predominantly African American, while focus group participants were primarily White. This imbalance may have influenced the range of perspectives captured. Implementation team member perspectives were collected several months after program delivery, which may not have captured immediate perspectives of program sustainability or technology acceptance. Nonetheless, the rich qualitative insights underscore the value of incorporating ongoing implementation team member evaluation and feedback into program design.

Conclusions

Technology-enhanced group exercise interventions hold considerable promise for reducing physical inactivity among underserved older adults and people with disabilities. The results of our CFIR-guided analysis suggest that, while adaptability, low cost, and social engagement facilitated uptake, the LOF and VR programs presented technical, infrastructural, and policy-related barriers. The findings of this study contribute to the growing literature on technology-enabled health interventions for aging populations and people with disabilities and offer insights into implementation strategies in underserved, resource-limited settings.

Acknowledgments

The authors wish to acknowledge Eric Evans for his initiative on the virtual reality portion of this project and to thank Ashley Wright for her expertise and assistance with program delivery. The authors also extend our appreciation to the Lakeshore Foundation fitness instructors and senior leadership, community volunteers, and staff at the Parks and Recreation Department. Without their time, dedication, commitment, and resources, this project would not have been possible. The authors thank Emily Delzell for her editorial review and assistance in the preparation of this manuscript.

Funding

This work was supported by the University of Alabama at Birmingham Center for the Study of Community Health University-Wide Interdisciplinary Research Center and the University of Alabama Center for Clinical and Translational Science.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Lakeshore Online Fitness or Get Active with Virtual Reality intervention participant interview guide.
[DOCX File, 34 KB - rehab_v13i1e79598_app1.docx]

Multimedia Appendix 2
Lakeshore Online Fitness or Get Active with Virtual Reality intervention focus group guide.
[DOCX File, 25 KB - rehab_v13i1e79598_app2.docx]

References

1. Martin Ginis KA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. Health Psychol Rev 2016 Dec;10(4):478-494. [doi: 10.1080/17437199.2016.1198240] [Medline: 27265062]

XSL•FO
RenderX

2. Okoro CA, Hollis ND, Cyrus AC, Griffin-Blake S. Prevalence of disabilities and health care access by disability status and type among adults—United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Aug 17;67(32):882-887. [doi: 10.15585/mmwr.mm6732a3] [Medline: 30114005]
3. Carroll DD, Courtney-Long EA, Stevens AC, et al. Vital signs: disability and physical activity—United States, 2009-2012. MMWR Morb Mortal Wkly Rep 2014 May 9;63(18):407-413. [Medline: 24807240]
4. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016 Jul 1;102(13):1009-1016. [doi: 10.1136/heartjnl-2015-308790] [Medline: 27091846]
5. Emerson E, Fortune N, Llewellyn G, Stancliffe R. Loneliness, social support, social isolation and wellbeing among working age adults with and without disability: cross-sectional study. Disabil Health J 2021 Jan;14(1):100965. [doi: 10.1016/j.dhjo.2020.100965] [Medline: 32843311]
6. Rimmer JH, Padalabalanarayanan S, Malone LA, Mehta T. Fitness facilities still lack accessibility for people with disabilities. Disabil Health J 2017 Apr;10(2):214-221. [doi: 10.1016/j.dhjo.2016.12.011] [Medline: 28143707]
7. Malone LA, Barfield JP, Brasher JD. Perceived benefits and barriers to exercise among persons with physical disabilities or chronic health conditions within action or maintenance stages of exercise. Disabil Health J 2012 Oct;5(4):254-260. [doi: 10.1016/j.dhjo.2012.05.004] [Medline: 23021736]
8. Kayes NM, McPherson KM, Taylor D, Schlüter PJ, Kolt GS. Facilitators and barriers to engagement in physical activity for people with multiple sclerosis: a qualitative investigation. Disabil Rehabil 2011;33(8):625-642. [doi: 10.3109/09638288.2010.505992] [Medline: 20695816]
9. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]
10. Vasudevan V, Vang PD, Fernandez-Baca D. An exploration of healthy eating and physical activity habits of Hmong high-school students by disability status: a pilot study. Disabil Health J 2019 Oct;12(4):694-698. [doi: 10.1016/j.dhjo.2019.06.012] [Medline: 31257033]
11. Cockerham WC, Hamby BW, Oates GR. The social determinants of chronic disease. Am J Prev Med 2017 Jan;52(1S1):S5-S12. [doi: 10.1016/j.amepre.2016.09.010] [Medline: 27989293]
12. Mahmoudi E, Meade MA. Disparities in access to health care among adults with physical disabilities: analysis of a representative national sample for a ten-year period. Disabil Health J 2015 Apr;8(2):182-190. [doi: 10.1016/j.dhjo.2014.08.007] [Medline: 25263459]
13. Evans EJ, Naugle KE, Owen T, Naugle KM. Active gaming: it is not just for young people. J Aging Phys Act 2020;28(5):731-739. [doi: 10.1123/japa.2019-0303] [Medline: 32422600]
14. Rowland JL, Malone LA, Fidopiastis CM, Padalabalanarayanan S, Thirumalai M, Rimmer JH. Perspectives on active video gaming as a new frontier in accessible physical activity for youth with physical disabilities. Phys Ther 2016 Apr;96(4):521-532. [doi: 10.2522/ptj.20140258] [Medline: 26316530]
15. Malone LA, Thirumalai M, Padalabalanarayanan S, Neal WN, Bowman S, Mehta T. Energy expenditure and enjoyment during active video gaming using an adapted Wii Fit balance board in adults with physical disabilities: observational study. JMIR Serious Games 2019 Feb 1;7(1):e11326. [doi: 10.2196/11326] [Medline: 30707098]
16. Lai B, Davis D, Narasaki-Jara M, et al. Feasibility of a commercially available virtual reality system to achieve exercise guidelines in youth with spina bifida: mixed methods case study. JMIR Serious Games 2020 Sep 3;8(3):e20667. [doi: 10.2196/20667] [Medline: 32880577]
17. Lindsay-Smith G, Eime R, O’Sullivan G, Harvey J, van Uffelen JGZ. A mixed-methods case study exploring the impact of participation in community activity groups for older adults on physical activity, health and wellbeing. BMC Geriatr 2019 Sep 2;19(1):243. [doi: 10.1186/s12877-019-1245-5] [Medline: 31477054]
18. Wang G, Xiang R, Yang X, Tan L. Digital technology empowers exercise health management in older adults: a systematic review and meta-analysis of the effects of mHealth-based interventions on physical activity and body composition in older adults. Front Public Health 2025 Oct 1;13:1661028. [doi: 10.3389/fpubh.2025.1661028] [Medline: 41103467]
19. Alley SJ, Waters KM, Parker F, et al. The effectiveness of digital physical activity interventions in older adults: a systematic umbrella review and meta-meta-analysis. Int J Behav Nutr Phys Act 2024 Dec 18;21(1):144. [doi: 10.1186/s12966-024-01694-4] [Medline: 39696583]
20. Dubbeldam R, Stemplewski R, Pavlova I, et al. Technology-assisted physical activity interventions for older people in their home-based environment: scoping review. JMIR Aging 2025 Sep 15;8:e65746. [doi: 10.2196/65746] [Medline: 40952779]
21. Dawson R, Oliveira JS, Kwok WS, et al. Exercise interventions delivered through telehealth to improve physical functioning for older adults with frailty, cognitive, or mobility disability: a systematic review and meta-analysis. Telemed J E Health 2024 Apr;30(4):940-950. [doi: 10.1089/tmj.2023.0177] [Medline: 37975811]
22. Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. J Geriatr Phys Ther 2018;41(1):49-61. [doi: 10.1519/JPT.0000000000000095] [Medline: 27362526]

XSL•FO
RenderX

23. Chen PJ, Hsu HF, Chen KM, Belcastro F. VR exergame interventions among older adults living in long-term care facilities: a systematic review with meta-analysis. Ann Phys Rehabil Med 2023 Apr;66(3):101702. [doi: 10.1016/j.rehab.2022.101702] [Medline: 36028201]
24. Donath L, Rössler R, Faude O. Effects of virtual reality training (exergaming) compared to alternative exercise training and passive control on standing balance and functional mobility in healthy community-dwelling seniors: a meta-analytical review. Sports Med 2016 Sep;46(9):1293-1309. [doi: 10.1007/s40279-016-0485-1] [Medline: 26886474]
25. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inform 2016 Jan;85(1):1-16. [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
26. Malone LA, Davlyatov GK, Padalabalanarayanan S, Thirumalai M. Active video gaming using an adapted gaming mat in youth and adults with physical disabilities: observational study. JMIR Serious Games 2021 Aug 26;9(3):e30672. [doi: 10.2196/30672] [Medline: 34435962]
27. Malone LA, Mehta T, Mendonca CJ, Mohanraj S, Thirumalai M. A prospective non-randomized feasibility study of an online membership-based fitness program for promoting physical activity in people with mobility impairments. Pilot Feasibility Stud 2024 Aug 2;10(1):104. [doi: 10.1186/s40814-024-01528-x] [Medline: 39095876]
28. Bauer MS, Kirchner J. Implementation science: What is it and why should I care? Psychiatry Res 2020 Jan;283:112376. [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]
29. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
30. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4(1). [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
31. Cardona MI, Monsees J, Schmachtenberg T, Grünewald A, Thyrian JR. Implementing a physical activity project for people with dementia in Germany-Identification of barriers and facilitator using consolidated framework for implementation research (CFIR): a qualitative study. PLoS One 2023;18(8):e0289737. [doi: 10.1371/journal.pone.0289737] [Medline: 37556503]
32. Bhardwaj A, FitzGerald C, Graham M, MacFarlane A, Kennedy N, Toomey CM. Barriers and facilitators to implementation of an exercise and education programme for osteoarthritis: a qualitative study using the consolidated framework for implementation research. Rheumatol Int 2024 Jun;44(6):1035-1050. [doi: 10.1007/s00296-024-05590-9] [Medline: 38649534]
33. Boxall C, Dennison L, Miller S, et al. Implementing a digital physical activity intervention for older adults: qualitative study. JMIR Aging 2025 Aug 21;8:e64953. [doi: 10.2196/64953] [Medline: 40840434]
34. Cooper J, Murphy J, Woods C, et al. Barriers and facilitators to implementing community-based physical activity interventions: a qualitative systematic review. Int J Behav Nutr Phys Act 2021 Sep 7;18(1):118. [doi: 10.1186/s12966-021-01177-w] [Medline: 34493306]
35. Anguera MT, Blanco-Villaseñor A, Losada JL, Sánchez-Algarra P, Onwuegbuzie AJ. Revisiting the difference between mixed methods and multimethods: is it all in the name? Qual Quant 2018 Nov;52(6):2757-2770. [doi: 10.1007/s11135-018-0700-2]
36. Mohanraj S, Malone LA, Mendonca CJ, Thirumalai M. Development and formative evaluation of a virtual exercise platform for a community fitness center serving individuals with physical disabilities: mixed methods study. JMIR Form Res 2023 Dec 15;7:e49685. [doi: 10.2196/49685] [Medline: 38100173]
37. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2021;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]
38. Braun V, Clarke V, Weate P. Using thematic analysis in sport and exercise research. In: Routledge Handbook of Qualitative Research in Sport and Exercise: Routledge; 2016:191-205.
39. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
40. Saxena S, Bhan A, Fleming W, Singh A, Tugnawat D. Acceptability, feasibility and appropriateness of care of patients with depression at Health and Wellness Centre in a district of India. Prev Med Res Rev 2024;1(1):41-46. [doi: 10.4103/PMRR.PMRR_77_23]
41. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
42. Smith B, Caddick N. Qualitative methods in sport: a concise overview for guiding social scientific sport research. Asia Pac J Sport Soc Sci 2012;1(1):60-73. [doi: 10.1080/21640599.2012.701373]
43. Neal W. Figure 3. Implementation barriers and facilitators coded into consolidated framework for implementation research (CFIR) domains and constructs. "–" indicates barriers, "+" indicates facilitators. bioRender. 2026. URL: https://BioRender.com/0rybw39 [accessed 2026-03-02]
44. Rimmer JH, Riley B, Wang E, Rauworth A, Jurkowski J. Physical activity participation among persons with disabilities: barriers and facilitators. Am J Prev Med 2004 Jun;26(5):419-425. [doi: 10.1016/j.amepre.2004.02.002] [Medline: 15165658]

XSL•FO
RenderX

45. Hung L, Mann J, Wallsworth C, et al. Facilitators and barriers to using virtual reality and its impact on social engagement in aged care settings: a scoping review. Gerontol Geriatr Med 2023;9:23337214231166355. [doi: 10.1177/23337214231166355] [Medline: 37020921]
46. Schroeder K, Deatrick JA, Klusaritz H, et al. Using a community workgroup approach to increase access to physical activity in an underresourced urban community. Health Promot Pract 2020 Jan;21(1):5-11. [doi: 10.1177/1524839919867649] [Medline: 31423845]
47. Zimmer C, McDonough MH, Hewson J, et al. Experiences with social participation in group physical activity programs for older adults. J Sport Exerc Psychol 2021 Jun 24;43(4):335-344. [doi: 10.1123/jsep.2020-0335] [Medline: 34167084]
48. Dal Bello-Haas VPM, O'Connell ME, Morgan DG, Crossley M. Lessons learned: feasibility and acceptability of a telehealth-delivered exercise intervention for rural-dwelling individuals with dementia and their caregivers. Rural Remote Health 2014;14(3):2715. [Medline: 25081991]
49. Khoshrounejad F, Hamednia M, Mehrjerd A, et al. Telehealth-based services during the COVID-19 pandemic: a systematic review of features and challenges. Front Public Health 2021;9:711762. [doi: 10.3389/fpubh.2021.711762] [Medline: 34350154]
50. Mitzner TL, Savla J, Boot WR, et al. Technology adoption by older adults: findings from the PRISM trial. Gerontologist 2019 Jan 9;59(1):34-44. [doi: 10.1093/geront/gny113] [Medline: 30265294]

Abbreviations

AIM: Acceptability of Intervention Measure
CFIR: Consolidated Framework for Implementation Research
FIM: Feasibility of Intervention Measure
IAM: Intervention Appropriateness Measure
LOF: Lakeshore Online Fitness
VR: virtual reality

Edited by S Munce; submitted 24.Jun.2025; peer-reviewed by C Idiong, E Dove, J Ambros-Antemate, O Abasilim; accepted 26.Jan.2026; published 24.Mar.2026.

Please cite as:
Neal WN, Malone LA
Technology-Driven Group Exercise Program Implementation in an Underserved Community: Multimethod Retrospective Evaluation Study
JMIR Rehabil Assist Technol 2026;13:e79598
URL: https://rehab.jmir.org/2026/1/e79598
doi:10.2196/79598

© Whitney N Neal, Laurie A Malone. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO RenderX

Appropriateness and Impact of a Vocal Cord Vibration Switch for Children with Complex Communication Needs: Case Series

Leslie Mumford[1], MHSc; Denise Guerriere[2], PhD; Tom Chau[1,3], PhD

[1]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, 150 Kilgour Road, Bloorview Research Institute, Toronto, ON, Canada
[2]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
[3]Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Tom Chau, PhD
Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, 150 Kilgour Road, Bloorview Research Institute, Toronto, ON, Canada

Abstract

Background: Communication is an essential component of participation. Communication impairment restricts full participation for children who have unintelligible speech. A vocal cord vibration switch offers an avenue for meaningful interaction to children who cannot rely on speech or voluntary limb movement but have some control of their vocal cords. Previous evaluations of the vocal cord vibration switch have been conducted primarily with adults and adolescents. However, implementation of a vocal cord vibration switch with younger, school-aged children in their natural environmental contexts can potentially foster the development of early communication skills.

Objective: This case series evaluated the appropriateness and impact of a vocal cord vibration switch, the “Hummingbird” (Holland Bloorview) with school-aged children who have complex communication needs and their mothers and teachers, using an individualized, collaborative, and iterative assistive device implementation protocol.

Methods: The Hummingbird was evaluated with 3 school-aged children, across educational and health-related contexts, over a 2-year period. Baseline, midterm, and final assessments took place at home or school in the first year with a follow-up assessment in the second year. In addition to field observations and device performance assessments by the research team, feedback from mothers and teachers was collected via questionnaires (Pictorial Children’s Effort Rating Table, Quebec User Evaluation of Satisfaction with Assistive Technology, and the Family Impact of Assistive Technology Scale) to ascertain the Hummingbird’s appropriateness and its impact.

Results: Appropriateness data indicated the suitability of the Hummingbird across settings. Compared to the child’s prestudy devices, mother and teacher participants reported that the physical effort required by all 3 children to use the Hummingbird was lower (scores on the Pictorial Effort Rating Table decreased from 8 and 11 to scores of 4 - 6). The switch efficacy assessment of the Hummingbird indicated moderate-to-high specificity and high sensitivity at midterm and high sensitivity (0.91 - 0.94) and specificity (0.92 - 1) at final assessments. Total satisfaction scores increased from baseline (prestudy device) to the 2-year assessment for all 3 children. While data on the impact on family and communication were incomplete for 1 participant, generally favorable effects were reported. The field notes underscored the value of an individualized protocol, where the implementation and evaluation phases were adapted to accommodate the health-related characteristics (eg, seizure disorder and sleep deprivation), evolving school contextual factors (new school and teacher), and unique family environments (involving the child participant’s toddler-sibling in Hummingbird sessions).

Conclusions: Overall, the Hummingbird was appropriate across home and school settings for our case study participants, all of whom had complex communication needs. The device was well-received by children and their mothers or teachers, providing an effective, setting-agnostic option for communication support. Modifications to both the device and its implementation process were required to address unanticipated health, family, and school challenges.

(JMIR Rehabil Assist Technol 2026;13:e75626) doi:10.2196/75626

KEYWORDS

complex communication needs; augmentative and alternative communication; vocal cord vibration switch; children with disabilities; access technology

XSL•FO
RenderX

Introduction

Communication impairment is associated with limited participation in children with complex disabilities [1,2]. Participation, defined as the engagement in all facets of life, has been identified as an essential outcome for children with disabilities [3,4]. Full and effective participation demands being on an equal basis with others and transcends home, school, health, and community environments. This engagement facilitates quality of life [5] and development [6], allowing children to realize their potential [4]. While communication is one of the core components of participation, communicative impairment has been associated with participation restrictions for children with unintelligible speech [1,3,4].

Augmentative and alternative communication (AAC) systems can enhance communication and, ultimately, participation for children with complex needs [7,8]. Electronic AAC systems produce speech via technology external to the child's body [9,10]. For children who have physical limitations that restrict direct access to electronic AACs via touch, an accompanying access technology is necessary. An access technology consists of a sensor that captures the child's intent from a physical movement, vocalization, or physiological indicator and a processing unit that translates the detected intent into a corresponding control signal [11]. A vocal cord vibration switch, one type of access technology, provides an avenue for meaningful interaction to children who cannot rely on speech or voluntary limb movement but have some control of their vocal cords. This switch consists of a sensor, positioned over the vocal cords, to detect periodic vibrations resulting from the child's vocalizations, which are translated into an output signal for computer control [12].

Previous evaluations of the vocal cord vibration switch have been conducted primarily with adolescents, adults, and children older than 8 years [12-15]. Intervening with AAC as early as possible allows for the opportunity to foster children's communication skills and establish their participation potential earlier in their development [10]. School-aged children, particularly those with complex disabilities, rely heavily on their family members and teachers to facilitate social engagement. As parents and teachers play an integral part in these children's participation by defining the opportunities to participate [3], engaging them in the implementation is essential [16]. Furthermore, in previous studies, the vocal cord vibration switch was evaluated over trajectories of 4 months or less. Offering children and their communication partners more time to use the vocal cord switch with support may encourage device adoption [17,18] and enable an evaluation of their experiences later in their trajectory. Finally, these assessments focused on the function of the Hummingbird and did not address the environments in which it would be used.

Accordingly, the purpose of this longitudinal case series was to evaluate the appropriateness and impact of the implementation of a vocal cord vibration switch with school-aged children who have complex communication needs in their natural environments (home and school), based on a collaborative, iterative, child- and family-centered protocol. Mothers' and teachers' perceptions of effort exertion, satisfaction, engagement, and psychosocial impact were measured. Device appropriateness was also assessed by the research team, and observations of process and progress were recorded in field notes. Temporal changes were measured across contexts (home and school) over a 2-year assessment period to allow for ongoing support and assessment after the initial introduction of the vocal cord vibration switch.

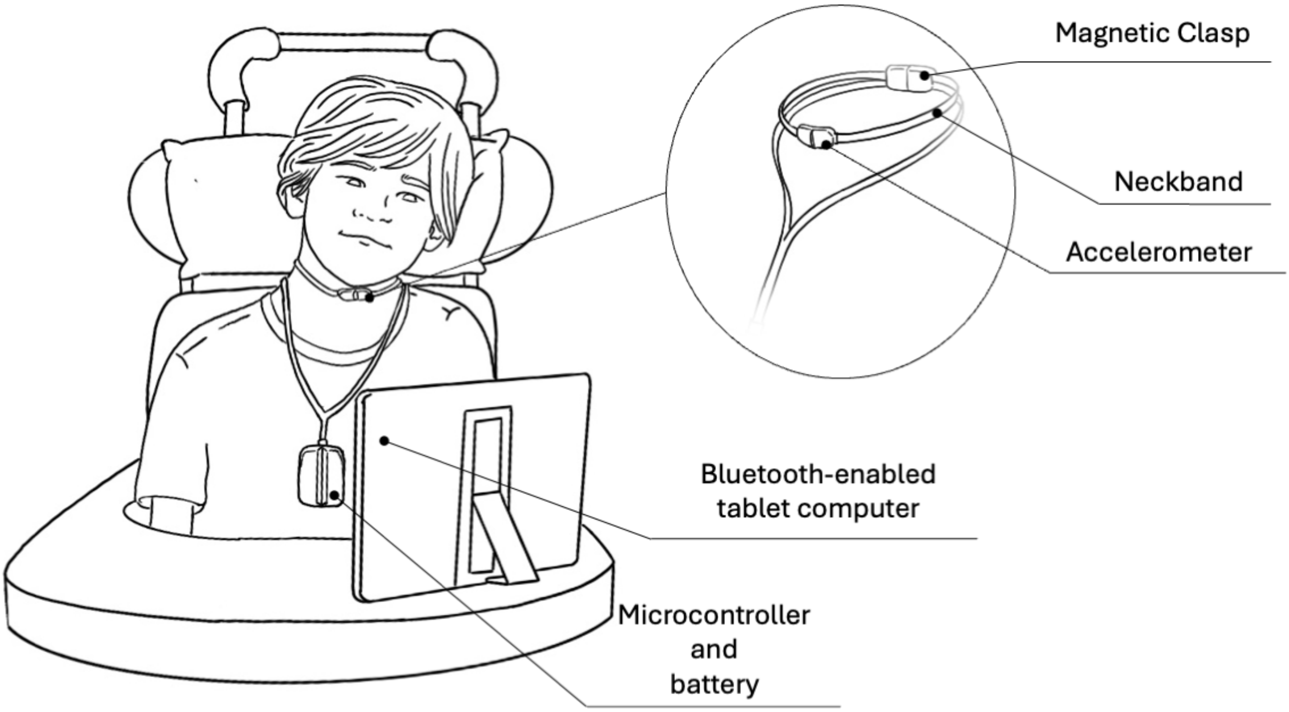
Methods

Overview

In this 2-year multiple case series, a vocal cord vibration switch, the "Hummingbird," was introduced to 3 school-aged children, guided by the Assistive Technology Delivery Protocol (ATDP) [19,20]. The ATDP informs the introduction and adoption of assistive technology for children who have severe disabilities and encourages device adoption by focusing on the compatibility of the technology with the needs of the users, while ensuring standardized training protocols, multidisciplinary inclusion, and an iterative assessment process. The Hummingbird consists of a dual-axis accelerometer, which is held in place on the child's neck proximal to the vocal cords using a neckband. The sensor is connected to a microcontroller box, where input vocal cord vibration signals are monitored and analyzed to detect short- or long-duration intended sounds (ie, voiced sounds or hums; Figure 1).

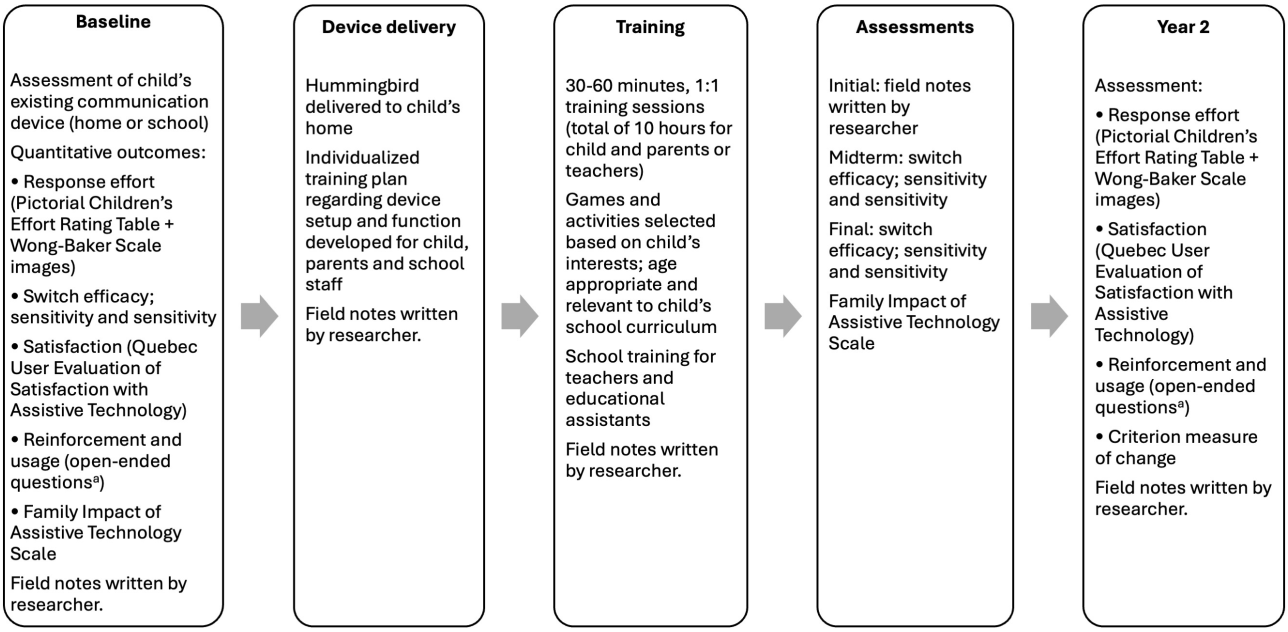
XSL•FO
RenderX

Figure 1. The Hummingbird device worn by a child, interacting with a computer through intentional vocal fold vibrations. The neckband is fastened by a magnetic clasp around the child’s neck. A control signal is transmitted to the tablet computer whenever an intentional vocal fold vibration is detected.



For this implementation, assessments across multiple time points were conducted, including a follow-up assessment (year 2; Figure 2). In the baseline session, assessments of the child’s existing communication device were conducted at home or the participant’s school, according to the preference of the family. Following the baseline assessment, the Hummingbird was delivered to families at home, and an individualized training plan was developed by the research team to include the children’s teachers and parents. The training plan was informed by the child’s family, caregivers (personal support workers, nurses, or additional family members), and hospital-based occupational therapists or speech-language pathologists. Device training for the children occurred in 30 to 60 minute, one-on-one sessions for a total of 10 hours over a 10 to 20-week time frame and involved games and activities that were age-appropriate and of interest to the child. A school integration phase consisted of training for teachers and/or educational assistants. No researcher involvement occurred in year 2 except for a check-in to determine if the children were still using the Hummingbird and to complete a subset of the outcome measures.

Figure 2. Overview of implementation and assessment sessions. ^{a}Open-ended questions for reinforcement: On average, how many attempts does the child make to activate the Hummingbird before he or she is successful? Is there any delay between when the Hummingbird is activated and when the child perceives a response from the technology? Is there any delay or lack of response from the communication partner?



Baseline

Assessment of child’s existing communication device (home or school)

Quantitative outcomes:

- Response effort (Pictorial Children’s Effort Rating Table + Wong-Baker Scale images)
- Switch efficacy; sensitivity and sensitivity
- Satisfaction (Quebec User Evaluation of Satisfaction with Assistive Technology)
- Reinforcement and usage (open-ended questions[a])
- Family Impact of Assistive Technology Scale

Field notes written by researcher.

Device delivery

Hummingbird delivered to child’s home

Individualized training plan regarding device setup and function developed for child, parents and school staff

Field notes written by researcher.

Training

30-60 minutes, 1:1 training sessions (total of 10 hours for child and parents or teachers)

Games and activities selected based on child’s interests; age appropriate and relevant to child’s school curriculum

School training for teachers and educational assistants

Field notes written by researcher.

Assessments

Initial: field notes written by researcher

Midterm: switch efficacy; sensitivity and sensitivity

Final: switch efficacy; sensitivity and sensitivity

Family Impact of Assistive Technology Scale

Year 2

Assessment:

- Response effort (Pictorial Children’s Effort Rating Table + Wong-Baker Scale images)
- Satisfaction (Quebec User Evaluation of Satisfaction with Assistive Technology)
- Reinforcement and usage (open-ended questions[a])
- Criterion measure of change

Field notes written by researcher.

XSL•FO
RenderX

Ethical Considerations

The study was approved by the research ethics board of the Holland Bloorview Kids Rehabilitation Hospital (eREB0291) and each of the participating school boards. To be eligible for the study, a participant had to meet the following criteria: (1) aged at least 4 years, (2) did not have an access pathway or had an existing access pathway that was deemed inappropriate as per International Organization for Standardization (ISO) 9241‐9, (3) understand cause and effect, and (4) have voluntary control over their vocal cords in a manner that generated a measurable vibration.

Additionally, one of the parents (or guardians) of the participant as well as his or her teacher had to possess English competency at no less than a grade 6 level and had to commit to the training workshops.

Participants were recruited by self-referral via the hospital website or through a therapist or teacher. A participant was deemed competent to provide consent if the participant had a reliable access pathway associated with a communication interface (AAC device, iPad with AAC application, or computer with AAC software) that allowed him or her to express his or her complete understanding or could use partner-assisted scanning to express his or her full understanding.

If the potential participant was deemed competent to provide consent, at the initial meeting, the principal investigator or co-investigator explained the study to the potential participant and his or her legal guardian. The potential participant was subsequently asked to explain to the researcher the purpose of the study and his or her involvement in his or her own words using his or her current mode of communication. If the potential participant appeared to understand and was able to provide responses with reasonable accuracy for all questions, he or she was asked to sign consent forms for participation in the study. If the potential participant could not provide answers that were accurate with his or her current mode of communication, he or she was asked to provide assent, and his or her legal guardian was asked to provide written consent. Assent was obtained by both a researcher and the family in order to ensure that the participant understood that he or she could opt out at any time. The participant was also reminded on a weekly basis that he or she did not have to participate if he or she did not want to.

If the potential participant did not meet competency criteria, the consent process followed as described above except that the participant was asked to provide assent, and his or her legal guardian was asked to provide written consent. Where applicable, the participant's teacher and educational assistant or assistants provided written informed consent to participate in structured communication partner training. All participants were compensated for their time according to institutional research ethics board guidelines.

Appropriateness and Impact

Four concepts were measured to evaluate the appropriateness of the vocal cord switch: (1) response effort; (2) switch efficacy; (3) mother or teacher satisfaction; and (4) reinforcement and usage. Response effort refers to the amount of physical and cognitive effort required to use the Hummingbird [19]. It was measured using a modified Pictorial Children's Effort Rating Table (PCERT), consisting of effort descriptors, and a 6-level scale ranging from 0 (very easy) to 5 (very hard), with total scale scores ranging from 0 to 20 [21] and the Wong-Baker scale images [22]. The PCERT has yielded valid and reliable measures of effort with typically developing children [21,23]. In previous studies, scale scores have been moderately to highly correlated with objective measures of effort (Pearson r=0.54‐0.87), demonstrating concurrent validity of the scale at various levels of effort intensity [21], and the scale has demonstrated sufficient test-retest reliability (intraclass correlation coefficient=0.77) [23]. The PCERT was modified to use the images from the Wong-Baker pain scale, as cartoon faces ranging from happy to sad were considered more meaningful to children in wheelchairs than depictions of a child climbing stairs.

Switch efficacy represents a measure of effectiveness as defined by ISO 9241‐9; it was assessed using a switch-accessible game that required low cognitive effort and had unambiguous correct and incorrect responses for both the mother or teacher and the child [18]. The game also provided clear auditory and visual feedback to the child. Each child had 3 training sessions with the game to familiarize themselves with the task. Switch efficacy was based on researchers manually recorded data during the game and calculated using two equations: (1) specificity=true negatives/(true negatives+false positives); and (2) sensitivity=true positives/(true positives+false negatives). The specificity of a vocal cord vibration switch reflects the extent to which the device rejects the user's unintended vocal sounds (eg, a cough) while the sensitivity is the device's ability to correctly identify intended vocalizations. Based on previous evaluations of the efficacy of the Hummingbird with comparatively older children (aged >9 y), we anticipated that the 3 school-aged children in our series would achieve similar performances of specificity between 90% and 100% and sensitivity between 65% and 80% [13-15].

The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) was used to assess caregivers' satisfaction with assistive technology on a 5-point Likert scale ranging from 1 (not satisfied at all) to 5 (very satisfied). A total scale score and 2 subscale scores (Assistive Device and Services) are generated [24,25]. The QUEST 2.0 has been used extensively with clients of all ages, and it has been used to elicit parents' [26] and health professionals' [27] perceptions of their satisfaction with their child's assistive technology. Psychometric evaluations of the QUEST 2.0 have yielded acceptable levels of internal consistency (Cronbach α=0.74‐0.79) [28]. Furthermore, moderate correlations between subscales of the QUEST 2.0 and the Psychosocial Impact of Assistive Devices Scale [29] have been reported (Pearson r=0.34‐0.45), supporting the concurrent validity of the scale [30].

To evaluate the performance of the Hummingbird, the rate of reinforcement (attempts required to activate the Hummingbird) and the immediacy of reinforcement (delay from technology or communication partner) were assessed by asking mothers and teachers open-ended questions (Figure 2). In addition, technology usage was determined by asking them to estimate

XSL•FO
RenderX

the frequency of usage of the Hummingbird. Responses were recorded by the researcher.

The impact of the Hummingbird was assessed using two questionnaires. The Family Impact of Assistive Technology Scale (FIATS-AAC) detects the multidimensional effect of assistive device use on families with children with disabilities [31]. The purpose of this parent-reported questionnaire is to evaluate the functional effects of AAC systems on the lives of children and their families. It consists of 89 items, comprising 13 domains (7 child-related domains and 6 parent-related domains). Item scores are averaged to yield a domain score between 1 and 7. Higher scores suggest positive functional effects within a specific domain, and overall total scale scores range between 13 and 91. It has high internal consistency (Cronbach α=0.91) and test-retest reliability (intraclass correlation coefficient =0.95) [32]. Furthermore, its concurrent validity was supported as demonstrated by moderate agreement with a measure of family functioning (Pearson r=0.54) [33].

The Criterion Measure of Change Questionnaire [34] measures respondents' rating of the change in their child's communication since the commencement of Hummingbird with 3 ordinal categories (worse, no change, and better), as well as the extent and importance of this change on 7 levels ranging from a score of 1 ("a tiny bit") to 7 ("a very great deal").

Field Notes

In addition to the administration of the quantitative measures, field notes were written. These notes, which were constructed by the researchers after each interaction with the children, mothers, and teachers, were unstructured and included observations of: the performance of the technology and the child, study and child or mother or teacher progress, communication with the child's mothers, teaching staff and clinicians, and protocol implementation. The notes were compiled and analyzed using a manifest content analysis [35,36], which relies on the face value of the words written by the researcher and requires the use of preconceived categories for initiating and guiding the analysis. These categories included the following: the child's health and communication characteristics and needs, the benefits and disadvantages of the child's existing communication technology at study entry, the child's experiences with using the Hummingbird (child's likes and dislikes, successful strategies for encouraging use, and activities performed), challenges in using the Hummingbird specific to the child and mother or teacher, problem-solving, and device training and Hummingbird use at year 2. The field notes were grouped into units (sentences or phrases written by the researcher) within each of these categories, and then further organized by concepts (eg, successful strategies) within the category (eg, experiences with using the Hummingbird). This deductive approach allowed for a rich description and provided context to the child's use of the Hummingbird at their home and school, over time.

The purpose of the field notes was to provide context, that is, to identify aspects of the technology or the environment that can improve and support the child's use of the technology. They are useful beyond the study context in that they provide perspective on the aspects of the implementation and evaluation process that may need to be customized for individual children in future studies.

The ATDP [19,20] was used as a framework to design and guide the implementation and assessment process for this case series. The choice of questionnaires and the conceptual anchors for the field notes were guided by the ATDP, specifically the extent to which the Hummingbird meets the overall communication needs of the children. According to this protocol, successful Hummingbird adoption relies on the extent to which the technology is appropriate for each individual child. Appropriateness is defined as the efficiency, effectiveness, and satisfaction associated with the Hummingbird. Furthermore, the impact of the Hummingbird on the child and the family is essential to its success. The extent to which the Hummingbird is aligned with the characteristics of the children, mothers, and teachers is ultimately dependent on achieving a good match between the child and the Hummingbird within their given environment (school and home). The likelihood of this match is strengthened by the guidance of the ATDP.

Results

Herein, the field notes are presented among the results from the quantitative measures to provide context and detail. The clinical and demographic characteristics of the 3 child participants appear in Table 1.

Table . Demographic characteristics of child participants.

Child	Age[a] (y)	Sex	Diagnoses	School	Sensory function (vision; hearing)	Cognitive function
Eve	7	Female	Hydrocephalus	Special education classroom	Some impairment; no impairment	Some impairment
Jameel	5	Male	Cerebral palsy spastic quadriplegia	Home-schooled	No impairment; no impairment	No impairment
Zach	6	Male	Global developmental delay and epilepsy	School	Cortical visual impairment; no impairment	Significant impairment

[a] At study enrollment.

XSL•FO
RenderX

Field Notes: Participant Description and Rationale for Hummingbird Introduction

At the start of the study, Eve was aged 7 years and attended a congregated public school for children with physical disabilities. In school, she used button-type switches (BIGmack button device) [31] with her hands, requiring a lot of assistance, demonstrating little accuracy, and experiencing a lot of fatigue. Her fatigue increased as the day progressed. Eve could control her head movements but had difficulty maintaining a steady head position, which resulted in her head dropping forward or to the side. She would vocalize when excited, happy, or upset. Accompanied by facial expressions, these vocalizations were used by parents and teachers to attempt to interpret what she was trying to communicate. She was able to verbalize the words “mama,” “papa,” and “hi,” and communicate with facial expressions and sounds that only familiar communication partners could understand.

Jameel enrolled in the study at the age of 5 years and was home-schooled. At that time, he had been attempting to use a Tobii eye-gaze system [32], but the high tone of his muscles made it difficult for him to hold his head in one position, even with physical head support. As such, he found it challenging to maintain a dwell long enough for the eye-gaze system to activate the desired choice on the screen. Jameel had good voluntary control of his voice. The Hummingbird was considered a potentially appropriate device for him, as it does not require head stability, and he possessed the vocal control for activation. It was hoped by his parents that he would be able to combine the eye-gaze system and the Hummingbird.

Zach was aged 6 years and attended a specialized school for children with disabilities at the start of the study. It was difficult for Zach to control his limbs and head, and he experienced frequent seizures. He had previously tried to use button-type switches (BIGmack button device) by controlling them with his hand and head, but he found them too tiring. He was able to indicate ‘yes’ using a simple non-word vocalization (‘ah’). This vocalization had been used to elicit answers using partner-assisted scanning and a communication book. However, his family reported that the pictures in the communication book were challenging for him to discern due to his cortical visual impairment. Zach’s family expressed a desire for an appropriate technology for him so that he could demonstrate his comprehension and answer questions in the classroom, as it was otherwise difficult for him to display his capacity for learning and understanding class material. Zach did not have the first-year final assessment for the FIATS because he was not using the Hummingbird at home, only at school during the first year. The Criterion of Change was also missing for this participant.

Field Notes: Observations During Sessions

Eve really enjoyed music and as such, it was used as the initial basis for her switch training. She began with errorless activities where she had to activate the Hummingbird to hear the next verse of a song and moved on to timed options where she would activate her Hummingbird to sing a repeating line from a song at the correct time. She often needed prompting to “wait” for the correct time to activate, but this need for prompting reduced over time. There were several interruptions to Eve’s participation in the study as she changed schools and school boards multiple times throughout the study. After consultation with one teacher, choice-making during circle time activities was chosen as the focus for classroom training for Eve.

Jameel progressed very rapidly from simple cause and effect exercises to games that required precise timing. At the start of the study, he was working with his mother (who was also his home-school teacher) and an educational assistant, with a home visit approximately once per week from a member of the research team. One of the main challenges for Jameel was finding activities that motivated him; he could easily become bored with an activity that was presented in a repetitive fashion. This made evaluation of the switch efficacy quite challenging—after just 1 or 2 trials of a game, he sometimes refused to play more. Despite his excellent performance with the Hummingbird, the study progressed more slowly for Jameel due to personal and family illnesses, leading to a hiatus of approximately 4 months. Even after the break, however, Jameel quickly regained proficiency in the use of the Hummingbird. In addition, Jameel’s family had been limiting their use of the Hummingbird to prevent Jameel’s toddler sibling from pulling and/or breaking the device wires. Jameel started using a wireless version of the Hummingbird, which became available midway through the study, and his use of the Hummingbird increased after this change. With the new wireless Hummingbird, Jameel was playing games in addition to controlling switch-activated toys that also engaged his sibling.

During Hummingbird training, Zach was motivated by music, astronauts and space, and any activities involving his family members. He was also motivated by helping others, so the research team would reframe simple games as him “helping” the character move through the story or activity. Auditory cues were used as often as possible to reduce his need to rely on his vision; bright colors on a dark background were used as these were easiest for him to see. Zach had sleep difficulties, which often required him to sleep during the day. Occasionally, he would fall asleep during his sessions or struggle to stay awake. On the days when he was well-rested, the time required for him to coordinate a response to scanning choices was 2 to 3 seconds, compared to 15 seconds when fatigued. In addition, he had frequent absent seizures. At times, the research team was unsure if he was processing his answer or temporarily unaware of the task due to seizure activity. Zach would begin a session in an alert state but could then become fatigued partway through. On good days, a quick “warm-up” using a short errorless game, followed by the actual task, seemed to give shorter response times. When he was fatigued at the start, the researchers tried to preserve his energy as much as possible for the actual task.

Appropriateness and Impact: Response Effort (PCERT)

Table 2 presents effort levels, as reported by a mother or teacher, for the baseline assessment when the child was using their baseline or non-Hummingbird device and those for the year 2 assessment. The reported perceived effort required to use the Hummingbird by all three of the children was lower than the effort required for their previous device. The largest temporal change was for the physical effort required by Jameel's mother, decreasing from the maximum effort of 5 to 2.

Table . Response effort scores[a].

	Eve		Jameel		Zach	
	Baseline[b]	Year 2	Baseline[c]	Year 2	Baseline[b]	Year 2
Respondent	Teacher	Teacher	Mother	Mother	Mother	Teacher
Item						
Physical effort for user (child)	3	1	3	1	4	2
Cognitive effort for user (child)	2	1	0	1	4	2
Physical effort for caregiver (mother or teacher)	2	2	5	2	0	1
Cognitive effort for caregiver (mother or teacher)	1	0	3	1	0	1
Total (0 - 20)	8	4	11	5	8	6

[a] By assessment time and item: 0 (very easy) to 5 (very hard), with total scale scores ranging from 0 to 20.

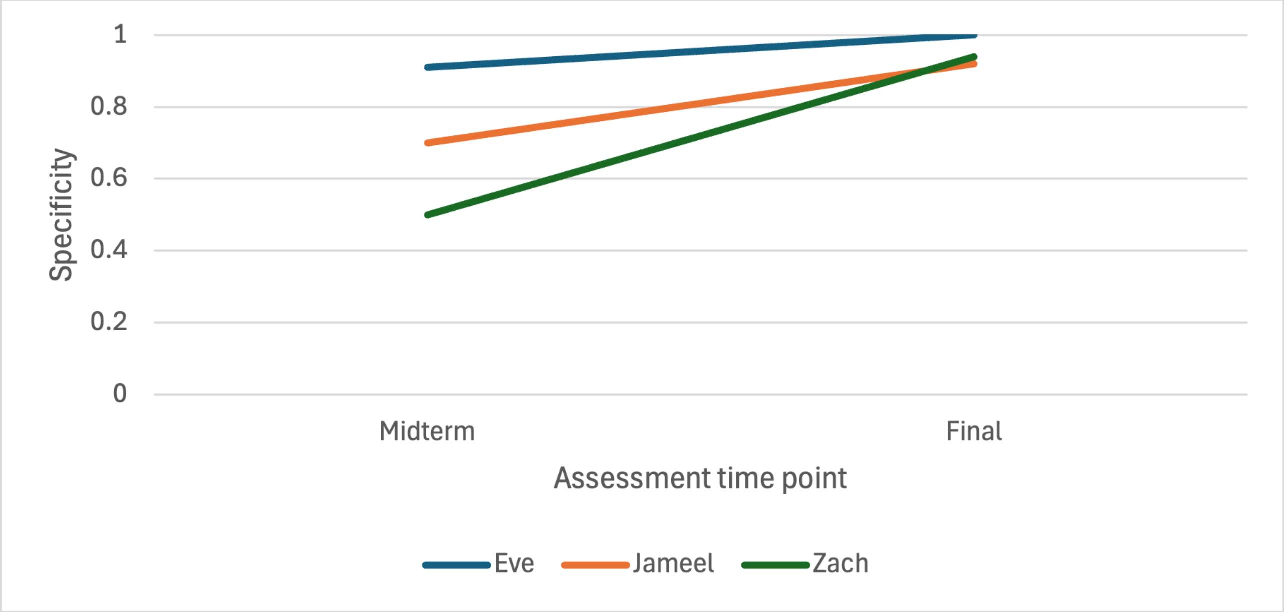
[b] BIGmack button device [37].

[c] Tobii eye-gaze [38].

Hummingbird-Switch Efficacy

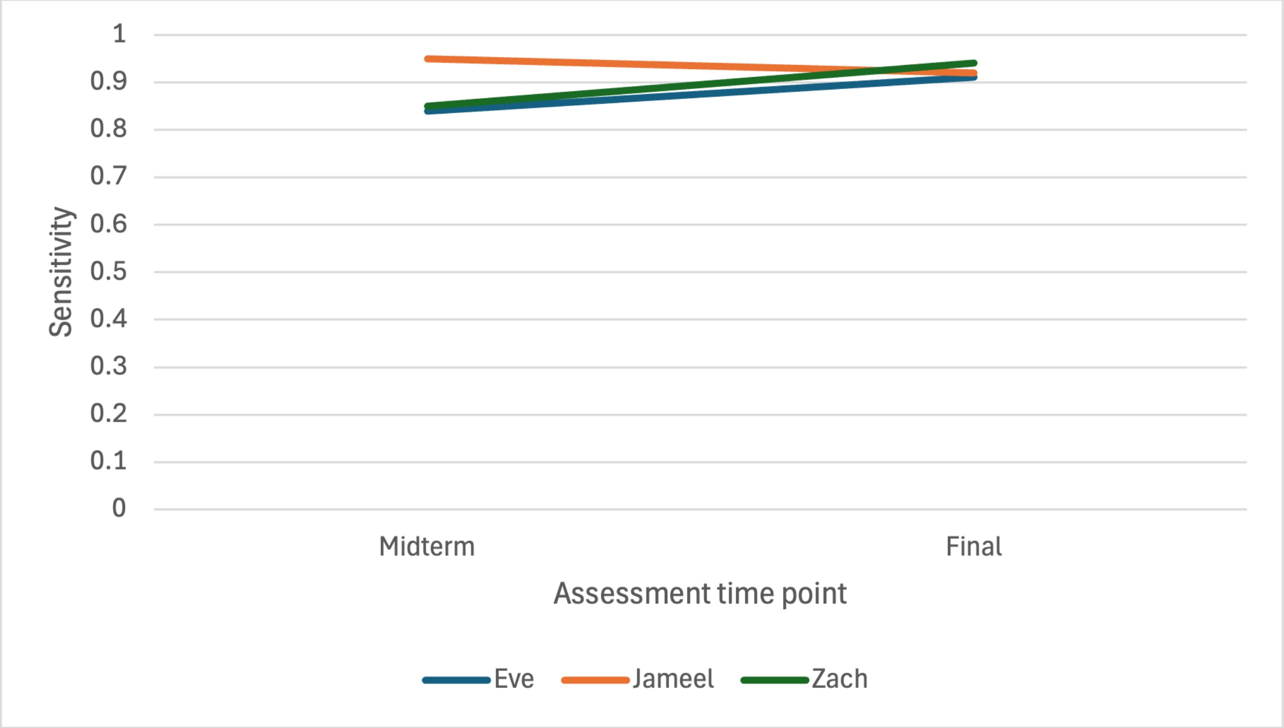
Figures 3 and 4 display switch efficacy at the midterm and final assessments for all 3 children. The Hummingbird demonstrated very high specificity (0.9 and 1) and sensitivity (0.8 and 0.9) for Eve at both the midterm and final assessments. For Jameel, specificity at the midterm was moderate (0.7) but increased in the final assessment (0.9). Finally, very high sensitivity at both midterm (0.9) and final (0.9) assessments was obtained for Zach, and a relatively low specificity at the midterm (0.5), which improved to a very high level for the final assessment (0.9).

Figure 3. Switch specificity at midterm and final assessments.



XSL•FO RenderX

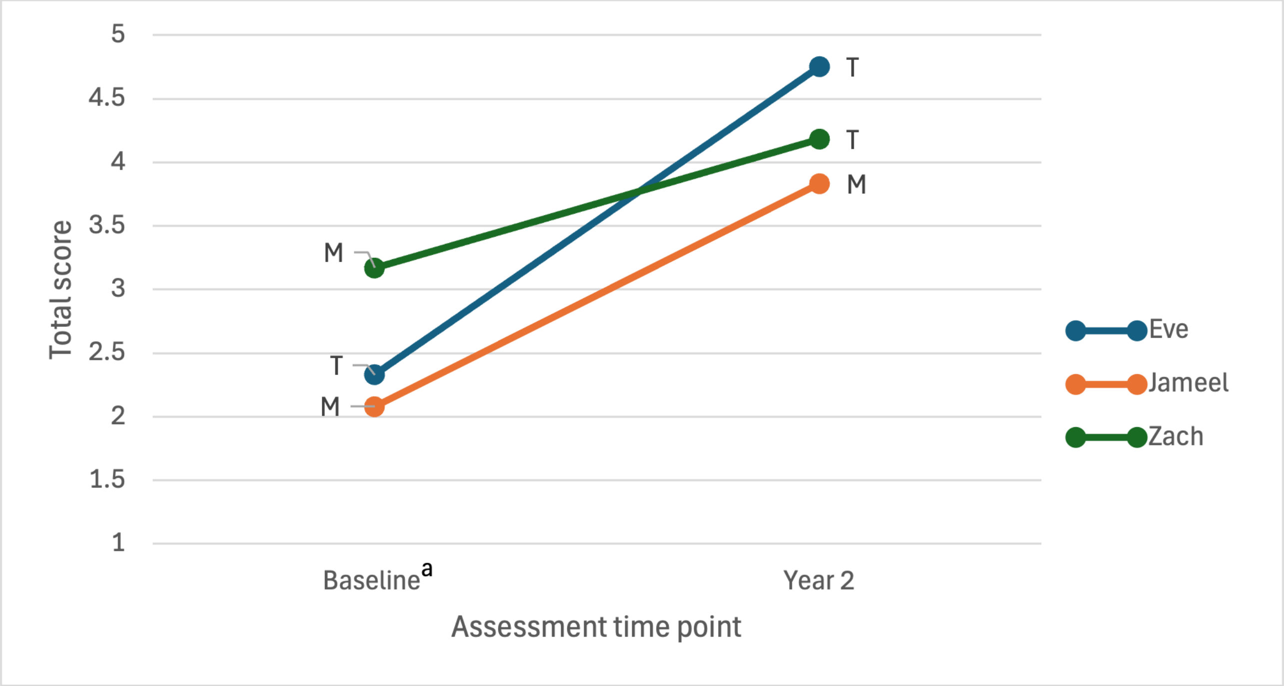
Figure 4. Switch sensitivity at midterm and final assessments.



Quebec User Evaluation of Satisfaction With Assistive Technology

Satisfaction ratings by teacher or mother participants increased from baseline (using a non-Hummingbird device) to year 2 for all 3 participants (Figure 5). All 4 respondents (2 teachers and 2 mothers) rated “easy to use” and “effectiveness” as being one of the top 3 most important items. “Professional service,” “comfort,” and “dimensions” were each reported by 2 respondents.

Figure 5. Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) total scores for baseline and year 2 (Hummingbird; 1=“not satisfied” to 5=“very satisfied”). [a]Eve and Zach using the BIGmack button device, and Jameel using the Tobii eye-gaze at baseline. T: teacher report; M: mother report.



Rate and Immediacy of Reinforcement

The immediacy of reinforcement with the Hummingbird was consistent across the 3 participants, with no delay from the technology during their year 2 evaluation (Table 3). The usage of the Hummingbird, as reported by mothers and teachers, was lower than that of her previous device for Eve. Due to inconsistency in reporting, comparative usage is difficult to

XSL•FO
RenderX

determine for Jameel and Zach. In Table 3, immediacy of response *from the technology* refers to whether there was a delay between activating the device and the intended output (ie, the Hummingbird responding via message output). On the other hand, immediacy of response *from the communication partner* refers to any delay between the time the child used the Hummingbird to communicate (eg, made a choice or requested something) and the time at which their communication partner acknowledged the message. In this latter context, no delay meant that the communication partner (eg, mother or teacher) acknowledged the message and desired action as soon as the child “says” it.

Table . Reinforcement and device usage at baseline and year 2 assessment.

	Eve		Jameel		Zach	
	Baseline (BIG-mack)	Year 2 (Hummingbird)	Baseline (Tobii)	Year 2 (Hummingbird)	Baseline (BIG-mack)	Year 2 (Hummingbird)
Respondent	• Teacher	• Teacher	• Mother	• Mother	• Mother	• Teacher
Usage	• 2 times per day for 30 minutes each • Weekly total: 420 minutes	• 3 days per week for 25–40 minutes each • Weekly range: 75–120 minutes	• 5 days per week for 10–30 minutes • Weekly range: 10–150 minutes	• Varied from: once daily for 5 days per week for months to not being used for months • Weekly rate: 5	• 2 times per day at home and school • Weekly rate: 14	• Minimum 1 time per day for a few minutes to share a short message • Weekly minimum: 35 minutes
Rate of reinforcement	• On some days, 2 attempts and other days, 3–4 attempts to hit switch—depends on mood+health	• Activates every time	• 1 attempt for playing games • Many attempts for communication	• With practice, 2 attempts	• Needs prompting or hand over hand to start, then uses without difficulty	• 1 attempt when device is working properly (ie, tuned to the user’s vocal fold vibration), otherwise several attempts
Immediacy of reinforcement	• No delay	• No delay from either technology or communication partner	• Delay from technology when used for communication • No delay from communication partner	• No delay from technology	• Delay in activating—quicker for the child to vocalize than hit the button	• No delay

Impact

Table 4 presents scores for Eve and Jameel for baseline and final assessments, as reported by their mothers. An increase in scores over time indicates gains (positive functional effects) in dimensions from baseline to final for Jameel. The “face-to-face communication” domain increased from 2.4/7 at baseline to 5.1/7 for the final assessment. The “doing activities” domain also increased from 2.6 to 3.8. No temporal change in the overall score was observed for Eve. Only 2 domains increased for Eve over time, namely, scores from 3.6 to 4.3 for “contentment” and scores from 3.8 to 4.3 for “family roles,” otherwise, her domain scores decreased (doing activities and security) or remained the same.

In year 2, the Criterion Measure of Change Questionnaire, which rates a child’s communication on 3 ordinal categories (worse, no change, and better), indicated that both Eve’s and Jameel’s communication was rated “better” than when they entered the study (baseline). The extent of this improvement was quantified as 2/7 and 4/7, and the importance of this change as 3/7 and 7/7 for Eve and Jameel, respectively.

XSL•FO
RenderX

Table . Family Impact of Assistive Technology Scale mean domain scores for Jameel and Eve over time, as reported by their mothers.

Domain	Jameel[a]		Eve[b]	
	Baseline	Final	Baseline	Final
Child-related				
Behavior	4.5	4.7	6	6
Caregiver relief	2.2	2.8	2.8	2.9
Contentment	3.4	4.9	3.6	4.3
Doing activities	2.6	3.8	4.2	3.8
Education	4.7	6	6	5.7
Energy	1.7	3	3.6	3.6
Face-to-face communication	2.4	5.1	2.8	2.6
Family-related				
Family roles	2.6	3.3	3.8	4.3
Finances	3.6	3	3.4	3
Security	3	3.9	2.3	3.1
Self-reliance	2.9	4.9	3.6	3
Social versatility	3.4	4.6	5	4.7
Supervision	2.1	3	2	2.6
Total score (range 13 - 91)	39.1	53	49.1	49.6

[a] Home-schooled.

[b] Special education classroom.

Field Notes: Observations at Year 1 Completion

By the end of the training phase, both Eve and her teacher were comfortable with choice-making activities. Eve really enjoyed vocalizing to talk or sing along, so the Hummingbird switch was only vigilant (ie, turned on) when she was being asked to make a choice. At those times, she could calm herself and wait for the option she wanted. This selective use of the Hummingbird was adopted to avoid the frustration of excessive accidental activations when she wanted to expressively communicate through her vocalizations. Eve would become excited for sessions and wanted to vocalize a lot. If she was in an excited state, the researcher would give her time at the beginning of the session to just play, activating her switch as much as she wanted, before asking her to relax and calm herself to do the activity. The teacher also waited patiently for her to reach her calm state before presenting her with choices.

By the end of the first year, Jameel was using the Hummingbird with the eye-gaze system and appeared to find it easier than using the Hummingbird alone with a scanning-based communication grid on a tablet. His mother reported that it was easier for him to combine the 2 access methods rather than relying solely on switch scanning with his communication grid on his tablet. He was exploring spelling and sentence building using predictive text, with the assistance of his educational assistant. Depending on his level of motivation for an activity, he could often use his set-up to select from multiple choices independently. At this point, Jameel's toddler sibling had also spontaneously begun to model for him, attempting to get his attention by humming and pointing at the Hummingbird while saying, "My turn!"

Zach's sleep difficulties and seizures made some of the evaluations difficult. When the research team was unsure whether Zach was processing his answer or temporarily unaware of the task due to seizure activity, more frequent indirect prompting was used when the time to produce a response seemed especially prolonged. Indirect prompting included presenting Zach with a visual cue or picture, telling him that something was expected (eg, "Now what?" or "What's next?") and using body language (eg, questioning hand motion). The team learned to carefully observe signs of increased effort, as Zach could start a session alert but become fatigued partway through. During some of the switch efficacy testing sessions, he struggled with fatigue and falling asleep; his behavior during those sessions did not reflect the team's observations in other sessions.

Discussion

Principal Findings

This case series evaluated the appropriateness and impact of a vocal cord vibration switch with children aged 5 to 7 years across home and school settings, according to the ATDP. Despite some challenges related to the children's health and environment such as fatigue, school changes, and illness, the results from this evaluation lend support for the use of the Hummingbird for school-aged children with complex communication needs. Overall, the implementation of the

XSL•FO
RenderX

Hummingbird with 3 school-aged children was successful, and mothers and teachers responded positively to supporting its use at home and school.

Appropriateness and Impact

In our case series, we observed higher sensitivity rates at both midterm and year 2, and similar percentages of specificity compared to 2 previous evaluations of the Hummingbird that were conducted over a shorter evaluation period with children who were primarily older. In 1 study with 2 children aged 5 and 9 years, the average specificity of the Hummingbird was 98.1% and 98.8%, respectively, and the average sensitivity was 73.3% and 80%, respectively [15]. Furthermore, previously reported sensitivity percentages for a child aged 12 years using the Hummingbird were 60% and 65% at 8 and 16 weeks, respectively, and specificity percentages similar to ours were reported at >90% [20]. Switch efficacy levels with children aged 5 to 7 years in our study are similar to those observed with older users, thus supporting the use of the Hummingbird with a variety of age groups. For the youngest participant, Jameel, who was aged 5 years at study entry, the Hummingbird correctly identified his intended vocalizations at a higher or similar rate as older users. As with any efficacy evaluation, the results can be affected by dynamic factors specific to the child or user at the time of any given assessment; for example, fatigue was such a factor for 1 participant in our case series as well as for 1 child in a previous implementation [15].

We observed that mothers' and teachers' response effort scores on the PCERT decreased for all 3 children from baseline, indicating that the perceived physical effort exerted by the child was lower when using the Hummingbird compared to their prestudy communication device. It is noteworthy that this decrease in effort was observed during the year 2 assessment when regular support and contact with the research team had concluded. Ease of use of an AAC has been reported in previous studies as an important determinant of parent satisfaction [39] and device use [40]. Furthermore, literature indicates that parents report device abandonment when the effort to use the device is too demanding [17].

Measuring parents' and teachers' satisfaction is essential, given that dissatisfaction with an AAC system is 1 factor leading to device abandonment [17]. Our satisfaction results suggest that the mothers and teachers in our evaluation had a higher level of satisfaction with the Hummingbird compared to the child's previous device. Teachers' and mothers' satisfaction ratings increased from baseline to the year 2 assessment for all 3 children (scores ranging from 3.8-4.8). Similarly, in a previous case study, an individual aged 19 years also reported higher satisfaction with the Hummingbird (score=4.6) compared to a sound-based device [12]. Furthermore, in another assessment of the Hummingbird with a girl aged 12 years with cerebral palsy, similar QUEST 2.0 scores (4.0 and 3.5 at 8- and 16-wk postdevice introduction, respectively) were reported by her occupational therapist [20]. Our data thus support the applicability and usability of the Hummingbird with younger children than previously documented.

Device reliability is an important factor in preventing device abandonment [41]. In this case series, the quality of reinforcement evaluation indicated that, overall, the children were able to activate the Hummingbird more reliably, with fewer attempts and no delay compared to their previous prestudy device. The assessments during year 2 of the study revealed that all 3 children had continued to use the Hummingbird after the main study period.

Family impact scores on some of the FIATS domains for Eve and Jameel did not change from baseline, using their non-Hummingbird device, to the final assessment after months of using the Hummingbird. Overall, Jameel's total scale scores indicated a moderate increase in the positive functional effects with the Hummingbird (from 40/91 to 53/91) and no change in total functional effects for Eve. Perhaps because Jameel was home-schooled, his mother may have observed more of an effect on family functioning. In addition, Jameel's sibling started to play games with him while he was using the Hummingbird, which may have resulted in a greater overall impact on family functioning.

It is critical that AAC systems are designed for the individual to optimally facilitate communication and promote sustained user engagement [18]. In this study, personalization involved changes to the Hummingbird itself as well as to the process aspects of the protocol. The field notes provided insight into the characteristics of the children and their environments and permitted an iterative approach whereby adjustments to the training, implementation, and evaluation processes were made during the study. For example, the researcher (or researchers) explored different activities with each child; for example, Zach excelled when engaging in activities that involved helping people. A wireless version replaced Jameel's original Hummingbird to address the needs of his family environment. These customizations encouraged engagement in subsequent sessions with the three children as well as offering a guide to the implementation process in future work.

Previous research in the broader AAC literature has supported the importance of training communication partners; effective partner training promotes the communication abilities of individuals with complex communication needs [18,42]. Despite this important element, most studies have excluded AAC training for parents and siblings of children with multiple disabilities.

Value of Multisite Inclusion and Individualized Implementation Approach

Challenges are inherent in the implementation process of a device for children who have complex, co-existing conditions that compound their existing activity limitations. In our case series, the child participants differed from one another as they had a variety of health, social, family, and school context characteristics. The successful longitudinal implementation of the Hummingbird with these 3 school-aged children resulted from a comprehensive, individualized approach, including adjusting and modifying both the technology and the protocol, to meet the needs of the child and the child's environment. For example, changing to a wireless device for Jameel was essential to encourage continued use of the technology when his brother was present.

XSL•FO
RenderX

As purported by the ATDP, training and support by the team ensure a child-specific approach that recognizes and adapts to the technology user, including elements of their physical and social or family environment. Of priority to our implementation was the inclusion of key players from each of the child's environments. Furthermore, based on the ATDP, device training and implementation were individualized and embedded into the participants' daily lives and natural environments, while evaluation was semipersonalized, such as affording the child the choice of a game. Inclusion of the school setting in the implementation, attention to training, and quality of service delivery are indicators of users' and caregivers' satisfaction as well as device adoption [42-46].

Contextualization Within Existing Literature

The development and initial evaluation of the vocal cord vibration switch have been described previously [13-15]; these previous evaluations were conducted over shorter time periods (0.5 - 4.0 mo), primarily with adolescents, adults, and children older than 9 years and focused on the physical aspects of the Hummingbird, its reliability, and the ease of set-up for parents and teachers. The results from these earlier assessments informed a more formalized switch training process, the design of the neckband and its positioning, and the decision to use a questionnaire assessing family impact in the case series. In contrast, this case series measured the impact of long-term use in the child's natural environment.

Fundamentally, our findings indicate that the Hummingbird can afford a pathway through which a child can access switch-controllable technology, inclusive of communication, mobility, and environmental devices. From the perspective of the International Classification of Functioning, Disability and Health for Children and Youth framework [3], our results suggest that via the Hummingbird (environmental facilitator), children's vocal cords (body structures) can be accessed by amplifying vocal fold vibrations (body function) to expand independent, intentional switch control (activities) within their given setting (environmental participation conditions). Environmental conditions included barriers such as a child's fatigue during sessions and school changes, as well as motivators such as music or a sibling joining in on a game. In turn, such technology-facilitated switch activations can augment children's participation in communication, interpersonal interactions, education, and directing their own care (specific participation domains).

Limitations

The assessment time points did not always follow the intended trajectory. For reasons of illness and family health issues, which reflect the reality of many families with children with multiple disabilities, some of the assessments were not conducted at the intended times and/or were not completed. Furthermore, some of the questionnaires were not completed at certain time points, resulting in missing data for the FIATS and the Criterion of Change for Zach. However, given the case series design, the results were interpreted individually rather than by making quantitative comparisons across participants. Furthermore, in instances where 1 assessment was missed, data from other time points were available, permitting nonetheless a temporal view of children's and families' experiences. Future studies may consider a different measure of family impact catered more toward children who are developing basic communication skills, such as expressing preferences.

For some data collection questionnaires, 2 different respondents (ie, mother and teacher) provided the assessment for a given child participant, based on who was present in the assessment session and where it took place (home or school). For both the response effort and QUEST 2.0 measures, scores were provided by a child's mother at 1 session and a teacher at the subsequent session. This may have confounded the potential temporal changes in an outcome. Nonetheless, we arrived at a comprehensive assessment of the Hummingbird by deploying a variety of measures to capture the perspectives of multiple key informants in the training, use, and assessment phases of our protocol. Subsequent studies may consider mixed methods designs and associated integrated analyses that allow for informant-specific data sources [47].

Usage reporting was varied. Mothers' reports regarding the usage of both the child's prestudy device and the Hummingbird were unstructured (Table 3). As a result, the usage data varied in terms of how mothers reported use by units of time. To overcome this limitation, future research may consider invoking automatic logging of Hummingbird usage to objectively capture time and duration of use, the number of activations, and the functional application.

Conclusions

Across home and school environments, the Hummingbird was appropriate for 3 school-aged children with complex communication needs as indicated by their low response effort, high scores on switch efficacy and immediacy of reinforcement, and strong caregiver satisfaction. The impact was generally favorable in familial and communication realms for the 2 participants with complete data. Participants encountered unanticipated health, family, and school challenges during the study, which necessitated modifications to the device and the implementation protocol. Future studies should consider a more expansive set of impact measures in light of these real-world circumstances.

Acknowledgments

The authors would like to thank the families and children for generously participating in the protracted protocol of this study.

Funding

This research was supported by the Accessible Technology Program, Ministry of Innovation, Science & Economic Development Canada (Grant #511898). The study was funded for a 2-year period for CAD $311,111.

XSL•FO
RenderX

Authors' Contributions

TC and LM designed the protocol. TC secured funding, oversaw all aspects of the study, and provided detailed edits to the manuscript and its revisions. LM recruited participants, trained them, collected all the data, managed day-to-day study-related tasks, and revised various versions of the manuscript. DG wrote the manuscript and performed data analysis.

Conflicts of Interest

None declared.

References

1. Allison KM, Doherty KM, Cerebral Palsy Research Network. Relation of speech-language profile and communication modality to participation of children with cerebral palsy. Am J Speech Lang Pathol 2024 Mar 7;33(2):1040-1050. [doi: 10.1044/2023_AJSLP-23-00267] [Medline: 38215219]
2. Connaghan KP, Baylor C, Romanczyk M, Rickwood J, Bedell G. Communication and social interaction experiences of youths with congenital motor speech disorders. Am J Speech Lang Pathol 2022 Nov 16;31(6):2609-2627. [doi: 10.1044/2022_AJSLP-22-00034] [Medline: 36215658]
3. International Classification of Functioning, Disability and Health: Children & Youth version: ICF-CY. : World Health Organization; 2007 URL: https://iris.who.int/server/api/core/bitstreams/81e04699-2930-4da2-ac97-76725c7dcc07/content [accessed 2026-03-04]
4. van Schie PEM, Siebes RC, Dallmeijer AJ, et al. Development of social functioning and communication in school-aged (5-9 years) children with cerebral palsy. Res Dev Disabil 2013 Dec;34(12):4485-4494. [doi: 10.1016/j.ridd.2013.09.033] [Medline: 24139717]
5. Davis E, Shelly A, Waters E, et al. Quality of life of adolescents with cerebral palsy: perspectives of adolescents and parents. Dev Med Child Neurol 2009 Mar;51(3):193-199. [doi: 10.1111/j.1469-8749.2008.03194.x] [Medline: 19191833]
6. Chantry J, Dunford C. How do computer assistive technologies enhance participation in childhood occupations for children with multiple and complex disabilities? A review of the current literature. Br J Occup Ther 2010 Aug;73(8):351-365. [doi: 10.4276/030802210X12813483277107]
7. Crowe B, Machalicek W, Wei Q, Drew C, Ganz J. Augmentative and alternative communication for children with intellectual and developmental disability: a mega-review of the literature. J Dev Phys Disabil 2022;34(1):1-42. [doi: 10.1007/s10882-021-09790-0] [Medline: 33814873]
8. O’Neill T, Light J, Pope L. Effects of interventions that include aided augmentative and alternative communication input on the communication of individuals with complex communication needs: a meta-analysis. J Speech Lang Hear Res 2018 Jul 13;61(7):1743-1765. [doi: 10.1044/2018_JSLHR-L-17-0132] [Medline: 29931287]
9. Iacono T, Trembath D, Erickson S. The role of augmentative and alternative communication for children with autism: current status and future trends. Neuropsychiatr Dis Treat 2016;12(12):2349-2361. [doi: 10.2147/NDT.S95967] [Medline: 27703354]
10. Romski MA, Sevcik RA. Augmentative communication and early intervention: myths and realities. Infant Young Child 2005 Jul;18(3):174-185. [doi: 10.1097/00001163-200507000-00002]
11. Tai K, Blain S, Chau T. A review of emerging access technologies for individuals with severe motor impairments. Assist Technol 2008;20(4):204-219. [doi: 10.1080/10400435.2008.10131947] [Medline: 19160907]
12. Chan J, Falk TH, Teachman G, Morin-McKee J, Chau T. Evaluation of a non-invasive vocal cord vibration switch as an alternative access pathway for an individual with hypotonic cerebral palsy - a case study. Disabil Rehabil Assist Technol 2010 Jan;5(1):69-78. [doi: 10.3109/17483100903278107] [Medline: 19941442]
13. Lui M, Falk TH, Chau T. Development and evaluation of a dual-output vocal cord vibration switch for persons with multiple disabilities. Disabil Rehabil Assist Technol 2012;7(1):82-88. [doi: 10.3109/17483107.2011.557467] [Medline: 21345001]
14. Falk TH, Chan J, Duez P, Teachman G, Chau T. Augmentative communication based on realtime vocal cord vibration detection. IEEE Trans Neural Syst Rehabil Eng 2010 Apr;18(2):159-163. [doi: 10.1109/TNSRE.2009.2039593] [Medline: 20071275]
15. Lu EC, Falk TH, Teachman G, Chau T. Assessing the viability of a vocal cord vibration switch for four children with multiple disabilities~!2009-10-06~!2010-02-19~!2010-04-02~!. Open Rehabil J 2010 Apr 15;3(1):55-61. [doi: 10.2174/1874943701003010055]
16. Marshall J, Goldbart J. “Communication is everything I think.” Parenting a child who needs augmentative and alternative communication (AAC). Int J Lang Commun Disord 2008;43(1):77-98. [doi: 10.1080/13682820701267444] [Medline: 17852533]
17. Moorcroft A, Scarinci N, Meyer C. “I’ve had a love-hate, I mean mostly hate relationship with these PODD books”: Parent perceptions of how they and their child contributed to AAC rejection and abandonment. Disabil Rehabil Assist Technol 2021 Jan;16(1):72-82. [doi: 10.1080/17483107.2019.1632944] [Medline: 31250678]

XSL•FO
RenderX

18. Brittlebank S, Light JC, Pope L. A scoping review of AAC interventions for children and young adults with simultaneous visual and motor impairments: clinical and research implications. Augment Altern Commun 2024 Sep;40(3):219-237. [doi: 10.1080/07434618.2024.2327044] [Medline: 38578299]
19. Mumford L, Lam R, Wright V, Chau T. An access technology delivery protocol for children with severe and multiple disabilities: a case demonstration. Dev Neurorehabil 2014 Aug;17(4):232-242. [doi: 10.3109/17518423.2013.776125] [Medline: 23869969]
20. Mumford L, Chau T. Application of an access technology delivery protocol to two children with cerebral palsy. Disabil Rehabil Assist Technol 2016 Feb;11(2):166-175. [doi: 10.3109/17483107.2015.1063017] [Medline: 26171580]
21. Yelling M, Lamb KL, Swaine IL. Validity of a pictorial perceived exertion scale for effort estimation and effort production during stepping exercise in adolescent children. Eur Phy Educ Rev 2002 Jun;8(2):157-175. [doi: 10.1177/1356336X020082007]
22. Wong-baker FACES® pain rating scale. Wong-Baker FACES Foundation. URL: https://wongbakerfaces.org/ [accessed 2024-07-03]
23. Marinov B, Mandadjieva S, Kostianev S. Pictorial and verbal category-ratio scales for effort estimation in children. Child Care Health Dev 2008 Jan;34(1):35-43. [doi: 10.1111/j.1365-2214.2007.00767.x] [Medline: 18171442]
24. Demers L, Weiss-Lambrou R, Ska B. Item analysis of the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST). Assist Technol 2000;12(2):96-105. [doi: 10.1080/10400435.2000.10132015] [Medline: 11508406]
25. Demers L, Weiss-Lambrou R, Ska B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): an overview and recent progress. Technol Disabil 2002;14(3):101-105. [doi: 10.3233/TAD-2002-14304]
26. Field DA, Livingstone RW. Parents' and therapists' satisfaction with four early childhood power mobility devices. Can J Occup Ther 2022 Dec;89(4):364-375. [doi: 10.1177/00084174221098879] [Medline: 35656731]
27. Eriksson M, Jylli L, Villard L, Kroksmark AK, Bartonek Å. Health-related quality of life and orthosis use in a Swedish population with arthrogryposis. Prosthet Orthot Int 2018 Aug;42(4):402-409. [doi: 10.1177/0309364618774059] [Medline: 29775129]
28. Aledda S, Galeoto G, Fabbrini G, et al. A systematic review of the psychometric properties of Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST). Disabil Rehabil Assist Technol 2024 May;19(4):1228-1235. [doi: 10.1080/17483107.2022.2161648] [Medline: 36645802]
29. Jutai J, Day H. Psychosocial Impact of Assistive Devices Scale (PIADS). Technol Disabil 2002;14(3):107-111. [doi: 10.3233/TAD-2002-14305]
30. Demers L, Monette M, Lapierre Y, Arnold DL, Wolfson C. Reliability, validity, and applicability of the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) for adults with multiple sclerosis. Disabil Rehabil 2002;24(1-3):21-30. [doi: 10.1080/09638280110066352] [Medline: 11827151]
31. Ryan SE, Shepherd TA, Renzoni AM, et al. Responsiveness of a parent-reported outcome measure to evaluate AAC interventions for children and youth with complex communication needs. Augment Altern Commun 2018 Dec;34(4):348-358. [doi: 10.1080/07434618.2018.1520296] [Medline: 30369273]
32. Delarosa E, Horner S, Eisenberg C, Ball L, Renzoni AM, Ryan SE. Family Impact of Assistive Technology Scale: Development of a measurement scale for parents of children with complex communication needs. Augment Altern Commun 2012 Sep;28(3):171-180. [doi: 10.3109/07434618.2012.704525] [Medline: 22946992]
33. Kron AT, Kingsnorth S, Wright FV, Ryan SE. Construct validity of the Family Impact of Assistive Technology Scale for augmentative and alternative communication. Augment Altern Commun 2018 Dec;34(4):335-347. [doi: 10.1080/07434618.2018.1518993] [Medline: 30369255]
34. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. Phys Ther 1996 Oct;76(10):1109-1123. [doi: 10.1093/ptj/76.10.1109] [Medline: 8863764]
35. Downe-Wamboldt B. Content analysis: method, applications, and issues. Health Care Women Int 1992;13(3):313-321. [doi: 10.1080/07399339209516006] [Medline: 1399871]
36. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]
37. AbleNet Inc. URL: https://www.ablenetinc.com/ [accessed 2024-06-20]
38. Tobii Dynavox. URL: https://ca.tobiidynavox.com/ [accessed 2024-06-20]
39. Starble A, Hutchins T, Favro MA, Prelock P, Bitner B. Family-centered intervention and satisfaction with AAC device training. Commun Disord Q 2005 Dec;27(1):47-54. [doi: 10.1177/15257401050270010501]
40. Baxter S, Enderby P, Evans P, Judge S. Barriers and facilitators to the use of high-technology augmentative and alternative communication devices: a systematic review and qualitative synthesis. Int J Lang Commun Disord 2012;47(2):115-129. [doi: 10.1111/j.1460-6984.2011.00090.x] [Medline: 22369053]
41. Howard J, Fisher Z, Kemp AH, Lindsay S, Tasker LH, Tree JJ. Exploring the barriers to using assistive technology for individuals with chronic conditions: a meta-synthesis review. Disabil Rehabil Assist Technol 2022 May;17(4):390-408. [doi: 10.1080/17483107.2020.1788181] [Medline: 32663110]

XSL•FO
RenderX

42. Elmquist M, Crowe B, Wattanawongwan S, et al. Caregiver-implemented AAC interventions for children with intellectual or developmental disabilities: a systematic review. Rev J Autism Dev Disord 2025 Jun;12(2):290-310. [doi: 10.1007/s40489-023-00394-2] [Medline: 40893926]
43. Larsson Ranada Å, Lidström H. Satisfaction with assistive technology device in relation to the service delivery process—a systematic review. Assist Technol 2019;31(2):82-97. [doi: 10.1080/10400435.2017.1367737] [Medline: 28892461]
44. Lidström H, Almqvist L, Hemmingsson H. Computer-based assistive technology device for use by children with physical disabilities: a cross-sectional study. Disabil Rehabil Assist Technol 2012 Jul;7(4):287-293. [doi: 10.3109/17483107.2011.635332] [Medline: 22612787]
45. Craddock G, McCormack L. Delivering an AT service: a client-focused, social and participatory service delivery model in assistive technology in Ireland. Disabil Rehabil 2002;24(1-3):160-170. [doi: 10.1080/09638280110063869] [Medline: 11827150]
46. Anderson KL, Balandin S, Stancliffe RJ. Alternative service delivery models for families with a new speech generating device: perspectives of parents and therapists. Int J Speech Lang Pathol 2015 Apr;17(2):185-195. [doi: 10.3109/17549507.2014.979876] [Medline: 25472834]
47. Dariotis JK, Mabisi K, Jackson-Gordon R, et al. Implementing adolescent wellbeing and health programs in schools: insights from a mixed methods and multiple informant study. Prev Sci 2023 May;24(4):663-675. [doi: 10.1007/s11121-022-01481-2] [Medline: 36630022]

Abbreviations

AAC: augmentative and alternative communication
ATDP: Assistive Technology Delivery Protocol
FIATS: Family Impact of Assistive Technology Scale
ISO: International Organization for Standardization
PCERT: Pictorial Children’s Effort Rating Table
QUEST: Quebec User Evaluation of Satisfaction with Assistive Technology

Edited by J Sarvestan, S Munce; submitted 07.Apr.2025; peer-reviewed by A Banayat, C Brum, D Ming; revised version received 09.Jan.2026; accepted 30.Jan.2026; published 31.Mar.2026.

Please cite as:
Mumford L, Guerriere D, Chau T
Appropriateness and Impact of a Vocal Cord Vibration Switch for Children with Complex Communication Needs: Case Series
JMIR Rehabil Assist Technol 2026;13:e75626
URL: https://rehab.jmir.org/2026/1/e75626
doi:10.2196/75626

© Leslie Mumford, Denise Guerriere, Tom Chau. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding the Origins and Factors of Burnout in Physical Medicine and Rehabilitation: Grounded Theory Analysis

Robert Simpson[1,2,3], MBChB, PhD; Eva Cohen[4], BSc; Stephanie Posa[1], MSc; Marina Wasilewski[1,5], PhD; Anthony Feinstein[1,5], MD, PhD; Mark Bayley[1,2], MD; Larry Robinson[1,5], MD; Sarah Munce[1,6], PhD; Carolyn Steele Gray[1,7], PhD; Kristina Kokorelias[1,7], PhD

[1]University of Toronto, Toronto, ON, Canada
[2]Toronto Rehabilitation Institute, Toronto, ON, Canada
[3]University of Glasgow, Glasgow, Glasgow City, United Kingdom
[4]Queen's University, Kingston, ON, Canada
[5]Sunnybrook Research Institute, Toronto, ON, Canada
[6]Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada
[7]Lunenfeld-Tanenbaum Research Institute, Toronto, ON, Canada

Corresponding Author:
Robert Simpson, MBChB, PhD
University of Toronto
1 King's College Circle
Toronto, ON, M5S 3K3
Canada
Phone: 1 (416) 864 5377
Email: robert.simpson@uhn.ca

Abstract

Background: Physician burnout is highly prevalent in Physical Medicine and Rehabilitation (PM&R), but its origins and drivers remain poorly understood.

Objective: This study aims to explore the factors contributing to burnout among Canadian physiatrists.

Methods: Using Charmaz’s Constructed Grounded Theory within a qualitative interpretivist paradigm, we interviewed 30 Canadian physiatrists about their experiences with burnout. Analysis was informed by Cooley’s looking-glass self theory.

Results: Burnout in PM&R in Canada stems from a medical culture prioritizing academic excellence over compassionate care. Canadian physiatrists report shame and self-criticism when unable to meet these high standards. Retrospective accounts from Canadian physiatrists suggest that burnout peaks during residency, where autonomy is low and demands are high. Participants also described feeling unprepared to handle patients’ emotional needs and experiencing moral distress when necessary care cannot be delivered due to systemic barriers. Health care bureaucracy further compounds burnout.

Conclusions: Addressing burnout in PM&R in Canada requires upstream systemic and contemporary cultural change.

(JMIR Rehabil Assist Technol 2026;13:e80499) doi:10.2196/80499

KEYWORDS

burnout; physical medicine and rehabilitation; physiatrists; grounded theory; medical culture; residency training; moral distress; health care bureaucracy; emotional labor; stigma

Introduction

Background

Burnout was originally described as an occupational stress syndrome associated with chronically stressful and emotionally intense work conditions [1]. In health care, burnout has been further classified as having 3 overlapping domains, including emotional exhaustion, diminished sense of achievement, and depersonalization, the latter of which refers to objectification of patients, or loss of compassion [2]. Of concern, burnout in physicians around the world is widespread [3], representing a crisis in health care globally. Physician burnout has significant individual, organizational, and societal costs through worse job

XSL•FO
RenderX

performance, workforce attrition, and lower patient satisfaction [4,5].

Among physician specialties in the United States, Physical Medicine and Rehabilitation (PM&R) consistently ranks among the most burned out [6] (odds ratio [OR] 1.30), with only Emergency Medicine (OR 2.53, 95% CI 2.18-2.93), Family Medicine (OR 1.46, 95% CI 1.29-1.65), and Radiology (OR 1.31, 95% CI 1.12-1.54) reporting higher odds of burnout. Burnout affects a significant proportion of PM&R specialists and trainees—ranging from 48% to 62% for specialists and up to 83.3% for trainees [7]. Putative factors contributing to burnout in PM&R physicians come mostly from US survey data, mirror those seen in physicians overall, and do not clarify the specialty-specific causes that drive or sustain burnout [6]. PM&R uniquely focuses on the longitudinal care of people with disabling conditions, aiming to treat impairments, support independence, enhance social participation, and improve quality of life [8]. These philosophical underpinnings are based on a biopsychosocial model of health [9]. This is the core work of PM&R specialists (“physiatrists”) [10].

Existing literature has highlighted several potential contributors to burnout among physiatrists, including high patient load, administrative burden, limited resources [10,11], and may also include the emotional toll of managing patients with chronic conditions [6,11,12], especially those subject to trauma [13]. However, these studies have largely not explored the personal experiences of physiatrists that lead to burnout in PM&R. To our knowledge, only one qualitative study has examined burnout in PM&R, focusing on personal strategies for occupational well-being during the COVID-19 pandemic; however, it used descriptive methods and did not develop a theory explaining burnout in the specialty.

Physiatrists’ work with people affected by disabling long-term conditions, coordination of multidisciplinary care, and prioritizing of functional recovery entails unique stressors not fully captured by burnout theories from other medical contexts [14,15]. This gap highlights the need for a PM&R-specific theoretical framework to clarify the unique stressors driving burnout and guide targeted interventions aimed at improving physiatrists’ well-being and patient care.

Aims

By capturing the lived experiences and perspectives of Canadian physiatrists, this study aims to provide a comprehensive understanding of the origins and factors contributing to burnout in PM&R through a grounded theory approach.

Methods

Design

A qualitative exploratory approach using grounded theory was chosen for its suitability in exploring underresearched areas and generating deeper insights into attitudes, cultures, and social processes [16], specifically Charmaz’s constructed grounded theory [17-19]. Grounded theory as a qualitative methodology is used when the focus is on the process and the creation of theory or a conceptual framework [17,18,20]. In this research, grounded theory was used to construct a conceptual understanding of how physiatrists comprehend and experience burnout. Charmaz’s [17-19] approach to grounded theory emphasizes the coconstruction of meaning between researchers and participants, aligning well with our theoretical grounding in Cooley’s [21] looking-glass self midrange theory. The looking-glass self theory posits that individuals form their self-concepts based on how they perceive others to view them. This idea was instrumental in shaping our analysis of how physiatrists internalize and respond to burnout—through both personal experiences and external reflections from colleagues, patients, health care institutions, and social systems overall [21]. Using the looking-glass self theory, this study explored how Canadian physiatrists’ burnout perceptions are shaped by interactions with and perceived judgments of others. This perspective enriched our grounded theory approach by highlighting relational dynamics such as how social comparison and external validation influence feelings of adequacy or support. It also guided the creation of visual and narrative materials, which complemented narrative data by capturing the interplay between internal experiences and external perceptions central to both the sensitizing theory and our subsequent analysis [22]. Visual materials were developed to reflect scenarios where physiatrists interacted with their environments, colleagues, and patients, illustrating how external judgments and expectations contribute to the emotional and psychological responses to burnout. In this context, the visuals served as a creative tool to depict how social comparisons, external validation, and perceived judgments play a role in shaping physiatrists’ experiences of burnout.

The study received approval from the Sunnybrook Research Ethics Board, ensuring voluntary participation, confidentiality, and the right to withdraw.

Data Collection

Recruitment and Sample

Between September 2023 and March 2024, 30 physiatrists from across Canada were recruited for the study. There are an estimated 501 physiatrists in Canada [23]. Eligibility required English proficiency, current employment as a physiatrist, and willingness to join an online interview. Recruitment used advertisements distributed by the Canadian Association of Physical Medicine and Rehabilitation, outreach by PM&R Division Directors, and via public licensing registries for remote areas. Potential participants received an initial email plus up to 2 reminders. The team monitored recruitment to ensure diversity in age, gender, race, geography, subspecialty, and experience with burnout, expanding outreach to PM&R leaders when needed. Participants provided consent before interviews and received a CAD $50 (US $36.05) Amazon gift card afterward.

Methods of Data Collection

Data collection occurred from September 2023 to March 2024 through in-depth qualitative interviews conducted by RS and a trained research assistant EC. Questions were designed to be open-ended, to explore participants’ experiences with burnout in PM&R, with the interviewer paying particular attention to views and actions, as well as emergent areas of potential theoretical interest. Audio and video recordings were made via

XSL•FO
RenderX

Zoom (Zoom Communications). The semistructured interview guide, focused on burnout experiences among Canadian physiatrists, is available in Multimedia Appendix 1.

Data Analysis

In this study, data analysis began alongside data collection, allowing for constant comparison between new data and emerging categories. This method ensured that each round of data collection was informed by the preceding analysis, allowing the research team to refine interview questions and explore areas that required deeper understanding. We used an inductive and iterative process to derive final insights [18]. Coding was done in multiple stages—open coding to identify initial concepts, axial coding to link these concepts together, and selective coding to integrate them into a coherent framework. First, two researchers open-coded the data in NVivo (Lumivero) using descriptive codes on participant responses [24]. The reviewers then met with the principal investigator to discuss the preliminary coding and developed an inductive codebook that was applied to the transcripts. The looking-glass self theory informed our coding process by encouraging us to attend to moments in the interviews where participants referenced feedback from their work environment—whether from peers, supervisors, or patients—as central to their feelings of burnout or professional fulfillment. This theoretical approach underscored the importance of external perceptions in shaping burnout, thus enriching our analysis of the interplay between personal identity, professional role expectations, health institutions, and systemic pressures. From this, the same researchers engaged in a constant comparison analysis where meetings were held to discuss the common higher-order themes implied by clusters of open codes, leading to the development of axial codes reflecting participants' insights on the conceptualization and understanding of burnout and experiences with this phenomenon [25,26]. The looking-glass self theory served as a sensitizing concept, framing how burnout is shaped through social interactions in PM&R and guiding the interpretation of nuanced reflections on acknowledgment by self, professional peers, and institutions. Two researchers applied axial coding to transcripts, discussing the process weekly [26]. Once team discussions suggested a preliminary understanding of a theoretical framework, the researchers engaged in a final reflexive process of selective (focused) coding [27]. Selective coding involved the deliberate examination of specific aspects of the preliminary framework within the transcripts to deepen our understanding of the framework, explore relationships between the data, and extract meaningful insights [19]. Categories were identified from comparative analysis, culminating in a core category that encapsulated the study's main theme [18]. Through team discussions, the researchers were able to extrapolate prominent themes within the data to inform an overarching conceptual understanding [19]. In between the weekly meetings, the researchers went independently back to the data to further refine their ideas against the emerging conceptual framework.

We systematically searched for disconfirming cases that challenged emerging themes, documenting and analyzing them to refine our findings. Analysis and team discussions continued until a conceptual framework achieving theoretical saturation was developed [28], where themes around how physiatrists understood and experienced burnout were consistently repeating, with no new dimensions or variations emerging, and the conceptual framework appeared sufficiently dense to explain the origins and factors sustaining burnout in PM&R [27].

Positionality

Our positionality as a team influenced this process, as we were mindful of how our own experiences and professional backgrounds shaped the way we coded and interpreted the data [29]. The lead author, a middle-aged Scots Canadian physiatrist with expertise in mindfulness and compassion, contributed insider insight on burnout in high-stress health care settings, framing it through the looking-glass self theory and emphasizing emotional and relational aspects. An interdisciplinary team, including a PhD-trained rehabilitation scientist and early-career researchers, incorporated structural and institutional factors and reflected on their roles, producing a nuanced, reflexive framework capturing the complexity of burnout in PM&R.

Ethical Considerations

This study was conducted in accordance with the World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Participants and prospectively approved by the Sunnybrook Research Ethics Board (project ID: SUN-5835). Informed consent was obtained electronically from all study participants after the nature and possible consequences of the study were explained. Respect for individual autonomy was emphasized, including that even following the granting of informed consent, participants could choose to withdraw from the study for any reason, without needing to give a reason and without adverse consequence. All participants who provided a semistructured interview received a CAD $50 (US $36.05) Amazon gift card in compensation. All reasonable efforts were made throughout to protect participant safety, dignity, privacy, and confidentiality. All personal identifying information was deidentified, and any interview quotes with potentially identifying information were not included in the study findings.

Results

Participant Characteristics

Most participants were women (19/30, 63%), with a mean age of 45.5 years. Most participants were recruited from Ontario (18/30, 60%). The most common racial identity was White European (12/30, 40%), followed by White North American (9/30, 30%), East Asian (6/30, 20%), South Asian (2/30, 7%), and Latin European (1/30, 3%). Participants worked in a wide range of PM&R settings, but most worked in an urban setting (28/30, 93%), full time (22/30, 73%), with 19 (63%) participants working 50 or more hours per week, with mixed inpatient and outpatient populations (24/30, 80%), and in most cases in neurological rehabilitation (13/30, 43%). No PM&R residents were interviewed (Table 1).

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of participants.

Characteristic	Values (N=30)
Sex/gender, n (%)	
Female/woman	19 (63.3)
Male/man	11 (36.7)
Age (years), mean (SD)	45.5 (10.5)
Number of years practicing, mean (SD)	13.6 (10.7)
Province of practice, n (%)	
Ontario	18 (60)
British Columbia	4 (13)
Quebec	3 (10)
Saskatchewan	2 (6.7)
Alberta	1 (3.3)
New Brunswick	1 (3.3)
Nova Scotia	1 (3.3)
Race/ethnicity[a] , n (%)	
White European	12 (40)
White North American	9 (30)
East Asian	6 (20)
South Asian	2 (6.7)
Mixed	2 (6.7)
Latin European	1 (3.3)
Specialty[a] , n (%)	
Neurorehabilitation	13 (43.4)
Electromyography	9 (30)
Other	8 (26.7)
Spinal rehabilitation	5 (16.7)
Musculoskeletal	4 (13.3)
General rehabilitation	3 (10)
Pain	3 (10)
Primary professional role, n (%)	
Clinician, full-time	22 (73.3)
Clinician, part-time	8 (11.8)
Secondary professional role[a] , n (%)	
Clinician-teacher	14 (46.7)
Clinician-scientist	13 (43.3)
Administrative roles	10 (33.3)
Clinician-quality improvement	1 (3.3)
Hours worked per week, n (%)	
50+	19 (63.3)
30-40	7 (23.3)
40-50	4 (13.3)
Time spent providing virtual care, n (%)	

XSL•FO
RenderX

Characteristic	Values (N=30)
0%-25%	27 (90)
25%-50%	3 (10)
Provides out-of-hours[b], n (%)	18 (60)
Clinic setting, n (%)	
Mixed inpatient and outpatient	24 (80)
Outpatient	6 (20)
Rurality, n (%)	
Urban	28 (93.3)
Urban and rural	2 (6.7)

[a] Participants could choose more than one option. Percentages may not add up to 100.

[b] Reflects the number and percentage of participants answering “yes” to this question.

Main Findings

The study findings present a conceptually grounded theory titled “Burnout in Physical Medicine and Rehabilitation is the Hierarchical Performance Paradox.” The theory comprises 3 interconnected concepts: hierarchical structure, perception of performance, and their interaction in PM&R. In this theory, which is based on retrospective accounts of the lived experience of practicing physiatrists, the origins of burnout in PM&R are sown early in the training of budding physiatrists. Specifically, the hierarchical structure of medical training and practice imposes escalating demands amidst constrained autonomy as Canadian physiatrists progress through their careers, whilst simultaneously the perceived necessity of perfection in performance—shaped by internalized appraisals by peers, patients, and regulatory authorities—foments feelings of imposter syndrome. This hidden curriculum prioritizes performance proficiency above all else, leading to a mindset of “Doctor first, human second” (Woman, aged 33 years) (the hidden curriculum here refers to the implicit lessons, values, and attitudes conveyed through the behaviors and expectations of mentors, colleagues, and the health care system, which are not formally part of the medical educational syllabus [30]). Burnout risk appears highest when autonomy is low and stress is high, fostering self-criticism and shame. Social comparison and perfectionism are reported by Canadian physiatrists to begin in medical school and intensify during PM&R residency, when autonomy and social support decline, and burnout peaks. Later, increased autonomy does not eliminate risk, as expanded responsibilities, emotional demands, health care system inadequacies, and bureaucratic pressures perpetuate moral distress and burnout throughout a Canadian physiatrist’s career. Illustrative quotes from participants are woven throughout the text. Figure 1 depicts the grounded theory process from coding to theory for burnout in Canadian physiatrists. Figure 2 depicts the hierarchical performance paradox.

Figure 1. Grounded theory process.

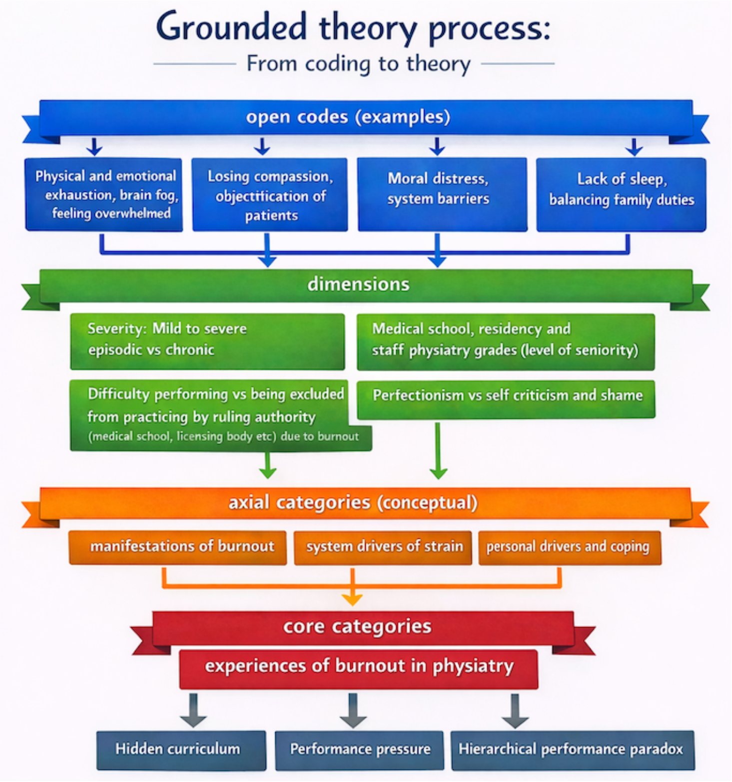
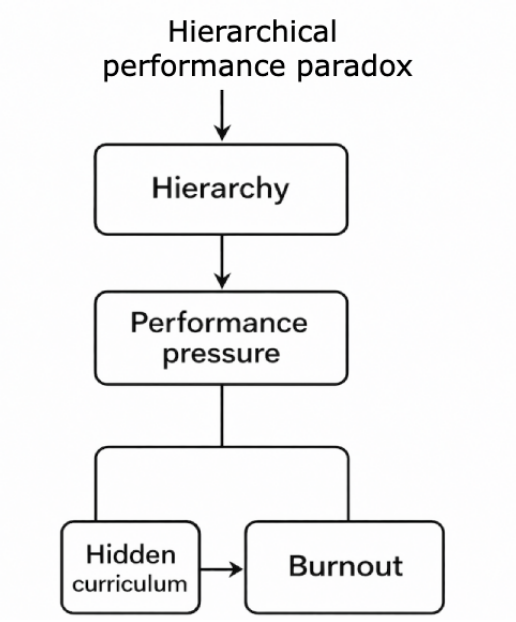


Figure 2. The hierarchical performance paradox and burnout in physiatry.



The Role of Hierarchy in Burnout in Physical Medicine and Rehabilitation

Retrospective accounts from practicing Canadian physiatrists indicate that burnout in PM&R originates from a hierarchical and performance-driven medical education system that inadvertently de-emphasizes compassionate care, leading to stress and emotional exhaustion:

> *Medicine is the stodgiest, most conformist profession that there is, and it has not adapted or evolved to the 20th century, and I think that creates a burnout environment.* [Man, aged 41 years]

Hierarchy in this context refers to the structured levels of authority and responsibility within medical education and practice, where individuals at lower levels, such as medical students and residents, are subject to the oversight, evaluation, and directives of those higher in the system, such as attending physicians and administrators. This hierarchical framework prioritizes academic performance, efficiency, and adherence to institutional norms over compassionate care, creating an environment where Canadian physiatrists face intense pressure,

XSL•FO
RenderX

reduced autonomy, and a heightened risk of burnout. A hidden curriculum in this context perpetuates these values, fostering emotional detachment and prioritizing efficiency and exceptional levels of academic performance. These factors, together with systemic constraints and relentless vertical oversight from within the hierarchy, and horizontal peer appraisal, contribute to burnout outcomes such as emotional exhaustion, depersonalization, and a diminished sense of achievement among many Canadian physiatrists.

Participant accounts indicated that the motivations for becoming a physiatrist are largely altruistic, reflecting core values of improving quality of life for people with disabling long-term conditions:

> *I try my hardest to be compassionate to patients because that’s why I wanted to get into what I wanted to do in physiatry. I feel that you have a longer relationship with people, so you need more of a compassionate approach versus a one-off here and there if you're doing surgery or something like that.* [Woman, aged 35 years]

However, participants also stressed that, in spite of altruistic motivations, the reality is that performance stratification determines who is eligible to apply for admission to medical school, emphasizing academic performance over other qualities:

> *I think being in medicine and needing the grades you need to get into medical school, things like that require a certain degree of perfectionism, so it’s something that’s been valued throughout your life.* [Woman, aged 35 years]

Participants iterated that once admitted to medical school, proficiency in academic performance is strongly emphasized. Failure to meet the bar is associated with adverse social comparisons, shame, and self-criticism:

> *We’re moving away from that shame-based teaching environment, but we still have a lot of that within our clinical training. Shame-based learning is by no means gone and a lot of the times it's just under the surface rather than directly somebody yelling at you for not being good enough.* [Woman, aged 33 years]

Compassion was described as not being a primary or explicit focus in this milieu; rather, the emphasis was on knowledge of disease and learning how the system really works:

> *Yes, I think part of it that I never learned that [self-compassion], it’s not part of training and medicine training magnifies mistakes. We have whole rounds about how you learn from how people, quote unquote, screwed up. So, I think that that’s part of it, is that the culture of medicine doesn’t necessarily lend itself to that.* [Man, aged 39 years]

Participants recalled how, upon graduating from medical school, as residents in PM&R, they often faced excessive workloads fulfilling the demands stipulated by their superiors, sacrificing their autonomy and sense of self:

> *I was a resident, so I had no control over. I was just told how much call I needed to take. It was different back then, there was payroll, which is the governing body to the residents, but I don't think it was as strict as it is now. The residents were fairly powerless* [Woman, aged 48 years]

Based on participant recollections, autonomy appeared inversely related to hierarchy within the medical education and practice they experienced. In a hierarchical structure such as this, individuals at lower levels, that is, residents, have limited control over their workload, schedules, and decision-making, as these are dictated by superiors. This lack of autonomy was described as resulting in residents prioritizing institutional demands over personal needs or values, leading to a diminished sense of self.

Only upon becoming fully qualified did a sense of autonomy return, at which point physiatrists in Canada often assume a leadership role in the multidisciplinary team:

> *How I really got better was I stopped being a resident and then I became a staff, and I had more autonomy. I think that is a factor to relieving burnout, being able to control your work schedules. Being able to say yes to some things and no to other things. I think that’s actually how it got helped* [Woman, aged 37 years]

However, participants recounted that the reality is that the job continues to be dictated by hierarchical health systems, regulatory body oversight, and peer appraisal of performance:

> *But I’ll say that even at [rehab hospital] there have certainly been instances where I feel like the response from the organisation in response to something that happened with an error that then caused harm was not, I think the kindest and most compassionate response* [Woman, aged 37 years]

Indeed, as Canadian physiatrists advance in their careers and move up the hierarchy, although in one sense autonomy is greater, on the other hand, their responsibilities and expectations also increase, creating a persistent tension between autonomy and normative performance standards.

Regardless, Canadian physiatrists of all stages struggle to provide the care they believe is necessary for their patients due to health system constraints and fear authoritarian reprisal for failing to meet expected high performance standards:

> *I know that in Ontario, people are always like, don’t mess with Ontario College of Physicians because it’s a big problem.* [Woman, aged 32 years]

Where burnout is concerned, regulatory bodies were seen as threatening due to their ability to “take away your license” to practice medicine:

> *The College of Physicians was not chill about it though. Because that was PGY2, right, and then when I graduated, you had to answer some questions about did you ever get time off because of this?...the College of Quebec they were just like, have you or are you in a situation where this will be a problem? There was no differentiating it and then they were like, we need a letter from your doctor saying that you are up to practise medicine* [Woman, aged 32 years]

XSL•FO
RenderX

This highly stressful situation may partly explain the difficulties Canadian physiatrists face in providing care in this hierarchical practice climate.

> *And now they're looking at even things like being rude in your private life as being a reportable offense. I mean, really? 24/7, we have to be models of perfection, or we could be reported and lose our livelihood, which would impact whether we can pay our bills. It's from all sides. To the point where I struggle with whether or not to encourage my own children to enter the profession. Yes.* [Woman, aged 49 years]

> *But the burden of care, I think more and more what happens earlier on is your lifeline is that licence to practise, and if that gets taken away, which seems always to be this looming threat, that if I get caught doing something that's not according to the College's mandate or whatever then basically I can't earn a living because my licence is off.* [Man, aged 60 years]

Participants recalled that, as medical students, despite altruistic motivations for entering PM&R, low hierarchical status, limited autonomy, and strict academic expectations fostered vulnerability and feelings of inadequacy, contributing to early burnout experiences:

> *Again, obviously I'm a woman so I have a bit of a lens there, but imposter syndrome, which I think is more common in women, but I still feel…* [Woman, aged 35 years]

Participants described that medical schools exert authority over those who were underperforming, dictating eligibility for module completion or exclusion from training. Self-care was not explicitly emphasized and seen as antithetical to the culture of medicine:

> *And rather than being harsh towards yourself when you make mistakes or whatever, it's being non-judgmental and kind to yourself pretty much…Well that's definitely not a part of medicine.* [Man, aged 39 years]

> *Well, I'm saying that I don't think that there is a curriculum in medicine that teaches people self-compassion and work-life balance.* [Woman, aged 58 years]

Participants described how this scenario continued into their PM&R residencies, where, as trainees, they had to meet performance standards based on long working hours and antisocial working patterns. These demands prevented them from spending time with friends and family, exercising, and attending to diet and sleep. Autonomy became restricted due to strict shift patterns, and as residents, they remained dependent on guidance from their superiors. During this stage, participants described how burnout loomed large as an ever-present threat:

> *I did burn out in PGY2 and that was the reason I left, is when I had no more compassion for my patients, like I had no empathy for them. I just didn't care. I think it was more me being exhausted and other things and then when my patients started pissing me off to be honest. It was just whenever they would talk to me about things, I was like, why are you complaining about this.* [Woman, aged 32 years]

Participants described that as they transitioned from being a trainee to becoming a fully qualified physiatrist, they took on a greater responsibility and became exposed to greater distress in their patients, "Being a physician, especially for a long time, especially dealing with nasty stuff, I think, if you do let that stuff get to you, it's going to be really hard to succeed in the long run" (Man, aged 41 years), and increasingly aware of health care system limitations, which was a source of moral distress:

> *I want to be compassionate for the people in front of me, whatever role I'm having, whether it's for students and residents or whether it's for the patients and families. But I'm coming up against either structural or system or contextual issues that don't let me do the thing that I think is the right thing to do, and that causes distress* [Woman, aged 43 years]

Risk of burnout was described by participants as ever-present as they struggle to balance growing personal and professional responsibilities. Constant oversight and stringent reporting requirements were seen to create a culture prioritizing protocols and metrics over patient-centered care, participants perceiving institutions to prioritize efficiency over quality in interpersonal interactions:

> *And when you're dealing with human beings, that's extremely difficult to individualize a program for somebody in a system where you're actually a datapoint.. success is measured by the number of patients sometimes just seen, not even healed, not even heard, not even spoken to.* [Woman, aged 49 years]

The chronic lack of resources, including inadequate staffing and time constraints, was described by participants as leaving them feeling pressured to prioritize logistical demands over holistic patient needs. Heavy administrative duties and systemic constraints were seen to undermine emotional openness and sensitivity to patient distress. In this context, participants described how they may come to view compassionate care as ancillary rather than a core element:

> *I think it's like going over and beyond for patients. And also, being there for them emotionally, even though that's not really part of your job description, but sometimes they just need a Kleenex and need somebody to give them a hug.* [Woman, aged 37 years]

Consequently, as Canadian physiatrists' advance in their careers within such a framework, there may be a gradual erosion of emphasis on compassionate care in favor of meeting normative peer appraisal, health care institution, and regulatory demands. Institutional wellness initiatives were perceived as misdirected, emphasizing personal responsibility for burnout and failing to address core cultural and systemic issues:

> *We can exercise, we can mediate. We know those things. Those are not the things that we need changes on. We've all made those changes already. We need*

XSL•FO
RenderX

> *the help in the change in the system. They’re the system factors that get us down* [Woman, aged 49 years]

The Hidden Curriculum and Its Damaging Effects in Physical Medicine and Rehabilitation

A hidden curriculum, rooted in medical school and reinforced through professional role modeling, was described by participants as shaping behaviors and attitudes that can drive burnout and erode well-being across a physiatrist’s career. Such a hidden curriculum was seen as subtly prioritizing professional achievement over compassion, system efficiency over patient-centered care, and physical treatments over addressing emotional distress.

> *I think compassion is really a hidden curriculum kind of thing, and depending on who you’re exposed to, you might or might not get any direct instruction on how it is to be compassionate to people. [Compassion] is not really emphasized a lot in medicine compared to clinical expertise. Our recruitment process and training is directed around shutting down self-compassion.* [Woman, aged 33 years]

These implicit lessons were described as conflicting with the altruistic motivations that draw many physiatrists into medicine:

> *Yes, of course. If you're a human being, you're going to be emotionally affected by other people's struggles, your patients' struggles. And that's why you end up in medicine, right, because you really want to help people.* [Man, aged 45 years]

Participants described that in the hidden curriculum, experiencing burnout is seen as a sign of weakness. To be successful in this arena, participants described how they must meet peer and professional standards at all costs:

> *And then, some of that worry about what other clinicians are going to think about my decision making and choices, and how that’s going to affect long-term relationships, both with the patient, but also with the other clinicians in the workplace.* [Woman, aged 33 years]

The stakes were described as high, margins for error small, and risk of shame apparent:

> *…If we mess up, it’s a big deal…the culture of we expect a lot from our students and learners in, I find it’s sometimes not realistic.* [Woman, aged 32 years]

Indeed, participants recounted how to meet the demands of their work, they often had to sacrifice parts of themselves:

> *I definitely have at different times throughout my career. Having relatively recently finished residency, trying to balance work as well as studying for the [regulatory board] exam. Trying to have any semblance of social life and I mean that very loosely, seeing your parents every now and then. It’s not like I see my friends or do anything really for myself all that often.* [Woman, aged 35 years]

Maintaining this level of professional performance was described as exhausting, physically and emotionally, and to deal with this, physiatrists take steps to minimize the stress they are exposed to in their work.

Participants also described that, to be a successful physiatrist in Canada, this often involved maintaining emotional distance from their patients to avoid burnout and manage the high demands of the profession:

> *I find those very challenging because I’m just trying to sling through and get through my day treating physical problems. But then when you have these mental problems that make it more challenging to treat the physical problem, I think that takes a lot of time and effort and just makes your job more difficult and it requires more empathy which is draining.* [Man, aged 32 years]

And, although participants recognized that distress is both physical and emotional, many did not feel they had been well prepared through their medical education to deal with emotional distress prevalent among their patients:

> *I was downright angry at my medical education early on.. we never even talked about stuff like this, let alone how feelings and emotions get in the way of healthcare.* [Man, aged 60 years]
>
> *I am generally treating physical problems and often times peoples’ psychological distress or mental health problems can worsen their physical symptoms and I just, that’s not a population that I do well with.* [Man, aged 32 years]

Thus, physiatrists may view emotional distress as not within their purview, leading to compartmentalization of care:

> *And I think a lot of my colleagues, especially on the outpatient side in Canada, end up seeing a lot of run-of-the mill chronic pain, which ends up being people complaining a lot. I don’t have to see a lot of that and so that helps me, because I actually see real suffering. Not real suffering, psychological suffering is still real, but I’m not trained to deal with the psychological suffering of things like chronic pain.* [Man, aged 41 years]

As such, Canadian physiatrists may become desensitized to the emotional needs of themselves and their patients, focusing primarily on efficiency and protocolized medicine. However, burnout can still result from the forces of this hidden curriculum, and results can be stark. For some, there is frustration and anger towards the system, as they feel there is no time to be compassionate due to the overwhelming demands of the job,

> *So, the things that get me burnt out would be like when I have a huge amount of work in a short period of time and I don't know how I'm going to be able to get it done. A lot of work and no time, and uncertainty about how I'm going to be able to get through.* [Man, aged 48 years]

XSL•FO
RenderX

Others might become disenfranchised completely, adopting the belief that fixing the broken health care system is not their responsibility:

> *No, I have no interest in doing that. It's too much of an uphill battle.. I would have to put all of my other interests and hobbies in life aside and have to focus on mobilising myself to participate in the process of improving Canadian healthcare, which I am completely unwilling to do.* [Man, aged 41 years]

Physiatrists Lead Multidisciplinary Teams and Focus on Adaptation Rather Than Cure

Participants in this study described the work that they do in PM&R as about leading teams, with a focus on adaptation rather than cure, a source of optimism, enthusiasm, and reward for physiatrists, with many iterating that this was where they derived most joy in their work, and felt most able to help:

> *I think that's why rehab is so rewarding. I know that [short-term wins] make a big difference… It's often those little extra moments that you take.* [Woman, aged 34 years]

Participants emphasized helping patients adapt to disabling conditions and improve quality of life, rather than "fix" a diagnosis. Participants described that involving patients in their care fosters control and confidence but requires physiatrists to have a deep understanding of individual circumstances, often eliciting profound and painful empathy in the process.

> *Really severe impairments. Patients that are so injured or have terrible conditions that are very debilitating. It's very easy to find compassion and empathy for those people. When you also see people who lose a limb, or have a, are paralysed or something. It's very easy to build your compassion up and feel for that person because it's terrible. It's subjectively an awful thing.* [Man, aged 32 years]

> *I also try not to absorb the terrible situations all the time and say it's a terrible situation, but I think subconsciously, because I see it all the time and they're all terrible situations. Usually people that are coming to our services are not coming because of a good situation.* [Woman, aged 47 years]

However, using the knowledge derived from a deep understanding of patient distress, participants described how, as physiatrists, they sought to create tailored care plans that optimize functional independence in their patients. This holistic approach was seen as fundamental to the conceptualization of PM&R, but also exposes physiatrists to difficult emotions:

> *I think the patients often express frustration in different ways… It's not as glamorous or as immediately effective as they had anticipated. And so, that often comes out to us as, 'why aren't you fixing this and why is that not working?* [Woman, aged 47 years]

Participants spoke about the need to boundary themselves in relation to the emotions of others, and to see the limitations of the health care system in which they work with realism and detachment, otherwise they might be quickly overcome by cynicism and burnout. Indeed, participants perceived that forbearance was expected from them in the face of abuse and that, in the predominant culture, their emotional well-being was not a priority:

> *We take a lot of verbal, emotional, sometimes even physical abuse. The expectation is often to just keep providing care and keep being compassionate. But that's not so simple in practice. I do find that can be a real challenge.* [Woman, aged 34 years]

> *We should be, somebody, at the end of the day has said, you know what, I heard about that patient, it was rough. How are you doing? That must be a really tough thing that you had to deal with, how are you feeling? I think that everybody just presumes that they're… Everybody else is stressed out, how is that any different than yours.* [Woman, aged 47 years]

Breaking bad news was another area of PM&R, which participants described as perennially difficult in an emotional sense, especially when patient and family expectations were unrealistic, or were not met, and this experience undermined physiatrists' sense of proficiency:

> *I think sometimes [patients] want a magic pill to make all their problems go away without any side effects of any work on their part… That is that part that I find the hardest. When you strive to go above and beyond and give people ten things they can do but they have excuses… they don't ever get better.* [Woman, aged 35 years]

As physiatrists coordinate and guide multidisciplinary specialist rehabilitation teams, including deliberating over access to care, participants described how they could often be a focal point for venting from other health care providers:

> *Eventually my team pushed me to act on it, like could you do something to potentially protect her, but me knowing that there really isn't an ethical or legal way to go about making this decision on behalf of the patient, I felt very uncompassionate towards my team in that situation.* [Man, aged 35 years]

Participants also spoke about a sense of not being appreciated broadly, by patients, institutions, system leaders, and even peers. They described how feeling unappreciated by others left them feeling isolated, disconnected, and disillusioned, "Nobody, from a physiatry perspective, has said 'we appreciate what you do.' For that reason, I've tended to work more collaboratively with my colleagues in other specialties rather than my own colleagues here, which is maybe unhealthy" (Woman, aged 47 years), and disillusioned with the health care system at large. When feeling this way, participants were less inclined to listen to the nuances in patient stories and more inclined to objectify patients:

> *seeing them as their diagnosis. Like, oh that's the cervical myelopathy, and seeing it as their list, the tick list of things that you've got to do on rounds, to follow up and you're just doing it to make sure that you've done your job.* [Woman, aged 47 years]

XSL•FO
RenderX

Bureaucracy Alongside Hierarchy

Bureaucracy and hierarchy were described by participants as contributing to burnout in PM&R by shifting focus from patient care to rules, paperwork, and complex approval processes, particularly for disability income support. This pervasive administrative burden was felt to foster powerlessness, frustration, and demoralization, subtly eroding autonomy and engagement more insidiously than direct authoritarian control.

> *Stupid rules that make no sense and that take a lot of time and effort to work around. So, dealing with the bureaucracy of stupid rules... And that causes a lot of frustration and wasted energy that could be used on other things. Stupid rules... we need to have some ways to be able to say, some direct lines to be able to call somebody and say, this is why this one shouldn't follow the rules and have someone at the other end being able to say okay. I get it. Yes.* [Woman, aged 55 years]

Participants perceived rigid rule enforcement as nonsensical, adding unnecessary distress to patients and physiatrists alike. In this context, participants described feeling disconnected from their initial sense of purpose, their professional identity overshadowed by the sheer volume of administrative work, which they felt obliged to do in their derived social and professional roles:

> *Oh, I think the paperwork. So, usually, rehab forms or disability forms are at least six to ten, maybe 20 pages long.* [Woman, aged 49 years]

This chronic exposure to perceived impersonal and inflexible bureaucratic hurdles, rather than direct authoritarian mandates, appeared to truly wear down physiatrists, contributing significantly to burnout:

> *Certainly, the burden of documentation, trying to get down all the documents and organize, sometimes some people, they'll need forms filled out, and those can be very time-consuming and at times frustrating. Because you feel like you're writing down things that are already in the record but you have to write them down for the completion of a form for a referral or completion of a form for a disability tax credit, I think documentation definitely.* [Man, aged 61 years]

Discussion

Summary of Main Findings

Using a constructed grounded theory approach, based on retrospective accounts of lived experience among Canadian physiatrists, this study developed a conceptual model of burnout in PM&R, highlighting how medical hierarchy—both explicit and via the hidden curriculum—insidiously shapes professional identity from medical school throughout the practice of PM&R. Unrealistic performance standards, emphasis on academic achievement over compassion, and lack of self-care training erode well-being, foster shame, and limit Canadian physiatrists' ability to respond empathetically. Coupled to this, excessive paperwork, complex administrative processes, and strict regulatory compliance wear Canadian physiatrists down over time.

While PM&R's focus on adaptation and holistic care is fulfilling, burnout undermines Canadian physiatrists' capacity to lead multidisciplinary teams effectively, negatively impacting both colleagues and patients.

Comparison With Existing Literature

No previous studies have used grounded theory to explore burnout in PM&R. However, findings from previous studies using grounded theory to explore burnout in health care providers share similarities with ours. For example, studies in primary care and stroke medicine also suggested systemic issues such as excessive workloads, shifting job roles, and misalignment between personal and institutional values, lead to burnout [31,32]. Likewise, professional dissonance in these contexts also led to feelings of demoralization and a sense of being undervalued [31]. Among nurses, distressing relational aspects of health care also featured in conceptualizations of burnout, particularly during personal interactions, shaping care provided [33], and social support played a mitigating role, affecting emotional responses and some components of attitudinal reactions [33]. In our study, Canadian physiatrists highlighted how they felt ill-prepared for the emotional demands of the work of PM&R, and professionally isolated amid a culture of expected forbearance in the face of emotional distress, where support for their own distress was limited.

Limited literature addresses the origins and sustaining factors of burnout in PM&R. A recent US qualitative study identified similar themes to ours, including the impact of perfectionism, stigma around disclosing distress, and burdensome administrative tasks. Participants advised younger physiatrists to maintain "agency," set boundaries, and practice self-compassion to prevent burnout. These comparisons are interesting but derive from the unusual context of the COVID-19 pandemic, and are not in relation to origins or factors sustaining burnout in Canadian physiatrists [34].

With regards to the overarching concept identified in our study, of hierarchy in medicine, this is a well-recognized phenomenon [35], but has not been described previously in relation to burnout in PM&R. Hierarchy in medicine has been described as a means of imbuing power to make and enforce decisions based on knowledge and role that allows for autonomy and self-governance in the profession. Hierarchies in medicine can be conceptualized in both negative, "dysfunctional," and positive, "functional," terms. In dysfunctional medical hierarchies, those of lower status in the hierarchy are subject to lower health [36], trainees lack psychological safety and agency, and can be subject to outright abuse. In this way, hierarchies can be harmful, not just to those in low-status positions, that is, medical students, residents, but also to patients, collaboration among multiprofessional teams, which is highly relevant in PM&R.

However, it is important to note that hierarchy may also be beneficial, at the macro level, with enhanced performance and evolvability of modular systems [37], but also for the meso (institution, team), and micro (individual) levels, with role and

XSL•FO
RenderX

expectation clarity, enabling a social order that serves a greater purpose, which can support a culture of efficiency, and patient safety where the most responsible clinician is the one making critical decisions regarding patient care, particularly when risk is high, and time is pressured. However, to do this in a way that supports social justice and compassionate care likely requires openness and clarity around normative values and conscious and accountable oversight [38]. If not, hierarchy can also be the basis upon which a hidden curriculum, that is, social norms, can develop, including those around punishment, reward, and progression upwards within the hierarchy. Indeed, medical students appear to actively seek to decipher which behaviors their supervisors value most and to preferentially cultivate these [35]. The implications for the development of compassion, or less desirable behaviors [39], at the individual and organizational levels thus seem clear.

Our theoretical framework, the looking-glass self, provided a helpful lens to explore the origins and factors sustaining burnout in PM&R. Physiatrists in our study relayed how the culture of academic elitism in medicine can lead to adverse social comparisons, imposter syndrome, self-criticism, and shame. In keeping with the looking-glass self precept, Gilbert defines shame as a form of undesired self and a form of social disconnection, but differentiates self-shame from social shame perceived to arise from the judgment (and rejection) of others [40]. Unfortunately, shame is a well-described phenomenon in medical education, which can relate to mistreatment, substandard academic performance, social comparisons, and can be reinforced by hidden curricula. Bynum [41] suggests that shame within medical education can be prevented by explicitly eliminating shame-based teaching strategies, creating a culture of inclusion, and facilitating a mindset of growth in learners. Among internal medicine trainees in the US, exposure to a hidden curriculum (unprofessional conduct, including humiliation by colleagues and patients, disrespectful behavior towards patients) has been found to correlate significantly with all 3 core domains of burnout [42]. Indeed, research into hidden curricula in medical education has previously identified that assessment criteria (which drive what medical students see as being prioritized) do not emphasize compassion as a competency being assessed [43]. We found that self-compassion is seen by Canadian physiatrists as antithetical to the medical training they received. This is unfortunate, as training in self-compassion can help health care providers better self-care, improve compassion for others, and may be protective against burnout with moderate to large effect sizes [44], whilst Compassionate Mind Training may be particularly beneficial in addressing shame-based disorders associated with burnout [45].

Strengths and Limitations

This is the first qualitative study examining the origins and sustaining factors of burnout in PM&R. Strengths include a large, diverse Canadian sample, rigorous methods, and varied team perspectives. Limitations include the majority of participants being full-time urban physiatrists from Ontario, a lack of regional analysis, and a cross-sectional design, which may miss formative early-career experiences and long-term patterns in burnout development. No PM&R residents being interviewed is a notable limitation because many of participants indicated origins of burnout date back to experiences in medical school and residency. We used Canadian PM&R leaders to identify participants, consistent with purposive sampling but potentially introducing selection bias. Self-reported interviews may also be affected by recall or social desirability biases. As a preliminary grounded theory, the framework is at an early stage, may not capture all nuances of burnout, and requires further empirical validation. Findings are specific to Canadian PM&R and may not generalize elsewhere.

Future Directions

Future research should explore the views and experiences of PM&R residents and specifically seek to delve deeper into how intersectional factors influence burnout among physiatrists from marginalized backgrounds within Canada and elsewhere. Research should consider how aspects such as immigration status, ethnic minority status, religious beliefs, and sexual orientation intersect with burnout experiences. Marginalized physiatrists might face unique stressors related to their identity that impact their professional experiences differently than those of their peers. For instance, immigrants might deal with additional challenges related to cultural integration and professional validation, while members of minority ethnic or religious groups might encounter systemic biases or discrimination. It would also be very useful to explore how regional variation in normative organizational culture, health care systems, and wellness practices influences the experiences of physiatrists with respect to hierarchy, hidden curricula, bureaucracy, and burnout.

Conclusions

The origins of burnout in Canadian physiatrists relate to a hierarchical performance paradox that appears to be seeded early in their careers. Budding Canadian physiatrists cite altruistic motivations for joining the specialty, but it is difficult to maintain compassion for oneself and others amidst a hidden curriculum that prioritizes academic elitism over compassionate care. Risk of burnout among Canadian physiatrists is described as peaking in residency but is also ever-present due to the emotionally demanding nature of the work for which they feel ill-prepared, fragmented health care systems, and societal bureaucracy around disability. Burnout can be a source of shame for Canadian physiatrists and is perceived as having the potential for severe professional repercussions. Canadian physiatrists value positive peer appraisal and emotional support (ie, a supportive community); this provides a potential inroad to addressing some of the factors that are responsible for burnout in the specialty.

Acknowledgments

Generative artificial intelligence was not used in any phase of this research study.

XSL•FO
RenderX

Data Availability

Study data will be made available on reasonable request to the corresponding author.

Funding

This work was funded by a research fellowship from Associated Medical Services (Canada; no grant number to report). The funder had no role in the study design; collection, analysis, and interpretation of data; writing of the paper; and decision to submit for publication.

Authors' Contributions

RS, AF, LR, MB, EC, CSG, SM, and KK conceptualized the study. EC conducted data collection. RS, EC, SP, and KK undertook all analyses. CSG, SM, and MB provided senior methodological guidance. RS and KK scripted the manuscript. All authors reviewed, edited, and approved the manuscript.

Conflicts of Interest

Coauthor SM is the current Editor in Chief of the JMIR Rehabilitation and Assistive Technology. Senior author, KK, is a current member of the editorial board of JMIR Rehabilitation and Assistive Technology. None of the authors has any personal financial interests related to the subject matter discussed in this manuscript.

Multimedia Appendix 1
Semistructured interview topic guide.
[DOCX File , 16 KB - rehab_v13i1e80499_app1.docx]

References

1. Freudenberger HJ. Staff burn-out. J Soc Issues 1974;30(1):159-165. [doi: 10.1111/j.1540-4560.1974.tb00706.x]
2. Maslach C. A multidimensional theory of burnout. In: Cooper CL, editor. Theories of Organizational Stress. England: Oxford University Press; 1998:68-85.
3. Rotenstein LS, Torre M, Ramos MA, Rosales RC, Guille C, Sen S, et al. Prevalence of burnout among physicians: a systematic review. JAMA 2018;320(11):1131-1150 [FREE Full text] [doi: 10.1001/jama.2018.12777] [Medline: 30326495]
4. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med 2018;283(6):516-529 [FREE Full text] [doi: 10.1111/joim.12752] [Medline: 29505159]
5. West CP, Dyrbye LN, Erwin PJ, Shanafelt TD. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. Lancet 2016;388(10057):2272-2281. [doi: 10.1016/S0140-6736(16)31279-X] [Medline: 27692469]
6. Makowski MS, Trockel M, Paganoni S, Weinstein S, Verduzco-Gutierrez M, Kinney C, et al. Occupational characteristics associated with professional fulfillment and burnout among US physiatrists. Am J Phys Med Rehabil 2023;102(5):379-388. [doi: 10.1097/PHM.0000000000002216] [Medline: 37076955]
7. Bateman EA, Viana R. Burnout among specialists and trainees in physical medicine and rehabilitation: a systematic review. J Rehabil Med 2019;51(11):869-874 [FREE Full text] [doi: 10.2340/16501977-2614] [Medline: 31608964]
8. Simpson R, Robinson L. Rehabilitation after critical illness in people with COVID-19 infection. Am J Phys Med Rehabil 2020;99(6):470-474 [FREE Full text] [doi: 10.1097/PHM.0000000000001443] [Medline: 32282359]
9. Negrini S, Selb M, Kiekens C, Todhunter-Brown A, Arienti C, Stucki G, 3rd Cochrane Rehabilitation Methodology Meeting participants. Rehabilitation definition for research purposes. a global stakeholders' initiative by cochrane rehabilitation. Neurorehabil Neural Repair 2022;36(7):405-414 [FREE Full text] [doi: 10.1177/15459683221093587] [Medline: 35574944]
10. International classification of functioning, disability and health: ICF. World Health Organization. 2001. URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2026-01-21]
11. O'Brien R, Wyke S, Guthrie B, Watt G, Mercer S. An 'endless struggle': a qualitative study of general practitioners' and practice nurses' experiences of managing multimorbidity in socio-economically deprived areas of Scotland. Chronic Illn 2011;7(1):45-59. [doi: 10.1177/1742395310382461] [Medline: 20974642]
12. Sliwa JA, Clark GS, Chiodo A, Kinney CL, Raddatz MM, Francisco GE, et al. Burnout in diplomates of the American board of physical medicine and rehabilitation-prevalence and potential drivers: a prospective cross-sectional survey. PM R 2019;11(1):83-89. [doi: 10.1016/j.pmrj.2018.07.013] [Medline: 30703291]
13. Teel J, Reynolds M, Bennett M, Roden-Foreman JW, McShan E, Hamilton R, et al. Secondary traumatic stress among physiatrists treating trauma patients. Bayl Univ Med Cent Proc 2019;32(2):209-214 [FREE Full text] [doi: 10.1080/08998280.2018.1559694] [Medline: 31191130]
14. Francisco GE, Chae JC, DeLisa JA. Physiatry as a primary care specialty. Am J Phys Med Rehabil 1995;74(3):186-192. [doi: 10.1097/00002060-199505000-00002] [Medline: 7779328]
15. Mann S. What life care planners need to know about the professional discipline of physician physiatrist. J Life Care Plann 2019;17(1). [doi: 10.70385/001c.151587]

XSL•FO
RenderX

16. Heath H, Cowley S. Developing a grounded theory approach: a comparison of Glaser and Strauss. Int J Nurs Stud 2004;41(2):141-150. [doi: 10.1016/s0020-7489(03)00113-5] [Medline: 14725778]
17. Charmaz K. Constructivist grounded theory. J Posit Psychol 2016;12(3):299-300. [doi: 10.1080/17439760.2016.1262612]
18. Charmaz K. Grounded Theory for Qualitative Research: A Practical Guide. Thousand Oaks, CA: Sage Publications, Inc; 2015:53-84.
19. Charmaz K. Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. Thousand Oaks, CA: Sage; 2006.
20. Chun Tie Y, Birks M, Francis K. Grounded theory research: a design framework for novice researchers. SAGE Open Med 2019;7:2050312118822927 [FREE Full text] [doi: 10.1177/2050312118822927] [Medline: 30637106]
21. Cooley CH. The looking-glass self. In: Human Nature and the Social Order. New York: Charles Scribner's Sons; 1902:126-128.
22. Corbin J. Grounded theory. J Posit Psychol 2017;12(3):301-302. [doi: 10.1080/17439760.2016.1262614]
23. Physical medicine and rehabilitation profile. Canadian Medical Association. 2019. URL: https://www.cma.ca/sites/default/files/2019-01/physical-med-rehab-e.pdf [accessed 2026-01-21]
24. Woike BA. Content coding of open-ended responses. In: Handbook of Research Methods in Personality Psychology. New York: The Guilford Press; 2007:292-307.
25. Belgrave LL, Seide K. Coding for grounded theory. In: The SAGE Handbook of Current Developments in Grounded Theory. Thousand Oaks, CA: Sage Publications, Inc; 2019:167-185.
26. Vollstedt M, Rezat S. An introduction to grounded theory with a special focus on axial coding and the coding paradigm. In: Compendium for Early Career Researchers in Mathematics Education. Cham, Switzerland: Springer; 2019:81-100.
27. Thornberg R, Charmaz K. Grounded theory and theoretical coding. In: The SAGE Handbook of Qualitative Data Analysis. Thousand Oaks, CA: SAGE Publications; 2014:153-169.
28. Aldiabat KM, Le Navenec CL. Data Saturation: the mysterious step in grounded theory method. TQR 2018;23(1):246-261. [doi: 10.46743/2160-3715/2018.2994]
29. Goundar PR. Researcher positionality: ways to include it in a qualitative research design. Int J Qual Methods 2025;24:16094069251321251. [doi: 10.1177/16094069251321251]
30. Hafferty FW, O'Donnell JF. The Hidden Curriculum in Health Professional Education. Hanover, NH: Dartmouth College Press; 2015.
31. Agarwal SD, Pabo E, Rozenblum R, Sherritt KM. Professional dissonance and burnout in primary care: a qualitative study. JAMA Intern Med 2020;180(3):395-401 [FREE Full text] [doi: 10.1001/jamainternmed.2019.6326] [Medline: 31904796]
32. Lee JV, Scott A, Osbun J, Zipfel G. Impact of stroke call on career satisfaction and burnout for academic neurointerventionalists: a grounded theory model. World Neurosurg 2021;151:e552-e564. [doi: 10.1016/j.wneu.2021.04.091] [Medline: 33933697]
33. Rafii F, Oskouie F, Nikravesh M. Factors involved in nurses' responses to burnout: a grounded theory study. BMC Nurs 2004;3(1):6 [FREE Full text] [doi: 10.1186/1472-6955-3-6] [Medline: 15541180]
34. Amano A, Makowski MS, Trockel MT, Menon NK, Wang H, Sliwa J, et al. A qualitative study of strategies to improve occupational well-being in physical medicine and rehabilitation physicians. Am J Phys Med Rehabil 2024;103(8):674-684. [doi: 10.1097/PHM.0000000000002555] [Medline: 38838100]
35. Vanstone M, Grierson L. Thinking about social power and hierarchy in medical education. Med Educ 2022;56(1):91-97. [doi: 10.1111/medu.14659] [Medline: 34491582]
36. Marmot MG. Status syndrome: a challenge to medicine. JAMA 2006;295(11):1304-1307. [doi: 10.1001/jama.295.11.1304] [Medline: 16537740]
37. Mengistu H, Huizinga J, Mouret J, Clune J. The evolutionary origins of hierarchy. PLoS Comput Biol 2016;12(6):e1004829 [FREE Full text] [doi: 10.1371/journal.pcbi.1004829] [Medline: 27280881]
38. de Zulueta P. Developing compassionate leadership in health care: an integrative review. J Healthc Leadersh 2016;8:1-10 [FREE Full text] [doi: 10.2147/JHL.S93724] [Medline: 29355200]
39. Ceñido JF, Obanor ON. Losing face: narcissism in medical training. J Psychiatry Psychiatr Disord 2020;04(04):188-190. [doi: 10.26502/jppd.2572-519x0105]
40. Gilbert P. Shame, humiliation, guilt, and social status: the distress and harms of social disconnection. In: Compassion Focused Therapy. UK: Routledge; 2022:122-163.
41. Bynum WE, Varpio L, Lagoo J, Teunissen PW. 'I'm unworthy of being in this space': the origins of shame in medical students. Med Educ 2021;55(2):185-197. [doi: 10.1111/medu.14354] [Medline: 32790934]
42. Billings ME, Lazarus ME, Wenrich M, Curtis JR, Engelberg RA. The effect of the hidden curriculum on resident burnout and cynicism. J Grad Med Educ 2011;3(4):503-510 [FREE Full text] [doi: 10.4300/JGME-D-11-00044.1] [Medline: 23205199]
43. Wright S, Boyd VA, Ginsburg S. The hidden curriculum of compassionate care: can assessment drive compassion? Acad Med 2019;94(8):1164-1169. [doi: 10.1097/ACM.0000000000002773] [Medline: 31033597]

XSL•FO
RenderX

44. Neff KD, Knox MC, Long P, Gregory K. Caring for others without losing yourself: an adaptation of the mindful self-compassion program for healthcare communities. J Clin Psychol 2020;76(9):1543-1562. [doi: 10.1002/jclp.23007] [Medline: 32627192]
45. Gilbert P. Compassionate mind training: key themes. In: Compassion Focused Therapy. Milton Park UK: Routledge; 2022:273-312.

Abbreviations

OR: odds ratio
PM&R: Physical Medicine and Rehabilitation

Edited by J Sarvestan; submitted 11.Jul.2025; peer-reviewed by P Puram, S Mitchell; comments to author 21.Nov.2025; accepted 15.Jan.2026; published 26.Feb.2026.

Please cite as:
Simpson R, Cohen E, Posa S, Wasilewski M, Feinstein A, Bayley M, Robinson L, Munce S, Steele Gray C, Kokorelias K
Understanding the Origins and Factors of Burnout in Physical Medicine and Rehabilitation: Grounded Theory Analysis
JMIR Rehabil Assist Technol 2026;13:e80499
URL: https://rehab.jmir.org/2026/1/e80499
doi:10.2196/80499
PMID:

©Robert Simpson, Eva Cohen, Stephanie Posa, Marina Wasilewski, Anthony Feinstein, Mark Bayley, Larry Robinson, Sarah Munce, Carolyn Steele Gray, Kristina Kokorelias. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

Original Paper

Design Recommendations for Virtual Reality–Based Upper Limb Exercises From People With Tetraplegia and Spinal Cord Injury Rehabilitation Specialists: Focus Group Study

Andrew Goodsell[1], BSc, MSc; Mariel Purcell[2], BCh, BAO, MB; Matthieu Poyade[3], MSc, MPhil, PhD; Louise Cownie[2], BSc, MSc; Co-designers[2,4]; Lorna Paul[1], MPhil, PhD

[1]School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom
[2]Queen Elizabeth National Spinal Injuries Unit, NHS Greater Glasgow and Clyde, Glasgow, United Kingdom
[3]School of Innovation & Technology, Glasgow School of Art, Glasgow, United Kingdom
[4]Spinal Injuries Scotland, Glasgow, United Kingdom

Corresponding Author:
Lorna Paul, MPhil, PhD
School of Health and Life Sciences
Glasgow Caledonian University
Cowcaddens Road
Glasgow, G4 0BA
United Kingdom
Phone: 44 141 331 8108
Email: Lorna.Paul@gcu.ac.uk

Abstract

Background: The global incidence of spinal cord injury (SCI) is between 10 and 80 new cases per million people each year. This equates to between 250,000 and 500,000 injuries worldwide per year. In the United Kingdom, approximately 4400 people per year sustain an SCI. People with tetraplegia report upper limb function as their highest priority for improvement after SCI. Using immersive virtual reality (VR) headsets, physical rehabilitation exercises can be completed in engaging digital environments. Immersive VR therefore has the potential to increase the amount of therapy undertaken, leading to improvements in arm and hand function. There is little evidence supporting immersive VR as exercise in SCI, especially while patients with SCI are undergoing acute rehabilitation. In SCI research, co-design of new interventions is not a widely adopted approach, yet people with tetraplegia want to contribute with their expert knowledge on their experiences of SCI.

Objective: This study aims to explore the lived experiences of people with tetraplegia and specialist SCI therapists related to acute upper limb rehabilitation and identify design considerations for VR-based interventions targeting the upper limb.

Methods: We conducted 7 online focus groups using Microsoft Teams: 4 with people with tetraplegia (n=15; age range, 36-65 years) and 3 with occupational therapists and physiotherapists specializing in SCI rehabilitation (n=11). Participants were asked to discuss their experiences and expertise about acute SCI upper limb rehabilitation and their opinions and ideas on the use of VR for upper limb rehabilitation. The transcripts were analyzed using content analysis, enabling the proposition of design characteristics of a VR-based intervention for upper limb exercise.

Results: The study identified 5 major themes describing the clinical features, treatment, and recovery of people with SCI during the acute stage of SCI, their motivations for participating in therapy, and suggestions for the design of a VR intervention in treating the upper limbs following SCI.

Conclusions: The themes identified in this study allow the elicitation of software requirements for a bespoke immersive VR platform for upper limb rehabilitation following SCI. They can also contribute to a better understanding of the advantages of using VR as an adjunct to upper limb rehabilitation. Additionally, participants used their expertise to suggest factors that would enable the development of a usable and effective intervention, as well as identifying potential pitfalls and software features to avoid during intervention development. These findings can be used to design accessible VR applications for use by people with tetraplegia and their therapists.

(JMIR Rehabil Assist Technol 2026;13:e66832) doi:10.2196/66832

XSL•FO
RenderX

KEYWORDS

acute/sub-acute tetraplegia; co-design; focus groups; secondary care; spinal cord injury; upper limb rehabilitation; virtual reality

Introduction

A spinal cord injury (SCI) results from an insult to the spinal cord and impacts nearly every aspect of a person's life. People with SCI usually have a degree of permanent neurological disability, including muscular paralysis and impaired sensory and autonomic function below the level of their injury [1]. The global incidence of SCI is between 10 and 80 new cases per million people each year [2], with 4400 new cases per year in the United Kingdom [3]. The lifetime costs of SCI are estimated to total £1.43 billion (US $1.97) for 1270 cases per year in the United Kingdom [4].

The level of SCI is a major factor determining the degree of independence a person will achieve following injury [5]. Typically, the higher the level of injury, the lower the level of independence [6]. Damage to the cervical spinal cord results in tetraplegia, where function of all 4 limbs is impaired [7]. As well as physical and sensory issues, people with SCI are at risk of pressure ulcers, urinary tract infections, spasticity, autonomic dysreflexia, depression, neuropathic pain, difficulty breathing, and circulatory problems [8]. This multitude of impairments following SCI is associated with lower quality of life [9].

Immediately following injury, patients with SCI require acute medical management. This is followed by a period of rehabilitation as an inpatient. The aim of rehabilitation post SCI is to optimize recovery, restore and maximize function by leveraging activity-dependent plasticity for neurological recovery, learn compensatory techniques for lost function, and prevent secondary complications [10]. Central nervous system plasticity, where neural pathways can be altered in response to an injury and promote recovery, can occur spontaneously and/or as a result of rehabilitation [11]. In the last 30 years, the primary aim of physical rehabilitation of SCIs has moved away from compensatory strategies to neurological recovery, particularly in people with incomplete SCI [12]. Incomplete SCI refers to the preservation of sensory and/or motor function below the level of injury, as opposed to complete injuries, which do not have preservation of sensation or movement [7]. Rehabilitation is lifelong, but it is recommended that rehabilitation starts early after injury to maximize functional status and clinical outcomes [10].

Reducing reliance on care and achieving higher levels of independence with respect to activities of daily living is a major goal for patients and therapists [13]. Improving motor function, especially of the upper limbs, through rehabilitation [14] enables patients to carry out tasks, such as dressing, bladder and bowel care, transferring in and out of a wheelchair, and feeding which may otherwise require a carer [15-17]. Even small improvements in upper limb function can have large effects on a patient's independence [18,19]. It is understandable, therefore, that improving upper limb function is a high priority for people with tetraplegia [13,20-22].

Improvements in upper limb function can be achieved through Activity-Based Therapy (ABT) [23]. ABT, a rehabilitation approach, refers to any intervention that involves high-intensity, repetitive exercises that target activity-dependent plasticity in spinal circuits [24]. The improvements in upper limb function from ABT have greater effects on quality of life when compared to traditional physical interventions targeted above the level of injury [19].

Virtual reality (VR) technology has the potential to deliver upper limb therapy by providing engaging interventions for patients as part of their rehabilitation [25,26]. Immersive VR delivers to the user a real-time digital environment that incorporates multimodal sensory paradigms, including proprioception, kinesthesia, stereopsis (the perception of depth via a stereoscopic display), and spatial audio, along with the ability to interact within the digital environment via head and hand tracking [27]. In these environments repetitive upper limb movements, like ABT, can be performed. VR-based therapy can facilitate greater adherence to therapy [28] and increase access to the most effective rehabilitation strategies for people with neurological disorders [27]. However, currently there are few studies that have investigated the use of VR in rehabilitation of the upper limbs, with most focusing on the chronic stage after injury.

VR-based upper limb therapy has the potential to provide engaging interactive environments within which patients can undertake the same type of movement repetitions as they would in conventional therapy. Due to its immersive qualities, patients may be more motivated to engage in therapy and may be able to complete more repetitions than they would compared to conventional treatment. A total of 6 reviews published between 2019 and 2024 have evaluated the evidence for VR for upper limb rehabilitation following SCI. These reviews included between 3 and 7 studies involving the upper limb. The overall findings from these reviews suggested that the evidence base is limited and that immersive VR appears to be more beneficial than non- or low-immersive VR [28-31]. However, these studies also suggested that exercises delivered in immersive VR could be used as an adjunct to conventional rehabilitation, which requires further investigation.

The VR devices, software, and methodology surrounding implementation used for upper limb therapy vary widely. VR-based interventions also vary in terms of their intended clinical use, fidelity, and cost, and the design and production of the software have rarely incorporated the lived experiences of SCI experts (people with SCI and health care professionals). The input of end users into the development of software, particularly in health care, is important in delivering successful products [32-34]. This process allows highly customized and bespoke projects to be developed, ensuring that the service or intervention is targeted toward the needs of patients [33].

This study forms the first part of a cocreation process for the design of a suite of VR-based exercises for upper limb rehabilitation in the acute SCI rehabilitation context. The aim of this study was to explore the lived experiences of people with tetraplegia and specialist SCI therapists related to acute upper

XSL•FO
RenderX

limb rehabilitation and to co-design immersive VR-based upper limb activities.

Methods

Overview

In this qualitative study, 7 focus groups were conducted online from August to October 2022 to gather opinions and insights from the lived experiences of people with tetraplegia and, separately, SCI specialist therapists with experience of upper limb rehabilitation.

Recruitment (Procedure)

A convenience sample of people with (1) tetraplegia and (2) therapists specializing in SCI rehabilitation were recruited. Although the VR intervention was to be developed for inpatient SCI rehabilitation, in order to fully explore lived experiences of participants with tetraplegia, the study recruited people who had been discharged from hospital having completed rehabilitation.

Participants with tetraplegia were recruited through the third sector organization Spinal Injuries Scotland. Recruitment material was also shared on social media channels (Twitter and LinkedIn). Inclusion criteria for participants with tetraplegia were a diagnosis of nonprogressive tetraplegia, aged ≥18 years, able to attend focus groups and contribute in spoken English, and having been discharged from hospital after receiving upper limb rehabilitation in a spinal injuries unit as part of their treatment. Exclusion criteria were any comorbidities that could preclude participants from attending and contributing to focus groups.

Therapists (occupational therapists and physiotherapists) were recruited through similar social media channels and a WhatsApp group of United Kingdom and Ireland–based SCI rehabilitation specialists. Inclusion criteria for therapists were being aged 18 years or older and able to attend and contribute to focus groups in spoken English.

Respondents to recruitment material emailed AG to register interest in participating in the study. Participant information sheets were provided to these interested parties by email. Participants provided informed consent prior to taking part in the study by either printing and physically signing a digital consent form or completing the form digitally. Participants without the physical ability to sign consent forms provided consent by designating a witness to sign on their behalf [35].

After providing informed consent, participants were asked to complete a demographic information form; for people with tetraplegia, this included their age, sex, Spinal Cord Independence Measure self-reported score [36], and for therapists, this included sex, occupation, and years of professional experience.

Focus Groups

Focus groups were chosen over individual interviews to promote discussion and ideation between focus group participants. Question schedules were created to enable semistructured discussion between participants (Multimedia Appendices 1 and 2). The questions were devised to enable future software design and software requirements specification. The difference between the question schedules and conduct of each focus group was minimal. Participants with tetraplegia focus groups aimed to explore the topic of VR through the lived experience of participants with tetraplegia, and therapist focus groups aimed to explore the topic through the lens of clinical practice. Each focus group had 2 facilitators, AG (male) and LP (female). A total of 3 focus groups of therapist participants and 4 focus groups of participants with tetraplegia were conducted. The duration of focus groups ranged between 45 and 75 minutes.

Participants with tetraplegia were asked to recall their experiences of the acute stage of inpatient SCI rehabilitation. They were then asked to discuss upper limb impairments caused by SCI and how these relate to prehension, strength, and range of motion. Therapists were instead asked about the current practice surrounding delivery of upper limb therapy. Both groups were introduced to the concept of VR as a commercial product for immersive entertainment. They were shown a video promoting a high-immersion head-mounted display being used by able-bodied people to play interactive games [37]. Both cohorts were then asked to discuss their opinions on VR as an adjunct to usual upper limb rehabilitation and the potential facilitators and barriers to its use in the acute rehabilitation setting. Finally, specifications for the design of a VR system to promote upper limb exercise were gathered from both groups.

Analysis of Transcripts

Focus groups were audio-recorded, and recordings were transcribed verbatim following pseudonymization of participants. Transcriptions were generated manually to enable familiarization with the data. Content analysis was used to count the responses and determine the meaning of participants' discussions via code generation [38]. Therapist and tetraplegia focus groups were analyzed separately. After codes were generated by AG, they were independently reviewed by LP, M Poyade, and M Purcell. AG, LP, M Poyade, and M Purcell discussed codes to come to a consensus on identification of codes and naming of themes in the theme hierarchy. Quotes are presented with participant identifiers and their focus group number. Data were managed using NVivo (version 20; Lumivero) [39].

Codes (units of meaning) were typically sentences, phrases, or longer portions of text. Shorter phrases and sometimes single words were coded if they related specifically to recommendations for the design of the VR games. Initially, codes were organized under categories deduced from the schedule of questions. Subthemes were abstractions of groups of similar codes and were created during coding. Major themes were then generated from groups of subthemes and replaced the organizational categories. Transcripts were not returned to participants for comments or corrections.

Reflexivity

Relationships with study participants were not established prior to study commencement. LP had previous experience facilitating focus groups and interviewing research participants, but AG did not. Neither facilitator worked in SCI rehabilitation and, as

XSL•FO
RenderX

such, reduced the likelihood of inhibiting frank discussion about patient care with participants [40].

Results

Participants (Demographics)

A total of 18 participants with tetraplegia and 14 therapists expressed interest in the study. Of these, 3 therapists and 3 participants with tetraplegia did not attend the focus groups; thus, 15 people with tetraplegia and 11 therapists took part (Table 1). We conducted 3 focus groups with therapists and 4 focus groups with participants with tetraplegia. A total of 7 participants with tetraplegia did not complete the demographic information forms. Reasons for not completing the information form included technical difficulty completing the form (reported only by participants with tetraplegia). A total of 3 therapists did not report their years of professional experience. Therapists reported not having enough time or opted out of form completion, with no reason given. Half of the participants with tetraplegia had experienced immersive VR at least once, but none had used VR extensively.

Table 1. Summary demographic information for participants with tetraplegia and therapist participants.

Characteristic	Value
Participants with tetraplegia (n=15)	
SCIM[a] score[b] (n=8), mean (SD)	17.9 (6.4)
Age (years; n=8), mean (SD)	54.5 (8.6)
Sex (n=13), n	
Male	10
Female	5
Experience with VR[c] (n=8), n	
None	4
Some	4
Therapist participants (n=11)	
Professional background (n=11), n	
Occupational therapists	7
Physiotherapists	4
Mean years of professional experience[d] (n=8)	14.8 (SD 8.6)
Sex (n=11), n	
Male	2
Female	9

[a] SCIM: Spinal Cord Independence Measure.

[b] Only 8 out of 15 participants with tetraplegia completed the SCIM and age information; 13 participants reported sex. Lower SCIM scores indicate lower independence (total score range, 0-75). Only 15 out of 19 subscales were included, comprising 6 self-care items and 9 mobility items.

[c] VR: virtual reality.

[d] A total of 3 therapists did not report their years of professional experience.

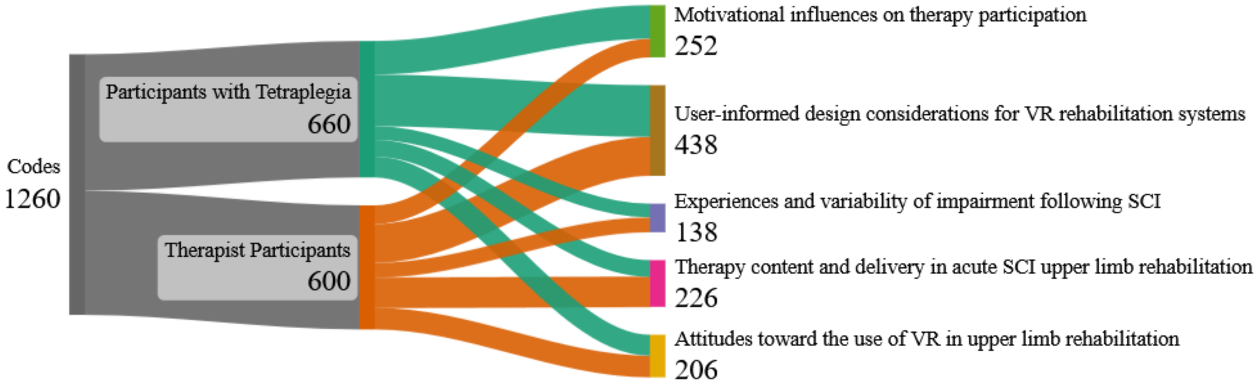
Major Themes and Subthemes

Overview

Analysis of the focus group transcripts identified 5 major themes shared by the therapists and participants with tetraplegia: descriptions of impairment, descriptions of acute rehabilitation, factors affecting participation in therapy, perspectives on VR, and suggestions for VR design. A total of 35 subthemes were generated from 660 and 600 individual codes from the tetraplegia and therapist groups transcripts, respectively (Figure 1).

XSL•FO
RenderX

Figure 1. Visual summary of the number of major themes (listed in the right column) identified from the 1260 codes (shown in the left column) generated by content analysis of focus groups with participants with tetraplegia and therapists. SCI: spinal cord injury; VR: virtual reality.



Major Theme 1: Experiences and Variability of Impairment Following Spinal Cord Injury

A total of 5 subthemes were identified that described impairments resulting from SCI immediately after injury and during the acute stage of rehabilitation (Table 2). Participants described profound loss of upper limb movement following injury, with variable patterns of recovery.

Table 2. Summary of subthemes for the major theme “experiences and variability of impairment following spinal cord injury.”

Subtheme	Therapist codes, n	Participant codes, n	Total, n
Impairment of upper limb movement	17	28	45
Other impairments	27	8	35
Impairment of function, including activities of daily living	9	19	28
Heterogeneous presentations and demographics	16	2	18
Impairment of upper limb strength	3	9	12

> *My arms, I wasn't able to move them at all after I had my accident.* [PwT27, FG 3]

Impairments were described as highly variable across individuals and influenced by the level and severity of SCI.

> *I suppose [the degree of upper limb impairment] depends on the severity of your tetraplegic-ness, you know. Some people might, might be really bad and some people might still have a bit of movement.* [PwT27, FG 3]

Both groups, but especially therapists, discussed “other impairments.” These included spasms, pain, and mental health issues, with therapists also raising age-related comorbidities and complications.

Therapists specifically raised the concept of “heterogeneous presentations and demographics” in acute SCI, with a growing prevalence of central cord syndrome in older adults, presenting as greater weakness in the upper limbs (than lower limbs) and greater involvement of distal muscles [7].

> *You could have somebody who’s got, you know, full shoulder movement, full elbow movement, but they’ve got some weakness in the hands. Or the opposite, you know you could have someone who has got good hands but reduced movement at the shoulder.* [TP3, FG 1]

Major Theme 2: Therapy Content and Delivery in Acute Spinal Cord Injury Upper Limb Rehabilitation

A total of 7 subthemes described the nature of inpatient rehabilitation following SCI (Table 3). Upper limb therapy was described as a combination of functional and remedial activities, with a strong focus on maintaining range of movement and developing task-specific skills.

XSL•FO
RenderX

Table 3. Summary of subthemes for the major theme “therapy content and delivery in acute spinal cord injury upper limb rehabilitation.”

Subtheme	Therapist codes, n	Participant codes, n	Total, n
Upper limb therapy	92	51	143
Patient-centered care	31	0	31
Assistive devices, orthotics, and neck stabilization	0	16	16
Barriers to effective therapy	12	0	12
Psychological adjustment and treatment	4	9	13
Bed rest and ventilation	0	5	5
Symptom management	6	0	6

> *[…] we do work on range of movement, you know shoulder, elbow, wrist, fingers.* [TP7, FG 2]
>
> *There were various things that we used—we had these games where we had little pegs that we had to move around the board, or also another one with a bucket full of the marbles.* [PwT16, FG 4]

Therapists also discussed upper limb functional improvement, particularly where patients have incomplete injuries:

> *If they’re an incomplete injury we might be looking for more of a normal movement. […] Whereas if they’re more of a complete, it’s obviously looking at function and how we can use what they’ve got—what muscles they’ve got working to functionally achieve the task.* [TP9, FG 2]

Both groups highlighted the common use of “assistive devices” during rehabilitation, which included arm and hand stabilization using orthotics and neck stabilization, which hindered movement.

> *I had to use the Active Hand, which are the Velcro gloves so I could hold things.* [PwT16, FG 4]
>
> *I was 12 weeks in a neck brace, so I ended up with a very, very weak neck.* [PwT13, FG 3]

Under therapist-only subthemes, patient-centered care was achieved through collaborative goal setting. Barriers to therapy delivery included delayed referral and limited resources. Psychological adjustment was described by both groups, with participants with tetraplegia emphasizing loss of independence and therapists focusing on expectation management.

> *We would set those goals collaboratively with the patient based on a little bit of, kind of advice and guidance from us as to what we anticipate their outcomes might be.* [TP5, FG 1]
>
> *Unfortunately, the later we get someone sometimes the more issues we can have with contractures and upper limb pain and things, which then makes it harder for us then to, to rehab them.* [TP9, FG 2]
>
> *Trying to deal with everything and I found it really hard to be having to rely on other people for your personal care just day in, day out for everything.* [PwT21, FG4]
>
> *In their mind they want [a goal] to be achievable. So it, again comes down to communication and managing those expectations of patients and their families.* [TH9, FG2]

“Bed rest and ventilation” and “symptom management” were less commonly raised in both groups.

Major Theme 3: Motivational Influences on Therapy Participation

A total of 10 subthemes described factors influencing participation in therapy. Subthemes were then classified as “motivational” or “demotivational” (Table 4). Motivational influences included task enjoyment, goal setting, peer interaction, and visible improvement. Participants with tetraplegia described greater engagement with game-like and functional activities compared to repetitive remedial exercises.

XSL•FO
RenderX

Table 4. Summary of subthemes for the major theme "motivational influences on therapy participation." Subthemes are categorized as either promoting motivation or reducing motivation.

Subtheme	Motivational or demotivational	Therapist codes, n	Participant codes, n	Total, n
Useful or enjoyable rehabilitation	Motivational	1	39	40
Goal setting and therapist support	Motivational	6	31	37
Understanding tasks and prognosis	Motivational	30	0	30
Peer support and competition	Motivational	5	11	16
Seeing improvement	Motivational	0	10	10
Boring, repetitive, or difficult rehabilitation	Demotivational	5	41	46
Limiting symptoms	Demotivational	22	6	28
Low mood	Demotivational	11	6	17
Downtime	Demotivational	0	15	15
Limited resources and lack of supervision	Demotivational	10	3	13

> *We had some games to play, which is quite good fun. Like maybe like playing dominoes. Uh, so fiddling about with them things and also trying to build things with Lego was quite motivational and we had a tower of blocks that was probably my favourite thing.* [PwT26, FG 4]

In contrast, "boring, repetitive, or difficult rehabilitation," which included pointless, tedious, and humiliating activities, was demotivational.

> *You know what it's like sometimes you go and do a bit of exercise, you know you're pulling on a TheraBand and [imitates snoring] 'boring!'.* [PwT1, FG 1]

Sometimes strategies that were intended to make therapy more enjoyable, such as creative activities using putty or interactive cycling, were frustrating for patients.

> *[…] most days doing the exercises with the putty and so on, it felt like a really kind of childish, I don't know. Like, 'what's this doing?'* [PwT21, FG 4]

> *And [patients] found that really—you know they were cycling [FES cycling] away, cycling away along this road but they never went anywhere. And they got really frustrated by that.* [TP9, FG 2]

Some activities were too difficult or too easy for participants with tetraplegia, which resulted in them disengaging from therapy.

> *I also had a board with different sizes of bolts and nuts that I had to screw and I hated that because I thought, you know, 'I can't do this! Why are they making me do this when I cannot do it? This is so humiliating'.* [PwT16, FG 4]

Therapist discussions described patients were more motivated if they understood their injury, prognosis, and purpose of therapy ("understanding tasks and prognosis").

> *[Patients need] to have a good understanding of their injury and what the task is working towards. If they don't then often they're: "What's the point?" or "I don't really know why I'm just rolling this putty out". Umm "I'm not gonna get any better" or "How's this gonna strengthen?". So how much they understand about their injury is really important in their engagement.* [TP5, FG 1]

Both groups also raised that "peer support and competition" could be motivational.

> *Some of our patients…. a kind of element of challenge or competition almost with themselves in some of the more repetitive remedial tasks, um, can be really beneficial—like if they get 10 pegs one day - the next time they want them to get 15.* [TP4, FG 1]

> *I think being in the group as well gave you a bit more, you know, just a bit more reason to sort of get on with it and people would be chatting and you had that, you know, nice social aspect as well.* [PwT27, FG3]

Demotivational factors included repetitive or infantilizing tasks, symptom burden, low mood, and comparisons with faster-progressing peers.

> *Other people were whizzing through, and you just sat there doing like one at a time. And it was just demoralizing.* [PwT21 FG 4]

> *[…] a lot of the time I was just knackered, it turned out that I had an underlying infection due to the metalwork in the back of my neck.* [PwT1, FG1]

> *A lot of pain can be difficult for them to kind of adhere to upper limb therapy (or any therapy).* [TP4, FG 1]

> *I mean, some people in the unit—when I was there for 12 months didn't do anything. And they just lay in bed. They gave up.* [PwT4, FG 2]

Disengagement could occur if therapists could not always be present during therapy ("limited resources and lack of supervision").

> *[…] you know people who never do what's asked of them unless someone's standing over them making them do it.* [TP12, FG 3]

> *You know, they did their best to keep things fresh and, and try introducing, but obviously they had a limited amount of resources available to them, but equally limited staff time.* [PwT13, FG 3]

Finally, participants with tetraplegia identified “downtime,” especially at the weekend, which meant there was nothing to do, and, as such, was demotivational.

> *The weekends at the spinal unit are, are very long and you just you know everything finishes on a Friday and you do nothing till Monday and you just feel I could I could be doing [something].* [PwT27, FG 3]

Major Theme 4: Attitudes Toward the Use of VR in Upper Limb Rehabilitation

A total of 8 subthemes captured attitudes toward VR (Table 5). Participants described VR as a potentially engaging and meaningful adjunct to therapy, particularly for individuals who struggle to engage with conventional rehabilitation.

Table 5. Summary of subthemes for the major theme “attitudes toward the use of virtual reality in upper limb rehabilitation.” Subthemes are categorized as having a positive or negative sentiment.

Subtheme	Positive or negative sentiment	Therapist codes, n	Participant codes, n	Total, n
VR[a] could/should be engaging and meaningful	Positive	14	31	45
VR could/should be used for rehabilitation	Positive	9	18	27
Younger patients will like VR	Positive	9	0	9
VR could alleviate symptoms	Positive	5	4	9
Consumer equipment is not suitable	Negative	29	15	44
VR could cause adverse effects/be inappropriate	Negative	17	33	50
Resources are limited	Negative	15	0	15
Training is required	Negative	7	0	7

[a] VR: virtual reality.

> *So I think VR would have a very big use in a spinal injury centre there’s a whole range of things it could cover. It’s exciting it would be fun to use with people.* [TP8, FG 3]
>
> *I think [VR] is great because as I said earlier, there are other people [in hospital] who, you know, their mindset is totally negative. […] they need that stimulation, that encouragement, and VR might be the way.* [PwT4, FG 2]
>
> *[VR could] make any training or rehab more, more interesting, more exciting, more varied. Yes, I think… huge potential.* [PwT13, FG 3]

Both groups said that VR could encourage engagement in therapy by reducing symptoms such as pain and improve mood (“VR could alleviate symptoms”).

> *And [VR is] good, actually help people to get involved in activities without even imagining the pain they are in.* [PwT19, FG 3]
>
> *I do quite like that element of escapism for them, I think. It could psychologically really help them as well.* [TP3 FG 3]

Conversely, concerns were raised from both groups regarding equipment suitability, adverse effects, staffing capacity, and training requirements. Head-mounted displays were described as potentially heavy and unsafe for individuals with high-level injuries.

> *If [devices] are too labour intensive or take a while to set up or need someone there then when we’re short-staffed as we have unfortunately been for quite a long time, um, you tend to go for things that are quicker and easier to use if you’ve only got quite a short treatment session.* [TH9, FG2]
>
> *The other thing as well is our very high injury levels, [who are] struggling just holding their head up—or that kind of movement with the head [TP8 demonstrates neck rotation] a lot—it can tire them.* [TP8, FG 3]

Both groups warned that “VR could cause adverse effects” such as nausea, or that the specific activities in VR that could remind them of their SCI.

> *I think when you’re already feeling self-conscious about yourself as an individual and how you move… VR might have a bit of a negative effect on people as well.* [TP5, FG 1]
>
> *When you take [the VR headset] off and you go ‘Oh shit. It’s true, this is where I am’.* [PwT21, FG 4]

Therapists speculated that their younger patients would be more suited to VR games (“younger patients will like VR”).

> *So I think the younger people will probably get [VR], you know they’ll understand a lot better how it works.* [TP4 FG 1]

Major Theme 5: User-Informed Design Considerations for VR Rehabilitation Systems

A total of 5 subthemes described participant recommendations for VR system design (Table 6). Suggested activities ranged from functional or vocational tasks, suggested mostly by therapists, to leisure, sporting, and creative environments, which were suggested predominantly by participants with tetraplegia.

XSL•FO
RenderX

Preferences varied widely, reflecting differences in interests and functional abilities.

Table 6. Summary of subthemes for the major theme “user-informed design considerations for virtual reality rehabilitation systems.”

Subtheme	Therapist codes, n	Participant codes, n	Total, n
Activities and scenarios	86	150	236
Usability and training	32	33	65
Measuring performance and achievement	38	27	65
Movements	21	23	44
Immersion, presence, and engagement	11	17	28

> *[…] it's down to the individual what interests they have and you know what would benefit them the most and… I mean, metalwork is obviously not gonna be the same as scuba diving and, you know, skiing, or playing the piano. One’s having loads of fun, one’s more sort of creative and you're constructing something in VR. Yeah. So yeah, just depends.* [PwT4, FG 2]

Therapists also discussed providing a range of different activities for their patients. Both groups recommended incorporating concepts such as relaxation and escapism into the design (“immersion, presence, and engagement”).

> *So we are really trying to provide a variety of different type of games that we offer but obviously when you’ve got a patient who’s had a long (and all of our spinal injury patients have quite a long) admission with us, it’s like the variety that’s key.* [TP14 FG 3]

> *What about making really realistic sound, […] soaring noises, painting noises of the paint going on to the wall and those kind of things where it actually feeds back extra reality to the experience?* [PwT16, FG 4]

Additionally, participants with tetraplegia and therapists suggested outdoor spaces such as hills, glens, mountains, underwater spaces, outer space, and jungles would be good environments for VR activities.

> *[I have] a static bicycle at home. I get a bit bored because I’m always in the same place, while if you could have the feeling that you're actually outside and going places that might make it more interesting.* [PwT16, FG 4]

Both groups said that VR games should enable a range of movements, from fine to gross (“movements”), including reaching, grasping, and wrist flexion and extension.

> *[VR] could be something that could be really useful for larger gross movements but also for really drilling down to repetition of very fine motor movements.* [TP12, FG 3]

Therapists also recommended the quality of movements in VR was of high importance.

> *Sometimes it’s not the number of movements, it’s quality as well. So, you know, we’d be looking at the quality of someone’s reach, not just that they are able to reach, but maybe not recruiting the right muscle groups to perform that task.* [TP9, FG 2]

Both groups gave recommendations for “measuring performance and achievement” of users. These included giving encouragement and providing feedback to increase engagement. Therapists discussed providing incremental levels of difficulty to accommodate progression and achievement.

> *[…] like you get a big tick or a big thumbs up or something. Not something annoying but something that shows you've reached the level of the activity that's actually doing something.* [PwT27, FG 4]

Therapists recommended incorporating easily comprehensible progress reports and feedback to users to help with motivation.

> *[…] and patients love seeing kind of that—we can put it—display it on a graph for them to kind of say like this is what you did like a month ago and this is what you’re doing now.* [TP8, FG 3]

Therapists and participants with tetraplegia recommended that the movements should be configurable for users, and that training should be provided to patients and therapists (“usability and training”).

> *So there’s gonna have to be sort of that adjustability with the system to make it relevant to that patient’s rehab and what muscles you want to work on and that, yeah.* [TP7, FG 2]

Furthermore, under this subtheme, it was highlighted that adding straps to the controllers might be required, and for those with limited hand function, participants with tetraplegia also requested interaction mechanisms using head-movement, gaze, hand tracking, and voice input.

> *I think you'd have to be really aware of even things like how people could grip any hand controls, you know, just the simple things like holding them, putting headsets on, how much assistance was needed to make it as practical as possible.* [PwT21, FG 4]

Ethical Considerations

Ethical approval for this study was granted by the Psychology, Social Work, and Allied Health Sciences (PSWAHS) Research Ethics Committee of Glasgow Caledonian University (reference: HLS/PSWAHS/21/247). Participants with tetraplegia were recruited via a gatekeeper (Spinal Injuries Scotland). Therapist participants were recruited via social media. Both groups were provided with participant information sheets that detailed the

XSL•FO
RenderX

aims of the study and, if they chose to take part, the requirements of participation and the information collected about them. This information would be stored on Glasgow Caledonian University's OneDrive for 5 years and shared only with the study team. All information would be anonymized after consent was given, and participants assigned an ID number to comply with the General Data Protection Regulation (GDPR). Prior to participation, all participants provided informed consent either by printing and physically signing a digital consent form or by completing the form digitally. Participants who were unable to physically sign the consent form provided consent through a designated witness who signed on their behalf [35]. The study abided by the guidelines set out in the Declaration of Helsinki. Participants were not compensated for their participation and were informed they were free to withdraw from the study at any time without giving a reason. Participants could have their information removed from the study but were informed that this would not be possible after anonymization.

Discussion

Principal Results

The aim of this qualitative study was to explore the lived experiences of people with tetraplegia and SCI specialist therapists in terms of upper limb rehabilitation to inform the design of a VR-based intervention. A total of 5 major themes and their associated subthemes were identified that enabled the specification of software requirements, which have subsequently been implemented in a prototype VR intervention. The focus groups indicated that VR has the potential to help improve function and build strength, as well as learning skills, by adapting to loss of function. This suggests that end users of potential VR-based interventions understand its possible benefits, as well as pitfalls. The results also show that end users see a place for co-designed interventions in the acute/subacute setting and showed interest in VR, which is critical for "buy-in" for the adoption of new technologies.

Major Themes 1 and 2

The findings indicate that VR-based interventions must accommodate substantial variability in upper limb impairment following SCI. To account for diverse user profiles and needs, software must adapt to the personal rehabilitation needs of the user. Although some therapists anticipated uptake among younger patients, and therefore specific targeting toward this demographic, restricting designs to this group would misalign with the clinical diversity of the population. Design should therefore prioritize broad accessibility and usability across impairment, age groups, and levels of technological expertise.

Major Theme 3

This study highlights how motivation and engagement are shaped by both individual and service-level factors during acute/subacute rehabilitation. The findings align with established barriers to engagement in therapy in the SCI literature, including comorbidities (eg, urinary tract infections, pressure sores, fatigue, pain, and low mood), alongside the resource limitations of health care providers, periods of "downtime," variation in individual rehabilitation needs, and patients' desire for more control over their therapy [41-44]. Patients with SCI were generally described by both groups as highly motivated, with engagement closely linked to perceived progress and the pursuit of greater independence. This finding may be influenced by sampling bias, as individuals willing to participate in research may be more motivated than the broader inpatient population. Motivational drivers suggest that interventions incorporating visible progress indicators and structured challenges may be particularly effective. Such features could also increase acceptability of novel interventions, including VR therapy, when used as an adjunct to conventional rehabilitation.

Major Theme 4

Participants with tetraplegia described VR as an assistive technology to enhance function and therefore improve independence, a factor shown to be of high importance to people with SCI [45]. In addition, VR was perceived as a tool to give participants with tetraplegia independence during their rehabilitation. Participants' emphasis on autonomy and meaningful activity suggests that VR may be most effective when it enables users to exercise choice and engage in personally relevant tasks. These factors could also enable therapists in administering or prescribing therapy by facilitating movements that are typically difficult to achieve in usual treatment or that are more appealing to people with tetraplegia.

The repetitive nature of upper limb rehabilitation presents an opportunity for VR to embed prescribed movements within more engaging tasks. VR could maintain motivation by using repetitive input to drive engaging game mechanics, such as firing a bow and arrow, hitting boxing targets, or completing puzzles. Progress-based feedback and structured challenges may further promote sustained participation.

Users could be engaged in gameplay and rewarded in immersive scenarios for repeating movements. This concept relies on the immersive capabilities of VR to provide positive distraction, which is a term inferred from participants' discussion regarding "escapism," from typical therapy while still enabling prescribed movement.

The acute context of rehabilitation is relevant, as patients must adapt to both sudden functional loss and prolonged hospitalization. Participants highlighted the importance of experiences that offer psychological relief from this environment via relaxation and escapism, which are achievable with immersive VR systems [46].

Resource limitations were identified as a barrier to therapy delivery by both groups. Participants implied that a VR-based system could mitigate these limitations by enabling self-guided, semi-independent practice. This could allow therapists to allocate time more efficiently while maintaining a supervisory role in therapy sessions.

The group-based nature of rehabilitation suggests that future VR systems could benefit from supporting multiuser networked environments to encourage peer-to-peer interactions.

Major Theme 5

Both therapists and participants with tetraplegia made many suggestions about what type of activities and upper limb

XSL•FO
RenderX

movements they would like to see implemented in VR. There was no consensus on particular types of VR scenarios with which patients with tetraplegia would engage, reflecting the heterogeneity of the tetraplegic population. This supports the development of flexible systems capable of offering varied activities and environments to their users.

VR is most likely to add value when repetitive therapeutic movements are embedded within adaptive, gamified tasks. These findings are consistent with existing research on rehabilitation during acute/subacute tetraplegia [45,47-50]. Participants' emphasis on independence, motivation, and shared decision-making reinforces the important of human-centered design principles. Unique to this population is the complexity of SCI, which remains a central factor for consideration during intervention development.

These findings add to the growing body of research concerning the recommendations for implementing VR for people with upper limb impairments and their rehabilitation [51-54]. VR interventions must be designed to take advantage of their attributes, which include real-time feedback and immersive environments. These designs must be informed by the specific populations who are intended to use these systems [55-58].

Recommendations for the Design of a VR Upper Limb Rehabilitation Intervention for People With Tetraplegia

The following recommendations outline key design considerations for VR-based upper limb rehabilitation interventions targeting people with tetraplegia. They are structured around practical software development domains to support systematic intervention design and informed by the themes identified in the focus group data.

Domain 1: Accessibility

VR applications should be usable by the widest possible range of people with tetraplegia and accommodate the wide variation in upper limb function of users for control of the system. They should be usable to the same extent for all users, irrespective of upper limb function, and support patient autonomy while maintaining the therapeutic role of clinicians. VR applications should be usable across age groups and comorbidities associated with tetraplegia, allow all prescribed movements to be specified and configured by the user, and must be comfortable to use. They should target as many movement domains as possible and support movement or activity that is difficult to achieve in standard upper limb treatment.

Domain 2: Implementation

VR applications should be feasible, fit into existing clinical workflows, and aim to alleviate resource limitations. They should support collaborative goal setting between patients and therapists and minimize physical and psychological adverse effects associated with, and attributable to, VR. Applications should consider the appropriate intensity and duration of an intervention and prevent overexertion during use. They should also be able to monitor usage and performance and be supported by appropriate training for end users to ensure safe implementation and effective integration into clinical services.

Domain 3: Immersion and Engagement

VR applications should increase and maintain engagement beyond what is achieved in usual treatment by encouraging repetitions of prescribed movements through meaningful virtual activities and environments. They should avoid factors that cause disengagement from therapy by providing an optimal degree of challenge for each user, ensuring that the experience is neither frustrating nor too difficult or too easy. Applications should provide motivational feedback on performance, such as scores and progress, and make use of the immersive and enjoyable qualities of VR technology to ensure users feel present within virtual environments and benefit from escapism or positive distraction.

Limitations

Participants with tetraplegia sometimes had difficulty recalling their time as inpatients, typically when years or decades had passed following discharge. However, it was important to prioritize the perspective of individuals with a full retrospective view of inpatient care.

Separate focus groups were conducted for therapists and participants with tetraplegia to remove the effect of the relationships or perceived power differences between patients and clinicians. For example, when discussing opinions about particular types of upper limb therapy, participants with tetraplegia may have been reluctant to express negative opinions about care if therapists were present. Interaction between participants with tetraplegia and therapists may, however, have been beneficial in terms of ideating suggestions for the design of VR scenarios.

Data saturation was not achieved as the final focus groups in both groups yielded additional insights, though at a lower frequency compared to earlier groups. However, the information power of the focus groups was improved by revising the question schedules after each focus group. In particular, follow-up questions and prompts were added to explore new areas of discussion that had not been covered by previous groups. The inclusion of different rehabilitation disciplines with experience of working in different specialist centers across the United Kingdom ensured diverse groups were assembled, which was a strength of this study and allowed for more varied discussion.

Conclusions

This study reported the perspectives and suggestions of people with tetraplegia and SCI therapists about the use of VR for upper limb rehabilitation in the acute setting. As a result, it presents valuable first-hand information about the design of VR technology for SCI rehabilitation. A total of 5 themes were generated that designers and developers of immersive applications for rehabilitation can use to create their applications. People with tetraplegia have a large variation in upper limb impairments, for which many different movements are prescribed by therapists. This variation must be accommodated by applications in a usable and accessible way. Applications must also implement features that motivate patients to attend and adhere to therapy, such as providing useful feedback and appealing to the broad range of patients' interests, and also

XSL•FO
RenderX

avoid factors that can cause disengagement from therapy. The findings of this study facilitated the cocreation of a novel immersive VR intervention for upper limb rehabilitation for people with tetraplegia while they are inpatients.

Future Work

The next part of this project is to use the results from this study to produce software requirements and a software architecture for a suite of VR games for upper limb therapy. This study enabled the development and deployment of the proposed software and its continued improvement.

Acknowledgments

The authors would like to thank all the participants (co-designers) who took part in this study, Spinal Injuries Scotland for helping recruit participants, and Stoke Mandeville Spinal Research for funding this PhD through their Assistive Technology for Upper Limb Rehabilitation PhD Studentship.

Authors' Contributions

The “Co-designers” members are Ray Brown, Lauren Cope, Graham Cullen, Sheena Egan, Amanda Howat, Donald Hutton, John Lynch, Samantha Maguire, and Sylvia Merino. AG, LP, M Poyade, and M Purcell designed and conceptualized this study. AG and LP facilitated the focus groups. AG and LC were responsible for project administration, including recruitment of focus group participants. AG, LP, M Poyade, and M Purcell performed the analysis. AG and LP wrote and prepared the original draft of the manuscript. AG, LP, M Purcell, M Poyade, and LC reviewed and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Schedule of Questions (PwT).
[DOCX File , 23 KB - rehab_v13i1e66832_app1.docx]

Multimedia Appendix 2
Schedule of Questions (Therapists).
[DOCX File , 25 KB - rehab_v13i1e66832_app2.docx]

References

1. Hutson TH, Di Giovanni S. The translational landscape in spinal cord injury: focus on neuroplasticity and regeneration. Nat Rev Neurol 2019;15(12):732-745. [doi: 10.1038/s41582-019-0280-3] [Medline: 31728042]
2. Thompson C, Mutch J, Parent S, Mac-Thiong J. The changing demographics of traumatic spinal cord injury: an 11-year study of 831 patients. J Spinal Cord Med 2015;38(2):214-223 [FREE Full text] [doi: 10.1179/2045772314Y.0000000233] [Medline: 25096709]
3. New spinal cord injury data revealed. Spinal Injuries Association. URL: https://www.spinal.co.uk/news/new-spinal-cord-injury-data-revealed/ [accessed 2026-01-27]
4. McDaid D, Park A, Gall A, Purcell M, Bacon M. Understanding and modelling the economic impact of spinal cord injuries in the United Kingdom. Spinal Cord 2019;57(9):778-788 [FREE Full text] [doi: 10.1038/s41393-019-0285-1] [Medline: 31086273]
5. Ditunno JF, Cohen ME, Hauck WW, Jackson AB, Sipski ML. Recovery of upper-extremity strength in complete and incomplete tetraplegia: a multicenter study. Arch Phys Med Rehabil 2000;81(4):389-393. [doi: 10.1053/mr.2000.3779] [Medline: 10768525]
6. Mateo S, Di Marco J, Cucherat M, Gueyffier F, Rode G. Inconclusive efficacy of intervention on upper-limb function after tetraplegia: a systematic review and meta-analysis. Ann Phys Rehabil Med 2020;63(3):230-240 [FREE Full text] [doi: 10.1016/j.rehab.2019.05.008] [Medline: 31233828]
7. Rupp R, Biering-Sørensen F, Burns S, Graves D, Guest J, Jones L. International standards for neurological classification of spinal cord injury: revised 2019. Top Spinal Cord Inj Rehabil 2021;27(2):1-22 [FREE Full text] [doi: 10.46292/sci2702-1] [Medline: 34108832]
8. Bickenbach J, Officer A, Shakespeare T, von Groote P, International Spinal Cord Society, World Health Organization. International Perspectives on Spinal Cord Injury. Geneva: World Health Organization; 2013.
9. Adriaansen J, Ruijs L, van Koppenhagen CF, van Asbeck FWA, Snoek G, van Kuppevelt D. Secondary health conditions and quality of life in persons living with spinal cord injury for at least ten years. J Rehabil Med 2016;48(10):853-860 [FREE Full text] [doi: 10.2340/16501977-2166] [Medline: 27834436]

XSL•FO
RenderX

10. Burns AS, Marino RJ, Kalsi-Ryan S, Middleton JW, Tetreault LA, Dettori JR. Type and timing of rehabilitation following acute and subacute spinal cord injury: a systematic review. Global Spine J 2017;7(3 Suppl):175S-194S [FREE Full text] [doi: 10.1177/2192568217703084] [Medline: 29164023]
11. Lynskey JV, Belanger A, Jung R. Activity-dependent plasticity in spinal cord injury. J Rehabil Res Dev 2008;45(2):229-240 [FREE Full text] [doi: 10.1682/jrrd.2007.03.0047] [Medline: 18566941]
12. Musselman KE, Shah M, Zariffa J. Rehabilitation technologies and interventions for individuals with spinal cord injury: translational potential of current trends. J Neuroeng Rehabil 2018;15(1):40 [FREE Full text] [doi: 10.1186/s12984-018-0386-7] [Medline: 29769082]
13. Anderson KD. Targeting recovery: priorities of the spinal cord-injured population. J Neurotrauma 2004;21(10):1371-1383. [doi: 10.1089/neu.2004.21.1371] [Medline: 15672628]
14. Roy RR, Harkema SJ, Edgerton VR. Basic concepts of activity-based interventions for improved recovery of motor function after spinal cord injury. Arch Phys Med Rehabil 2012;93(9):1487-1497. [doi: 10.1016/j.apmr.2012.04.034] [Medline: 22920448]
15. Savic G, Frankel HL, Jamous MA, Soni BM, Charlifue S. Participation restriction and assistance needs in people with spinal cord injuries of more than 40 year duration. Spinal Cord Ser Cases 2018;4:28 [FREE Full text] [doi: 10.1038/s41394-018-0056-9] [Medline: 29619249]
16. Lu X, Battistuzzo CR, Zoghi M, Galea MP. Effects of training on upper limb function after cervical spinal cord injury: a systematic review. Clin Rehabil 2015;29(1):3-13. [doi: 10.1177/0269215514536411] [Medline: 25575932]
17. Kramer JLK, Lammertse DP, Schubert M, Curt A, Steeves JD. Relationship between motor recovery and independence after sensorimotor-complete cervical spinal cord injury. Neurorehabil Neural Repair 2012;26(9):1064-1071 [FREE Full text] [doi: 10.1177/1545968312447306] [Medline: 22647878]
18. Kalsi-Ryan S, Beaton D, Curt A, Popovic MR, Verrier MC, Fehlings MG. Outcome of the upper limb in cervical spinal cord injury: profiles of recovery and insights for clinical studies. J Spinal Cord Med 2014;37(5):503-510. [doi: 10.1179/2045772314Y.0000000252] [Medline: 25229734]
19. Quel de Oliveira C, Refshauge K, Middleton J, de Jong L, Davis GM. Effects of activity-based therapy interventions on mobility, independence, and quality of life for people with spinal cord injuries: a systematic review and meta-analysis. J Neurotrauma 2017;34(9):1726-1743. [doi: 10.1089/neu.2016.4558] [Medline: 27809702]
20. Huh S, Ko H. Recovery target priorities of people with spinal cord injuries in Korea compared with other countries: a survey. Spinal Cord 2020;58(9):998-1003. [doi: 10.1038/s41393-020-0457-z] [Medline: 32246088]
21. Simpson LA, Eng JJ, Hsieh JT, Wolfe DL, Spinal Cord Injury Rehabilitation Evidence Scire Research Team. The health and life priorities of individuals with spinal cord injury: a systematic review. J Neurotrauma 2012;29(8):1548-1555 [FREE Full text] [doi: 10.1089/neu.2011.2226] [Medline: 22320160]
22. Namrata Y. Priorities of spinal cord injured population – a survey. Am J Appl Psychol 2017;6(6):183. [doi: 10.11648/j.ajap.20170606.17]
23. Behrman A, Ardolino E, Harkema S. Activity-based therapy: from basic science to clinical application for recovery after spinal cord injury. J Neurol Phys Ther 2017;41 Suppl 3(Suppl 3 IV STEP Spec Iss):S39-S45. [doi: 10.1097/NPT.0000000000000184] [Medline: 28628595]
24. Kazim SF, Bowers CA, Cole CD, Varela S, Karimov Z, Martinez E. Corticospinal motor circuit plasticity after spinal cord injury: harnessing neuroplasticity to improve functional outcomes. Mol Neurobiol 2021;58(11):5494-5516. [doi: 10.1007/s12035-021-02484-w] [Medline: 34341881]
25. Scalise M, Bora TS, Zancanella C, Safa A, Stefini R, Cannizzaro D. Virtual reality as a therapeutic tool in spinal cord injury rehabilitation: a comprehensive evaluation and systematic review. J Clin Med 2024;13(18):5429 [FREE Full text] [doi: 10.3390/jcm13185429] [Medline: 39336916]
26. Vinciguerra C, Federico A. Virtual rehabilitation in spinal cord injury patients: realities and future perspectives: a systematic review. Curr Phys Med Rehabil Rep 2024;12(4):425-434. [doi: 10.1007/s40141-024-00470-6]
27. Schiza E, Matsangidou M, Neokleous K, Pattichis CS. Virtual reality applications for neurological disease: a review. Front Robot AI 2019;6:100 [FREE Full text] [doi: 10.3389/frobt.2019.00100] [Medline: 33501115]
28. Leemhuis E, Esposito RM, De Gennaro L, Pazzaglia M. Go virtual to get real: virtual reality as a resource for spinal cord treatment. Int J Environ Res Public Health 2021;18(4):1819 [FREE Full text] [doi: 10.3390/ijerph18041819] [Medline: 33668438]
29. De Miguel-Rubio A, Rubio MD, Alba-Rueda A, Salazar A, Moral-Munoz JA, Lucena-Anton D. Virtual reality systems for upper limb motor function recovery in patients with spinal cord injury: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020;8(12):e22537. [doi: 10.2196/22537] [Medline: 33270040]
30. Orsatti Sánchez BA, Diaz Hernandez O. Efficacy of virtual reality in neurorehabilitation of spinal cord injury patients: a systematic review. Rev Mex Ing Bioméd 2021;42(2):90-103 [FREE Full text] [doi: 10.17488/RMIB.42.2.8]
31. de Araújo AVL, Neiva JFDO, Monteiro CBDM, Magalhães FH. Efficacy of virtual reality rehabilitation after spinal cord injury: a systematic review. Biomed Res Int 2019;2019:7106951 [FREE Full text] [doi: 10.1155/2019/7106951] [Medline: 31828120]

XSL•FO
RenderX

32. Bird M, McGillion M, Chambers EM, Dix J, Fajardo CJ, Gilmour M. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021;7(1):12 [FREE Full text] [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]
33. Bourke JA, Bragge P, River J, Sinnott Jerram KA, Arora M, Middleton JW. Shining a light on the road towards conducting principle-based co-production research in rehabilitation. Front Rehabil Sci 2024;5:1386746 [FREE Full text] [doi: 10.3389/fresc.2024.1386746] [Medline: 38660394]
34. Bate P, Robert G. Experience-based design: from redesigning the system around the patient to co-designing services with the patient. Qual Saf Health Care 2006;15(5):307-310 [FREE Full text] [doi: 10.1136/qshc.2005.016527] [Medline: 17074863]
35. The Medicines and Healthcare products Regulatory Agency, Health Research Authority. Joint statement on seeking consent by electronic methods. UK Government & HRA/MHRA. 2018. URL: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/hra-mhra-econsent-statement-sept-18.pdf [accessed 2026-01-27]
36. Fekete C, Eriks-Hoogland I, Baumberger M, Catz A, Itzkovich M, Lüthi H. Development and validation of a self-report version of the spinal cord independence measure (SCIM III). Spinal Cord 2013;51(1):40-47. [doi: 10.1038/sc.2012.87] [Medline: 22890418]
37. SteamVR featuring the HTC Vive. Valve Corporation. 2016. URL: https://youtu.be/qYfNzhLXYGc [accessed 2026-01-27]
38. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
39. NVivo qualitative data analysis software. QSR International Pty Ltd. 2018. URL: https://lumivero.com/products/nvivo/ [accessed 2026-01-27]
40. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
41. Vissers M, van den Berg-Emons R, Sluis T, Bergen M, Stam H, Bussmann H. Barriers to and facilitators of everyday physical activity in persons with a spinal cord injury after discharge from the rehabilitation centre. J Rehabil Med 2008;40(6):461-467 [FREE Full text] [doi: 10.2340/16501977-0191] [Medline: 18509562]
42. Tallqvist S, Eskola K, Täckman A, Kauppila A, Koskinen E, Anttila H. Facilitators and barriers in the rehabilitation process described by persons with spinal cord injury: a deductive-inductive analysis from the Finnish spinal cord injury study. Ann Med 2023;55(2):2303398 [FREE Full text] [doi: 10.1080/07853890.2024.2303398] [Medline: 38232950]
43. Whalley Hammell K. Experience of rehabilitation following spinal cord injury: a meta-synthesis of qualitative findings. Spinal Cord 2007;45(4):260-274. [doi: 10.1038/sj.sc.3102034] [Medline: 17310257]
44. Lindberg J, Kreuter M, Taft C, Person L. Patient participation in care and rehabilitation from the perspective of patients with spinal cord injury. Spinal Cord 2013;51(11):834-837. [doi: 10.1038/sc.2013.97] [Medline: 23999110]
45. Sousa SS, Andrade MJ, Fernandes CS, Barbeiro SR, Teixeira VT, Pereira RS. Healthcare experience of people with acute spinal cord injury: a phenomenological study. Nurs Rep 2023;13(4):1671-1683 [FREE Full text] [doi: 10.3390/nursrep13040138] [Medline: 38133114]
46. Williamson SD, Aaby AO, Ravn SL. Psychological outcomes of extended reality interventions in spinal cord injury rehabilitation: a systematic scoping review. Spinal Cord 2025;63(2):58-65. [doi: 10.1038/s41393-024-01057-7] [Medline: 39789357]
47. Kirchberger I, Sinnott A, Charlifue S, Kovindha A, Lüthi H, Campbell R. Functioning and disability in spinal cord injury from the consumer perspective: an international qualitative study using focus groups and the ICF. Spinal Cord 2010;48(8):603-613. [doi: 10.1038/sc.2009.184] [Medline: 20065983]
48. Krysa JA, Gregorio MP, Pohar Manhas K, MacIsaac R, Papathanassoglou E, Ho CH. Empowerment, communication, and navigating care: the experience of persons with spinal cord injury from acute hospitalization to inpatient rehabilitation. Front Rehabil Sci 2022;3:904716 [FREE Full text] [doi: 10.3389/fresc.2022.904716] [Medline: 36188987]
49. Hearn JH, Finlay KA, Fine PA, Cotter I. Neuropathic pain in a rehabilitation setting after spinal cord injury: an interpretative phenomenological analysis of inpatients' experiences. Spinal Cord Ser Cases 2017;3:17083 [FREE Full text] [doi: 10.1038/s41394-017-0032-9] [Medline: 29423289]
50. Kerstin W, Gabriele B, Richard L. What promotes physical activity after spinal cord injury? An interview study from a patient perspective. Disabil Rehabil 2006;28(8):481-488. [doi: 10.1080/09638280500211932] [Medline: 16513581]
51. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018;10(11):1237-1251.e1 [FREE Full text] [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
52. Herrera V, Albusac J, Schez-Sobrino S, Reyes-Guzmán A, Glez-Morcillo C. Dynamic adjustment of interactive objects in virtual environments for upper limb rehabilitation: a patient-centred solution. : The Eurographics Association; 2024 Presented at: Proceedings of the Spanish Computer Graphics Conference (CEIG); June 17-19, 2024; A Coruña. [doi: 10.1007/s10055-024-01078-w]
53. Herrera V, Albusac J, Castro-Schez JJ, González-Morcillo C, Monekosso DN, Pacheco S. Creating adapted environments: enhancing accessibility in virtual reality for upper limb rehabilitation through automated element adjustment. Virtual Reality 2025;29(1). [doi: 10.1007/s10055-024-01078-w]

XSL•FO
RenderX

54. Lewis GN, Rosie JA. Virtual reality games for movement rehabilitation in neurological conditions: how do we meet the needs and expectations of the users? Disabil Rehabil 2012;34(22):1880-1886. [doi: 10.3109/09638288.2012.670036] [Medline: 22480353]
55. Creed C, Al-Kalbani M, Theil A, Sarcar S, Williams I. Inclusive AR/VR: accessibility barriers for immersive technologies. Univ Access Inf Soc 2023;23(1):59-73. [doi: 10.1007/s10209-023-00969-0]
56. Mott M, Cutrell E, Gonzalez Franco M, Holz C, Ofek E, Stoakley R. Accessible by design: an opportunity for virtual reality. Proc IEEE Int Symp Mix Augment Real Adjunct 2019:451-454. [doi: 10.1109/ismar-adjunct.2019.00122]
57. Dudley J, Yin L, Garaj V, Kristensson PO. Inclusive immersion: a review of efforts to improve accessibility in virtual reality, augmented reality and the metaverse. Virtual Reality 2023;27(4):2989-3020. [doi: 10.1007/s10055-023-00850-8]
58. Vlahovic S, Suznjevic M, Skorin-Kapov L. A framework for the classification and evaluation of game mechanics for virtual reality games. Electronics 2022;11(18):2946. [doi: 10.3390/electronics11182946]

Abbreviations

ABT: Activity-Based Therapy
GDPR: General Data Protection Regulation
PSWAHS: Psychology, Social Work, and Allied Health Sciences
SCI: spinal cord injury
VR: virtual reality

Edited by A Scano; submitted 08.Apr.2025; peer-reviewed by MM Ahmed Al Nattah, J Thrift, H Heppner; comments to author 01.Oct.2025; revised version received 17.Dec.2025; accepted 17.Jan.2026; published 11.Feb.2026.

Please cite as:
Goodsell A, Purcell M, Poyade M, Cownie L, Co-designers, Paul L
Design Recommendations for Virtual Reality–Based Upper Limb Exercises From People With Tetraplegia and Spinal Cord Injury Rehabilitation Specialists: Focus Group Study
JMIR Rehabil Assist Technol 2026;13:e66832
URL: https://rehab.jmir.org/2026/1/e66832
doi:10.2196/66832
PMID:

©Andrew Goodsell, Mariel Purcell, Matthieu Poyade, Louise Cownie, Co-designers, Lorna Paul. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telerehabilitation Following Stroke: Development of Training Content and Evaluation of an App-Based Training Program

Carina Ziller[1,2], MSc; Szabina Gäumann[1], MSc; Silya Lüscher[1], MSc; Nele Paulissen[1,2], MSc; Frank Behrendt[1,3], PhD; Zorica Suica[1,2], MSc; Björn Crüts[4], MSc; Luana Gamerschlag[4], BA; Katrin Parmar[1,5], PhD; Hans Ulrich Gerth[1,6], PhD; Leo H Bonati[1,7,8], PhD; Corina Schuster-Amft[1,3,9], PhD

[1]Research Department, Reha Rheinfelden, Salinenstrasse 98, Rheinfelden, Switzerland
[2]Physiotherapy Department, Reha Rheinfelden, Rheinfelden, Switzerland
[3]School of Engineering and Computer Science, Bern University of Applied Sciences, Biel, Switzerland
[4]Blended Clinic AI GmbH, Nuremberg, Germany
[5]Translational Imaging in Neurology (ThINk) Basel, Departments of Head, Spine and Neuromedicine and Biomedical Engineering, University Hospital, University of Basel, Basel, Switzerland
[6]Department of Medicine, University Hospital Münster, Münster, Germany
[7]Stroke Center and Department of Neurology, University Hospital Basel, Basel, Switzerland
[8]Department of Clinical Research, University of Basel, Basel, Switzerland
[9]Department for Sport, Exercise and Health, University of Basel, Basel, Switzerland

Corresponding Author:
Frank Behrendt, PhD
Research Department, Reha Rheinfelden, Salinenstrasse 98, Rheinfelden, Switzerland

Abstract

Background: Enhancing rehabilitation methods for patients with stroke is essential, particularly during the transition from inpatient to outpatient care. Digital applications are being developed to provide telerehabilitation programs. The existing virtual blended care platform Blended Clinic (Blended Clinic AI GmbH) offers app-based training for patients after a stroke and comprises 3 main components, including training, coaching, and monitoring.

Objective: This study assesses the usability and user experience of the novel Swiss Tele-Assisted Rehabilitation and Training (START) program within the Blended Clinic platform in patients after stroke and therapists.

Methods: The START program was developed within 3 workshops and an online survey. It contains 10 6-week exercise programs tailored to the levels of the modified Rankin Scale (mRS), 5 infographics, and 5 podcasts. Eight patients after stroke and 10 therapists took part in a single-center usability study. All participants were introduced to the Blended Clinic app and subsequently used it independently. The Blended Clinic platform, including the START program, was evaluated by both user groups based on the System Usability Scale (SUS) and the Mobile App Rating Scale (MARS-G). Additionally, feedback was collected, observations were documented, and program adherence metrics were calculated.

Results: The mean SUS scores were 87.2 (SD 10.8) for 8 patients and 83.3 (SD 11.3) for 10 therapists. The MARS-G scores were 3.9 (SD 0.5) for patients and 4.1 (SD 0.4) for therapists for categories A-D. User experience was rated 4.1 (SD 0.5) for patients, while device usability was 3.8 (SD 0.8) for patients and 4.2 (SD 0.5) for therapists. Adherence to the training schedule varied among patients (16.7%-80% of the planned sessions) and was rather low for many of the patients.

Conclusions: The START program, delivered via the Blended Clinic platform, was considered user-friendly and received good usability ratings from patients with stroke and therapists. Recommendations to enhance compliance are provided.

(JMIR Rehabil Assist Technol 2026;13:e77090) doi:10.2196/77090

KEYWORDS

telerehabilitation; stroke; app-based training; usability; blended care

Introduction

Stroke is a leading cause of acquired disability in adults worldwide, significantly impacting individuals' quality of life due to persistent motor and cognitive impairments [1]. Importantly, these impairments in patients with stroke are often accompanied by other issues like fatigue and depressive symptoms, significantly impacting quality of life and independence [2]. Stroke is a significant contributor to cognitive decline and dementia, being the second most common cause after Alzheimer disease [2,3].

Rehabilitation strategies for individuals after a stroke primarily focus on restoring impaired motor control to enhance the independence of those affected [4]. An interdisciplinary approach that considers individual needs is essential for the success of rehabilitation and the overall quality of life for patients [5]. Thus, the advancement of options for improved, sustainable, and effective follow-up care in stroke rehabilitation is urgently needed [6]. Technological innovations can play an essential role in efforts to maintain the motor and cognitive abilities of affected individuals at the highest possible level. By integrating new technologies into rehabilitation programs, recovery outcomes can be enhanced, while early intervention and rehabilitation are crucial for improving outcomes and enhancing recovery after a stroke [7].

In addition to face-to-face rehabilitation, telerehabilitation has emerged due to advancing digitalization in the health care sector. It has been found that telerehabilitation in physiotherapy is comparable to in-person rehabilitation for conditions such as stroke [8,9]. In line with that, studies have shown that telerehabilitation can effectively deliver therapeutic interventions, maintain patient engagement, and achieve similar outcomes to traditional therapy, making it a valuable option for patients who may have difficulty attending in-person sessions [10].

Various mobile health apps have been developed for patients after a stroke, mostly focusing on upper extremity function, secondary prevention, or language skills [11]. Additionally, according to a review by Rintala et al [12], most studies investigating these apps focused on patients with chronic stroke, who are mildly impaired [13]. Although existing mobile health apps have shown positive effects on various stroke-related impairments, a substantial gap remains in the development of apps specifically targeting goal-oriented exercises and also integrating educational components [14].

The newly developed Blended Clinic platform (Blended Clinic AI GmbH) consists of a mobile app for patients (Blended Clinic, Version 1.0.1 [iOS]; 1.3 [Android]), a web-based portal for therapists, and a database. The system is designed to provide home-based exercise programs for different patient groups by integrating telerehabilitation with a blended care approach. Blended care itself refers to a health care approach that combines traditional in-person services with digital or remote care options. It aims to provide a more flexible and comprehensive approach to health care delivery, enabling diverse patient populations, including individuals after stroke in the subacute and chronic phases, to receive care through multiple modalities [15]. To date, an earlier iteration of a training program for patients after a stroke, in conjunction with a previous version of the Blended Clinic platform, has been evaluated for user interaction with the app and for acceptability among patients with chronic stroke [16]. It was found to be highly acceptable, with strong user engagement, good task adherence, and small clinical improvements in quality of life [16].

Based on these positive findings, the new Swiss Tele-Assisted Rehabilitation and Training (START) program was developed. It builds on the same core technology but extends the intervention and addresses the research gap in 2 ways: first, by further developing the content with a stronger focus on the needs of patients with moderate-to-severe impairments; and second, by evaluating its use in a new clinical context, the subacute phase of stroke rehabilitation, particularly the transition phase from inpatient to outpatient rehabilitation.

The aim of the present study was to describe the development process of the new START program and to assess the perceptions of both therapists and patients after a stroke regarding the usability and functionality of certain components of the START program, provided via the updated Blended Clinic platform, during test sessions.

Methods

Hardware and Software Components of the Blended Clinic Platform

The Blended Clinic platform consists of a web-based portal for therapists and a mobile app for patients. It is intended to be installed on the mobile devices of patients and comprises components for training, coaching, and monitoring. The content for these components is provided by the respective program modules, such as START, which was developed and tested within this project.

1. Training: in conjunction with a targeted training program, exercises can be performed daily, individually, and to the necessary and feasible extent. Through the app, patients are provided with a personal training plan tailored to their individual impairments and rehabilitation goals, including exercise videos. The exercise content can be adjusted to the achieved functional level.
2. Coaching: the app ensures that individual training is supported by a personal therapist to maintain high motivation levels. Patients’ questions about training and general topics related to stroke can be promptly addressed, both on-site and primarily through a digital chat function, with therapists available daily during working hours on weekdays.
3. Monitoring: the app also facilitates the collection of health-related, measurable parameters to support the ongoing individual optimization of the training plan. In this study, patient-reported outcome measures (PROMs), self-conducted blood pressure measurements, and activity tracking with a wrist-worn device (GENEActiv, Activinsights) were applied.

The START Program

The main objective of START was to create a training program to be used with the Blended Clinic app specifically designed for patients with moderate to severe motor impairments during the early subacute stages of rehabilitation. We developed the START content in line with recent exercise recommendations [17-20] and stroke guidelines [21,22]. While these guidelines offer valuable recommendations, they do not specify explicit exercise instructions or dosages.

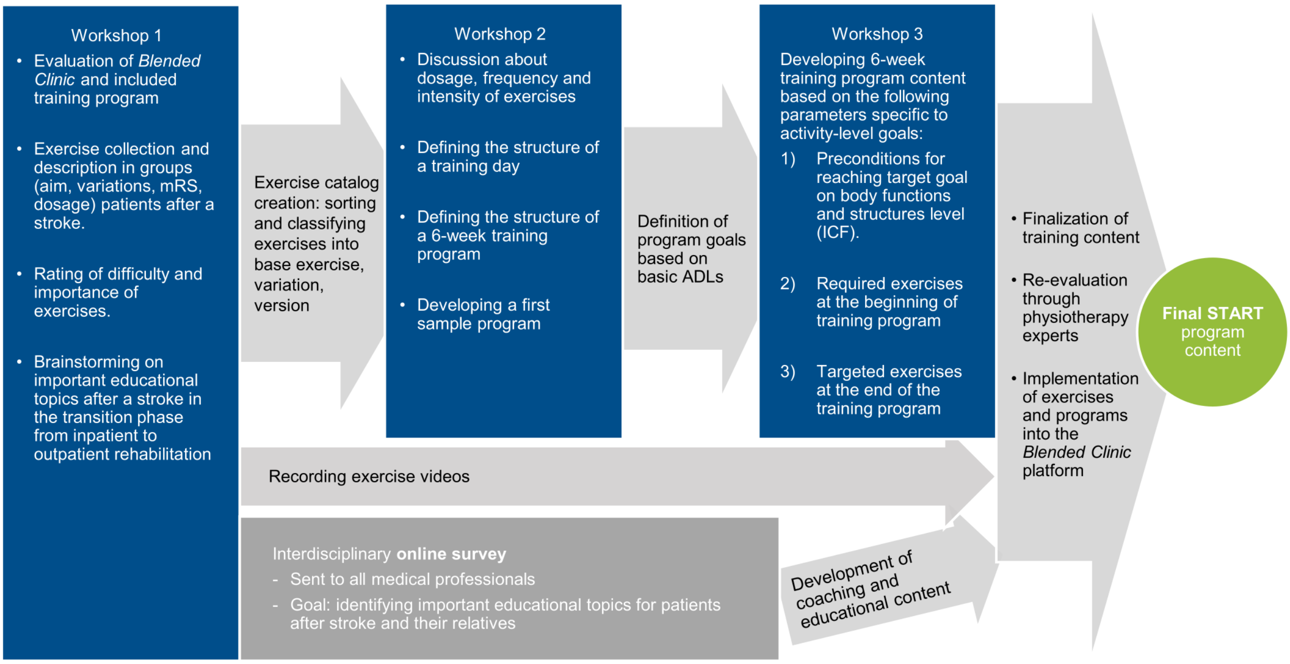
Three workshops were performed with a group of 7 clinicians, including 2 clinical physiotherapy neurological experts with over 25 years of professional experience and one research

XSL•FO
RenderX

physiotherapist with a doctoral degree, 3 clinical physiotherapists with at least 3 years of neurological experience, and 1 sports scientist. The participating therapists served as clinical instructors in the clinic and, in part, as lecturers and researchers.

In these workshops, important exercises and exercise variations for patients after stroke were identified and classified into different disability levels based on the modified Rankin Scale (mRS) [23,24] scores 1-5 (Figure 1). Exercises were further categorized into lower extremity, trunk exercises, balance and coordination exercises, and breathing exercises. In total, this resulted in the creation of 224 new exercises, resulting in a pool of around 300 exercises. In addition, for the presentation within the Blended Clinic platform, an introductory text outlining the aim of each exercise and a video instruction demonstrating how to perform the exercise were developed. Please refer to the supporting information for further details regarding the exercises (Multimedia Appendix 1).

Figure 1. Development process of the Swiss Tele-Assisted Rehabilitation and Training (START) program content and the main focus of each workshop. ADL: activities of daily living; ICF: International Classification of Functioning, Disability and Health; mRS: modified Rankin Scale; START: Swiss Tele-Assisted Rehabilitation and Training



Patient exercises were compiled to create 10 6-week programs, each with an overarching goal. The activity goals of such a program were based on abilities that are necessary to perform activities of daily living (ADL). The experts identified key components necessary for establishing individualized program goals, such as the ability to transfer to the toilet. For instance, a target exercise could involve standing for at least 60 seconds while engaging in arm activities. For this, a preliminary exercise might consist of "standing independently for 30 seconds without any arm activity."

Within each 6-week program, exercises progressed in terms of the number of sets, number of repetitions, and difficulty level. The exercise sessions, each lasting around 15 minutes and conducted daily, consisted of 3 phases, including a warm-up that typically involved a breathing or sensory stimulation exercise; a main component focusing on balance, strength, or endurance; and a cool-down that incorporated flexibility exercises. Additionally, care was taken to alternate strength, endurance, and balance training each day. Patients received information regarding training safety, content, meaning, and purpose of each exercise through the app.

In addition to the training programs, informational and educational materials for patients were developed. Relevant topics were identified by conducting an in-house interdisciplinary survey using the Findmind tool (Findmind). Based on the results from 90 health care professionals, several topics were identified, and content was developed. Podcasts were recorded covering the topics of fatigue, sleep, training, nutrition, and psychological well-being, along with the creation of 5 infographics. In addition to the regular motivational and informational messages sent via chat, this content constitutes an essential component of patient coaching and education and the blended care approach.

The START program was designed to be introduced in an inpatient rehabilitation setting and continue during the transition phase from inpatient to outpatient rehabilitation and then continue in the community. In the context of this usability study, not all the described content was used and investigated; rather, our focus was limited to an inpatient setting and specific components of START provided via the Blended Clinic platform.

Evaluation of Usability and Functionality of START

Participants

The evaluation of the usability and functionality of the app-based training was conducted as a single-center study primarily with patients after a stroke in their subacute stage and secondarily with therapists. The study took place in a neurological

XSL•FO
RenderX

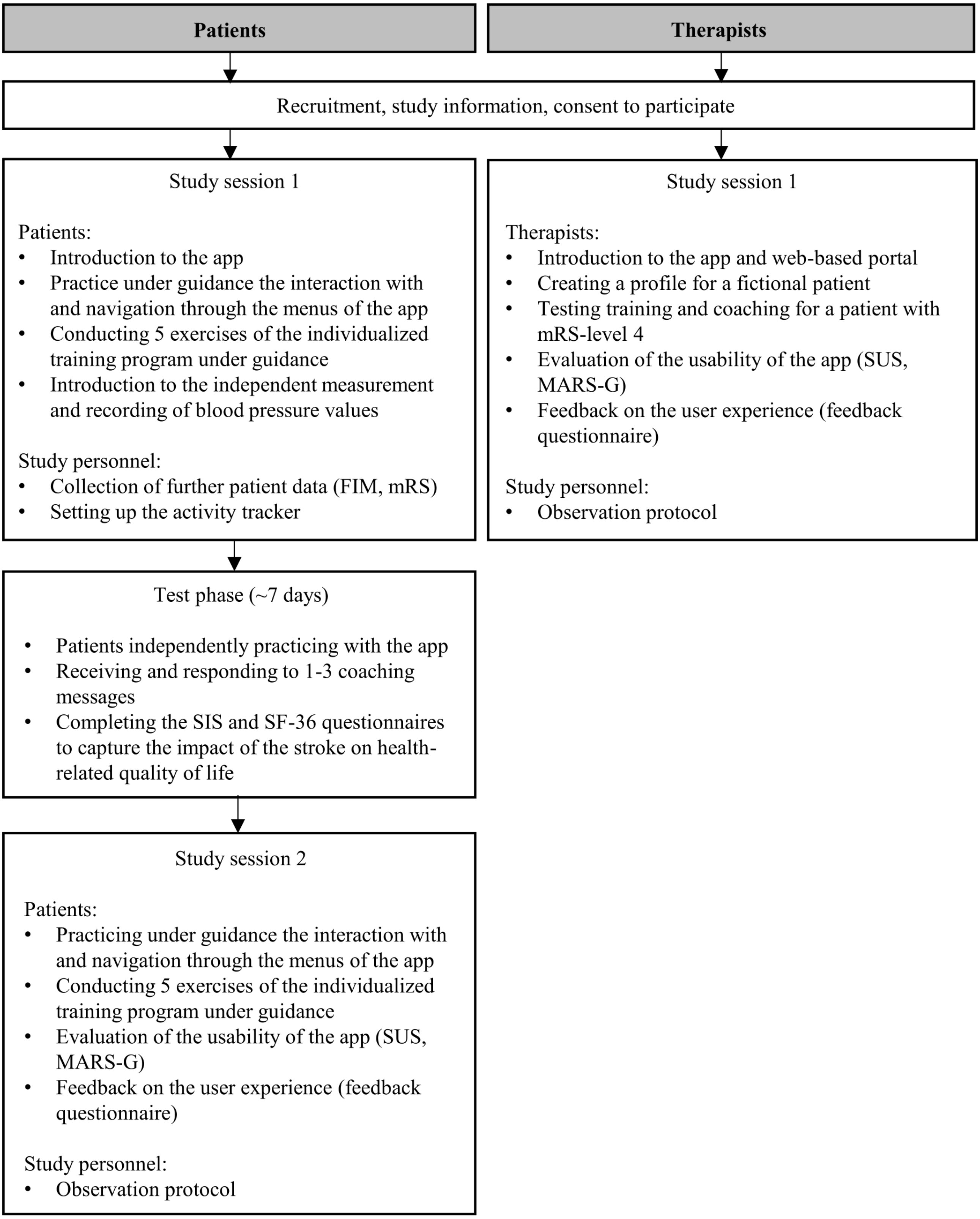
rehabilitation clinic in northwestern Switzerland. We planned to include 10 patients and 10 therapists, starting in December 2023. Inpatients and outpatients attending the clinic’s day center were recruited by screening internal patient lists against the predefined eligibility criteria. The criteria comprised age ≥18 years, mRS 3-5, no severe aphasia, severe visual, neurological, cardiopulmonary, psychiatric, or orthopedic impairments that would prevent a patient from following the instructions, and no recent surgery or fractures. Therapists were included in the study if they held at least a bachelor’s degree in physiotherapy, occupational therapy, speech and language therapy, or sports and exercise science, and if they had at least one year of experience in stroke rehabilitation. Therapists were invited via email to participate in the study.

Procedures

For the course of the study, please refer to the study flow chart (Figure 2). As demonstrated here, the therapists participated in a single session, whereas patients underwent 2 sessions, with a one-week testing phase in between.

XSL•FO
RenderX

Figure 2. Study flowchart. FIM: Functional Independence Measure, MARS-G: Mobile App Rating Scale; mRS: modified Rankin Scale; SF-36: 36-Item Short Form Survey Instrument; SIS: Stroke Impact Scale; SUS: System Usability Scale.



Study session 1

Patients:
- Introduction to the app
- Practice under guidance the interaction with and navigation through the menus of the app
- Conducting 5 exercises of the individualized training program under guidance
- Introduction to the independent measurement and recording of blood pressure values

Study personnel:
- Collection of further patient data (FIM, mRS)
- Setting up the activity tracker

Study session 1

Therapists:
- Introduction to the app and web-based portal
- Creating a profile for a fictional patient
- Testing training and coaching for a patient with mRS-level 4
- Evaluation of the usability of the app (SUS, MARS-G)
- Feedback on the user experience (feedback questionnaire)

Study personnel:
- Observation protocol

Test phase (~7 days)

- Patients independently practicing with the app
- Receiving and responding to 1-3 coaching messages
- Completing the SIS and SF-36 questionnaires to capture the impact of the stroke on health-related quality of life

Study session 2

Patients:
- Practicing under guidance the interaction with and navigation through the menus of the app
- Conducting 5 exercises of the individualized training program under guidance
- Evaluation of the usability of the app (SUS, MARS-G)
- Feedback on the user experience (feedback questionnaire)

Study personnel:
- Observation protocol

Patients had the opportunity to familiarize themselves with the app under supervision. During the introductory session with the study personnel, patients were instructed in the exercises assigned according to their mRS score and were trained in safe exercise execution. Besides, patients were also instructed on blood pressure measurements and on how to complete 2 questionnaires through the app (Stroke Impact Scale [SIS; 25,26], 36-Item Short Form Survey Instrument [SF-36; 27,28]). Patients subsequently tested the functionality of the app independently for a duration of one week. Before starting a daily exercise session, they received an automatic message reminding them to pay attention to their own safety during training. Furthermore, patients were provided with 2 information sheets, one outlining the criteria for discontinuing the exercises (eg, feeling unwell) and the other containing general information on the training environment (eg, ensuring that a table and chair were available for support). Patients wore an activity tracker (GENEActiv, Activinsights) for the week. During the second session following the testing phase, patients were asked to perform predefined tasks within the app under the guidance of the study personnel, for instance, to follow exercise instructions and to write a message to the coach. Patients were observed to document this process, to evaluate their performance, and thus the comprehensibility of the necessary interactive steps within the app.

Therapists were also introduced to the use of the app and had the opportunity to test it at the beginning of the session. They subsequently were instructed to perform specific tasks on the web-based dashboard platform, such as selecting and sending a questionnaire. Their performance was also observed and documented. Finally, the therapists themselves evaluated the usability of the system.

Outcome Measures

For the subsequent assessment of usability, the System Usability Scale (SUS) [29], the German version of Mobile App Rating Scale (MARS-G) [30-32] and a customized feedback questionnaire (Multimedia Appendix 2) were used.

The widely used SUS is a questionnaire that assesses the usability of a system through a 10-item survey. Each item is rated on a 5-point Likert scale, allowing users to express their level of agreement or disagreement with various statements about the system.

MARS-G assesses the multidimensional quality of mobile health apps and comprises 4 components, including engagement (Section A), functionality (Section B), esthetics (Section C), and information quality (Section D). Each domain is rated on a 5-point Likert scale, where 1 indicates “insufficient,” 2 denotes “poor,” 3 represents “acceptable,” 4 signifies “good,” and 5 reflects “excellent.” Additionally, the MARS-G includes 2 sections for the subjective evaluation of characteristics within the domains of subjective app quality (Section E) and perceived impact of the app (Section F). The German version of MARS-G that we used has been supplemented with a further section (Section T) that focuses on the therapeutic benefits associated with the app [32]. For all sections, the maximum achievable score is 5 points [32]. The therapists were asked to complete the entire MARS-G, while the patients were only provided with Sections A, B, C, D, E, and F for evaluation.

The training experiences with and the usability of the app were assessed using a self-constructed, standardized feedback questionnaire (Multimedia Appendix 2). This questionnaire used a 5-point scale and included specific questions regarding the app and was pilot tested with one patient. The feedback questionnaire aimed to provide deeper insights into the participants’ opinions about the training platform. For patients, the questionnaire contained 8 questions related to user experience, along with a section comprising 3 questions regarding the use of the blood pressure monitor, activity tracker, and smartphones or tablets. In contrast, therapists completed a questionnaire consisting of 9 questions about user experience.

During the study sessions for patients and therapists, the study personnel also completed an observation protocol (see Results for task details). This protocol facilitated the evaluation of the usability and user experience for both patients and therapists by an external observer, assessing the extent to which the app is useful and applicable, as well as identifying any existing issues. The study personnel also evaluated the emotional reactions by observation of the patients while they were navigating through the app and while performing the exercises, using the following categories (present and not present): attention, joy, curiosity, astonishment, boredom, nervousness, confusion, anger, and rejection.

Data Analysis

The adherence of the participating patients to the START program protocol was derived from the execution of the planned training sessions, the number of blood pressure measurements, the completion of the questionnaires, and the duration of activity tracker usage. For all parameters, adherence was assessed in relation (ratio) to the intended use, with the following classifications defined: adherence of ≥75% of the planned activities was categorized as high adherence, 50% - 74% as moderate adherence, 25% - 49% as low adherence, and <25% as no adherence [33]. In this study, the tracker was used solely to assess whether the patients would tolerate wearing it, with a view to planning further studies and monitoring activity levels over a longer period.

- Training: number of completed training sessions versus planned training sessions. Patients received an exercise program consisting of 5 exercises each day. An adherent training day was defined as having performed at least 3 out of 5 exercises, quantified by spending a minimum of 20 seconds and a maximum of 5 minutes for each exercise on the app screen.
- Monitoring:
  - Blood pressure measurements: patients were instructed to conduct 3 blood pressure measurements each morning. Days on which the respective patient entered data themselves (eg, blood pressure) were classified as adherent.
  - Activity tracker: an adherent day was defined as a day on which the activity tracker was worn for a minimum of 5 hours on the first and last day or for at least 10 hours on each intervention day, consistent with established practices [34]. In this study, the focus was solely on understanding whether the trackers were accepted and worn, in preparation for a larger study.

All data were analyzed descriptively using JASP (version 0.18.0, The Jasp Team), R Studio (version 2023.06.0; Posit PBC) and Microsoft Excel (Excel 2019; Microsoft Corporation).

Ethical Considerations

The study received approval from the independent ethics committee of Northwestern and Central Switzerland (EKNZ, reference number: 2023 - 01783). All participants provided written informed consent and all participant data were de-identified. The participating therapists received compensation for participation of approximately $190.

XSL•FO
RenderX

Results

Overview

From December 2023 to March 2024, a total of 161 patients with stroke were screened for eligibility (Figure 3). Eight patients (6 females) completed the study. The patient group had a mean age of 61.5 (SD 12.7) years with an average duration since the stroke of 180 days (265 days, range: 15-618 days) and a Functional Independence Measure (FIM) [24,35] score of 93.8 (SD 12.8; see Table 1).

Figure 3. Study patient flowchart.

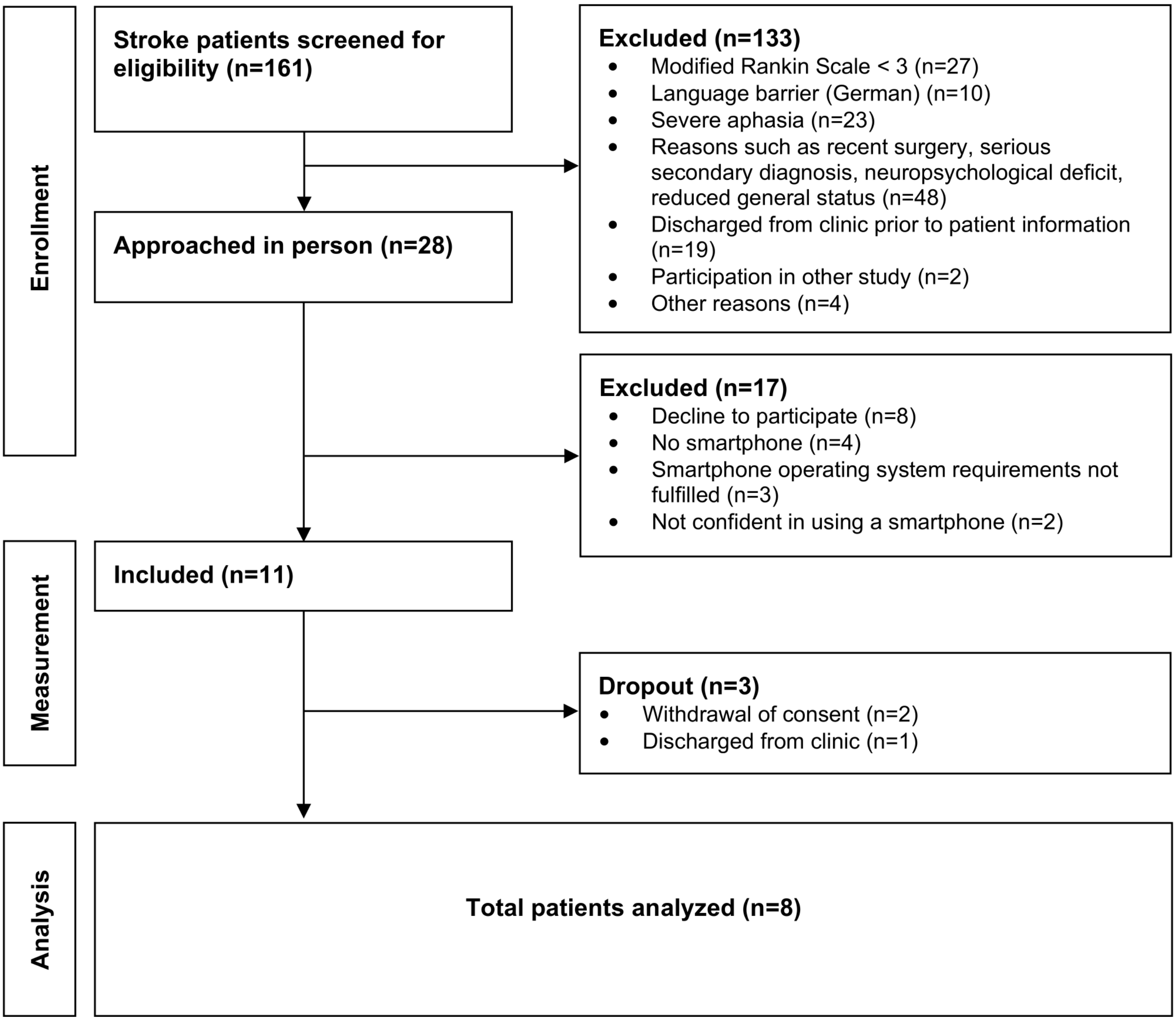


Table . Patient characteristics.

Patient demographics and clinical characteristics	Patient number							
	1	2	3	4	5	6	7	8
Sex	Female	Male	Female	Female	Female	Female	Male	Female
Age (years)	77	57	72	74	56	59	38	59
Stroke type	Ischemic	Ischemic	Ischemic	Ischemic	Ischemic	Hemorrhagic	Hemorrhagic	Ischemic
Time since stroke (days)	26	58	28	15	48	44	618	600
FIM[a] total score (18-126)	96	103	119	90	84	90	91	77
FIM-motor (13-91)	65	75	86	64	59	65	68	—[b]
FIM-cognitive (5-35)	31	28	33	26	25	25	23	—
Modified Rankin Scale (0-5)	4	4	3	3	4	3	3	4
Smartphone operating system	iOS	iOS	Android	Android	iOS	Android	iOS	iOS

[a] FIM: Functional Independence Measure.

[b] Not available.

Ten therapists (8 females) with a mean age of 35.3 (SD 11.2) were enrolled. They were mostly physiotherapists (n=7) and occupational therapists (n=3).

Outcome Measures

The patients rated the usability of the app, including training content and monitoring, with an average SUS score of 86.6 out of 100, which is slightly higher than the score of the therapists, who reported a score of 83.3 out of 100 (Tables 2 and 3).

Table . Results from the System Usability Scale for all patients.

SUS score	Patient number								Mean (SD)	Median (IQR)
	1	2	3	4	5	6	7	8		
SUS[a] (total score)	97.5	87.5	80	100	95	70	87.5	75	86.6 (10.9)	87.5 (20.6)
SUS[a] (per item)										
Item 1	5	5	4	5	5	3	4	5	4.5 (0.8)	5 (1)
Item 2	2	2	2	1	1	3	1	5	2.1 (1.4)	2 (1.8)
Item 3	5	4	5	5	5	3	5	5	4.6 (0.7)	5 (0.8)
Item 4	1	1	3	1	1	1	1	1	1.3 (0.7)	1 (0)
Item 5	5	4	5	5	5	4	3	3	4.3 (0.9)	4.5 (1.8)
Item 6	1	1	2	1	1	3	1	5	1.9 (1.5)	1 (1.8)
Item 7	5	5	5	5	3	5	3	5	4.5 (0.9)	5 (1.5)
Item 8	1	2	1	1	1	1	1	1	1.1 (0.4)	1 (0)
Item 9	5	5	4	5	5	3	5	5	4.6 (0.7)	5 (0.8)
Item 10	1	2	3	1	1	2	1	1	1.5 (0.8)	1 (1)

[a] SUS: System Usability Scale.

XSL•FO
RenderX

Table . Results from the System Usability Scale (SUS) for all therapists.

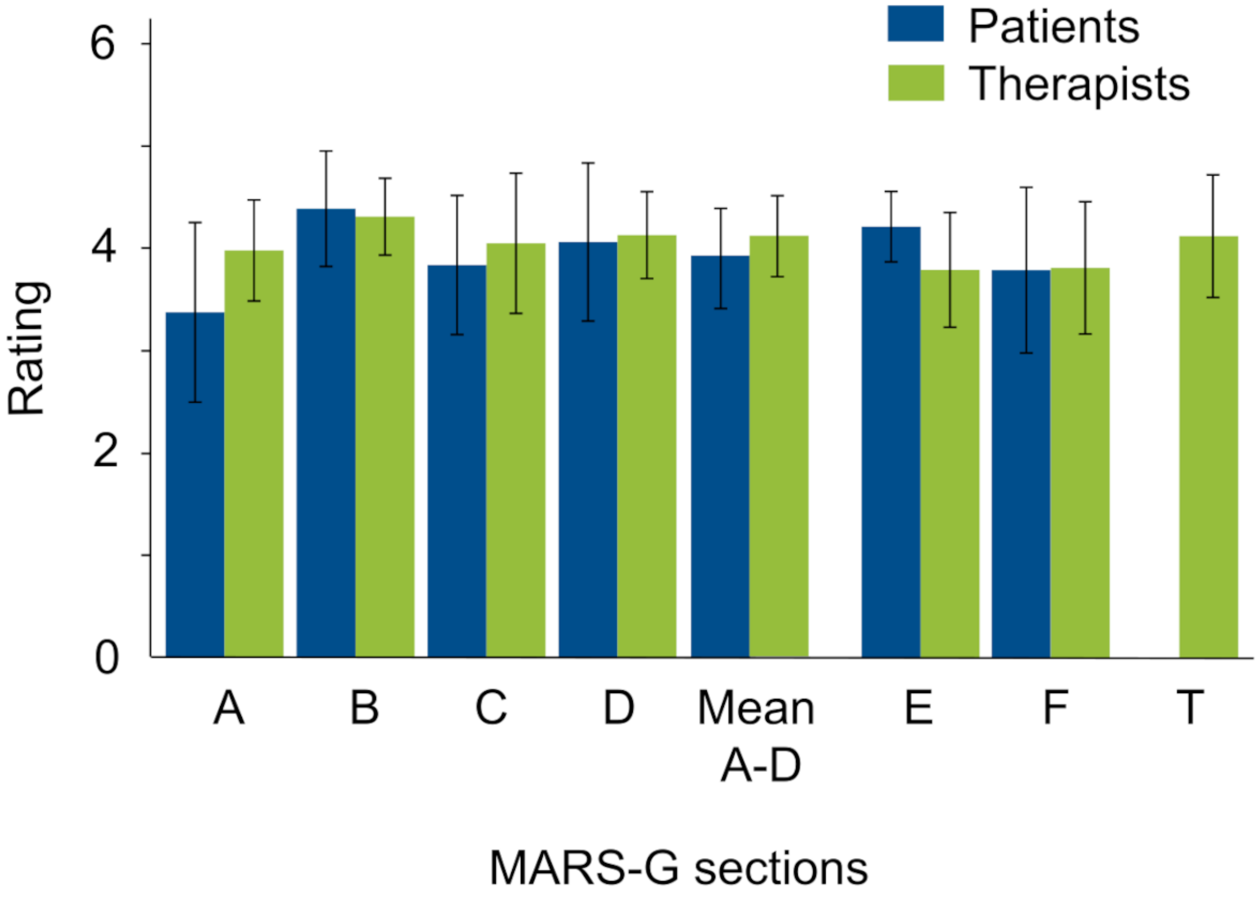
SUS[a] score	Therapist number										Mean (SD)	Median (IQR)
	1	2	3	4	5	6	7	8	9	10		
SUS[a] (total score)	95	97.5	62.5	90	90	80	77.5	85	65	90	83.3 (12.0)	87.5 (16.9)
SUS[a] (per item)												
Item 1	5	5	4	5	4	3	4	3	3	4	4.0 (0.8)	4 (2)
Item 2	1	1	2	1	1	1	2	1	1	1	1.2 (0.4)	1 (0.3)
Item 3	5	5	3	4	5	4	4	5	2	4	4.1 (0.9)	4 (1.3)
Item 4	1	1	1	1	1	1	1	1	3	1	1.2 (0.6)	1 (0)
Item 5	4	5	5	4	4	4	4	3	5	5	4.3 (0.7)	4 (1)
Item 6	1	1	4	1	2	2	1	1	1	2	1.6 (1.0)	1 (1)
Item 7	5	5	5	4	5	5	4	4	3	5	4.5 (0.7)	5 (1)
Item 8	1	1	3	1	1	2	2	1	1	1	1.4 (0.7)	1 (1)
Item 9	4	4	2	4	4	4	3	4	2	4	3.5 (0.8)	4 (4)
Item 10	1	1	4	1	1	2	2	1	3	1	1.7 (1.1)	1 (1)

[a] SUS: System Usability Scale.

For the MARS-G, both patients and therapists generally rated the user experience and usability of the app as acceptable to good, with the mean scores of Sections A-D being 3.9 for patients and 4.1 for therapists, as well as from the mean scores of the individual sections (Figure 4). The therapists rated no section below 4.0 points, thus categorizing all sections as good. Notably, they tended to rate the app slightly higher than the patients, particularly in the sections “Engagement” and “Aesthetics.” The section on “Functionality” received the highest ratings from patients and therapists. A table with the scores from patients and therapists is provided in the supporting information (Multimedia Appendix 3).

XSL•FO
RenderX

Figure 4. The German version of Mobile App Rating Scale (MARS-G) ratings of patients and therapists. A: engagement; B: functionality; C: esthetics; D information quality; E subjective quality; F: perceived impact; MARS-G: German version of Mobile App Rating Scale; T: therapeutic benefits.



The results of the feedback questionnaire indicated a generally positive perception of the app and the devices. The patients had mixed opinions regarding the handling of the devices in terms of self-monitoring blood pressure and wearing the activity tracker. The latter was partly perceived as somewhat bothersome. The operation of the app on the tablet or smartphone was deemed to be relatively straightforward.

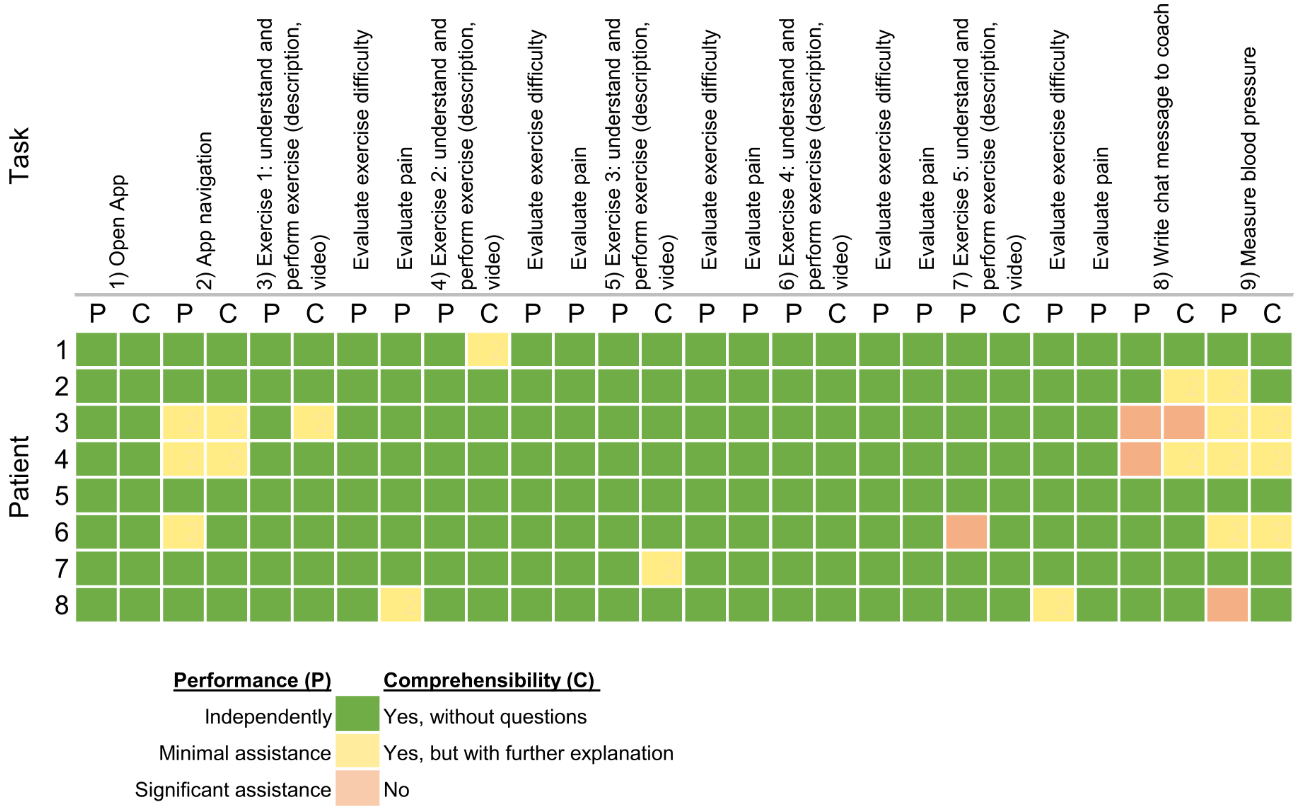
Among the therapists, there was a generally positive evaluation of the system. The question regarding the self-explanatory nature of the web-based dashboard for therapists received the lowest rating, while the question regarding technical functionality received the highest rating (Multimedia Appendix 4 for patients and therapists). Overall, all involved therapists expressed a desire to incorporate the app into their therapy planning.

Observation Protocols

Some patients encountered minor challenges in navigating the app's menus. A few needed substantial assistance in using the chat function to communicate with the coach or had follow-up questions. Patients required the most assistance with blood pressure measurement, with the majority needing at least some help during the measurement process (Figure 5). The emotions of attention, joy, curiosity, and astonishment were repeatedly mentioned. Except for one individual, all patients were engaged and attentive. Two-thirds of the patients exhibited joy throughout the observation. Curiosity and amazement were also noted in approximately half of the patients. In contrast, emotions such as rejection, anger, confusion, and nervousness were rarely observed, with 2 patients exhibiting these feelings.

XSL•FO
RenderX

Figure 5. Evaluation of patients’ task performance.



Task	1) Open App		2) App navigation		3) Exercise 1: understand and perform exercise (description, video)		Evaluate exercise difficulty	Evaluate pain	4) Exercise 2: understand and perform exercise (description, video)		Evaluate exercise difficulty	Evaluate pain	5) Exercise 3: understand and perform exercise (description, video)		Evaluate exercise difficulty	Evaluate pain	6) Exercise 4: understand and perform exercise (description, video)		Evaluate exercise difficulty	Evaluate pain	7) Exercise 5: understand and perform exercise (description, video)		Evaluate exercise difficulty	Evaluate pain	8) Write chat message to coach		9) Measure blood pressure	
Patient	P	C	P	C	P	C	P	P	P	C	P	P	P	C	P	P	P	C	P	P	P	C	P	P	P	C	P	C
1	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
2	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟨	🟩
3	🟩	🟩	🟨	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧	🟧	🟨	🟨
4	🟩	🟩	🟨	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧	🟨	🟨	🟨
5	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
6	🟩	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟧	🟩	🟩	🟩	🟩	🟩	🟨	🟨
7	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
8	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟩	🟩	🟩	🟧	🟩

Performance (P)		Comprehensibility (C)
Independently	🟩	Yes, without questions
Minimal assistance	🟨	Yes, but with further explanation
Significant assistance	🟧	No

All therapists were able to perform the assigned tasks. Only a few needed some further explanations (Figure 6).

Figure 6. Evaluation of therapists’ task performance.

Task	Login web-based dashboard		Set up patient profile		Select and assign a course		Write and send a private chat message		Write and send a broadcast message		Select and send questionnaire		Monitor fictional patient data		Search predefined exercises and create a training plan	
Therapist	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A
1	🟩	🟩	▨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟩
2	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟨	🟩	🟩	▨	🟩
3	🟩	🟩	▨	🟩	🟨	🟨	🟩	🟩	▨	🟩	🟩	🟩	🟨	🟨	▨	🟨
4	🟩	🟩	▨	🟩	🟩	🟩	▨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
5	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
6	🟩	🟩	🟩	🟩	⬛	🟩	🟩	🟩	🟨	🟨	🟩	🟩	🟩	🟩	🟩	🟩
7	🟩	🟩	🟩	🟨	🟨	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨
8	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
9	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟨	🟩	🟩	▨	🟨
10	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	▨	🟩	🟩	🟩	🟩	🟩	🟩	🟩

Performance (P)		Assistance (A)
Correct	🟩	No
Hesistant behavior	▨	
Difficulties	🟨	Yes, with further explanation
Incorrect	🟧	Yes, by study personnel
	⬛	Missing value

Adherence to the START Program Protocol

Overall, patients’ adherence to the protocol varied and was moderate for the test phase between study sessions. Adherence regarding training and monitoring differed distinctively among individual patients (Multimedia Appendix 5). Some patients consistently wore the activity tracker but did not adhere to the planned exercises or blood pressure measurements. There were no adverse events reported during the one-week trial phase.

- Training: adherence to training was low on average. Only one patient trained on most days (≥75% of planned training sessions), while 3 patients showed moderate adherence (between 50% and 74% of the planned sessions). One patient revealed low adherence to the training plan (between

XSL•FO
RenderX

  25% and 49%), and 3 had no adherence (<25%). Two patients could not perform exercises on one day or 2 days due to technical problems. Furthermore, some patients skipped part of or all exercises in the app, as noticed by the short time spent on the exercise screens.
- Monitoring:

- Blood pressure measurements: only one patient performed the blood pressure measurement every day. Three patients measured blood pressure on more than half of the planned days.
- Activity tracker: most patients (n=6) showed high adherence (≥75% of planned days) wearing the activity tracker. The adherence of one patient was low (25% - 49%), and one patient did not wear the tracker at all. In the case of this latter patient, however, the reason was a skin allergic reaction. There was no reason for the patients to remove the tracker, whether for charging, sleeping, or bathing, which may have contributed to the generally high adherence regarding the wearing of the trackers.

Discussion

Principal Findings

The aim of this study was to introduce our newly developed START program and initially determine how patients after a stroke and therapists evaluate the usability and functionality of START within the Blended Clinic platform. The following key findings were identified:

- This study demonstrated that both patients and therapists consider the Blended Clinic platform, including the START program, to be user-friendly.
- Regarding the usability of the app for patients after a stroke, the study yielded generally positive results. However, there remains potential for improvement to enhance long-term compliance with the program by patients and to ensure ease of use for therapists in the future.

Usability Evaluation

Evaluating usability is crucial to engage both patients and therapists early in the development process and to incorporate their valuable feedback into ongoing product refinement [36]. The initial insights into the usability of the telerehabilitation platform Blended Clinic in combination with the START program demonstrated good user-friendliness. Most of the given tasks were completed by the patients without assistance, which reassures us that the intended independent use of the Blended Clinic app is achievable. This is supported by the high SUS scores (83.3 among therapists and 87.2 among the participating patients with stroke), indicating the Blended Clinic platform with the START program as user-friendly. In a meta-analysis of digital health apps, the benchmark was a mean score of 68 with an SD of 12.5 [37], which is considerably exceeded in this study.

The evaluation based on the MARS-G allows for the assessment of the overall quality of the app-based training program, as well as the identification of both strengths and weaknesses through the mean scores of the individual sections [30]. In general, both therapists and patients rated the user experience and usability of the app as acceptable to good. Notably, therapists tended to rate the app slightly higher than the patients, particularly in the sections on “Engagement” and “Aesthetics.” Patients appear to be somewhat more critical in their evaluation of the app. This discrepancy may be attributed to the differing requirements and expectations of the 2 groups regarding the app [38].

Although the generalizability and credibility of the results from the feedback questionnaire are limited, it provided valuable insights, for example, concerning the handling of the blood pressure monitor, which can be applied in future studies. The results of the feedback questionnaire also indicated a generally positive perception of the app and the devices by the patients, although several areas for potential improvement have been identified. The technical functionality was regarded as having development potential. However, technical problems are quite common in studies investigating mobile health apps, as described in a scoping review on problems and barriers related to the use of these apps [39].

The therapists’ feedback additionally provided valuable insights into the practical implementation of the app. The therapists expressed a consistently positive response to the question of whether they would like to use the system in their therapy. The ratings indicated a high level of acceptance, overall satisfaction, and the potential of the system as a telerehabilitation option during the rehabilitation process. Our results are consistent with previous research. Physiotherapists are generally open to digital health technologies but frequently emphasize barriers to their implementation. In the study by Martinsen et al [37], therapists expressed concerns particularly about technical aspects, reimbursement issues, and establishing and maintaining a therapeutic relationship with patients [40]. The Blended Clinic platform with the START program is intended to address these issues through a blended care approach, which combines digital elements and personal interaction. The therapists in our study highlighted other important considerations, such as integration into existing therapy workflows, the need for flexibility in scheduling, and potential barriers related to staffing or digital infrastructure.

Patient Adherence to Protocol

The results of the one-week test phase indicated low adherence to the START training exercises among patients, despite the exercises being designed to be manageable and requiring only 10 minutes of daily commitment. Several factors may have contributed to the low adherence rates.

Participation in the study was conducted alongside the patients’ regular therapy regimen, which may have impeded their motivation to fully adhere to the study protocol. In addition, the exercises were rather easy for some patients, which might have influenced their adherence. Exercise difficulty and progression should be evaluated in further investigations.

While the presence of instructional videos and access to a physiotherapist via chat provided valuable resources, patients may still have struggled with their training. It is important to consider that impaired cognitive abilities can hinder protocol adherence. A study involving patients with stroke showed that factors such as age and disarticulation significantly influence

XSL•FO
RenderX

rehabilitation exercise adherence [41]. From our perspective, it would be beneficial to assess and individualize the amount and method of external support and motivation required. Further strategies may have been needed to enhance patient engagement and adherence to the exercise program, for example, family involvement.

Regarding low adherence to blood pressure measurements, patients may have perceived little added value in measuring their blood pressure themselves since these were routinely taken by the care personnel during inpatient stay. Not all patients were equally affected after a stroke, and some may have had further difficulties using the device due to motor function limitations, uncontrolled movements, or even hemiplegia, which could explain the issues encountered during blood pressure measurement.

On the contrary, the adherence to wearing the activity tracker was quite high. Such trackers have high potential to deliver meaningful data for the health care professionals, for example, step counts or time spent on specific activity levels. This activity tracker data could also be included directly in the Blended Clinic platform to increase the motivation and physical activity of patients. Such an approach is described in the pilot study by Paul et al [39], where a gamified approach was used to enhance step count goals [42].

Grau-Pellicer et al [40] conducted a randomized controlled trial using a similar program in conjunction with a smartphone app [43]. The authors could show positive results, for example, an improved activity level in the intervention group, but there was a low adherence to app use. They stressed the importance of developing technologies adapted to patients with stroke that are easy to use in order to increase adherence [43]. In our study, we tested only part of the START program, a one-week period, with a selected patient group (mRS 3 - 5), as usability was of particular importance to us at this stage. In contrast, adherence rates are generally high (≥80%) in telerehabilitation studies for patients with stroke, although this aspect is often underreported [44].

However, a randomized controlled trial by Emmerson et al [42] found no superiority of a home-based exercise program delivered via electronic tablet compared to a paper-based version, underscoring the importance of remote therapeutic support [12]. Consistent with this, promising results have been reported when remote support was provided to enhance adherence to home exercises in individuals with musculoskeletal conditions [45].

Usability evaluations for a wide range of mobile health apps have been conducted across various target populations [46], including patients after a stroke [47], demonstrating beneficial effects for patients. In physical stroke rehabilitation, similar mobile apps were developed and evaluated, for example, ARMStrokes (Towson University) for upper extremity training [48], an app to increase physical activity [42], or a program involving caregivers [49]. These and other technologies show promising results, especially for improving upper extremity function [47].

However, despite the already available apps for patients after stroke, most apps have been developed outside Europe [11]. Furthermore, additional equipment, for example, a balance disc [50] or a virtual reality system [51], is often necessary. With START, we focused on individualized physiotherapeutic exercises that require no special equipment. Notably, our usability study deliberately included patients with moderate to severe disabilities—a group often excluded from similar research—thereby providing valuable insights into the applicability and accessibility of the intervention for a broader spectrum of stroke survivors.

Limitations

Our usability study has several limitations that should be considered when interpreting the results. The sample size was small; however, we could include the planned patients after a stroke and therapists to evaluate the START program.

Furthermore, the intervention period for patients was only one week, which may not have been sufficient to draw conclusions about long-term adherence for the holistic START program. Extended usage periods are warranted to gain a deeper understanding of the factors influencing adherence to digital home-based exercise programs.

We also lack more detailed information about the patients, such as cognitive function or digital literacy. In our study, patients and therapists were asked to engage with a technical system and evaluate its usability and user experience. For study participation, patients were required to own a smartphone or tablet. Therefore, we assume that patients possess at least basic technical knowledge and feel confident in their ability to operate the technology. On the other hand, these requirements proved to be a challenge, as we experienced in a prolonged recruitment phase. However, the proportion of older individuals who own a smartphone is expected to rise [52], ensuring that access to mobile apps will be available for the majority of patients after a stroke. It can therefore be assumed that in the future, a significant proportion of patients after stroke will possess the necessary prerequisites to use these digital opportunities.

The feedback collected from patients and therapists was quite positive, but we cannot exclude the possibility of socially desired responses when interpreting the subjective outcomes. Ried et al [51] propose 8 preventive measures to address this issue; one of them is indirect questioning. Future studies should consider such measures to minimize this bias.

These limitations underscore the need for further research. Based on the findings from our study, we have designed a feasibility study aimed at a more comprehensive evaluation of the entire START program with a larger sample size. This evaluation will encompass the transition between inpatient and outpatient care, as well as the holistic content of the program covering all mRS levels. Additionally, we have established other focal points concerning the outcome measures, such as performance-based outcome measures.

Conclusions

In summary, the Blended Clinic platform, including the START program, was initially considered user-friendly, with good

XSL•FO
RenderX

usability ratings from both patients with stroke and therapists. The available data provide an insight into patient adherence and the underlying challenges. Future efforts should focus on optimizing technical functionality, enhancing user motivation, and integrating training into the daily routines of patients to potentially improve adherence. This study's results and adjustments aimed at enhancing the motivational component and paving the way for a long-term evaluation within a multiweek clinical study, in which the transition between hospital stay and the use of the system in a home setting will be tested.

Acknowledgments

We sincerely thank all the patients and therapists who participated in this study.

Funding

This study is part of the "TeleRehaStroke" project funded by the Canton Aargau, Department of Health and Social Care, Division of Health. There is no reference number available.

Data Availability

Data collected during the usability study can be made available by the first author upon reasonable request.

Conflicts of Interest

BC is the lead developer and founder of the software application used in this study. LHB received a grant from the Canton of Aargau, Switzerland, for the conduct of the present study.

Multimedia Appendix 1
Exercise catalog.
[PDF File, 354 KB - rehab_v13i1e77090_app1.pdf]

Multimedia Appendix 2
Feedback questionnaires.
[PDF File, 188 KB - rehab_v13i1e77090_app2.pdf]

Multimedia Appendix 3
Mobile App Rating Scale (MARS-G) results.
[PDF File, 129 KB - rehab_v13i1e77090_app3.pdf]

Multimedia Appendix 4
Feedback questionnaire results of patients and therapists.
[PDF File, 130 KB - rehab_v13i1e77090_app4.pdf]

Multimedia Appendix 5
Adherence metrics for patients.
[PDF File, 93 KB - rehab_v13i1e77090_app5.pdf]

References

1. Mendis S. Stroke disability and rehabilitation of stroke: World Health Organization perspective. Int J Stroke 2013 Jan;8(1):3-4. [doi: 10.1111/j.1747-4949.2012.00969.x] [Medline: 23280261]
2. Pantoni L, Poggesi A, Inzitari D. Cognitive decline and dementia related to cerebrovascular diseases: some evidence and concepts. Cerebrovasc Dis 2009;27 Suppl 1(suppl 1):191-196. [doi: 10.1159/000200459] [Medline: 19342851]
3. Erkinjuntti T. Vascular cognitive deterioration and stroke. Cerebrovasc Dis 2007;24 Suppl 1(1):189-194. [doi: 10.1159/000107395] [Medline: 17971655]
4. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
5. Thelwell S. Rehabilitation for patients after stroke. Nurs Times 2013;109(35):20-22. [doi: 10.33667/2078-5631-2019-4-39(414)-35-39] [Medline: 24266239]
6. Duncan PW, Bushnell C, Sissine M, et al. Comprehensive stroke care and outcomes. Stroke 2021 Jan;52(1):385-393. [doi: 10.1161/STROKEAHA.120.029678] [Medline: 33349012]

XSL•FO
RenderX

7. Bernhardt J, Godecke E, Johnson L, Langhorne P. Early rehabilitation after stroke. Curr Opin Neurol 2017 Feb;30(1):48-54. [doi: 10.1097/WCO.0000000000000404] [Medline: 27845945]
8. Shem K, Irgens I, Alexander M. Getting started: mechanisms of telerehabilitation. In: Telerehabilitation: Principles and Practice: Elsevier; 2022. [doi: 10.1016/B978-0-323-82486-6.00002-2]
9. Laver KE, Adey-Wakeling Z, Crotty M, Lannin NA, George S, Sherrington C. Telerehabilitation services for stroke. Cochrane Database Syst Rev 2020 Jan 31;1(1):CD010255. [doi: 10.1002/14651858.CD010255.pub3] [Medline: 32002991]
10. Seron P, Oliveros MJ, Gutierrez-Arias R, et al. Effectiveness of telerehabilitation in physical therapy: a rapid overview. Phys Ther 2021 Jun 1;101(6):pzab053. [doi: 10.1093/ptj/pzab053] [Medline: 33561280]
11. Mendes Pereira C, Matos M, Carvalho D, et al. Building bridges between people with stroke, families, and health professionals: development of a blended care program for self-management. J Clin Med 2024 Jan 4;13(1):300. [doi: 10.3390/jcm13010300] [Medline: 38202307]
12. Rintala A, Päivärinne V, Hakala S, et al. Effectiveness of technology-based distance physical rehabilitation interventions for improving physical functioning in stroke: a systematic review and meta-analysis of randomized controlled trials. Arch Phys Med Rehabil 2019 Jul;100(7):1339-1358. [doi: 10.1016/j.apmr.2018.11.007] [Medline: 30529323]
13. Bindschedler A, Ziller C, Gerber EY, et al. Feasibility of an application-based outpatient rehabilitation program for stroke survivors: acceptability and preliminary results for patient-reported outcomes. Bioengineering (Basel) 2024 Jan 29;11(2):135. [doi: 10.3390/bioengineering11020135] [Medline: 38391621]
14. Saunders DH, Sanderson M, Hayes S, et al. Physical fitness training for stroke patients. Cochrane Database Syst Rev 2020 Mar 20;3(3):CD003316. [doi: 10.1002/14651858.CD003316.pub7] [Medline: 32196635]
15. Billinger SA, Arena R, Bernhardt J, et al. Physical activity and exercise recommendations for stroke survivors: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2014 Aug;45(8):2532-2553. [doi: 10.1161/STR.0000000000000022] [Medline: 24846875]
16. Liguori G, editor. ACSM’s Guidelines for Exercise Testing and Prescription, 11th edition: Lippincott Williams & Wilkins; 2021.
17. Kim Y, Lai B, Mehta T, et al. Exercise training guidelines for multiple sclerosis, stroke, and Parkinson disease: rapid review and synthesis. Am J Phys Med Rehabil 2019 Jul;98(7):613-621. [doi: 10.1097/PHM.0000000000001174] [Medline: 30844920]
18. National Clinical Guideline for Stroke for the UK and Ireland. National Clinical Guideline for Stroke for the UK and Ireland. 2023. URL: https://www.strokeguideline.org [accessed 2025-10-09]
19. Living clinical guidelines for stroke management. Stroke Foundation. URL: https://informme.org.au/guidelines/living-clinical-guidelines-for-stroke-management [accessed 2025-10-09]
20. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988 May;19(5):604-607. [doi: 10.1161/01.str.19.5.604] [Medline: 3363593]
21. Bruno A, Akinwuntan AE, Lin C, et al. Simplified modified rankin scale questionnaire: reproducibility over the telephone and validation with quality of life. Stroke 2011 Aug;42(8):2276-2279. [doi: 10.1161/STROKEAHA.111.613273] [Medline: 21680905]
22. Mulder M, Nijland R. Stroke impact scale. J Physiother 2016 Apr;62(2):117. [doi: 10.1016/j.jphys.2016.02.002] [Medline: 26947003]
23. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. Stroke 1999 Oct;30(10):2131-2140. [doi: 10.1161/01.str.30.10.2131] [Medline: 10512918]
24. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [doi: 10.1097/00005650-199206000-00002] [Medline: 1593914]
25. Morfeld M, Kirchberger I, Bullinger M. SF-36 Fragebogen Zum Gesundheitszustand: Deutsche Version Des Short Form-36 Health Survey: Hogrefe Verlag; 2011. [doi: 10.1007/978-3-642-21700-5_14]
26. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry: Taylor & Francis; 1996:189-194. [doi: 10.1201/9781498710411]
27. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72. [doi: 10.2196/mhealth.5849] [Medline: 27287964]
28. Terhorst Y, Philippi P, Sander LB, et al. Validation of the mobile application rating scale (MARS). PLoS ONE 2020;15(11):e0241480. [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
29. Messner EM, Terhorst Y, Barke A, et al. The German version of the mobile app rating scale (MARS-G): development and validation study. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e14479. [doi: 10.2196/14479] [Medline: 32217504]
30. Hawley-Hague H, Horne M, Skelton DA, Todd C. Review of how we should define (and measure) adherence in studies examining older adults’ participation in exercise classes. BMJ Open 2016 Jun 23;6(6):e011560. [doi: 10.1136/bmjopen-2016-011560] [Medline: 27338884]
31. Chan A, Chan D, Lee H, Ng CC, Yeo AHL. Reporting adherence, validity and physical activity measures of wearable activity trackers in medical research: a systematic review. Int J Med Inform 2022 Apr;160:104696. [doi: 10.1016/j.ijmedinf.2022.104696] [Medline: 35121356]

XSL•FO
RenderX

32. Granger CV, Hamilton BB, Keith RA, Zielezny M, Sherwin FS. Advances in functional assessment for medical rehabilitation. Top Geriatr Rehabil 1986 Apr;1(3):59-74. [doi: 10.1097/00013614-198604000-00007]
33. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
34. Hyzy M, Bond R, Mulvenna M, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
35. Saaei F, Klappa SG. Rethinking telerehabilitation: attitudes of physical therapists and patients. J Patient Exp 2021;8:23743735211034335. [doi: 10.1177/23743735211034335] [Medline: 34377773]
36. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
37. Martinsen L, Østerås N, Moseng T, Tveter AT. Usage, attitudes, facilitators, and barriers toward digital health technologies in musculoskeletal care: survey among primary care physiotherapists in Norway. JMIR Rehabil Assist Technol 2024 Sep 16;11:e54116. [doi: 10.2196/54116] [Medline: 39283661]
38. Xing F, Liu J, Mei C, et al. Adherence to rehabilitation exercise and influencing factors among people with acute stroke: a cross-sectional study. Front Neurol 2025;16:1554949. [doi: 10.3389/fneur.2025.1554949] [Medline: 40083453]
39. Paul L, Wyke S, Brewster S, et al. Increasing physical activity in stroke survivors using STARFISH, an interactive mobile phone application: a pilot study. Top Stroke Rehabil 2016 Jun;23(3):170-177. [doi: 10.1080/10749357.2015.1122266] [Medline: 27077973]
40. Grau-Pellicer M, Lalanza JF, Jovell-Fernández E, Capdevila L. Impact of mHealth technology on adherence to healthy PA after stroke: a randomized study. Top Stroke Rehabil 2020 Jul;27(5):354-368. [doi: 10.1080/10749357.2019.1691816] [Medline: 31790639]
41. Stephenson A, Howes S, Murphy PJ, et al. Factors influencing the delivery of telerehabilitation for stroke: a systematic review. PLoS ONE 2022;17(5):e0265828. [doi: 10.1371/journal.pone.0265828] [Medline: 35544471]
42. Emmerson KB, Harding KE, Taylor NF. Home exercise programmes supported by video and automated reminders compared with standard paper-based home exercise programmes in patients with stroke: a randomized controlled trial. Clin Rehabil 2017 Aug;31(8):1068-1077. [doi: 10.1177/0269215516680856] [Medline: 27920262]
43. Lambert TE, Harvey LA, Avdalis C, et al. An app with remote support achieves better adherence to home exercise programs than paper handouts in people with musculoskeletal conditions: a randomised trial. J Physiother 2017 Jul;63(3):161-167. [doi: 10.1016/j.jphys.2017.05.015] [Medline: 28662834]
44. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil 2019 Jan;100(1):115-127. [doi: 10.1016/j.apmr.2018.07.439] [Medline: 30171827]
45. Guo J, Smith T, Messing D, Tang Z, Lawson S, Feng JH. ARMStrokes: a mobile app for everyday stroke rehabilitation. Presented at: ASSETS ’15: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility; Oct 26-28, 2015; Lisbon, Portugal p. 471-472. [doi: 10.1145/2700648.2811337]
46. van den Berg M, Crotty M Prof, Liu E, Killington M, Kwakkel G Prof, van Wegen E. Early supported discharge by caregiver-mediated exercises and e-Health support after stroke: a proof-of-concept trial. Stroke 2016 Jul;47(7):1885-1892. [doi: 10.1161/STROKEAHA.116.013431] [Medline: 27301941]
47. Burns SP, Terblanche M, Perea J, et al. mHealth intervention applications for adults living with the effects of stroke: a scoping review. Arch Rehabil Res Clin Transl 2021 Mar;3(1):100095. [doi: 10.1016/j.arrct.2020.100095] [Medline: 33778470]
48. Aphiphaksakul P, Siriphorn A. Home-based exercise using balance disc and smartphone inclinometer application improves balance and activity of daily living in individuals with stroke: a randomized controlled trial. PLoS ONE 2022;17(11):e0277870. [doi: 10.1371/journal.pone.0277870] [Medline: 36409753]
49. Sheehy L, Taillon-Hobson A, Sveistrup H, et al. Home-based nonimmersive virtual reality training after discharge from inpatient or outpatient stroke rehabilitation: parallel feasibility randomized controlled trial. JMIR Rehabil Assist Technol 2025 Mar 28;12:e64729. [doi: 10.2196/64729] [Medline: 40153779]
50. Seifert A. Digitale transformation in den haushalten älterer menschen. Z Gerontol Geriat 2022 Jul;55(4):305-311. [doi: 10.1007/s00391-021-01897-5] [Medline: 33909129]
51. Ried L, Eckerd S, Kaufmann L. Social desirability bias in PSM surveys and behavioral experiments: considerations for design development and data collection. Journal of Purchasing and Supply Management 2022 Jan;28(1):100743. [doi: 10.1016/j.pursup.2021.100743]
52. Gäumann S, Ziller C, Paulissen N, et al. START-the Swiss tele-assisted rehabilitation and training program to support transition from inpatient to outpatient care in the subacute phase after a stroke: feasibility, safety and performance evaluation. Front Digit Health 2024;6:1496170. [doi: 10.3389/fdgth.2024.1496170] [Medline: 39959919]

Abbreviations

ADL: activities of daily living
MARS-G: German version of Mobile App Rating Scale

XSL•FO
RenderX

mRS: modified Rankin Scale
SF36: 36-Item Short Form Survey Instrument
SIS: Stroke Impact Scale
START: Swiss Tele-Assisted Rehabilitation and Training program
SUS: System Usability Scale

Edited by S Munce; submitted 07.May.2025; peer-reviewed by L Sheehy, TA Wani, T Shimamoto; revised version received 19.Dec.2025; accepted 23.Dec.2025; published 31.Mar.2026.

Please cite as:
Ziller C, Gäumann S, Lüscher S, Paulissen N, Behrendt F, Suica Z, Crüts B, Gamerschlag L, Parmar K, Gerth HU, Bonati LH, Schuster-Amft C
Telerehabilitation Following Stroke: Development of Training Content and Evaluation of an App-Based Training Program
JMIR Rehabil Assist Technol 2026;13:e77090
URL: https://rehab.jmir.org/2026/1/e77090
doi:10.2196/77090

© Carina Ziller, Szabina Gäumann, Silya Lüscher, Nele Paulissen, Frank Behrendt, Zorica Suica, Björn Crüts, Luana Gamerschlag, Katrin Parmar, Hans Ulrich Gerth, Leo H Bonati, Corina Schuster-Amft. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telerehabilitation Trends in Australian Physiotherapy and an Exploration of Factors That Influence Use After COVID-19 Restrictions: Qualitative Content Analysis

Megan H Ross[1], PhD; Joshua Simmich[1], PhD; Belinda J Lawford[2], PhD; Kim L Bennell[2], PhD; Rana S Hinman[2], PhD; Trevor Russell[1], PhD

[1]RECOVER Injury Research Centre, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Surgical, Treatment and Rehabilitation Service (STARS), 296 Herston Rd, Herston, Queensland, Australia

[2]Centre for Health, Exercise and Sports Medicine, Department of Physiotherapy, School of Health Sciences, Faculty of Medicine, Dentistry & Health Sciences, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Megan H Ross, PhD
RECOVER Injury Research Centre, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Surgical, Treatment and Rehabilitation Service (STARS), 296 Herston Rd, Herston, Queensland, Australia

Abstract

Background: Telerehabilitation is a safe and effective means of delivering physiotherapy services, but implementation in clinical practice has not been widespread.

Objective: This study aimed to explore the shifts in telerehabilitation use throughout the COVID-19 pandemic and the key factors that influenced telerehabilitation caseload after restrictions were eased.

Methods: Between September and November 2023, physiotherapists practicing in Australian private practice, hospital outpatient, or community settings completed an online survey. Data were collected regarding participants' use of telerehabilitation before, during, and after the COVID-19 pandemic restrictions to in-person physiotherapy. Qualitative content analysis of open-text questions was performed to garner more nuanced information about the use of telerehabilitation in clinical practice, and quantitative data were analyzed descriptively.

Results: The proportion of participants using telerehabilitation rose from 30% (44/148) before the pandemic to 94% (138/147) when restrictions to in-person physiotherapy were in place. Although 82% (118/144) of the sample continued to deliver telerehabilitation after COVID-19 restrictions were eased, telerehabilitation accounted for only 14% of the total caseload. Exploratory analyses suggest that despite increased confidence, satisfaction, and perceptions about the effectiveness of telerehabilitation, reduced patient demand, physiotherapists' perceptions about patient preference for in-person consultations, and the perception that in-person physiotherapy is easier continue to influence the use of telerehabilitation in the post-COVID era.

Conclusions: Despite increased uptake during the pandemic, telerehabilitation caseload after restrictions were eased was low. Physiotherapists' perceptions about telerehabilitation in clinical practice remain a substantial barrier to sustained adoption.

(JMIR Rehabil Assist Technol 2026;13:e81008) doi:10.2196/81008

KEYWORDS

telehealth; COVID-19; videoconferencing; survey; physiotherapy; qualitative

Introduction

Background

Evidence suggests that telerehabilitation is a safe, feasible, and effective means of delivering physiotherapy care that is at least as good as in-person physiotherapy in terms of patient outcomes [1,2]. Despite two decades of evidence supporting the effectiveness of telerehabilitation in the management of musculoskeletal [3], neurological [4], cardiorespiratory [5], and postsurgical rehabilitation [6], clinician acceptance and adoption have been low [7]. While telerehabilitation is a viable alternative to traditional in-person physiotherapy with the potential to overcome geographical barriers, improve access, and facilitate continuity of treatment, integration into routine physiotherapy practice before the COVID-19 pandemic remained limited [8].

Before the COVID-19 pandemic, physiotherapists were slow and reluctant to adopt telerehabilitation as a standard model of care [9,10]. Several barriers contributed to this limited uptake of telerehabilitation in physiotherapy practice, including limited acceptance of and low confidence in using telehealth technology,

XSL•FO
RenderX

perceived limitations in conducting physical assessments remotely, and reduced capacity to deliver hands-on interventions that are central to traditional physiotherapy practice [8,11,12]. Additional challenges included perceptions that telerehabilitation was less effective for certain clinical presentations and concerns about developing rapport and patient engagement [11,12].

The COVID-19 pandemic brought unprecedented challenges to health care systems worldwide, including the delivery of physiotherapy services. In response to the pandemic-related restrictions on in-person consultations, many physiotherapy practices turned to telerehabilitation [13] as an alternative method to continue providing necessary care [14]. Although physiotherapy was recognized as an essential health care service in Australia during the COVID-19 pandemic, practice was still notably restricted. Stringent infection control measures, such as mandatory use of personal protective equipment, rigorous patient screening, physical distancing requirements, density limits within clinical spaces, and group size limitations on group therapy sessions [15], impacted the delivery of care across multiple clinical settings. During initial lockdowns, some states, such as Victoria, further restricted in-person physiotherapy services, permitting face-to-face consultations only for urgent cases [16]. Community and aged care physiotherapy faced further barriers, including restrictions on therapists attending multiple sites and outright bans on external providers entering residential facilities [17]. Consequently, physiotherapy practice during the pandemic was markedly disrupted, forcing providers to rapidly transition to providing telerehabilitation services to adhere to public health guidelines and ensure continuity of care.

With the rapid transition to telerehabilitation in response to the pandemic came changes in regulatory frameworks to fund telerehabilitation [18], position statements advocating for the use of telerehabilitation [19], and increased infrastructure and clinical training to support the integration of telerehabilitation into clinical care [20]. During this period, uptake of telerehabilitation increased substantially, reflecting the necessity to maintain continuity of care. Research conducted at the time suggested that physiotherapists intended to continue offering services via telerehabilitation after the easing of restrictions to in-person physiotherapy [20,21]. However, international evidence suggests that uptake and usage have generally decreased from the pandemic peak [22].

Objectives

With the rapid transition to telerehabilitation in response to the pandemic came changes in regulatory frameworks to fund telerehabilitation [18], position statements advocating for the use of telerehabilitation [19], and increased infrastructure and clinical training to support the integration of telerehabilitation into clinical care [20]. During this period, uptake of telerehabilitation increased substantially, reflecting the necessity to maintain continuity of care. Research conducted at the time suggested that physiotherapists intended to continue offering services via telerehabilitation after the easing of restrictions to in-person physiotherapy [20,21]. However, international evidence suggests that uptake and usage have generally decreased from the pandemic peak [22].

The aim of this study was to investigate the use of telerehabilitation in Australian physiotherapy clinical practice throughout the COVID-19 pandemic, with a focus on telerehabilitation use after restrictions were eased.

The specific research questions for this study were as follows: (1) How did the use of telerehabilitation vary in physiotherapy clinical practice in Australia before, during, and after COVID-19 restrictions to in-person consultations? and (2) What are the key factors that influence physiotherapists' telerehabilitation caseload in the postrestrictions period?

Methods

Design

A descriptive, cross-sectional survey was conducted online with physiotherapists currently practicing in Australia. The study was primarily quantitative, with a small qualitative component to supplement descriptive analyses.

Ethical Considerations

The study was approved by The University of Queensland Human Research Ethics Committee (approval number: 2023/HE001802) and reported following the consensus-based CROSS (Checklist for Reporting of Survey Studies) [23]. Participants provided electronic informed consent after reviewing an information sheet and before completing the survey. Participants were entered into a draw for a AUD $1000 (US $667) gift voucher upon completion of the survey. Participants' privacy and confidentiality were maintained by storing nonidentifiable survey data separately from contact details on the University of Queensland Research Data Management System.

Participants

Participants were physiotherapists recruited from the community via online advertisements on social media (eg, Facebook, X, and LinkedIn), via targeted emails, and through Australian Physiotherapy Association member communications (eg, eComms). Physiotherapists were eligible to participate if they were registered with the Australian Health Practitioner Regulation Agency and currently practicing in an Australian private practice, hospital, or community setting. Participants who had not delivered telerehabilitation services were eligible to complete a short version of the questionnaire to explore reasons for not engaging with telerehabilitation and the circumstances that might influence uptake.

Procedure

An online survey was designed to capture information that was relevant to stakeholders and ensure readability and credibility (Multimedia Appendix 1). The survey was developed by the authors using Bennell [20] as a guide and adapted to capture information relevant to the different phases of the COVID-19 pandemic restrictions. The 3 phases were "Prior to the pandemic restrictions," "During the period of restrictions to in-person physiotherapy" (from the introduction of restrictions in 2020 to 2022), and "After restrictions were eased" (2022 onward). Questions were primarily multiple choice (checkbox questions), numerical rating scales (0 - 10), and 5-point Likert scales.

XSL•FO
RenderX

Respondents were asked to estimate their telerehabilitation caseload (individual video, group video, and telephone) for each phase using a sliding scale (0% - 100%). Free-text responses were sought for some questions to ascertain more nuanced and in-depth information about physiotherapists’ perceptions of using telerehabilitation in clinical practice.

The survey was administered via an online secure platform (Qualtrics, LLC) and hosted by The University of Queensland. Participants were first invited to complete the online consent form and screening and, if eligible, proceeded to the survey. Participants were asked to provide demographic information, details of clinical practice, and experience with telerehabilitation. The second primary section of the survey comprised questions pertaining to the use of telerehabilitation during each phase of the COVID-19 pandemic restrictions (before, during, and after). All data were collected between September 15 and November 8, 2023.

Data Analysis

Data were exported from the online platform for analysis in R (version 4.3.3; R Core Team). Descriptive statistics, including frequencies (percentages) and means and standard deviations, were used to summarize the data. All responses (including partial responses) meeting eligibility criteria were included in analyses. When an “other” field was provided for additional response options, 2 researchers reviewed free-text responses and either aligned them with existing response options or designated them as unique responses that were added to the final list of response options. Any discrepancies in coding were resolved via discussion.

Responses to free-text questions were analyzed qualitatively using inductive content analysis in Microsoft Excel [24]. First, 2 researchers (MHR and JS) independently read the entire dataset, conducted open coding, and identified topics and initial patterns. The unit of analysis was meaning units, identified within individual responses. Codes were subsequently categorized and combined to form main categories or themes (abstraction), with both authors returning to the dataset to check that codes made sense in relation to the raw data. The 2 authors then met to compare and discuss their coding frameworks, and discrepancies were resolved through discussion. An audit trail was maintained to document coding decisions and category development. Themes with the highest number of individual data points were identified, reported, and described. To enhance trustworthiness, reflexivity was considered throughout the process, and attention was paid to credibility and transparency in coding and interpretation.

To explore which factors influenced physiotherapists’ use of telerehabilitation in the postpandemic restrictions period, the total proportion of videoconferencing telerehabilitation caseload (individual and group consultations) was examined. Specifically, this proportion was plotted against the following five key postpandemic variables: confidence, satisfaction, and perceived effectiveness of telerehabilitation; physiotherapists’ perception about how much patients like telerehabilitation; and how often patients are requesting it. Locally estimated scatterplot smoothing curves were fitted using the full span of the data (span=1). These smoothed trends, along with their corresponding 95% CIs, were used to visually explore apparent associations. No statistical correlation or regression analyses were performed on these trends.

Results

Sample Characteristics

A total of 222 physiotherapists responded to the survey, with 152 (68%) meeting eligibility criteria and providing sufficient data to be included in analyses (58/222, 26%, excluded for not being an Australian Health Practitioner Regulation Agency–registered physiotherapist currently practicing in an eligible setting [eg, private practice, hospital outpatient, or community] and 12/222, 5% not providing sufficient data to determine eligibility). Most participants (107/152, 70%) completed the survey in less than 20 minutes.

Respondents were primarily women (87/152, 57%); working in musculoskeletal (105/152, 69%) private practice (84/152, 55%) in Queensland (42/152, 28%), Victoria (42/152, 28%), or New South Wales (38/152, 25%); and held either a Bachelor’s (70/152, 46%) or Master’s (58/152, 38%) degree in physiotherapy. Physiotherapists primarily used Zoom (66/142, 47%), a telephone (59/142, 42%) or Microsoft Teams (47/142, 33%) to conduct telerehabilitation consultations. Only 40% (n=56) of respondents indicated that they had participated in telerehabilitation training. Additional participant characteristics are provided in Table 1.

XSL•FO
RenderX

Table . Participant characteristics (total N=152 unless otherwise specified).

Characteristic	Values, n (%)[a]
Gender	
Woman	87 (57)
Man	63 (41)
Prefer not to say	2 (1)
State or territory	
Queensland	42 (28)
Victoria	42 (28)
New South Wales	38 (25)
Western Australia	13 (9)
South Australia	10 (7)
Australian Capital Territory	6 (4)
Tasmania	1 (1)
Area of practice	
Private practice (primary care)	83 (55)
Public health outpatient center	45 (30)
Community health center	29 (19)
Private hospital	10 (7)
Other	20 (13)
Clinical focuses	
Musculoskeletal or orthopedic	105 (69)
Sports and exercise	43 (28)
Neurology	30 (20)
Gerontology	24 (16)
Pediatric	12 (8)
Other	35 (23)
Highest education	
Bachelor’s degree	70 (46)
Master’s by coursework	58 (38)
Masters by research	3 (2)
Postgraduate diploma	11 (7)
PhD	6 (4)
Other	4 (3)
Prior training in telehealth	
No	96 (63)
Yes, <6 mo ago	5 (3)
Yes, between 6 and 12 mo ago	6 (4)
Yes, between 12 mo and 2 y ago	17 (11)
Yes, between 2 and 3 y ago	17 (11)
Yes, longer than 3 y ago	11 (7)
Telerehabilitation software (recently used; n=142)	
Zoom	66 (47)
Telephone	59 (42)

XSL•FO
RenderX

Characteristic	Values, n (%)[a]
Microsoft Teams	47 (33)
Physitrack	30 (21)
Other	89 (63)

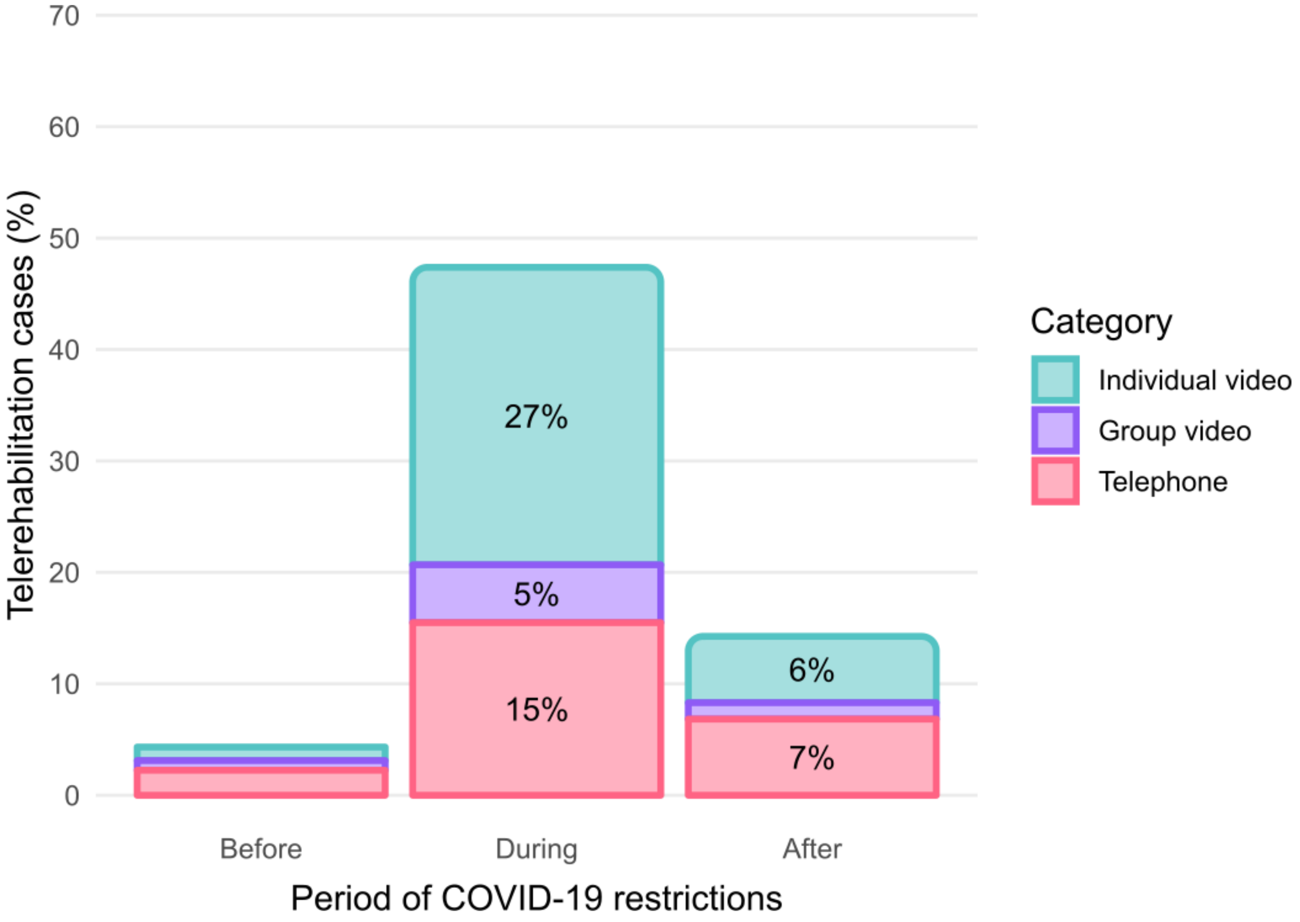
[a] Percentages may not sum to 100% because respondents could select multiple options.

Shifts in Telerehabilitation Use Through the Phases of the COVID-19 Pandemic

Thirty percent (44/148) of respondents indicated that they were using telerehabilitation in clinical practice before the pandemic. This rose to 94% (138/147) during the period of COVID-19 restrictions and reduced to 82% (118/144) after restrictions were lifted. Only 3% (4/152) of the sample indicated that they had never provided telerehabilitation consultations (individual or group videoconferencing, or telephone consultations). Total telerehabilitation caseload rose to account for almost 47% of the total caseload during the period of restrictions but dropped substantially to 14% once restrictions were lifted, but still remained above the prepandemic level of 4% (Figure 1). This pattern was fairly consistent across areas of practice over the COVID-19 pandemic (Figure S1 in Multimedia Appendix 2).

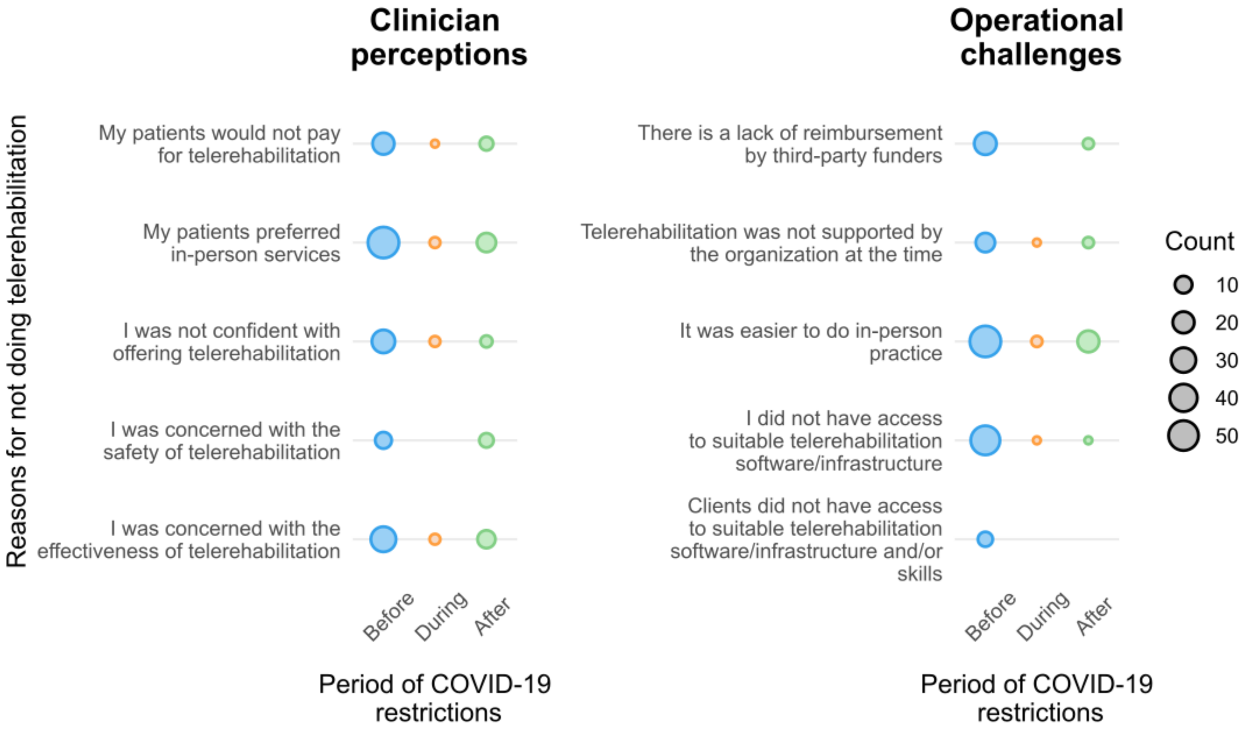
Reasons for not providing telerehabilitation consultations during each phase of the pandemic are provided in Figure 2. Across all 3 phases, the primary reasons were the perception that patients prefer in-person consultations (83/139, 60%) and that it was easier to do in-person consultations (55/139, 55%; Figure 2).

Figure 1. Shift in estimated telerehabilitation caseload before, during, and after the COVID-19 pandemic restrictions (values <2% are plotted without labels).



XSL•FO
RenderX

Figure 2. Reasons for not providing telerehabilitation throughout the phases of the COVID-19 pandemic, grouped by clinician perceptions and operational challenges. The area of circles represents the total count of respondents listing that reason for that point in time.



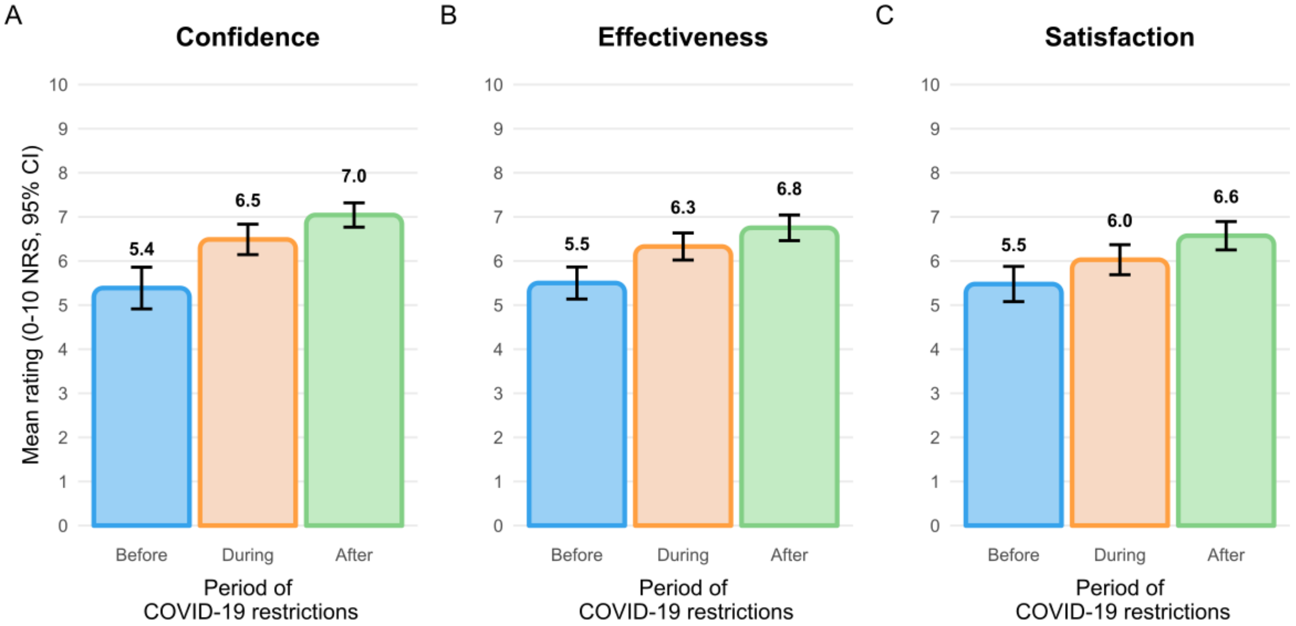
Before the pandemic, additional reasons for not offering telerehabilitation were primarily the perception that there was no need for telerehabilitation (57/104, 55%) or that physiotherapists did not have access to suitable telerehabilitation software or infrastructure (47/104, 45%; Table S1 in Multimedia Appendix 3). After restrictions were eased, the primary reasons for not providing telerehabilitation services were that respondents were concerned about the effectiveness of telerehabilitation (13/26, 50%) and did not like providing care via telerehabilitation (11/26, 42%; Table S1 in Multimedia Appendix 3).

Shifts in Confidence, Effectiveness, and Satisfaction With Telerehabilitation

Physiotherapist ratings of confidence in providing care via telerehabilitation, perceived effectiveness of telerehabilitation, and satisfaction with telerehabilitation progressively increased from before, during, to after restrictions associated with the COVID-19 pandemic (Figure 3). Almost 85% (120/142) of respondents indicated that providing telerehabilitation had become easier over time.

XSL•FO
RenderX

Figure 3. Participant ratings of (A) confidence, (B) perceived effectiveness, and (C) satisfaction with telerehabilitation across the pandemic. NRS: numerical rating scale.

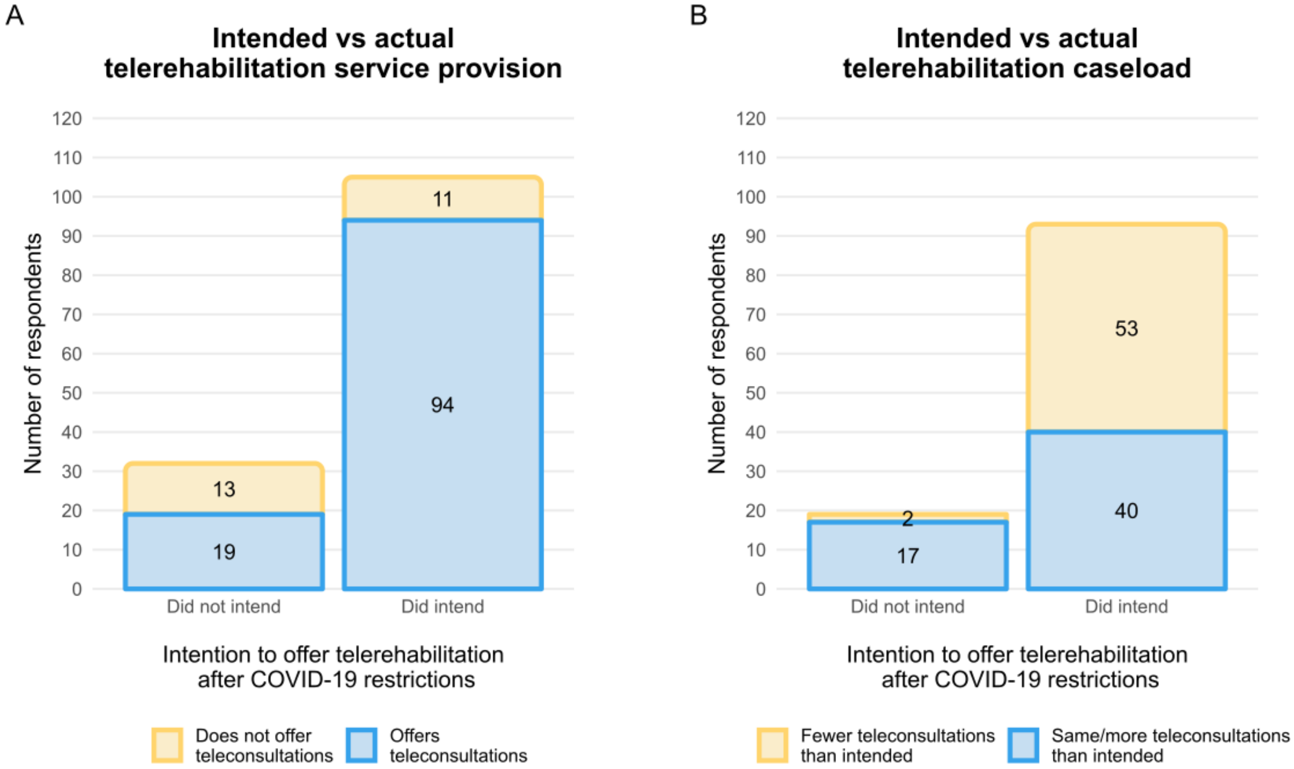


Intended Versus Actual Telerehabilitation Use

Most respondents (105/137, 77%) intended to offer telerehabilitation after the easing of COVID-19 restrictions (Figure 4). Of these, only 10% (11/105) did not offer telerehabilitation despite intending to do so (Figure 4A). Primary reasons for not intending to offer telerehabilitation were because it was “easier to do in-person” (23/32, 72%) and because “patients prefer in-person” (22/32, 69%; Table S2 in Multimedia Appendix 3). Of those who did not intend to, more than half (19/32, 59%) did continue to offer telerehabilitation after restrictions were eased (Figure 4A). Approximately 50% of respondents who intended to continue offering telerehabilitation consultations (53/93) or who actually continued offering them (58/118) after restrictions were eased were providing fewer consultations than initially intended (Figure 4B). Respondents indicated that this was because patients “prefer in-person services” (44/58, 76%), “patient demand reduced more than expected” (32/38, 55%) and because it was “easier to do in-person consultations” (27/38, 47%; Table S3 in Multimedia Appendix 3). Primary reasons for continuing to offer telerehabilitation services included that telerehabilitation allowed physiotherapists to offer services to patients who would not usually be able to attend their clinic (84/118, 71%), that patients like the option of receiving care via telerehabilitation (76/118, 64%), and that patients find telerehabilitation convenient (74/118, 63%; Table S4 in Multimedia Appendix 3).

XSL•FO
RenderX

Figure 4. Intended versus actual telerehabilitation service provision and caseload. (A) Number of respondents who intended to offer telerehabilitation after easing of pandemic restrictions compared to whether they do offer telerehabilitation now. (B) Number of respondents who intended to offer telerehabilitation after easing of pandemic restrictions compared to whether the frequency of telerehabilitation met intentions.



Factors That Influence Telerehabilitation Use Postpandemic Restrictions

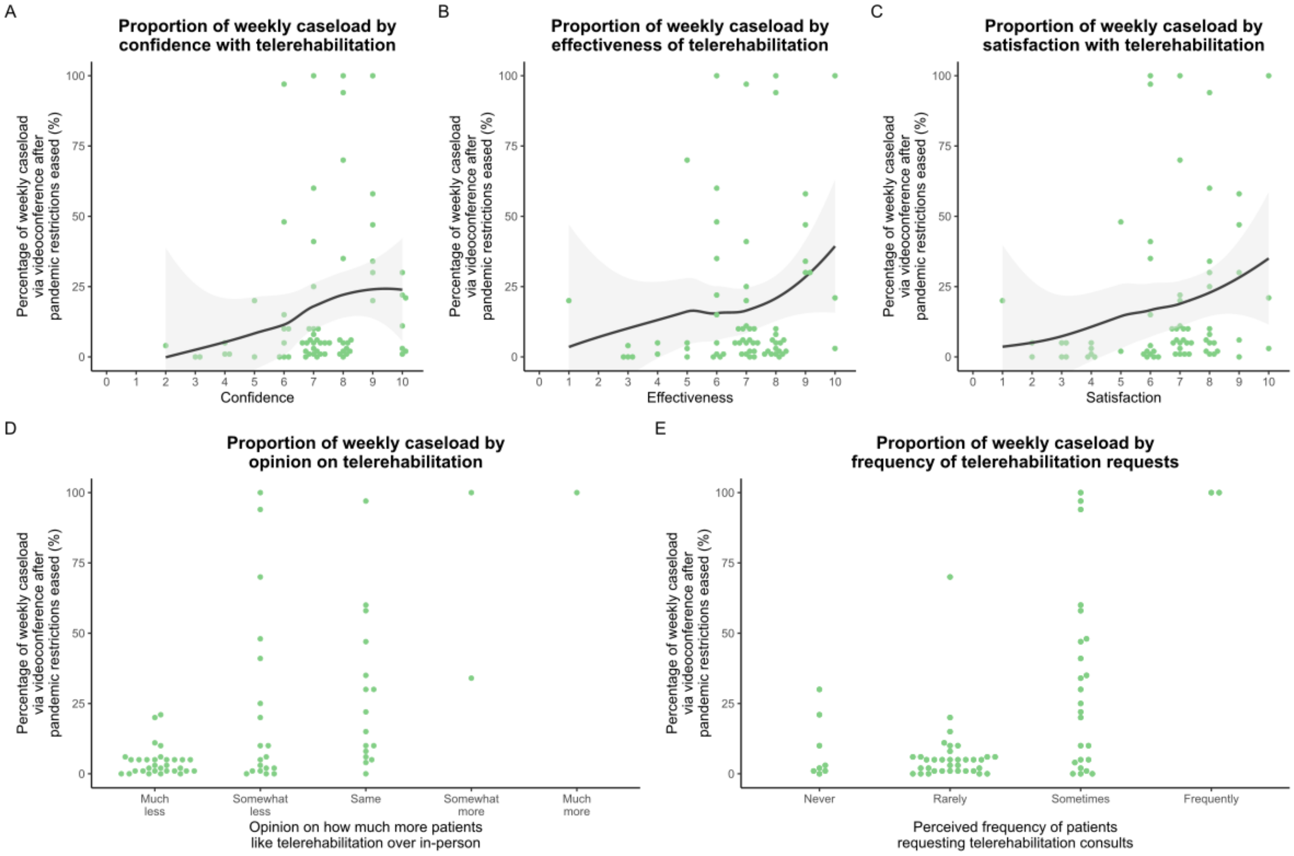
Positive correlations were noted between a higher proportion of weekly caseload conducted via telerehabilitation and higher ratings of confidence in using telerehabilitation (Figure 5A), perceived effectiveness of telerehabilitation (Figure 5B), and satisfaction with telerehabilitation (Figure 5C).

Almost half of the respondents (69/142, 49%) indicated that patients were “rarely” requesting telerehabilitation since the easing of restrictions, and in the opinion of approximately half of the respondents (70/142, 49%), patients like telerehabilitation “much less than in-person consultations.” Physiotherapists who believed that patients liked telerehabilitation much less than in-person consultations appeared to have a lower proportion of their weekly caseload conducted via telerehabilitation (Figure 5D). Similarly, physiotherapists who reported that their patients requested telerehabilitation at least sometimes seemed more likely to have a higher proportion of weekly cases conducted via telehealth (Figure 5E).

The median (IQR) percentage of patients considered unsuitable for telerehabilitation by the respondents was 50.5% (50). Patient complexity and conditions requiring hands-on treatment were the primary reasons that patients were “often” considered unsuitable (Figure S2 in Multimedia Appendix 4 and Table S5 in Multimedia Appendix 3). Additional reasons respondents provided for deeming patients unsuitable are provided in Table S6 in Multimedia Appendix 3, with the largest proportion being patient preference for in-person consultations (6/30, 20%), physical examination being indicated (4/30, 13%), and complex patient presentations (6/30, 14%).

XSL•FO
RenderX

Figure 5. noTotal weekly telerehabilitation caseload versus (A) confidence, (B) effectiveness, (C) satisfaction, (D) patient liking for telerehabilitation, and (E) patient requests for telerehabilitation. Dark gray lines represent locally estimated scatterplot smoothing fits (using the whole span of the data) with 95% CIs (light gray shading).



Telerehabilitation Clinical Practice Considerations

Although only 5% (8/142) of respondents reported never experiencing technical issues themselves, most (120/142, 84%) indicated encountering these issues rarely or sometimes, and just 9% (14/142) experienced them often. Likewise, only one respondent (1/142, 1%) reported that their patients had never encountered technical issues, whereas the majority (112/142, 79%) reported that patients experienced technical issues rarely or sometimes, and 19% (29/142) reported that patients often experienced technical issues. When technical issues were encountered, 74% (105/142) reported only moderate or less disruption to the consultation. Only 5% (7/142) reported that technical issues were extremely disruptive, and just 4% (6/142) reported often having to cancel or reschedule appointments due to technical issues (Figure S3 in Multimedia Appendix 5). To support the delivery of telerehabilitation consultations, physiotherapists used text message reminders (109/142, 77%); written or digital educational material about the condition (67/142, 47%); and written instructions, diagrams, or booklets (63/142, 44%; Table S7 in Multimedia Appendix 3).

Almost three-quarters (104/142, 73%) of respondents indicated that they used similar parameters of care for telerehabilitation as for in-person consultations (eg, similar consultation frequency, duration, and similar content). For the respondents indicating that parameters of care were different (38/142, 27%), the primary reasons were that physical assessment or treatment was limited via telerehabilitation (15/38, 39%), consultations were shorter (10/38, 26%), and consultations were more focused on exercise or education (8/38, 21%). Additional reasons are provided in Table S8 in Multimedia Appendix 3. Similarly, most respondents (128/142, 90%) indicated that telerehabilitation consultations were about the same duration or shorter than in-person consultations (Table S9 in Multimedia Appendix 3) and that consultation frequency was “about the same” as in person (72/142, 51%) or less often than in person (53/142, 37%; Table S9 in Multimedia Appendix 3).

Considering the cost of telerehabilitation consultations, almost three-quarters of physiotherapists indicated that they charged “about the same” as an in-person consultation (104/142, 73%), with very few respondents (5/142, 4%) charging more than in-person consultations. Responses about the cost to the business of providing telerehabilitation were similar, with 50% (71/142) of respondents considering telerehabilitation to cost about the same and 36% (51/142) indicating that telerehabilitation consultations cost the business less than in-person consultations (Table S9 in Multimedia Appendix 3).

The median proportion of patients offered hybrid care in a current weekly caseload was 5% (minimum=0, Q1=1, Q3=20, maximum=100). In hybrid models of care, 42% (48/113) of physiotherapists indicated that patients typically receive many more in-person than telerehabilitation consultations, 27% (30/113) receive the same, and 18% (21/113) receive fewer in-person visits compared to telerehabilitation consultations (Table S9 in Multimedia Appendix 3). Other ways in which telerehabilitation models of care differ from in-person models were coded qualitatively and provided in Table S10 in

XSL•FO
RenderX

Multimedia Appendix 3. When respondents used a hybrid model, 27% (8/27) offered telerehabilitation only after an initial in-person consult, 15% (4/27) described limiting the physical assessment or treatment component of consultations, and 11% (3/27) said telerehabilitation consultations in hybrid models had a greater case management focus (3/27, 11%).

Inductive content analysis of free-text responses identified four key themes that reflected respondents' perspectives on using telerehabilitation in clinical practice postpandemic (Table S11 in Multimedia Appendix 3): (1) concerns about telerehabilitation, (2) perceived benefits of telerehabilitation, (3) how telerehabilitation is used in practice, and (4) physiotherapists' willingness to provide telerehabilitation services.

Theme 1: Concerns About Telerehabilitation (n=28)

Physiotherapists expressed a range of concerns about the suitability and practicality of telerehabilitation in postpandemic physiotherapy care. The most commonly reported issue was that clients prefer or actively seek in-person consultations (n=12, 43%). For example, one participant said that despite telerehabilitation remaining available for their patients, they "*often prefer face-to-face*" and that "*people wanted to revert back to the 'usual' ways and leave the changes of COVID behind them moving forward once restrictions eased*" (musculoskeletal physiotherapist). Some respondents (n=4, 14%) emphasized that telerehabilitation is not suitable for all clients, particularly those with complex conditions, communication difficulties, or low digital literacy, and that they were selecting suitable clients for telerehabilitation, and "*not offering [it] for those not 'tech-savvy'*" (neurological physiotherapist). Concerns were also raised about the limitations of assessment via videoconference (n=2, 7%) and challenges related to internet connectivity and software reliability (n=3, 11%). Additional issues included payment and reimbursement barriers (n=2, 7%), difficulties building rapport remotely (n=2, 7%), and reduced referrals and attendance for telerehabilitation compared to in-person care.

Theme 2: Perceived Benefits of Telerehabilitation (n=20)

Despite concerns, respondents acknowledged several advantages of using telerehabilitation in their clinical practice postpandemic restrictions. The most frequently cited benefit was that telerehabilitation improved patient access to care, particularly for those in rural or remote areas or those with difficulties with travel or limited time (n=8, 40%). One musculoskeletal physiotherapist said that telerehabilitation "*has made physiotherapy much more accessible to a wider population and allows people greater flexibility with appointments*." Participants also noted an increased acceptance of telerehabilitation (among patients and providers; n=6, 30%), with some suggesting that it "*has become common practice now*" (musculoskeletal physiotherapist) and that it can be effective for certain presentations (eg, chronic musculoskeletal conditions; n=3, 15%), for supporting patient self-management (n=2, 10%) and providing greater flexibility in service delivery (n=1, 5%).

Theme 3: How Telerehabilitation Is Used in Practice (n=8)

Participants described integrating telerehabilitation into their clinical practice for subsequent consultations following initial in-person visits (n=2, 25%), for triaging (n=1, 12.5%) and case management (n=1, 12.5%), and as a tool for exercise prescription (n=1, 12.5%). Some participants (n=2, 25%) indicated that videoconferencing was preferred over telephone, and 1 (12.5%) participant noted that at times additional support is required at the patient end to effectively deliver telerehabilitation services. One cardiorespiratory, hospital-based physiotherapist described that telerehabilitation "*consultations have been effective in triaging patients and determining the appropriate level of care required*," whereas another described that telerehabilitation "*has been a great option for follow-up appointments, especially when you have already build rapport with patients...[and it]...has been a great way to check in with people who have busy schedules or live far away and find it difficult coming in*" (pelvic health and musculoskeletal physiotherapist).

Theme 4: Physiotherapists' Willingness to Provide Telerehabilitation Services (n=23)

Many participants (n=16, 70%) were willing to continue providing telerehabilitation services, driven by the perceived benefits and uses of telerehabilitation. For example, one private practice, musculoskeletal and mental health physiotherapist said that "*for the provision of exercise and movement based interventions, telerehabilitation has worked better than in-person as it provides easier access to more people given I live in a regional area*." Despite this willingness, some participants expressed low satisfaction with telerehabilitation (n=2, 9%) or a preference for in-person consultations (n=2, 9%). For example, a sports, exercise, and musculoskeletal physiotherapist working in private practice said that despite telerehabilitation "*opening up my practice to lots of different people around Australia and internationally... I still prefer to consult in-person...*" The need for ongoing education about the utility of telerehabilitation in physiotherapy was noted (n=1, 4%), despite the perception that education about how to deliver telerehabilitation had improved during the pandemic (n=1, 4%).

Discussion

Principal Findings

Despite an initial increase due to the COVID-19 pandemic restrictions and physiotherapists' intentions to continue offering telerehabilitation services, many physiotherapists were offering fewer telerehabilitation consultations than anticipated once restrictions were lifted. This was primarily due to a preference for in-person consultations, concerns about the effectiveness of telerehabilitation, and the perception that physiotherapy consultations are easier to conduct in person.

International postpandemic data across both physiotherapy and other health services show a similar "peak-to-plateau" pattern, where telehealth usage increased substantially during restrictions before falling and stabilizing at a lower level rather than returning to prepandemic levels. In a Polish national dataset,

XSL•FO
RenderX

telehealth in both outpatient health and rehabilitation services (excluding mental health) rose from prepandemic levels near zero to peak in 2020 before subsequently stabilizing at approximately one-fifth and one-third of their respective peak volumes [25]. Similarly, musculoskeletal physical therapists in the United States reported reduced telerehabilitation usage postpandemic, albeit at levels higher than prepandemic [22]. Likewise, across the US health system, overall telehealth usage peaked in 2020 and then declined but stabilized by 2023 [26,27]. Notably, however, telehealth usage was better sustained for services less dependent on hands-on care (eg, behavioral health and psychiatry) and in states where policies were put in place to ensure payment parity with comparable in-person services [26].

Although the perceived effectiveness of telerehabilitation had increased over the course of the pandemic, more than half of participants still identified concerns about the effectiveness of telerehabilitation as a primary reason to stop offering telerehabilitation consultations once able to resume in-person services. This is consistent with other studies conducted during the pandemic, where physiotherapists indicated concerns about the effectiveness of telerehabilitation for physiotherapy assessment and/or management [13,28].

Research indicates that outcomes for telerehabilitation are the same, if not better, than in-person physiotherapy for a range of conditions. For example, systematic reviews and randomized controlled trials in musculoskeletal, cardiac, and pulmonary populations demonstrate the noninferiority of telerehabilitation [29] and good validity for assessment conducted via telerehabilitation [30]. Physiotherapists’ perceptions may be centered around occupational self-efficacy [31] or their own personal clinical experience of “effectiveness” rather than evidence of effectiveness in the published literature. However, our results suggest that physiotherapists were not concerned about their own ability to deliver services via telerehabilitation. Both perceived satisfaction with telerehabilitation and confidence in delivering telerehabilitation trended upward from the prepandemic to the postpandemic period, and “I was not confident with telerehabilitation” was not a key factor in our findings after restrictions were lifted (3/26, 12%) . However, large proportions of respondents who did not offer telerehabilitation at this stage said it was easier to do in-person consultations instead (20/26, 70%); they were concerned with the effectiveness of telerehabilitation (13/26, 50%), and they did not like providing care via telerehabilitation (11/26, 42%). These findings suggest that there are additional factors influencing physiotherapists’ perceptions about the superiority of “hands-on” or “in-person” physiotherapy [32-34] that have not been comprehensively explored, such as the professional identity of a physiotherapist [31,35].

A qualitative study describing a successful, rapid transition to telerehabilitation during the pandemic challenges the perception that physiotherapy requires “hands-on” approaches and needs to be in person [36]. This study identified that physiotherapists’ readiness and willingness to modify their approach influenced the success of telerehabilitation. In our study, physiotherapists preferred in-person consultations themselves and perceived that their patients also preferred in-person consultations, which is likely to influence whether they offer telerehabilitation to patients. While systematic reviews suggest that patient satisfaction with telerehabilitation is comparable to and often higher than in-person care [37,38], many patients report a preference for in-person physiotherapy if given a choice [37,39]. Although physiotherapists might have thought during the pandemic that patient demand for telerehabilitation would remain (eg, explaining their intention to offer it), if patient demand for it decreased (as 55% of our sample indicated), physiotherapists would likely perceive that patients prefer in-person care (and indeed 76% of our sample did).

Clinician preferences for providing in-person physiotherapy have also been explored and reported on in the literature. Despite high levels of clinician satisfaction when providing telerehabilitation in clinical trials [40,41], this does not appear to be the case for *in-practice* preference for, or satisfaction with, telerehabilitation [21,22,28]. Although satisfaction and confidence with telerehabilitation increased over time, participants in this study still perceived in-person physiotherapy to be easier. The rigorous planning or structured training required for telerehabilitation delivery in a randomized clinical trial, rather than day-to-day clinical practice, may explain this difference in perceptions, highlighting a need for training specific to the clinical implementation of telerehabilitation. Studies examining barriers to implementing telerehabilitation in routine physiotherapy practice consistently identify insufficient training for conducting telerehabilitation consultations as a primary concern [42]. To address these challenges, international clinical practice guidelines provide evidence-based recommendations and strategies for overcoming barriers, guiding the training of clinicians and facilitating effective implementation of telerehabilitation into physiotherapy practice [43].

Clinicians have long identified the technological illiteracy of clients as a barrier to the adoption of telerehabilitation in physiotherapy [42]. Despite advances in technology infrastructure, when transitioning to telerehabilitation during the COVID-19 pandemic period, clinicians still identified “technology concerns” (including clinician concerns about client ability to use technology) as a barrier to telerehabilitation use in clinical practice [28,34,36,42,44]. In our study, concerns about technical issues or patients being unable to use or access technology were not identified as primary reasons physiotherapists determined patients were unsuitable for telerehabilitation. Additionally, technical issues were only slightly or moderately disruptive to consultations. This is consistent with findings from an evaluation of consultations delivered in a randomized controlled trial, which found that technical issues occurred but were infrequent and minimally disruptive [45]. This could potentially be because, at the time data were collected for this study (2023), physiotherapists and clients had greater experience with and exposure to the technology required for telerehabilitation and had become more comfortable over time [46]. In other studies where data were collected earlier in the pandemic, it is possible that fewer people were familiar with telerehabilitation technology, hence it being a bigger barrier to delivering telerehabilitation services at the time [20,21,36].

XSL•FO
RenderX

Strengths and Limitations

The findings of this study should be interpreted with the following limitations in mind. First, this was a small convenience sample, and findings may have been skewed by self-selection bias (with those with strong opinions, either positive or negative, electing to complete the survey). Second, we asked participants to recall what they were doing before and during the pandemic several years after the fact. Therefore, it should be acknowledged that participants' responses may have been influenced by recall bias. However, our data pertaining to before and during the pandemic were consistent with other studies conducted during the pandemic and their intentions to continue (eg, 69% in our study said that during the pandemic, they intended to offer telerehabilitation after restrictions were eased). In a study by Bennell et al [20], 81% intended to continue offering telerehabilitation consultations after the pandemic, and in a study by Peng et al [28], 55% and 68% intended to continue offering phone and videoconferencing, respectively. If the opportunity arises (ie, another period of restrictions to in-person consultations), researchers should consider using prospective study designs. Moreover, because our survey encompassed both phone calls and videoconferencing, our findings may not reflect modality-specific differences in perceptions reported elsewhere [28]. This survey also only sampled physiotherapists operating within Australia's health care system, so its findings may not fully translate to other countries with different telerehabilitation policies, funding models, or cultural attitudes toward remote care. Finally, due to an error, questions about confidence, satisfaction, and perceived effectiveness after the pandemic restrictions were eased were misworded and instead asked about experiences during the pandemic. It is likely that, given that all questions before this were about easing restrictions, most respondents still answered according to the intention of the question, but we cannot discount that some answered more literally, thereby skewing the data.

Conclusions

Although telerehabilitation use surged with pandemic restrictions, it has subsequently decreased significantly, with telerehabilitation accounting for only a small proportion of the total caseload. Despite increased confidence and satisfaction with telerehabilitation, clinician preference, and physiotherapists' perceptions of patient preference for in-person care, reduced demand and the ease of in-person practice influence the use of telerehabilitation postrestrictions and suggest persistent barriers to frequent use. Addressing these barriers is crucial to enhance the long-term viability and effectiveness of telerehabilitation physiotherapy in Australia.

Funding

This work was supported by the National Health and Medical Research Council (NHMRC Project Grant 1157977); RSH is supported by an NHMRC Investigator grant (2025733); KLB is supported by a NHMRC Investigator Grant (1174431); BJL is supported by a University of Melbourne CR Roper Fellowship and Dame Kate Campbell Fellowship. For the purposes of open access, the author has applied a CC BY public copyright licence to any author-accepted manuscript version arising from this submission. The funders have no role in conduct, analysis, or reporting of this study.

Data Availability

The datasets generated during and analyzed in this study are not publicly available due to lack of ethical clearance to disclose data to third parties.

Authors' Contributions

Conceptualization: MHR, JS, BL, KLB, RSH, TR Data curation: MHR, JS Formal analysis: MHR, JS Visualization: JS Writing – original draft: MHR, JS Writing – review & editing: MHR, JS, BL, KLB, RSH, TR Funding acquisition: KLB, RSH, TR

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey.
[DOCX File, 59 KB - rehab_v13i1e81008_app1.docx]

Multimedia Appendix 2
Telerehabilitation caseload over the COVID-19 pandemic for each physiotherapy area of practice.
[PNG File, 220 KB - rehab_v13i1e81008_app2.png]

Multimedia Appendix 3
Detailed quantitative (n, %) and qualitative survey findings on physiotherapists' telerehabilitation practice across COVID-19 restrictions: reasons for not offering, reducing, ceasing, or continuing telerehabilitation; reasons patients were considered unsuitable

XSL•FO
RenderX

and the resources used to support videoconferencing; differences in consultation parameters and hybrid models of care compared with in-person practice; and postrestriction perspectives and willingness to provide telerehabilitation.
[DOCX File, 43 KB - rehab_v13i1e81008_app3.docx]

Multimedia Appendix 4
Patient suitability for telerehabilitation.
[PNG File, 41 KB - rehab_v13i1e81008_app4.png]

Multimedia Appendix 5
Technical issues.
[PNG File, 42 KB - rehab_v13i1e81008_app5.png]

References

1. Brown RC, Coombes JS, Rodriguez KJ, Hickman IJ, Keating SE. Effectiveness of exercise via telehealth for chronic disease: a systematic review and meta-analysis of exercise interventions delivered via videoconferencing. Br J Sports Med 2022 Sep;56(18):1042-1052. [doi: 10.1136/bjsports-2021-105118]
2. Dias JF, Oliveira VC, Borges PRT, et al. Effectiveness of exercises by telerehabilitation on pain, physical function and quality of life in people with physical disabilities: a systematic review of randomised controlled trials with GRADE recommendations. Br J Sports Med 2021 Feb;55(3):155-162. [doi: 10.1136/bjsports-2019-101375] [Medline: 33060156]
3. Dario AB, Cabral AM, Almeida L, et al. Effectiveness of telehealth-based interventions in the management of non-specific low back pain: a systematic review with meta-analysis. Spine J 2017 Sep;17(9):1342-1351. [doi: 10.1016/j.spinee.2017.04.008] [Medline: 28412562]
4. Laver KE, Adey-Wakeling Z, Crotty M, Lannin NA, George S, Sherrington C. Telerehabilitation services for stroke. Cochrane Database Syst Rev 2020 Jan 31;1(1):CD010255. [doi: 10.1002/14651858.CD010255.pub3] [Medline: 32002991]
5. Hwang R, Bruning J, Morris NR, Mandrusiak A, Russell T. Home-based telerehabilitation is not inferior to a centre-based program in patients with chronic heart failure: a randomised trial. J Physiother 2017 Apr;63(2):101-107. [doi: 10.1016/j.jphys.2017.02.017] [Medline: 28336297]
6. Zhou Z, Zhou X, Cui N, et al. Effectiveness of tele-rehabilitation after total hip replacement: a systematic review and meta-analysis of randomized controlled trials. Disabil Rehabil 2024 Oct;46(20):4611-4616. [doi: 10.1080/09638288.2023.2280070] [Medline: 37990882]
7. Wade V, Soar J, Gray L. Uptake of telehealth services funded by Medicare in Australia. Aust Health Rev 2014 Nov;38(5):528-532. [doi: 10.1071/AH14090] [Medline: 25219655]
8. Gray LC, Edirippulige S, Smith AC, et al. Telehealth for nursing homes: the utilization of specialist services for residential care. J Telemed Telecare 2012 Apr;18(3):142-146. [doi: 10.1258/jtt.2012.SFT105] [Medline: 22362837]
9. Brewster L, Mountain G, Wessels B, Kelly C, Hawley M. Factors affecting front line staff acceptance of telehealth technologies: a mixed-method systematic review. J Adv Nurs 2014 Jan;70(1):21-33. [doi: 10.1111/jan.12196] [Medline: 23786584]
10. Gregory WJ, Burchett S, McCrum C. National survey of the current clinical practices of the UK rheumatology physiotherapists: a short report. Musculoskeletal Care 2021 Mar;19(1):136-141. [doi: 10.1002/msc.1516] [Medline: 32975369]
11. Cottrell MA, Hill AJ, O’Leary SP, Raymer ME, Russell TG. Service provider perceptions of telerehabilitation as an additional service delivery option within an Australian neurosurgical and orthopaedic physiotherapy screening clinic: a qualitative study. Musculoskelet Sci Pract 2017 Dec;32:7-16. [doi: 10.1016/j.msksp.2017.07.008] [Medline: 28787636]
12. Lawford BJ, Bennell KL, Kasza J, Hinman RS. Physical therapists’ perceptions of telephone- and internet video-mediated service models for exercise management of people with osteoarthritis. Arthritis Care Res (Hoboken) 2018 Mar;70(3):398-408. [doi: 10.1002/acr.23260] [Medline: 28437566]
13. Reynolds A, Awan N, Gallagher P. Physiotherapists’ perspective of telehealth during the COVID-19 pandemic. Int J Med Inform 2021 Dec;156:104613. [doi: 10.1016/j.ijmedinf.2021.104613] [Medline: 34688969]
14. Bettger JP, Thoumi A, Marquevich V, et al. COVID-19: maintaining essential rehabilitation services across the care continuum. BMJ Glob Health 2020 May;5(5):e002670. [doi: 10.1136/bmjgh-2020-002670] [Medline: 32376777]
15. COVID-19 submission 325. : Australian Physiotherapy Association; 2020 URL: https://www.aph.gov.au/DocumentStore.ashx?id=52e7027c-1187-4d62-8f2b-53fcbcceb939&subId=685789 [accessed 2025-12-31]
16. COVID-19 (coronavirus disease 2019). Victorian Government Department of Health. 2020 Apr 1. URL: https://www.health.vic.gov.au/infectious-diseases/covid-19-coronavirus-disease-2019 [accessed 2025-12-31]
17. Aged care royal commission impact of COVID-19 on aged care supplementary submission. : Australian Physiotherapy Association; 2020 URL: https://australian.physio/sites/default/files/submission/APA_Supplementary_Submission_Impact_of_COVID-19_on_Aged_Care_August_2020.pdf [accessed 2025-12-31]

XSL•FO
RenderX

18. Report from the telehealth working group. : Medicare Benefits Schedule Review Taskforce; 2020 URL: https://www.health.gov.au/sites/default/files/documents/2021/06/final-report-from-the-telehealth-working-group.pdf [accessed 2025-12-31]
19. Telehealth guidelines – Response to COVID-19. : Australian Physiotherapy Association; 2020 URL: https://australian.physio/sites/default/files/APATelehealthGuidelinesCOVID190420FA.pdf [accessed 2025-12-31]
20. Bennell KL, Lawford BJ, Metcalf B, et al. Physiotherapists and patients report positive experiences overall with telehealth during the COVID-19 pandemic: a mixed-methods study. J Physiother 2021 Jul;67(3):201-209. [doi: 10.1016/j.jphys.2021.06.009] [Medline: 34147399]
21. Malliaras P, Merolli M, Williams CM, Caneiro JP, Haines T, Barton C. “It’s not hands-on therapy, so it’s very limited”: telehealth use and views among allied health clinicians during the coronavirus pandemic. Musculoskelet Sci Pract 2021 Apr;52:102340. [doi: 10.1016/j.msksp.2021.102340] [Medline: 33571900]
22. McLaughlin K, Minick KI, Fritz JM, et al. Physical therapists’ perceptions of telerehabilitation for patients with musculoskeletal conditions in a post-pandemic world. medRxiv. Preprint posted online on Jan 18, 2025. [doi: 10.1101/2025.01.17.25320739] [Medline: 39867367]
23. Sharma A, Minh Duc NT, Luu Lam Thang T, et al. A consensus-based Checklist for Reporting of Survey Studies (CROSS). J Gen Intern Med 2021 Oct;36(10):3179-3187. [doi: 10.1007/s11606-021-06737-1] [Medline: 33886027]
24. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
25. Krawczyk A, Marszałek M. A comparative analysis of telerehabilitation and telemedicine utilization during the COVID-19 pandemic in Poland: trends, patterns, and implications. Int J Telerehabil 2024;16(1):e6627. [doi: 10.5195/ijt.2024.6627] [Medline: 39022440]
26. Gage AD, Knight MA, Bintz C, et al. Disparities in telemedicine use and payment policies in the United States between 2019 and 2023. Commun Med (Lond) 2025 Feb 26;5(1):52. [doi: 10.1038/s43856-025-00757-2] [Medline: 40011624]
27. Mandal S, Wiesenfeld BM, Mann DM, Nov O. The “new” new normal: changes in telemedicine utilization since COVID-19. Am J Manag Care 2025 Mar 1;31(3):e74-e78. [doi: 10.37765/ajmc.2025.89700] [Medline: 40053411]
28. Peng TH, Eng JJ, Harris A, et al. A survey of the experiences of delivering physiotherapy services through telerehabilitation during the COVID-19 pandemic. Front Rehabil Sci 2024;5:1486801. [doi: 10.3389/fresc.2024.1486801] [Medline: 39512760]
29. Suso-Martí L, La Touche R, Herranz-Gómez A, Angulo-Díaz-Parreño S, Paris-Alemany A, Cuenca-Martínez F. Effectiveness of telerehabilitation in physical therapist practice: an umbrella and mapping review with meta-meta-analysis. Phys Ther 2021 May 4;101(5):pzab075. [doi: 10.1093/ptj/pzab075] [Medline: 33611598]
30. Bernhardsson S, Larsson A, Bergenheim A, et al. Digital physiotherapy assessment vs conventional face-to-face physiotherapy assessment of patients with musculoskeletal disorders: a systematic review. PLOS ONE 2023;18(3):e0283013. [doi: 10.1371/journal.pone.0283013] [Medline: 36943857]
31. Roitenberg N, Pincus T, Ben Ami N. Physiotherapy services during the COVID-19 pandemic: a mediated model of physiotherapists’ self-efficacy, tele-physiotherapy role stressors, and motivation to provide tele-physiotherapy. Physiother Theory Pract 2024 Jun;40(6):1140-1149. [doi: 10.1080/09593985.2022.2138662] [Medline: 36305357]
32. Hasani F, Malliaras P, Haines T, et al. Telehealth sounds a bit challenging, but it has potential: participant and physiotherapist experiences of gym-based exercise intervention for Achilles tendinopathy monitored via telehealth. BMC Musculoskelet Disord 2021 Feb 4;22(1):138. [doi: 10.1186/s12891-020-03907-w] [Medline: 33541314]
33. Fernandes LG, Oliveira RFF, Barros PM, Fagundes FRC, Soares RJ, Saragiotto BT. Physical therapists and public perceptions of telerehabilitation: an online open survey on acceptability, preferences, and needs. Braz J Phys Ther 2022;26(6):100464. [doi: 10.1016/j.bjpt.2022.100464] [Medline: 36410257]
34. Sia LL, Sharma S, Kumar S, Ajit Singh DK. Exploring physiotherapists’ perceptions of telerehabilitation for musculoskeletal disorders: insights from focus groups. Digit Health 2024;10:20552076241248916. [doi: 10.1177/20552076241248916] [Medline: 38665882]
35. Hawley-Hague H, Gluchowski A, Lasrado R, et al. Exploring the delivery of remote physiotherapy during the COVID-19 pandemic: UK wide service evaluation. Physiother Theory Pract 2024 Oct;40(10):2241-2255. [doi: 10.1080/09593985.2023.2247069] [Medline: 37610255]
36. Ross MH, Nelson M, Parravicini V, et al. Staff perspectives on the key elements to successful rapid uptake of telerehabilitation in medium-sized public hospital physiotherapy departments. Physiother Res Int 2023 Jul;28(3):e1991. [doi: 10.1002/pri.1991] [Medline: 36540908]
37. Zischke C, Simas V, Hing W, Milne N, Spittle A, Pope R. The utility of physiotherapy assessments delivered by telehealth: a systematic review. J Glob Health 2021;11:04072. [doi: 10.7189/jogh.11.04072] [Medline: 34956637]
38. Simmich J, Ross MH, Russell T. Real-time video telerehabilitation shows comparable satisfaction and similar or better attendance and adherence compared with in-person physiotherapy: a systematic review. J Physiother 2024 Jul;70(3):181-192. [doi: 10.1016/j.jphys.2024.06.001] [Medline: 38879432]
39. Lawford BJ, Bennell KL, Kimp A, Campbell PK, Hinman RS. Understanding negative and positive feelings about telerehabilitation in people with chronic knee pain: a mixed-methods study. J Orthop Sports Phys Ther 2024 Sep;54(9):594-607. [doi: 10.2519/jospt.2024.12383] [Medline: 39207737]

XSL•FO
RenderX

40. Hinman RS, Nelligan RK, Bennell KL, Delany C. "Sounds a bit crazy, but it was almost more personal:" a qualitative study of patient and clinician experiences of physical therapist-prescribed exercise for knee osteoarthritis via Skype. Arthritis Care Res (Hoboken) 2017 Dec;69(12):1834-1844. [doi: 10.1002/acr.23218] [Medline: 28217864]
41. Amin J, Ahmad B, Amin S, Siddiqui AA, Alam MK. Rehabilitation professional and patient satisfaction with telerehabilitation of musculoskeletal disorders: a systematic review. Biomed Res Int 2022;2022:7366063. [doi: 10.1155/2022/7366063] [Medline: 35958819]
42. Sia LL, Sharma S, Ing JBM, Kumar S, Singh DKA. Physiotherapists' perceptions, readiness, enablers, and barriers to use telerehabilitation: a scoping review. J Back Musculoskelet Rehabil 2024;37(6):1441-1454. [doi: 10.3233/BMR-240009] [Medline: 38905032]
43. Lee AC, Deutsch JE, Holdsworth L, et al. Telerehabilitation in physical therapist practice: a clinical practice guideline from the American Physical Therapy Association. Phys Ther 2024 May 1;104(5):pzae045. [doi: 10.1093/ptj/pzae045] [Medline: 38513257]
44. Albahrouh SI, Buabbas AJ. Physiotherapists' perceptions of and willingness to use telerehabilitation in Kuwait during the COVID-19 pandemic. BMC Med Inform Decis Mak 2021 Apr 8;21(1):122. [doi: 10.1186/s12911-021-01478-x] [Medline: 33832473]
45. Ross MH, Russell T, Bennell KL, et al. Technical issues occur but are infrequent and have little impact on physiotherapist-delivered videoconferencing consultations for knee osteoarthritis: a descriptive study. Musculoskelet Sci Pract 2023 Aug;66:102782. [doi: 10.1016/j.msksp.2023.102782] [Medline: 37269590]
46. Tousignant M, Boissy P, Moffet H, et al. Patients' satisfaction of healthcare services and perception with in-home telerehabilitation and physiotherapists' satisfaction toward technology for post-knee arthroplasty: an embedded study in a randomized trial. Telemed J E Health 2011 Jun;17(5):376-382. [doi: 10.1089/tmj.2010.0198] [Medline: 21492030]

Abbreviations

CROSS: Checklist for Reporting of Survey Studies

Edited by S Munce; submitted 21.Jul.2025; peer-reviewed by D Ceprnja, I Wilson; accepted 23.Dec.2025; published 27.Jan.2026.

Please cite as:
Ross MH, Simmich J, Lawford BJ, Bennell KL, Hinman RS, Russell T
Telerehabilitation Trends in Australian Physiotherapy and an Exploration of Factors That Influence Use After COVID-19 Restrictions: Qualitative Content Analysis
JMIR Rehabil Assist Technol 2026;13:e81008
URL: https://rehab.jmir.org/2026/1/e81008
doi:10.2196/81008

© Megan H Ross, Joshua Simmich, Belinda J Lawford, Kim L Bennell, Rana S Hinman, Trevor Russell. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Linguistic Validation and Cross-Cultural Adaptation of the Shoulder Telehealth Assessment Tool for Filipino Patients with Musculoskeletal Shoulder Condition: Cross-Sectional Study

Jeffrey Arboleda[1], PTRP, MD, MBA; Sharon Ignacio[1,2], MD; Jose Alvin Mojica[1], MD, MHPEd; Carl Froilan Leochico[2,3], PTRP, MD, MHSc

[1]Department of Rehabilitation Medicine, Philippine General Hospital, University of the Philippines Manila, Taft Avenue, Ermita, Metro Manila, Philippines

[2]Department of Physical Medicine and Rehabilitation, St. Luke's Medical Center, Metro Manila, Philippines

[3]Department of Medicine, Division of Physical Medicine and Rehabilitation, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

Corresponding Author:
Jeffrey Arboleda, PTRP, MD, MBA
Department of Rehabilitation Medicine, Philippine General Hospital, University of the Philippines Manila, Taft Avenue, Ermita, Metro Manila, Philippines

Abstract

Background: Telerehabilitation has been widely adopted to meet the growing rehabilitation demand, but it is often limited by unstable internet connection, poor audiovisual resolution, and difficult virtual assessment. The Shoulder Telehealth Assessment Tool (STAT), a comprehensive, patient-led, preconsultation shoulder physical examination pictorial guide, was developed to address these limitations by easing the communication of instruction during the consultation and potentially removing the need for video calls.

Objective: This study aimed to develop a linguistically valid and culturally appropriate Filipino version of STAT and to evaluate its content validity, internal consistency, understandability, and ease of use.

Methods: A cross-sectional study on the Filipino STAT was conducted in three phases: (1) linguistic validation by experts, (2) cross-cultural adaptation through pretesting of 12 participants diagnosed with a musculoskeletal shoulder condition at the Philippine General Hospital, and (3) pilot study on 47 participants of the same population.

Results: The Filipino STAT had an excellent content validity (scale validity index=0.80 - 0.97), excellent interrater reliability (κ coefficient=0.82 - 1.00), and good internal consistency (Cronbach α=0.87). Understandability was found to be excellent for pain and activity (98%), good for range of motion and special tests (85%), and poor for strength (37%). However, 24% (11/46) of participants perceived the tool difficult to understand with the use of some Tagalog words as the primary barrier, followed by non-familiarity with the tool and difficulty reading the text.

Conclusions: Development of the Filipino STAT through a rigorous linguistic validation and cultural adaptation has produced a culturally appropriate, valid, and reliable tool. Pain and activity, range of motions, and special test subdomains are suitable for clinical assessment, while strength subdomain needs further improvement in understandability.

(JMIR Rehabil Assist Technol 2026;13:e67974) doi:10.2196/67974

KEYWORDS

cross-cultural adaptation; linguistic validation; musculoskeletal rehabilitation; shoulder examination; Shoulder Telehealth Assessment Tool; telerehabilitation; STAT

Introduction

Telemedicine is the use of telecommunication technologies to deliver health care, public health, and health education services remotely [1]. Telerehabilitation is a branch of telemedicine specifically aimed at delivering rehabilitation services. Being an archipelago with limited health care professionals and facilities, the Philippines uses these innovative health care delivery services to improve health access for Filipinos. Telerehabilitation in the Philippines was first adopted in the context of community-based rehabilitation in 2017 [2] and was widely used in 2020 during the COVID-19 pandemic [3]. With the improving COVID-19 situation, telerehabilitation remains a viable solution to delivering rehabilitation services in far-flung areas of the country [4].

However, telerehabilitation is not without limitations, such as unstable internet connection, lack of confidence in establishing clinical diagnosis virtually, limited time allotment per patient, and poor audiovisual resolution [5]. Specifically, in the administration of virtual physical examination, rehabilitation professionals face difficulties in the conduct of special tests, range of motion (ROM) assessment, and strength testing, among others [3]. These challenges have opened opportunities for innovation in telerehabilitation.

One of these innovations was the development of the Shoulder Telehealth Assessment Tool (STAT), which is a comprehensive patient-led shoulder physical examination pictorial guide done prior to the actual teleconsultation, aimed to improve clinical efficiency [6]. It is the first published patient-reported outcome measure to simulate the performance of in-person physical examination, which includes special tests for screening of different shoulder pathologies [7]. A validated visual analog scale (VAS) [8], single assessment numeric evaluation (SANE) [9], and motion analysis and range of motion studies on activities of daily living [10] have also been integrated into the tool. The first subcategory of the STAT is composed of 3 questions for pain and activity. Both VAS and SANE are patient-reported. The VAS ranges from 0=no pain to 10=maximum pain, while the SANE is scored from 0% to 100% of normal. Meanwhile, the current level of daily activity is a nominal score (ie, unaffected sleep, full work, and full recreation or sport). The ROM testing has 9 questions that are answered with either yes (movement completed) or no (movement not completed). There are 5 items for the shoulder special tests answerable by yes or no, and finally 3 items for strength that are answered with either painful, weak, both, or none. Overall, there are a total of 20 questions in the STAT that can be completed in 30 to 45 minutes. In the Philippines and numerous other countries where Filipinos are found, the STAT can potentially improve the accuracy of virtual physical examination techniques for the shoulder, provide better rehabilitation care for patients who do not have stable internet connections or video call capacity, and optimize the delivery of telerehabilitation services. This may also be used by providers from various health care disciplines, such as physiatry, physical therapy, occupational therapy, orthopedics, rheumatology, and pain management, among others, which routinely conduct musculoskeletal examination of the shoulder.

In the Philippines, shoulder conditions have been found to be prevalent among Filipino office workers [11], as well as migrant workers [12]. Currently, the STAT is available in English [6] and has not been translated into any other language. Language and cultural differences call for a careful adaptation of health outcome measures to accurately reflect the cultural nuances and context of the target language version [13]. Hence, this study aimed to develop a linguistically valid and culturally appropriate Filipino version of the STAT and determine its content validity, internal consistency, understandability, and ease of use.

Methods

Study Design

This cross-sectional study used a mixed methods research design on the linguistic validation and cross-cultural adaptation of the Filipino STAT.

Participants

The target population was Filipino adults (aged at least 18 years) with unilateral shoulder pain for at least 6 weeks who consulted in-person or through telemedicine at the Philippine General Hospital (PGH)—Outpatient Department of Rehabilitation Medicine from January 2023 to June 2023. Based on the Department census during that period, there were 304 patients diagnosed with shoulder pathology, such as adhesive capsulitis, rotator cuff injury, and myofascial pain syndrome. The participants should have access to a stable internet and a device with video call capacity and be able to understand written instructions in Filipino or Tagalog, which is the Philippines' most commonly used language. Participants with a history of trauma, suspicion or diagnosis of upper extremity fracture, severe cognitive impairment, known psychiatric comorbidity, cerebrovascular disease, cervical radiculopathy, brachial plexus injury, upper extremity peripheral nerve injury, complex regional pain syndrome, and prior shoulder, neck, and breast surgery were excluded. The participants were allowed to withdraw from the study for any reason.

Sampling

The sample size for the pretest was based on a generally recommended sample size of 12 from the target population based on feasibility and precision of estimates for succeeding studies [14]. The sample size for a pilot study to determine internal consistency through Cronbach α was 46 [15]. Systematic randomized sampling was done where every seventh patient was selected.

Study Procedure

Overview

Adapting the recommendations on linguistic validation and cross-cultural adaptation of self-reported measures, the study procedure was divided into two phases [16,17]: (1) translation of the STAT from the original English to the Filipino version; and (2) cross-cultural adaptation through pretesting. An additional phase was added for internal consistency testing of the final version of the Filipino STAT (Figure 1). Meetings, where necessary, were all done virtually, and permission to record each online meeting was sought from all participants.

XSL•FO
RenderX

Figure 1. This is a flowchart adapted from Lorca et al [17] for linguistic validation and cross-cultural adaptation [6]. STAT: Shoulder Telehealth Assessment Tool.

XSL•FO
RenderX

Phase 1: Linguistic Validation

Step 1: Permission From the Instrument's Authors

Permission for linguistic validation and cross-cultural adaptation into the Filipino language was obtained from the STAT developers prior to initiation of translation.

Step 2: Forward Translation

The forward translation was independently done by 2 bilingual translators. The first forward translator (T1) was a licensed physical therapist with a Master's degree in physical therapy who provided a "reliable equivalence from a clinical perspective," while the second forward translator (T2) was a university instructor and a representative from the *Sentro ng Wikang Filipino* (National Center of the Filipino Language) with no medical background who served to "reflect the language used by the population."

Step 3: Review and Synthesis

The translators, together with a recording observer, convened to synthesize their versions and arrive at a consensus version. Translation and synthesis forms were used throughout the linguistic validation process by the translators and recording observer to document the translated version, rationale for changes, and any disputes or challenges in doing the translation or synthesis.

Step 4: Backward Translation

Backward translation was independently done by another pair of bilingual translators (B1 and B2), who happened to be high school Filipino language teachers and were not familiar with the original version. This step was done to prevent information bias and to detect any errors in the forward translation.

Step 5: Content Validation

The first forward- and backward-translated version of the Filipino STAT was then assessed for content validity independently by clinical experts using a content validation form. The experts were composed of 6 board-certified local physiatrists with expertise in the fields of sports medicine, musculoskeletal ultrasound, interventional physiatry, telerehabilitation, clinical anatomy, and kinesiology. The form was composed of a yes or no scale for clarity, 4-point Likert scale for relevance, and a comment section for each item.

Step 6: Synthesis and Conciliation

All versions of the Filipino STAT available so far and the completed content validation forms (Multimedia Appendix 1) were then reviewed by a consensus committee of experts to arrive at a prefinal version. The consensus committee consisted of all 4 translators (T1, T2, B1, and B2) from the previous steps and the 6 clinical experts who participated in the content validation step. They all met virtually and collectively decided and agreed on each item through the Expert Consensus Committee Guide on Equivalence Form (Multimedia Appendix 2), wherein semantic, idiomatic, experiential, and conceptual equivalence were assessed between the forward- and backward-translated English and Filipino versions. Semantic equivalence pertains to the use of words that have similar, unique meaning in both cultures. Idiomatic equivalence pertains to the use of equivalent colloquialisms and idioms in both cultures. Conceptual equivalence pertains to the use of words or phrases with equivalent conceptual meanings in both cultures, and experiential equivalence pertains to experiences elicited that are consistent or equivalent in both cultures. The form contained a yes or no scale on each type of equivalence, a column for rewording suggestions, and another column for other comments and suggestions to meet cultural equivalence.

Phase 2: Cross-Cultural Adaptation

Step 1: Pretesting

A pretest on 12 adult Filipino patients with shoulder pain was done using the prefinal version of the translated tool through video consultation to simulate the actual STAT procedure and to avoid unnecessary risk of contagion during the pandemic (time of study).

A think-aloud protocol during each video call with the principal investigator or research assistant using an encrypted platform, such as Zoom (Zoom Communications, Inc) or Viber (whichever the participant preferred), was used. The think-aloud protocol entailed each participant to read, perform, and answer the Filipino STAT while verbalizing their thoughts. This was done twice for each question: first without the pictorial guide; and second with the pictorial guide. The trial without the pictorial guide was done to assess the feasibility of the Filipino STAT Text Version for patients with no access to smartphones. An observational checklist (Multimedia Appendix 3) was filled out by the principal investigator to document the participant response, understandability of the tool, and ease and accuracy in performing the Filipino STAT, noting participants' quality of movements and compensatory movements. The principal investigator also asked each participant open-ended interview questions (Multimedia Appendix 4) on their experience (including perceived barriers and facilitators) in using the tool and their suggestions in improving it.

Step 2: Review of Results

General and question-specific errors and suggestions from the pre-test participants were reviewed by the committee in developing the final version of the Filipino STAT.

Step 3: Proofreading

The resulting version was then proofread by a Filipino language teacher to correct any spelling and grammatical errors. All translation forms and STAT versions were sent to the original authors for their review and feedback.

Phase 3: Internal Consistency Testing

Finally, all 46 participants underwent the same process as pretesting using the final version of the Filipino STAT for the pilot study to determine the internal consistency of the tool.

Data Analysis

Content validation was determined through computation of item-level and scale-level content validity indices based on average and universal agreement methods. Sociodemographic data were reported using descriptive statistics. The internal consistency of participant responses was determined through Cronbach α. This was evaluated separately for items in VAS,

XSL•FO
RenderX

SANE, ROM, strength tests, and special tests using STATA 16.1/IC (StataCorp LLC). Considering ≥90% as excellent, 89% to 70% as good, 69% to 50% as fair, and ≤49% as poor, understandability was assessed using the worst-performing item in the observational checklist.

Interview responses were uploaded to NVivo 12 (Lumivero) for organization and thematic analysis of unstructured data pertaining to understandability. An inductive approach was employed where the data content directed the development of themes. The members of the research team were all physiatrists and had broad experience in musculoskeletal evaluation and telerehabilitation in clinical practice and research. JA was a rehabilitation medicine resident trained during the height of telerehabilitation during the COVID-19 pandemic. CFL spearheaded several research endeavors on telerehabilitation in the Philippines. SI and JM were the chairs of the Department of Rehabilitation Medicine with extensive teaching, research, and clinical experiences.

Ethical Considerations

The study was approved by the University of the Philippines Manila Research Ethics Board (UPMREB Code 2022-0516-01). Informed consent was secured through Google Forms, written and explained in Filipino either by the principal investigator or research assistant. Study introduction, objectives, procedure, duration, risks, benefits, incentives, and contact information were discussed with the participants. Participants were also assured of the confidentiality of their information, the voluntary nature of participation, and their right to refuse at any point of the study. The translators, recording observers, and Filipino teacher were appropriately compensated with Philippine 1000 each (US $20). All phase 2 participants were compensated with Philippine 200 (US $4) worth of prepaid load as reimbursement for internet fee for the video call interview and answering of online forms. The clinical experts did not receive remuneration and were acknowledged in this paper.

Results

Forward Translation

Both forward translators had challenges in translating the strength and special tests questions due to absence of direct translation of some English words to Filipino. Hence, both translators resorted to rephrasing some English sentences in Filipino. During synthesis, T1 shared the technical context of the subtests, while T2 suggested the use of understandable albeit literal translations of some words. For item Q9 (ROM) (Table 1), the use of the phrase "*itupi ang siko*" was selected over "*itupi ang braso*" based on T1's input to pertain to the correct anatomical joint that does flexion. On the other hand, for item Q12 (special test), "*idiin ang kamay sa tiyan*" was selected over "*pindutin ang tiyan*" based on T2's input on the context of pressing the belly. Items that resulted in differences between T1 and T2, as well as the final terminology agreed upon during synthesis, are presented in Table 1.

XSL•FO
RenderX

Table . Results of the forward translation and synthesis steps.

Item	T1[a]	T2[b]	Prefinal STAT[c]
Pain and activity			
Q1	*Lebel*	*Antas*	*Lebel*
	Pinakamasakit na naramdaman sa iyong buhay	*Pinakamasakit*	*Pinakamasakit*
Q3	*Dalas ng pisikal na aktibidad* • *Hindi apektado ang pagtulog* • *Nagtatrabaho nang buong araw* • *Nagagawa ang panlibangang aktibidad o isports*	*Antas ng pang-araw-araw na aktibidad* • *Di naaabalang pagtulog* • *Panay trabaho* • *Panay lingan/ isports*	*Antas ng pang-araw-araw na aktibidad* • *'Di naaabalang pagtulog* • *Nakakapagtrabaho nang buong araw* • *Nagagawa ang panlibangang aktibidad o isports*
Range of motion	*Saklaw ng galaw*	*Saklaw ng paggalaw*	*Saklaw ng paggalaw*
Instructions	*Kilos*	*Paggalaw*	*Paggalaw*
Q4	*Taas ng ulo*	*Tuktok ng ulo*	*Taas ng ulo*
Q7	*Bulsa ng pantalon sa parehong panig*	*Bulsa sa likod*	*Bulsa sa likod ng apektadong balikat*
Q8	*Ibabang likod*	*Ibabang bahagi ng iyong likuran*	*Ibabang bahagi ng iyong likuran*
Q9	*Braso ay nakaharap sa pader*	*Nakatagilid sa pader*	*Nakatagilid sa pader*
	Itupi ang siko	*Itupi ang braso*	*Itupi ang braso*
	Ilapat	*Idikit*	*Ilapat*
Strength			
Instructions	*Bigyan ng pwersa*	*Habang nilalabanan ang bigat*	*Bigyan ng pwersa*
	Mahina ba ang pakiramdam?	*Nanghihina ba?*	*Mahina ba ang pakiramdam?*
	Wala sa mga nabanggit	*Wala*	*Wala*
Q10	*Gamit ang kamao*	*Nakakuyom ang kamao*	*Gamit ang kamao*
	Idiin ang palad ng kamay sa masakit na braso	*Idiin ang mga ito sa isa't isa*	*Idiin ang palad ng kamay sa masakit na braso*
Q11	*Hilahin*	*Kapitan*	*Hilahin*
	Direksyon ng kabilang braso	*Gumalaw pakanan at pakaliwa*	*Direksyon ng normal na braso*
Q12	*Labanan*	*Kapitan*	*Kapitan*
	Ilayo	*Palayo*	*Ilayo*
Special tests			
Instructions	*Nakasaad na kilos at sasabihin mo*	*Maniobra na makakapagsabi sa amin*	*Mga kilos na makakapagsabi sa amin*
Q13	*Iangat ang buong braso*	*Itaas ang kamay*	*Itaas ang kamay*
Q14	*Umabot lagpas sa kabilang balikat*	*Ilagay ang inyong apektadong braso sa harap ng inyong dibdib*	*Umabot lagpas sa kabilang balikat*
Q15	*Pindutin*	*Idiin ang kamay*	*Idiin ang kamay*
Q16	*Ilapat*	*Idikit*	*Ilapat*
Q17	*Itaas ang braso na parang may hawak na plato*	*Iangat ang braso sa harap ng inyong dibdib at ilahad ang palad*	*Itaas ang braso na parang may hawak na plato*
	Dahan-dahang itulak pababa	*Bahagyang diinan*	*Dahan-dahang itulak pababa*

[a] T1: forward translator 1.

[b] T2: forward translator 2.

[c] STAT: Shoulder Telehealth Assessment Tool.

XSL•FO
RenderX

Backward Translation

Similar to the forward translators, the back translators did not apply word-for-word translation in strength and special test questions to simplify the items and make them easier to understand. The back translation versions were generally consistent with the original English version, with no significant differences, as agreed upon by the consensus committee.

Content Validation of Prefinal STAT

The instructions in all subdomains, except for strength, were clear to the experts. The items on pain and activity (Q1) and range of motion (Q4, Q5, and Q6) were clear to all experts. Content validity through the scale validity index was measured using the item-level content validity index (0.97) and using universal agreement (0.80). κ values across all items ranged from 0.82 to 1.00, indicating excellent inter-rater reliability.

Cultural Equivalence

Pertinent findings from the discussion of the consensus experts on equivalence are summarized in Table 2.

Table . Pertinent results from the expert consensus discussion on equivalence[a].

	SE[b]	IE[c]	CE[d]	EE[e]	Comments
Pain and activity					
Q1	✓	✓	✓	✓	• Inconsistent: not a question
Q2	✓	✓	✓	✓	• Inconsistent: not a question
Q3	✓	✓	✓	X	• Inconsistent: not a question • Vague: choices not in one spectrum • EE: did not include informal work
Range of motion					
Q7	✓	✓	X	✓	• CE: not relatable to everyone
Special test					
Instruction	✓	X	✓	✓	• IE: Test demands active patient response
Q13	✓	✓	✓	✓	• Safety: for severely painful shoulder • Unclear: confusing literal translation
Q17	✓	✓	X	✓	• CE: incomplete instruction for the test

[a] Checks (✓) indicate equivalence, while cross marks (X) indicate non-equivalence.

[b] SE: semantic equivalence.

[c] IE: idiomatic equivalence.

[d] CE: conceptual equivalence.

[e] EE: experiential equivalence.

To facilitate consistency of the translations throughout the tool, items Q1 to Q3 (pain and activity) were converted from declarative statements into questions. Choices for item Q3 (sleep, work, and recreation or sports) were deemed to pertain to different aspects of activities and were thus converted to stand-alone questions for each activity (Q3A, Q3B, and Q3C). To facilitate clarity, the literal translation of item Q13 was rephrased for easier understanding. To ensure patient safety when using the tool, precautions for severely painful shoulders were added to the instructions for the special test subdomain.

XSL•FO
RenderX

To facilitate experiential equivalence, item Q3B’s use of the term “*nakakapagtrabaho*” (able to work) was supplemented by the modifier “*gawaing bahay*” (house chores) to be more encompassing of the different types of work in the Filipino culture. To facilitate conceptual equivalence, item Q7’s (ROM) use of “*bulsa sa likod ng apektadong balikat*” (back pocket on the affected side) was changed to “*pigi*” (buttock) as it pertains to the same area and is more relatable and easily understandable. Likewise, item Q17 was clarified with the phrase “*paharap sa lebel ng braso*” (forward raising to arm level) to demonstrate shoulder flexion to 90 degrees as in the Speed test. To facilitate idiomatic equivalence, instruction for the special test subdomain was improved from “*na makakapagsabi sa amin*” (that may tell us) to “*upang malaman namin*” (so that we will know) as the test demands active patient response. Finally, the expert committee all suggested taking new pictures for the tool with a Filipino-looking model for the pictorial guide for it to be more relatable and also to facilitate minor improvements in the angle of the shots and designation of the task of each arm.

Pretest

There were 12 participants in the pretest, with the majority being females (10/12, 83%), aged 50 - 59 years (5/12, 42%) with a mean age of 53 (SD 12) years, married or cohabiting (5/12, 42%), and having finished tertiary education (8/12, 67%; Table 3). Most were unemployed (6/12, 50%), while those employed were engaged in nonhealth-related work (8/12, 67%). The monthly family income of participants was positively skewed, with the majority (7/12, 58%) earning Philippine 5000 to 10,000 (US $100 - $200).

XSL•FO
RenderX

Table . Sociodemographic profile of the participants.

Characteristics	Pretest (n=12)	Pilot (n=47)
Age group (y), n (%)		
19 - 29	0 (0)	2 (4)
30 - 39	2 (2)	5 (11)
40 - 49	2 (2)	3 (6)
50 - 59	5 (42)	21 (45)
≥60	3 (25)	16 (34)
Age (y), mean (SD)	53 (12)	55 (14)
Sex, n (%)		
Female	10 (83)	35 (75)
Male	2 (17)	12 (26)
Civil status, n (%)		
Single	4 (33)	10 (21)
Married or cohabiting	5 (42)	29 (62)
Separated or divorced	1 (8)	2 (4)
Widowed	2 (17)	6 (13)
Educational status, n (%)		
Primary	0 (0)	4 (9)
Secondary	4 (33)	15 (32)
Tertiary	8 (67)	26 (55)
Postgraduate	0 (0)	2 (4)
Employment, n (%)		
Student	0 (0)	2 (4)
Employed	4 (33)	15 (32)
Unemployed	6 (50)	23 (49)
Retired	2 (17)	7 (15)
Type of work, n (%)		
Health-related work	4 (33)	7 (15)
Nonhealth-related work	8 (67)	40 (85)
Family monthly income (Philippine [US $]), n (%)		
<5000 (<100)	2 (17)	21 (45)
5000-10,000 (100 - 200)	7 (58)	9 (19)
10,001-20,000 (201 - 400)	2 (18)	8 (17)
20,001-50,000 (401 - 1000)	1 (8)	6 (13)
>50,000 (>1000)	0 (0)	3 (6)

Items on pain and activity (Q1 and Q3A-Q3C), ROM (Q2-Q6), and special tests (Q15) were performed correctly by at least 75% (9/12) of the participants independently. Up to 25% (3/12) of the participants needed cueing in the whole pain and activity subdomain and special tests (Q14). Less than 25% of the participants were able to perform ROM items Q8 and Q9, all strength items, and half of the items in the special test subdomain.

Six out of 12 (50%) participants reported an increase in pain upon performance of certain maneuvers (eg, items Q4, Q5, Q7, Q9, and Q14), although all were still able to complete the tasks. They were advised to seek outpatient consultation at the Rehabilitation Medicine Outpatient Clinic if the pain persisted.

Six (50%) participants found the tool easy to follow, while 5 (42%) found it difficult. Visual aids in the form of pictures and arrows were most helpful in making the tool easily

XSL•FO
RenderX

understandable. The presence of a caregiver was helpful for two of the participants.

The use of some Tagalog words (such as *pigi* or buttocks in English) was the primary barrier for understanding the tool (Table 4). Five (42%) participants were content with the prefinal version of the tool and had no suggested changes. One (8%) participant suggested that the presence of a physician or a caregiver (at the least) was still necessary to guide patients and ensure accuracy and safety in following the tool. Finally, 6 (50%) participants reiterated the need for pictures, and 1 (8%) participant suggested improvement in the portrayal of movement in the pictures.

Table . Results from the thematic analysis related to understandability of the Filipino Shoulder Telehealth Assessment Tool (STAT) and factors that make it easy or difficult to understand.

	Pretest (n=12), n (%)	Pilot (n=46), n (%)
Ease of understanding		
Tool is easy	6 (50)	13 (28)
Use of pictures and arrows	10 (83)	11 (24)
Presence of caregiver	2 (17)	3 (6)
Tool is difficult	5 (42)	11 (24)
Use of some Tagalog words	3 (25)	7 (15)
Nonfamiliarity with the tool	1 (8)	3 (6)
Difficulty reading the text	2 (17)	3 (6)
Poor internet connection	2 (17)	2 (4)
Poor audiovisual setup	1 (8)	1 (2)
Not adept with technology	1 (8)	1 (2)
Absence of caregiver	0 (0)	1 (2)
Nonideal venue	1 (8)	1 (2)
Use of arrows	1 (8)	0 (0)

Lessons From the Pretest and Development of the Final Filipino STAT

Difficulty reading the text of the tool and tendency to skip a specific question (Q3A) in a predominantly older adult population suggested inappropriate user interface design to the study team. While readers with advancing age have varying levels of possible age-related cognitive and visual decline, inclusivity was ensured in redesigning the tool [18]. These lessons were applied to the final version of the Filipino STAT (Multimedia Appendix 5) as follows: breaking down chunks of texts into bullet points, using Sans Serif typefaces with at least a 16-point font size, and providing white spaces and appropriate contrasts [18,19].

The conduct of special tests revealed an average of 73% (SD 14%) positive test findings (painful) that might have contributed to the increase in shoulder pain among pretest participants. This was addressed in the final version by placing the special test subdomain last as in clinical examination so any pain would not get in the way of the conduct of the other parts of the assessment.

Specific pretest errors could be classified as (1) misinterpretations observed in the subdomains of pain and activity (items Q2 and Q3C) and ROM (Q8); (2) differences in the semantic understanding of specific body parts, evident in ROM items Q1, Q4, Q5, Q7, and Q8; (3) difficulty understanding long instructions, which were evident in ROM item Q9 and all strength and special test items; and (4) inconsistent experiential equivalence in special test item Q17. Text qualifiers were added to misinterpreted items such as Q2, labeling 0=“normal” and 100=“hindi normal” (not normal), to avoid reversal of the scale. Items with different potential semantic meanings were significantly improved when pictorial guides were provided. Long, unclear instructions were broken down into short, step-by-step instructions with pictorial guides in each step. Finally, in item Q17 (Speed test), instead of holding a plate forward, most participants held a plate on the side, as a waiter would. The phrase was therefore changed from “*parang may hawak na plato*” (like holding a plate) to “*parang nanghihingi habang nakaunat ang siko*” (like reaching out or asking for something with an outstretched arm) to improve participants’ accuracy in performing the task and the relatability of the task in Filipino culture.

Final Filipino STAT Pilot Study

There were 47 participants in the pilot study; one of them did not complete the video consultation and withdrew due to technical difficulties (ie, unstable internet) on the patient’s end. The majority of the participants were females (35/47, 75%), aged 50 - 59 years (21/47, 45%) with a mean age of 55 (SD 14) years, married or cohabiting (29/47, 62%), and finished tertiary education (26/47, 55%; Table 3). Most were unemployed (23/47, 49%), while those employed were mostly engaged in nonhealth-related work (40/47, 85%). The monthly family income of the participants was positively skewed, with the majority (21/47, 45%) earning less than Philippine 5000 (US $<100).

XSL•FO
RenderX

For each of the items under the pain and activity subdomains, 35 (75%) to 43 (91%) participants were able to perform the tasks correctly. Meanwhile, some needed cueing from their caregiver to perform the tasks correctly, particularly for item Q2, where 11 (23%) participants needed help. With items Q3A and Q3C, 1 out of the 47 (2%) participants was not able to perform the task completely.

The ROM, strength, and special test subdomains were tested both without and with pictorial guides. For ROM-related items, a lot of participants were not able to perform the tasks correctly without a picture (Q2: chin, Q7: back pocket, Q8: lower back, and Q9: wall touching). In contrast, for those items with picture guides, there was a greater number of patients who were able to perform the tasks correctly, with little to no need of cueing from their caregivers. Four (8%) participants showed compensatory movements with some items, yet they answered yes when asked if they could do the tasks. Hence, they were considered incorrect task performances, with item Q8 being the most common item that resulted in compensatory movements.

Without pictorial guides, more than half of the participants were not able to perform all the tasks under the strength subdomain. Slightly more participants correctly performed the tasks for items Q11 and Q12 when shown pictorial guides. Moreover, the same scenario was observed for all items in the special test subdomain—having pictorial guides improved the number of participants able to perform the tasks correctly from 34 (74%) to 39 (87%) participants, while 2 (4%) to 6 (13%) participants needed cueing from their caregivers to be able to perform the tasks correctly.

Among the 46 participants who shared their experience with the final Filipino STAT, 13 (28%) found the tool easy to understand, while 11 (24%) found it difficult. The pictures and arrows were found to be most helpful for participants. Participant 47 remarked, "*maayos po binigay ang panuto at dahil sa larawan ay naintindihan po nang mabuti*" (the instructions were given properly, and because of the pictures, they were better understood). The use of some Tagalog words, however, was considered the primary barrier by the majority of the participants, followed by their nonfamiliarity with the tool and difficulty reading the text (Table 4). The same participant also remarked, "*May kahinaan po ako sa Pilipino, tulad po ng pigi po, hindi ko siya naintindihan*" (I have some difficulty in understanding Tagalog, for example, I did not understand the Tagalog word for buttock").

Content Validity and Internal Consistency Testing of the Final Filipino STAT

Content validity through scale validity indices was deemed excellent at 0.97 using the item-level content validity index and 0.80 using the universal agreement. The κ coefficients were excellent across all items, supporting that the degree of agreement among the experts was beyond chance [20]. A Cronbach α score of 0.87 indicated that all test items were unidimensional, or that performance on the items could be explained in terms of a single underlying factor (Multimedia Appendix 6).

Discussion

Principal Findings

The Filipino STAT has an excellent content validity (scale validity index=0.80 - 0.97), excellent interrater reliability (κ coefficient=0.82 - 1.00), and good internal consistency (Cronbach α=0.87). Its understandability is excellent for pain and activity (98%), good for ROM and special tests (85%), and poor for strength (37%).

Comparison With Prior Work

The findings of this study are consistent with the existing literature on shoulder teleconsultation psychometric properties. Internet assessment of rehabilitation outcomes such as pain, ROM, muscle strength, and functional assessment had good concurrent validity. There is a strong agreement between virtual and in-person examination for ROM (87.4% agreement; x^2=30.782; P<.001), and diagnosis (85.1% agreement; κ=0.82, 95% CI) [21-23]. Further studies on the Filipino STAT compared against in-person examination or imaging modalities may be done to ascertain concurrent validity of the tool.

Strengths and Limitations

Since a direct translation of a tool from an original English version may not reflect the cultural nuances and context of another culture [13], this study observed a rigorous and standard linguistic validation and cross-cultural adaptation to ensure a culturally appropriate, valid, and reliable translation. Although the original STAT has not been validated as a whole tool, the Filipino-translated tool has been subjected to content validity and internal consistency testing in this study. The collective inputs from the consensus committee of experts, the understandability questionnaire, and qualitative data from the participants ensured that the final tool can be used in the clinical setting.

There are some limitations to this study. First, the assessment of ease of understanding of the tool using the open-ended questionnaire may have been confounded by the consecutive performance of the tool without and with a pictorial guide, as well as the videoconsultation study setup. The test-taking that was twice as long, with no visual aid provided at the start, could have made the consultation more difficult for the participants. The barriers to the videoconsultation study setup were consistent with previous literature, such as unstable internet connection and poor audiovisual resolution [5]. In the envisioned clinical performance of the Filipino STAT, only the pictorial guide will be performed before the actual consultation, thereby eliminating these concerns.

Second, understandability per subdomain was excellent for pain and activity, good for ROM and special tests, and poor for strength. The incidence of compensatory or trick movement was accounted for in this study and was most frequently seen in the ROM subdomain, particularly item Q8 (50 degrees of internal rotation).

Finally, such as in an in-person physical examination, incorrect performance of a virtual test, from poor understandability or compensatory movement, makes its results invalid and

XSL•FO RenderX

questionable. This shall serve as a reminder that the Filipino STAT is not meant to replace actual in-person consultations and should just aid the clinician in assessing the patient and make clinical work efficient. The clinician must be prudent in confirming the initial Filipino STAT findings during the actual consultation, as necessary. Further studies may consider the use of instructional videos to improve understandability of the tool, especially for the strength subdomain.

Future Directions

Both pretest and pilot study findings reveal significant improvements in the correct performance of the Filipino STAT tasks with pictorial guides, compared to none. The majority of the participants from both phases of the study also shared a positive perception of the use of pictures and arrows and the presence of a caregiver. Thus, it is the study's recommendation to use the Filipino STAT with a pictorial guide, as intended in the original tool [6]. The association of the presence of a caregiver in the successful performance of the Filipino STAT was beyond the scope of this present study. Nonetheless, having a caregiver around may help with the conduct of the Filipino STAT but should not discourage those who do not have an available caregiver.

Conclusions

The development of the Filipino STAT through a rigorous linguistic validation and cultural adaptation ensured a culturally appropriate, valid, and reliable translation. Understandability and ease of understanding by end-users are also critical to assess in patient-reported outcome measures. The pain and activity, ROM, and special test subdomains of the Filipino STAT may be used for clinical assessment, while the strength subdomain needs further improvement on understandability.

Acknowledgments

The authors would like to acknowledge the clinical experts Dorothy Dy Ching Bing-Agsaoay, Monalisa Lim-Dungca, Christopher Constantino, Nathaniel Napa, Mitchelle Gabuya, and Frances Carlos, as well as our Filipino STAT model Nancy Cheng, for their generous contributions to this study. They would also like to acknowledge the STAT original authors, Gregory Sprowls, Jaycen Brown, and Brett Robin, for allowing translation of their tool.

Funding

This research study was funded by the Philippine General Hospital—Expanded Hospital Research Office.

Data Availability

The data that support the findings of this study are uploaded online as supplementary files and may be requested from the corresponding author, JA, upon reasonable request.

Authors' Contributions

Conceptualization: JA, SI, JAM, CFL
Data curation: JA
Funding acquisition: SI
Implementation: JA
Methodology: JA, SI, JAM, CFL
Supervision: SI, JAM, CFL
Writing – original draft: JA
Writing – reviewing & editing: SI, JAM, CFL

Conflicts of Interest

None declared.

Multimedia Appendix 1
The content validation form.
[PDF File, 2546 KB - rehab_v13i1e67974_app1.pdf]

Multimedia Appendix 2
The expert consensus committee guide on equivalence.
[PDF File, 114 KB - rehab_v13i1e67974_app2.pdf]

Multimedia Appendix 3
The observational checklist.
[PDF File, 50 KB - rehab_v13i1e67974_app3.pdf]

XSL•FO
RenderX

Multimedia Appendix 4
The open-ended interview questions.
[PDF File, 48 KB - rehab_v13i1e67974_app4.pdf]

Multimedia Appendix 5
The final Filipino Shoulder Telehealth Assessment Tool (STAT).
[PDF File, 9454 KB - rehab_v13i1e67974_app5.pdf]

Multimedia Appendix 6
Supplemental dataset to a full manuscript published in the *J Med Internet Res*.
[DOCX File, 29 KB - rehab_v13i1e67974_app6.docx]

References

1. Bos WH, van Tubergen A, Vonkeman HE. Telemedicine for patients with rheumatic and musculoskeletal diseases during the COVID-19 pandemic; a positive experience in the Netherlands. Rheumatol Int 2021 Mar;41(3):565-573. [doi: 10.1007/s00296-020-04771-6] [Medline: 33449162]
2. D Leochico CF. Adoption of telerehabilitation in a developing country before and during the COVID-19 pandemic. Ann Phys Rehabil Med 2020 Nov;63(6):563-564. [doi: 10.1016/j.rehab.2020.06.001] [Medline: 32544528]
3. Leochico CFD, Mojica JAP, Rey-Matias RR, Supnet IE, Ignacio SD. Role of telerehabilitation in the rehabilitation medicine training program of a COVID-19 referral center in a developing country. Am J Phys Med Rehabil 2021 Jun;100(6):526-532. [doi: 10.1097/PHM.0000000000001755] [Medline: 33998606]
4. Leochico CFD, Montemayor GAM, Obeles AJT, Ong BAG. Telerehabilitation in a developing country toward the tail end of the COVID-19 pandemic: is it here to stay? Acta Med Philipp 2022 Nov. [doi: 10.47895/amp.vi0.5905]
5. Leochico CFD, Espiritu AI, Ignacio SD, Mojica JAP. Challenges to the emergence of telerehabilitation in a developing country: a systematic review. Front Neurol 2020;11:1007. [doi: 10.3389/fneur.2020.01007] [Medline: 33013666]
6. Sprowls GR, Brown JC, Robin BN. The Shoulder Telehealth Assessment Tool in transition to distance orthopedics. Arthrosc Tech 2020 Nov;9(11):e1673-e1681. [doi: 10.1016/j.eats.2020.07.008] [Medline: 33294325]
7. McLean JM, Awwad D, Lisle R, Besanko J, Shivakkumar D, Leith J. An international, multicenter cohort study comparing 6 shoulder clinical scores in an asymptomatic population. J Shoulder Elbow Surg 2018 Feb;27(2):306-314. [doi: 10.1016/j.jse.2017.08.016] [Medline: 29221757]
8. Reed MD, Van Nostran W. Assessing pain intensity with the visual analog scale: a plea for uniformity. J Clin Pharmacol 2014 Mar;54(3):241-244. [doi: 10.1002/jcph.250] [Medline: 24374753]
9. Retzky JS, Baker M, Hannan CV, Srikumaran U. Single Assessment Numeric Evaluation scores correlate positively with American Shoulder and Elbow Surgeons scores postoperatively in patients undergoing rotator cuff repair. J Shoulder Elbow Surg 2020 Jan;29(1):146-149. [doi: 10.1016/j.jse.2019.05.039] [Medline: 31401127]
10. Oosterwijk AM, Nieuwenhuis MK, van der Schans CP, Mouton LJ. Shoulder and elbow range of motion for the performance of activities of daily living: a systematic review. Physiother Theory Pract 2018 Jul;34(7):505-528. [doi: 10.1080/09593985.2017.1422206]
11. Capio CM, Maranion DZ, Pascua HE. Musculoskeletal pain and dysfunction among University of the Philippines Manila personnel. Univ Philippines Manila J 2001;6(4):25-23 [FREE Full text]
12. Labao HC, Faller EM, Bacayo MFD. “Aches and pains” of Filipino migrant workers in Malaysia: a profile of work-related musculoskeletal disorders. Ann Glob Health 2018 Aug;84(3):474-480. [doi: 10.29024/aogh.2331] [Medline: 30835403]
13. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993 Dec;46(12):1417-1432. [doi: 10.1016/0895-4356(93)90142-n] [Medline: 8263569]
14. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
15. Bonett DG. Sample size requirements for testing and estimating coefficient alpha. J Educ Behav Stat 2002 Dec;27(4):335-340. [doi: 10.3102/10769986027004335]
16. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. SPINE (Phila Pa 1976) 2000 Dec;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
17. Lorca LA, Torres-Castro R, Ribeiro IL, et al. Linguistic validation and cross-cultural adaptation of the post–COVID-19 functional status scale for the Chilean population. Am J Phys Med Rehabil 2021 Apr 1;100(4):313-320. [doi: 10.1097/PHM.0000000000001706] [Medline: 33496442]
18. Office of Disease Prevention and Health Promotion. Section 5.3 use a readable font that’s at least 16 pixels. Health Literacy Online, 3rd Edition. 2025. URL: https://odphp.health.gov/healthliteracyonline/design-easy-scanning/use-readable-font-thats-least-16-pixels [accessed 2025-12-01]

XSL•FO
RenderX

19. Hou G, Anicetus U, He J. How to design font size for older adults: a systematic literature review with a mobile device. Front Psychol 2022;13:931646. [doi: 10.3389/fpsyg.2022.931646] [Medline: 35978796]
20. Rodrigues IB, Adachi JD, Beattie KA, MacDermid JC. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. BMC Musculoskelet Disord 2017 Dec;18(1):540. [doi: 10.1186/s12891-017-1914-5] [Medline: 29258503]
21. Steele L, Lade H, McKenzie S, Russell TG. Assessment and diagnosis of musculoskeletal shoulder disorders over the internet. Int J Telemed Appl 2012;2012:945745. [doi: 10.1155/2012/945745] [Medline: 23193395]
22. Rabin A, Dolkart O, Kazum E, et al. Shoulder assessment by smartphone: a valid alternative for times of social distancing. Arch Orthop Trauma Surg 2022 Jun;142(6):979-985. [doi: 10.1007/s00402-021-03762-x] [Medline: 33439302]
23. Mani S, Sharma S, Omar B, Paungmali A, Joseph L. Validity and reliability of internet-based physiotherapy assessment for musculoskeletal disorders: a systematic review. J Telemed Telecare 2017 Apr;23(3):379-391. [doi: 10.1177/1357633X16642369] [Medline: 27036879]

Abbreviations

ROM: range of motion
SANE: single assessment numeric evaluation
STAT: Shoulder Telehealth Assessment Tool
VAS: visual analog scale

Edited by N Cahill; submitted 25.Oct.2024; peer-reviewed by T Lienou, V Prieto; revised version received 04.Nov.2025; accepted 11.Nov.2025; published 20.Jan.2026.

Please cite as:
Arboleda J, Ignacio S, Mojica JA, Leochico CF
Linguistic Validation and Cross-Cultural Adaptation of the Shoulder Telehealth Assessment Tool for Filipino Patients with Musculoskeletal Shoulder Condition: Cross-Sectional Study
JMIR Rehabil Assist Technol 2026;13:e67974
URL: https://rehab.jmir.org/2026/1/e67974
doi:10.2196/67974

© Jeffrey Arboleda, Sharon Ignacio, Jose Alvin Mojica, Carl Froilan Leochico. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Role of Medical Caption Technology to Support Physician-Patient Communication for Patients With Hearing Loss: Mixed Methods Pilot Study

Sarah E Hughes[1*], BA; Liang-Yuan Wu[2*], MS; Lindsay J Ma[1], MD; Dhruv Jain[2], PhD; Michael M McKee[3,4], MD, MPH

[1]University of Michigan Medical School, Ann Arbor, MI, United States
[2]Department of Computer Science and Engineering, University of Michigan, Ann Arbor, MI, United States
[3]Department of Family Medicine, University of Michigan, 1018 Fuller St., Ann Arbor, MI, United States
[4]Department of Physical Medicine and Rehabilitation, University of Michigan Medical School, Ann Arbor, MI, United States
[*]these authors contributed equally

Corresponding Author:
Michael M McKee, MD, MPH
Department of Family Medicine, University of Michigan, 1018 Fuller St., Ann Arbor, MI, United States

Abstract

Background: Speech recognition technology is widely used by individuals who are Deaf/deaf and hard-of-hearing (DHH) in everyday communication, but its clinical applications remain underexplored. Communication barriers in health care can compromise safety, understanding, and autonomy for individuals who are DHH.

Objective: This study aimed to evaluate a real-time speech recognition system (SRS) tailored for clinical settings, examining its usability, perceived effectiveness, and transcription accuracy among users who are DHH.

Methods: We conducted a pilot study with 10 adults who are DHH participating in mock outpatient encounters using a custom SRS powered by Google's speech-to-text application programming interface. We used a convergent parallel mixed-methods design, collecting quantitative usability ratings and qualitative interview data during the same study session. These datasets were subsequently merged and jointly interpreted. Participants completed postscenario surveys and structured exit interviews assessing distraction, trust, ease of use, satisfaction, and emotional response. Caption accuracy was benchmarked against professional communication access real-time translation transcripts using word error rate (WER). Because WER assigns equal weight to all tokens, it does not differentiate between routine transcription errors and those involving safety-critical clinical terms (eg, medications or diagnoses). Therefore, WER may underestimate the potential impact of certain errors in medical contexts.

Results: Across 29 clinical scenario simulations, 86% (25/29) of participants found captions nondistracting, 90% (26/29) reported them easy to follow and trustworthy, and 76% (22/29) were satisfied with the experience. Participants described the SRS as intuitive, emotionally grounding, and preferable to lip reading in masked settings. WER ranged from 12.7% to 22.8%, consistent with benchmarks for automated SRSs. Interviews revealed themes of increased confidence in following clinical conversations and staying engaged despite masked communication. Participants reported less anxiety about missing critical medical information and expressed a strong interest in expanding the tool to real-world settings, especially for older adults or those with cognitive impairments.

Conclusions: Our findings support the potential of real-time captioning to enhance accessibility and reduce the cognitive and mental burden of communication for individuals who are DHH in clinical care. Participants described the SRS as both functionally effective and personally empowering. While accuracy for complex medical terminology remains a limitation, participants consistently expressed trust in the system and a desire for its integration into clinical care. Future research should explore real-world implementation, domain-specific optimization, and the development of user-centered evaluation metrics that extend beyond transcription fidelity to include trust, autonomy, and communication equity.

(JMIR Rehabil Assist Technol 2026;13:e79073) doi:10.2196/79073

KEYWORDS

health communication; hearing loss; deafness; speech recognition software; usability testing; health care accessibility

XSL•FO
RenderX

Introduction

Effective communication is foundational to safe, equitable, and high-quality health care [1]. However, individuals who are Deaf/deaf and hard-of-hearing (DHH) often face communication barriers that compromise understanding and autonomy [2]. These barriers contribute to poor health outcomes and reduced patient engagement in real-time clinical settings [2]. The scale of this issue highlights the need to understand which communication support tools are available and provided, and to whom. In the United States, an estimated 48 million people live with some degree of hearing loss (HL), and 1 in 3 adults older than 65 years experiences disabling age-related hearing loss [3,4]. Despite this growing population, access to communication supports remains inconsistent [5,6].

Deaf individuals who use American Sign Language often receive interpreter services [7]. In contrast, oral communicators with people with HL who normally rely on spoken English are less likely to receive accommodations such as captioning, assistive listening devices, or environmental modifications [7]. Especially in clinical workflows, interpreter services are systematically implemented, whereas accommodations for oral communicators are likely not [8-12]. This gap persists despite longstanding mandates under the Americans with Disabilities Act, which mandates effective communication in health care [13]. As a result, many patients who are DHH still receive incomplete or delayed health information [5]. These gaps undermine informed decision-making, autonomy, and overall care outcomes [14,15]. Far from logistical oversights, these structural inequities perpetuate persistent disparities in care for individuals who are DHH.

These long-standing disparities became even more visible during the COVID-19 pandemic [14]. Universal masking eliminated lip reading and facial cues, which were essential supports for many individuals who are DHH and rely on oral communication [16]. This shift underscored the need for scalable solutions to maintain accessible communication in high-stakes settings [14,17].

Real-time captioning is 1 solution for improving communication access for individuals who are DHH when traditional strategies (eg, lip reading or interpreters) are unavailable [18]. Captioning tools can be deployed quickly and readily support both in-person and virtual communication [19]. However, captioning accuracy of clinical conversations may be affected by terminology unique to the medical field or speaker attribution and is understudied [19,20]. This has left a critical gap in the development of effective and equitable access tools.

By allowing both conversation partners to see each other's faces while reading the same captions, transparent or dual-visibility captioning preserves the natural flow of spoken interaction and is a promising solution for clinical communication. Prior work, such as See-Through Captions [21], See-Through Captions in a Museum Guided Tour [22], and Wearable Subtitles [23], has primarily focused on general or educational settings. Our study extends this line of research into medical contexts, where communication accuracy can directly affect patient safety and outcomes. It also emphasizes the emotional and psychological impact of captioning during clinical interactions and addresses the unique technical challenges posed by medical vocabulary and workflow integration.

In summary, we developed and evaluated a real-time captioning tool using Google's speech-to-text engine to generate live captions during simulated clinical encounters. We tested this system in dynamic, medically relevant scenarios designed to simulate typical ambulatory care encounters. In this pilot study, we explored how individuals who are DHH experienced the captioning system in these simulated encounters, focusing on usability, accuracy, and communication access.

Methods

Background

The pilot took place in a patient room at one of the Department of Family Medicine clinics. The primary goal was to assess the feasibility and acceptability of a real-time captioning tool in a clinical setting. Secondary objectives included evaluating ease of use, distraction, trust, and satisfaction, factors critical to determining whether the tool supports communication access. Quantitative and qualitative data were collected concurrently within the same study session using a convergent parallel mixed-methods design. Participants completed postscenario surveys and a brief structured exit interview during the same visit, allowing us to analyze both datasets in parallel before merging findings during the interpretation phase.

Recruitment

We recruited participants who self-identified as DHH through internal email lists compiled from prior studies, social media, and snowball sampling. Inclusion criteria included people who were DHH, preferred to communicate in spoken English, and were at least 18 years old. Recruitment materials explained that the study evaluated a real-time captioning system in simulated medical scenarios.

Mock Clinical Scenarios

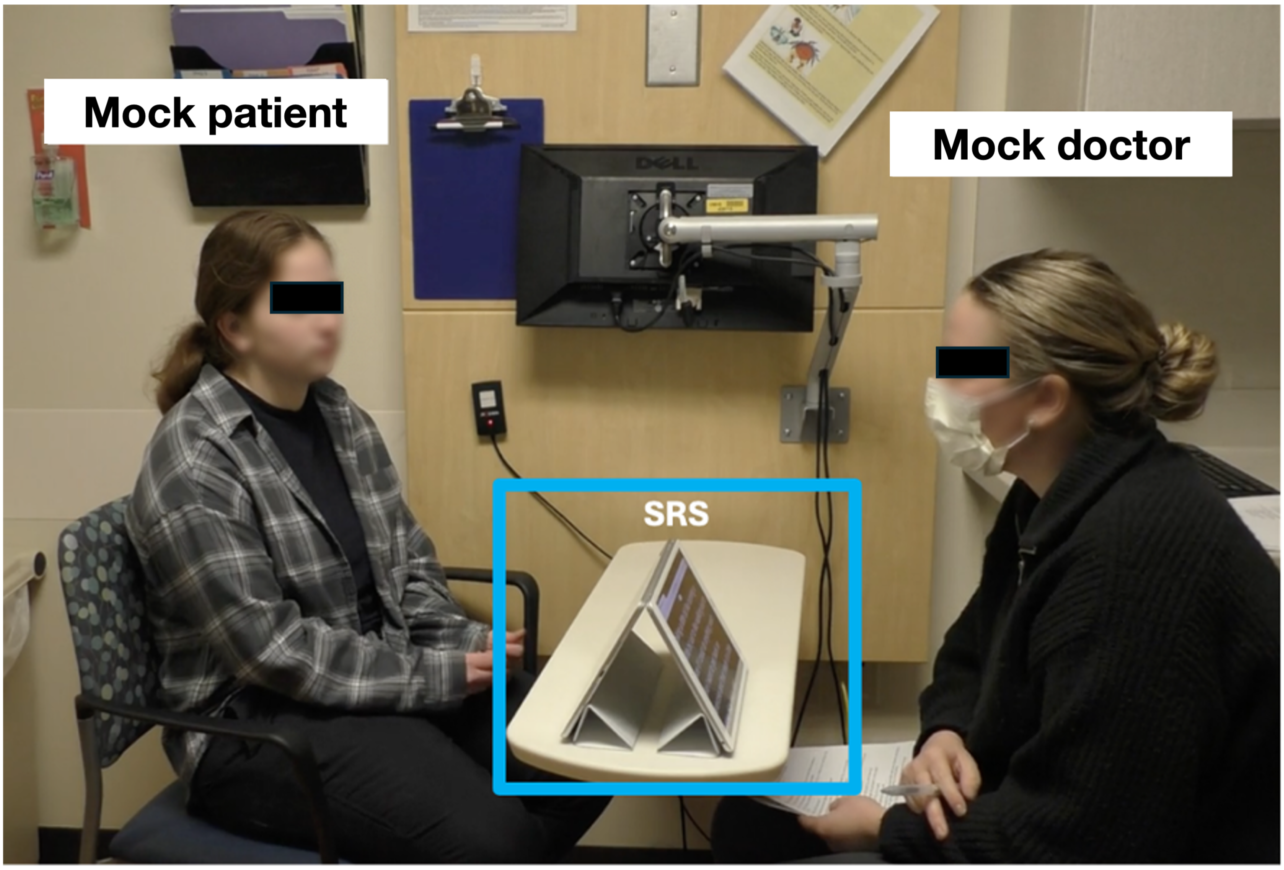
Participants completed 3 mock clinical scenarios using the automated speech recognition system (SRS) which was developed by us. The SRS used Google's speech-to-text application programming interface to transcribe speech to text with low latency and competitive accuracy [24]. The setup included 2 iPads arranged in a tented position so that each device faced either the participant or the mock doctor. Both iPads displayed the generated captions simultaneously (Figure 1).

Before each experiment, we used a random number generator to assign scenario order for each participant. Two team members (both medical students) alternated between serving as the mock doctor (administering scenarios) or facilitator (administering postscenario surveys and exit interviews).

The scenarios were based on commonly reported primary care concerns: (1) back pain, (2) headache, and (3) high blood pressure. Scenario scripts were designed by trained medical students and a clinical faculty member to closely replicate real clinical conversations. The mock doctors wore surgical masks to simulate real-life communication barriers, such as muffled sound and loss of visual cues.

XSL•FO
RenderX

Figure 1. An example of a mock clinical scenario with the real-time speech recognition system set up on a table between the participant (left) and the mock doctor (right). A microphone on the iPad facing the mock doctor detects audio during interviews. Transcripts are displayed on both iPads in real time. SRS: speech recognition system.



Postscenario SRS Assessments

Following each scenario, participants provided feedback on the captioning system, rating it across 4 domains: distraction, ease of use, trust, and overall satisfaction (Multimedia Appendix 1). Scenario-specific questions included: "In your discussion with the mock doctor, how distracting were the captions?" "How easy or difficult was it to watch the caption while talking with the mock doctor?" "How much did you trust the accuracy of the generated captions?" "In this scenario, how satisfied were you with the captioning technology?" To reduce response bias, we alternated the direction of the scales: ease of use and trust rated from 1 (strong agreement) to 5 (strong disagreement), and satisfaction rated from 1 (strong disagreement) to 5 (strong agreement). Distraction was scaled separately from 1 (strong disagreement) to 3 (strong agreement).

Participant Survey Questions

To evaluate user experience with the SRS, participants completed a structured exit interview consisting of 9 questions (5 scalar and 4 open-ended items; Multimedia Appendix 1). To ensure accessibility, a study team member read all questions aloud while they were displayed on an iPad (Apple Inc). We audio-recorded and transcribed responses verbatim using a third-party service, then deidentified the transcripts. We reviewed audio files to clarify unclear segments. Given the brief interviews, we organized and analyzed responses in Microsoft Excel (version 16.77).

Open-ended responses were reviewed using a structured framework aligned with predefined domains: ease of use, comfort, satisfaction, trust, emotional response, and the captioning system's ability to support or replace lip reading. Overall, 3 team members (SEH, LJM, and LW) independently applied initial codes to a subset of transcripts. Coding discrepancies were resolved through discussion, and the codebook was refined iteratively. Consistency was maintained through regular team meetings, and reflexive discussions were used to address potential bias.

Themes were identified based on frequency, relevance to study aims, and salience across participants. Representative participant comments were selected to illustrate key insights. Thematic saturation was reached when no new concepts emerged from successive interviews.

Mixed Methods Integration

To integrate quantitative and qualitative data, we used a convergent parallel approach in which both datasets were collected during the same phase, analyzed separately, and then merged during the interpretation phase. Integration occurred through (1) narrative weaving of findings across domains and (2) construction of a joint display that juxtaposed quantitative ratings with representative qualitative insights to generate meta-inferences. This approach allowed identification of areas of convergence and divergence between usability ratings and participants' lived communication experiences.

XSL•FO
RenderX

Closed Captioning Accuracy

In addition to participant feedback, we analyzed the accuracy of the system's transcriptions. We compiled all transcripts generated by the mock doctors and compared them to professional communication access real-time translation transcripts.

We used word error rate (WER), a standard metric in automatic speech recognition (ASR) that calculates errors as the ratio of insertions, deletions, and substitutions required to align the system output with the reference [25,26]. We implemented WER calculations using the Python-based *jiwer* library, which provides standardized scoring for automated SRSs. This approach allowed us to assess how closely the SRS-generated captions matched professional-level transcription, validating the system's effectiveness in realistic use cases.

Statistical Analysis

We performed univariate analyses on demographic data and postscenario survey responses. Because of the small sample size, the study was not powered to detect subgroup differences. For transcript analysis, we segmented transcripts from mock sessions into 3 distinct scenarios. To focus on the primary use case, captioning clinician speech, we excluded utterances from participants who are DHH and analyzed only the mock doctors' speech.

Ethical Considerations

This study was approved by the University of Michigan Institutional Review Board (IRB; HUM00240244). All participants provided informed consent prior to participation. Participants were informed of the study purpose, procedures, potential risks, and their right to withdraw at any time without penalty. All study data were deidentified prior to analysis, and transcripts were reviewed to remove personally identifiable information. Audio recordings and transcripts were stored on secure, password-protected institutional servers accessible only to the study team. Participants received a US $25 Amazon gift card for their participation. The individuals depicted in the figure provided explicit written consent for publication of their images. The individuals shown in Figure 1 provided explicit written consent for their images to be published.

Results

Participant Characteristics

Overall, 11 participants who are DHH enrolled and participated in the pilot study. Due to equipment failure resulting in complete data recording loss with Participant 5, this participant was excluded from the analysis. The 10 remaining participants had an even distribution of genders (Table 1).

Table . Study participant demographics.

ID	Age (years)	Sex	Identity	Hearing loss levels[a]	Wearable technology[b]	Lip reader
P01	61	Female	HoH[c]	Severe	Yes	All of the time
P02	66	Male	HoH	Severe	Yes	Sometimes
P03	66	Male	HoH	Severe	Yes	No
P04	66	Female	HoH	Moderately severe	Yes	Sometimes
P06	43	Female	Deaf	Profound	Yes	All of the time
P07	21	Female	deaf	Moderately severe	Yes	Sometimes
P08	39	Female	HoH	Severe	Yes	All of the time
P09	24	Male	Deaf	Profound	Yes	No
P10	56	Male	deaf	Profound	Yes	Sometimes
P11	20	Male	HoH	Mild	Yes	Sometimes

[a] Hearing loss levels were self-identified, and all participants reported equal hearing loss levels bilaterally.

[b] Wearable technology includes hearing aids and cochlear implants.

[c] HoH: hard of hearing.

The mean age of the participants was 46.2 (SD 19.3) years. Six participants identified as "hard of hearing," 22 as "Deaf," and 2 as "deaf." Seven participants self-reported severe to profound HL, and all participants had bilateral HL. Five participants self-reported congenital HL, 2 reported childhood onsets of HL (<12 y old), and 2 reported HL as adults (>18 y old). Hearing aids were used by 8 participants, and 2 participants used cochlear implants. Seven participants used captioning services in the past. Seven also incorporated smartphone-based hearing assistive technology. Three used "other" tools, including using cupped hands behind ears to assist in hearing. Eight participants reported varying degrees of dependence on lip reading, but 5 participants depended sometimes on lip reading and 5 depended fully on lip reading.

Postscenario SRS Assessments

There were 29 postscenario SRS assessment surveys, 3 survey responses each from 9 participants and 2 survey responses from 1 participant. One survey response from participant P11 was not collected due to a technician error. Overall, participants found the captioning technology not distracting in 86% (25/29) of scenarios (Table 2). In 90% (26/29) of scenarios, participants

XSL•FO
RenderX

trusted the accuracy of generated transcription and felt the captions were easy to watch while conversing with the mock doctor. In 76% (22/29) of scenarios, participants were satisfied with the captioning technology. The technology was least satisfying to participants in the back pain scenarios (70% satisfaction) compared to the high blood pressure (78% satisfaction) and headache (80% satisfaction) scenarios.

Table . Summary of participant assessments regarding live captioning technology compiled from all 3 scenarios and dichotomized.

Questions[a]	Assessments	Values, n (%)
In your discussion with the mock doctor, how distracting were the captions?	Not distracting[b]	25 (86)
How easy or difficult was it to watch the caption while talking with the mock doctor?	Easy[c]	26 (90)
How much did you trust the accuracy of the generated captions?	Trusted[d]	26 (90)
In this scenario, how satisfied were you with the captioning technology?	Satisfied[e]	22 (76)

[a] For all 4 questions, n=29 since 1 of the 10 participants did not participate in 1 of the 3 scenarios.

[b] Not distracting: not at all distracting.

[c] Easy: very easy + somewhat easy.

[d] Trusted: completely trusted + somewhat trusted.

[e] Satisfied: very satisfied + somewhat satisfied.

Participant Experience Surveys

All 10 participants completed structured exit interviews following the captioning scenarios, providing reflections on their overall experience with the SRS (Table 3). Interview responses were analyzed using a predefined framework aligned with domains explored in the postscenario ratings (eg, ease of use, comfort, satisfaction, trust, emotional impact, and support for lip reading). This section summarizes participant perspectives and provides representative quotes to contextualize the quantitative results described above.

Table . Representative participant reflections by theme.

Themes	Relevant quotes[a]	Interpretation
Ease of use	"At first I wasn't sure what to expect, but after a few lines of text I stopped even thinking about it—it just worked. That made me feel more in control." (P04)	Participants found the system intuitive and accessible.
Comfort	"I didn't have to strain or overthink. It just flowed naturally and I didn't even realize how relaxed I was until the end." (P08)	Technology reduced cognitive effort and fostered emotional ease.
Satisfaction	"I was happy. I wish all the doctors would have something like this. It made me feel like my experience mattered." (P03)	Participant expresses satisfaction and a sense of being valued.
Safety and trust	"Because it's live, it feels very safe. You're not left guessing, and I felt confident nothing important was missed." (P01)	Real-time functionality enhanced user confidence and perception of safety.
Emotional response	"I didn't realize how much stress I usually carry during appointments. This made me feel heard and like I could finally breathe." (P09)	System reduced communication-related anxiety and supported emotional well-being.
Support or replace lip reading	"With the mask on, it would have been extremely difficult to follow—and with the captioning, it was just leaps better. I wasn't exhausted from trying to read lips the whole time." (P07)	The technology was viewed as a vital alternative to lip reading, especially in masked settings.

[a] Relevant quotes from individual participants illustrating each core theme, including insight into perceived usability, comfort, satisfaction, emotional impact, and the role of real-time captions in supporting communication.

Most participants (9/10) described the system as easy to use, frequently using phrases like "very easy" or "easier than usual." One participant remarked,

XSL•FO
RenderX

> *After a few lines of text I stopped even thinking about it—it just worked. That made me feel more in control.*

Another noted,

> *It was easier than usual because we don't have captioning. It's always nice to have it just in case you miss something.*

Participants also reported high comfort with the system. Descriptions included "very comfortable," "easy to work with," and "high comfortability*.*"

As 1 participant shared,

> *It just flowed naturally, and I didn't even realize how relaxed I was until the end.*

Satisfaction was also high across interviews. While 76% of scenario ratings reflected satisfaction, all participants described themselves as satisfied or very satisfied in exit interviews. One stated,

> *I was happy. I wish all the doctors would have something like this.*"

Another shared,

> *I was pretty satisfied, and the captioning was spot-on.*

When asked about trust in the system, participants frequently described the captions as reliable. One participant reflected,

> *Because it's live, it feels very safe. You're not left guessing.*

A few raised questions about data privacy, with one noting,

> *I would also want to know what happens to the transcript and who has access to it.*

Participants also described emotional benefits from the technology. In total, 9 of 10 participants used words like "reassured," "relaxed," and "comfortable" to describe how the SRS made them feel. One participant shared,

> *This made me feel heard and like I could finally breathe.*

Perceptions of the captions' ability to support or replace lip reading were more varied. Several participants described the system as a helpful supplement or improvement, particularly in masked settings. As one noted,

> *With the mask on, I definitely depended on it more.*

Another stated,

> *I think it's better than lip reading.*

Others expressed that lip reading remained important, with one participant saying,

> *Not going to replace lip reading... captions help, but I still rely on visual cues.*

Beyond these predefined domains, participants spontaneously shared reflections on broader applications of the SRS. Several expressed enthusiasm for expanding its use in real-world clinical settings, with one stating,

> *I would like to see that in many doctor's offices tomorrow.*

Others suggested the system may be particularly helpful for patients who are older, have cognitive impairments, or use interpreters. A few noted that having real-time captions reduced the pressure to maintain constant visual attention, allowing for more natural communication and less fatigue.

Closed Captioning Accuracy

We collected and preprocessed transcripts from 10 mock clinical sessions. Due to varying levels of verbosity among the participants, the total transcript lengths varied substantially, ranging from 1144 to 4704 words.

Overall, participants found the SRS to be sufficiently accurate (Table 4). For instance, P04 noted that the system was *"more accurate than the phone captions"* she typically uses in daily conversations. Similarly, P06 commented on the system's effectiveness compared to human captioners, stating,

> *A lot of the captions I had were court reporters—they caption fast, but sometimes they make mistakes. ... And this one [the SRS], it's more accurate and I see words better.*

Nonetheless, participants expressed concerns about the system's ability to handle more complex or specialized medical vocabulary. For example, P10 questioned *"how it would be with more complex medical terminologies,"* in real clinical settings where more technical jargon and medication names were frequently used.

XSL•FO
RenderX

Table . The word error rate for each scenario, along with the accumulated word error rate for each participant across all 3 scenarios. These word error rate scores specifically reflect the accuracy of the automated speech recognition system in transcribing the mock doctors' speech.

ID	Mock doctor	Scenario 1[a]	Scenario 2[a]	Scenario 3[a]	Accumulated (range: 0.127-0.167)
P01	M1	0.136	0.133	0.125	0.131
P02	M2	0.193	0.141	0.133	0.153
P03	M1	0.129	0.122	0.133	0.127
P04	M2	0.228	0.128	0.151	0.167
P06	M2	0.137	0.152	0.135	0.144
P07	M1	0.127	0.134	0.132	0.133
P08	M2	0.155	0.142	0.152	0.149
P09	M1	0.136	0.131	0.141	0.137
P10	M1	0.127	0.126	0.133	0.129
P11	M2	0.147	0.185	0.134	0.151

[a] The scenario-level word error rates ranged between 0.122 and 0.228.

Joint Display of Integrated Findings

To illustrate convergence between quantitative usability ratings and qualitative interview themes, we constructed a joint display summarizing merged findings and resulting meta-inferences across key domains (Table 5).

Table . Joint display of integrated quantitative and qualitative findings

Domain	Quantitative result	Representative quote	Integrated meta-inference
Ease of use	90% rated captions "easy"	"After a few lines of text I stopped even thinking about it—it just worked."	High usability with minimal cognitive load Captions supported natural conversational flow
Comfort	Not directly measured	"It just flowed naturally, and I didn't realize how relaxed I was."	Technology reduced strain and fostered emotional ease during communication
Satisfaction	76% satisfied	"I wish all the doctors would have something like this."	Satisfaction tied to both functional value and feeling understood and supported
Safety and trust	90% trusted accuracy	"Because it's live, it feels very safe. You're not left guessing."	Real-time display strengthened perceived safety and reliability despite minor errors
Emotional response	Not directly measured	"This made me feel heard and like I could finally breathe."	Captions enhanced psychological safety and reduced anxiety—benefits not captured numerically
Support or replace lip reading	Not directly measured	"With the mask on, I depended on it more… it was leaps better."	Captions supplemented or replaced lip reading, reducing fatigue in masked settings

Discussion

Principal Results

To successfully deploy SRS in clinical settings, it is essential that the system accurately captures and reflects clinicians' speech. Our findings show that although the SRS output was not flawless, its WERs fell between 0.10 and 0.20, a range generally considered acceptable for real-world ASR use [19,26]. Furthermore, participants understood the captions with relative ease, suggesting that transcription quality was sufficient to support comprehension in simulated outpatient scenarios. However, stricter accuracy standards may be required in high-stakes contexts, such as discussions of medications or treatment options, where small errors can have serious consequences.

Although WER is widely used to evaluate ASR performance, it weighs all error types equally, regardless of their impact on comprehension [27]. Prior work has proposed alternative evaluation approaches that aim to capture semantic accuracy or user-centric measures of intelligibility and usefulness [20]. In clinical communication, we support developing evaluation metrics that align more closely with safety-critical requirements. Such metrics would be instrumental in determining when ASR

XSL•FO
RenderX

systems are truly ready for deployment in health care environments. In clinical settings, misrecognition of medical terminology can have consequences far more serious than common transcription errors, especially when involving medication names, diagnoses, or treatment instructions. Because of this, future work should consider safety-critical evaluation frameworks that go beyond traditional WER. Approaches, such as semantic error analysis, comprehension-based scoring, or accuracy, weighting for medically significant terms could better capture the real-world implications of captioning errors in health care communication.

Our participants represented a variety of ages, genders, HL levels, and degrees of dependence on lip reading. However, most participants had previously used captioning technology as an accommodation, so our usability findings may be less generalizable to individuals who are DHH with no prior captioning experience. Also, only 2 participants preferred written communication with hearing people. Therefore, satisfaction with our captioning technology may be higher than our results suggest for people who are DHH and depend more on written communication. Regardless of the scenario, most participants were satisfied with the SRS, trusted its accuracy, found it easy to watch, and were not distracted.

Participants trusted the captioning system despite occasional transcription errors, which embodies the concept of trust-in-automation frameworks, where user reliance is shaped by perceived system reliability and predictability [28]. Exit interviews revealed that beyond meeting technical expectations, the captioning system also meaningfully supported emotional connection, trust, and autonomy during clinical interactions. Participants described the captions as easy to use and grounding. They also noted reduced stress, lower cognitive fatigue, better understanding, and a stronger sense of being heard. Encouragingly, the observed reduction in stress and fatigue is consistent with prior work where assistive technology helped manage cognitive effort during information processing [29,30]. These findings suggest that accessibility tools should be evaluated not only by their accuracy but by their ability to support psychological safety and communication equity [31].

Additionally, although participants generally trusted the captioning system, a few raised concerns about transcript privacy and data handling. These concerns highlight the ethical need for transparency when implementing automated captioning in health care. This pilot used secure, locally stored recordings without identifiable data, but clinical deployment will require Health Insurance Portability and Accountability Act (HIPAA)-compliant encryption and explicit consent protocols. Adding user controls, such as options to delete transcripts or disable storage, could further strengthen trust among users who are DHH and other vulnerable populations. Nevertheless, participants recommended broader adoption of SRS, particularly for older adults and others facing progressive hearing-related communication barriers, underscoring the system’s potential to improve care for a heterogeneous population of DHH patients.

Limitations

While our findings are promising, this study has several limitations. Most participants were experienced caption users and had prior familiarity with assistive communication technologies, which may have positively influenced usability and satisfaction ratings. As a result, these findings may not fully represent the experiences of individuals who are DHH and are less familiar with captioning or other accessibility tools or primary American Sign Language users. Future research should include participants with varying levels of captioning experience and a broader demographic range to better assess generalizability and identify barriers for first-time users. This study was conducted in controlled, simulated settings, which may not fully reflect the complexity and spontaneity of real-world medical encounters. Because these mock scenarios involved medical students rather than practicing clinicians, the communication dynamics may differ from authentic physician–patient interactions. Future work should therefore include real-world clinical deployments to evaluate how captioning systems perform in active care settings and adapt to diverse communication styles and environmental conditions.

Second, although our participant pool included individuals with diverse hearing identities and varying degrees of familiarity with assistive technologies, it does not capture the full range of experiences within the broader community who are DHH. Future work should include longitudinal application in various clinical settings and recruitment of a more diverse participant population to better assess long-term usability and impact.

In addition, our SRS was not specifically optimized for medical vocabulary. This limitation was evident in the system’s tendency to misrecognize medical terminology, words that are infrequent in everyday speech yet crucial for accurate clinical communication. Furthermore, while we used WER as a standard quantitative evaluation metric, it does not fully capture how users who are DHH interpret and understand captions, particularly in high-stakes contexts. Future research should explore the development of domain-specific SRS trained on medical speech and adopt evaluation metrics that better reflect comprehension and user experience among individuals who are DHH. Finally, since WER assigns equal weight to all tokens, it does not differentiate between routine transcription errors and those involving safety-critical clinical terms (eg, medications or diagnoses). Therefore, WER may underestimate the potential impact of certain errors in medical contexts.

Future Directions

Improving SRS accuracy for medical terminology remains a key technical priority for clinical use. Strategies may include (1) speech recognition models on deidentified clinical audio to capture the acoustic variability of real-world medical speech [32], (2) embedding domain-specific medical dictionaries and medication name libraries into the language model of the SRS systems to reduce substitution errors [33], (3) leveraging context-aware large language models that can infer meaning from partial or uncertain input [34], and (4) integrating clinician feedback loops for rapid correction of recurring misinterpretations [35]. These enhancements would not only improve accuracy for technical vocabulary but also strengthen user trust and perceived reliability in clinical environments.

Building on these preliminary findings, future work should also explore integration with medical-domain ASR models to

XSL•FO
RenderX

enhance accuracy for specialized terminology and complex clinical dialog. Longitudinal studies will be valuable for assessing maintained usability, user trust, and performance over time. Additionally, testing captioning systems in broader clinical contexts, such as emergency care, geriatrics, and among patients with cognitive impairment, will help determine their adaptability and impact across diverse care settings.

Implications for Clinical Workflow Integration

Our findings demonstrate that real-time captioning is usable and beneficial in clinical settings for patients who are DHH, aligning with prior evidence that captioning improved recall of anesthesia-related consent conversations [36]. Given this demonstrated value, practical integration of captioning tools into clinical workflows will require thoughtful design to minimize disruption while enhancing accessibility. Participants envisioned use cases in which SRS displays could be embedded within existing electronic health record systems or mirrored on clinician tablets to preserve natural eye contact and conversational flow. Integration will also depend on clear institutional protocols for activating captioning on demand, ensuring confidentiality, and providing clinician training on how to engage with patients who are DHH using this technology. Establishing these processes could enable captioning to function as a routine accessibility feature rather than an exception, supporting both efficiency and equitable communication in care delivery.

Conclusions

This pilot study demonstrates that artificial intelligence–enhanced captioning can meaningfully improve communication experiences for individuals who are DHH in clinical settings. Participants found the system intuitive, emotionally supportive, and effective in bridging common communication barriers, especially those worsened by face masks and unfamiliar environments. While traditional captioning tools often fall short in medical contexts, integrating large language models into the speech recognition process offers a promising path toward more coherent, accurate, and human-centered accessibility. By centering on user perspectives, this study highlights the importance of evaluating assistive technologies not only for transcription quality, but for their impact on trust, inclusion, and psychological safety. Future research should build on these early insights to further refine captioning systems, examine their use in real-world clinical care, and ensure that patients who are DHH are active partners in the design of accessible digital health solutions.

Acknowledgments

The authors thank the Deaf/deaf and hard-of-hearing community members who participated in this study for generously sharing their insights and experiences. We are grateful to the Dexter Family Medicine clinic at the University of Michigan for their support in facilitating pilot testing.

Funding

This work was supported by funding from the Blue Cross Blue Shield of Michigan Foundation (2024010111) and by the University of Michigan’s e-HAIL (e-Health and Artificial Intelligence Laboratory) program.

Data Availability

All data generated or analyzed during this study are included in this published article. Further inquiries can be directed to the corresponding author.

Authors' Contributions

Methodology (equal), participant recruitment (equal), investigation (equal), data curation (lead), formal analysis (lead), writing – original draft (lead), and writing – review and editing (equal): SEH
Methodology (equal), participant recruitment (equal), technology development (lead), investigation (equal), data curation (equal), formal analysis (equal), writing – original draft (equal), and writing – review and editing (equal): LYW
Methodology (equal), project administration (lead), investigation (equal), data curation (equal), formal analysis (equal), writing – original draft (equal), and writing – review and editing (equal): LJM
Conceptualization (equal), technology development (supporting), methodology (supporting), writing – review and editing (equal), and supervision (equal): DJ
Conceptualization (equal), methodology (equal), writing – original draft (supporting), writing – review and editing (equal), and supervision (lead): MMM

Conflicts of Interest

None declared.

Multimedia Appendix 1
Structured exit interview questions.
[DOCX File, 17 KB - rehab_v13i1e79073_app1.docx]

XSL•FO
RenderX

References

1. Guttman OT, Lazzara EH, Keebler JR, Webster KLW, Gisick LM, Baker AL. Dissecting communication barriers in healthcare: a path to enhancing communication resiliency, reliability, and patient safety. J Patient Saf 2021 Dec 1;17(8):e1465-e1471. [doi: 10.1097/PTS.0000000000000541] [Medline: 30418425]
2. McKee M, James TG, Helm KVT, et al. Reframing our health care system for patients with hearing loss. J Speech Lang Hear Res 2022 Oct 17;65(10):3633-3645. [doi: 10.1044/2022_JSLHR-22-00052] [Medline: 35969852]
3. Quick statistics about hearing, balance, & dizziness. NIH. URL: https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing [accessed 2025-05-26]
4. Culturally affirmative and linguistically accessible services. National Association of the Deaf. URL: https://www.nad.org/resources/health-care-and-mental-health-services/mental-health-services/culturally-affirmative-and-linguistically-accessible-services/ [accessed 2025-05-10]
5. James TG, Coady KA, Stacciarini JMR, et al. “They’re not willing to accommodate deaf patients”: communication experiences of deaf american sign language users in the emergency department. Qual Health Res 2022 Jan;32(1):48-63. [doi: 10.1177/10497323211046238] [Medline: 34823402]
6. Marlow NM, Samuels SK, Jo A, Mainous AG. Patient-provider communication quality for persons with disabilities: a cross-sectional analysis of the Health Information National Trends Survey. Disabil Health J 2019 Oct;12(4):732-737. [doi: 10.1016/j.dhjo.2019.03.010] [Medline: 30995967]
7. Diamond L, Izquierdo K, Canfield D, Matsoukas K, Gany F. A systematic review of the impact of patient-physician non-english language concordance on quality of care and outcomes. J Gen Intern Med 2019 Aug;34(8):1591-1606. [doi: 10.1007/s11606-019-04847-5] [Medline: 31147980]
8. Barnett S, McKee M, Smith SR, Pearson TA. Deaf sign language users, health inequities, and public health: opportunity for social justice. Prev Chronic Dis 2011 Mar;8(2):A45. [Medline: 21324259]
9. Barnett SL, Matthews KA, Sutter EJ, et al. Collaboration with deaf communities to conduct accessible health surveillance. Am J Prev Med 2017 Mar;52(3 Suppl 3):S250-S254. [doi: 10.1016/j.amepre.2016.10.011] [Medline: 28215374]
10. Pollard RQ, Barnett S. Health-related vocabulary knowledge among deaf adults. Rehabil Psychol 2009 May;54(2):182-185. [doi: 10.1037/a0015771] [Medline: 19469608]
11. Mitra M, Akobirshoev I, McKee MM, Iezzoni LI. Birth outcomes among U.S. women with hearing loss. Am J Prev Med 2016 Dec;51(6):865-873. [doi: 10.1016/j.amepre.2016.08.001] [Medline: 27687529]
12. Alexander A, Ladd P, Powell S. Deafness might damage your health. Lancet 2012 Mar 17;379(9820):979-981. [doi: 10.1016/S0140-6736(11)61670-X] [Medline: 22423872]
13. ADA.gov. URL: https://www.ada.gov [accessed 2025-05-27]
14. McKee M, Moran C, Zazove P. Overcoming additional barriers to care for deaf and hard of hearing patients during COVID-19. JAMA Otolaryngol Head Neck Surg 2020 Sep 1;146(9):781-782. [doi: 10.1001/jamaoto.2020.1705] [Medline: 32692807]
15. Iezzoni LI, O’Day BL, Killeen M, Harker H. Communicating about health care: observations from persons who are deaf or hard of hearing. Ann Intern Med 2004 Mar 2;140(5):356-362. [doi: 10.7326/0003-4819-140-5-200403020-00011] [Medline: 14996677]
16. Moreland CJ, Ruffin CV, Morris MA, McKee M. Unmasked: how the COVID-19 pandemic exacerbates disparities for people with communication-based disabilities. J Hosp Med 2021 Mar;16(3):185-188. [doi: 10.12788/jhm.3562] [Medline: 33617440]
17. Bernard A, Weiss S, Rahman M, et al. The impact of COVID-19 and pandemic mitigation measures on persons with sensory impairment. Am J Ophthalmol 2022 Feb;234:49-58. [doi: 10.1016/j.ajo.2021.06.019] [Medline: 34197781]
18. Kawas S, Karalis G, Wen T, Ladner RE. Improving real-time captioning experiences for deaf and hard of hearing students. Presented at: ASSETS ’16: Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 23-26, 2016. [doi: 10.1145/2982142.2982164]
19. Kafle S, Huenerfauth M. Evaluating the usability of automatically generated captions for people who are deaf or hard of hearing. Presented at: ASSETS ’17: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 20 to Nov 1, 2017. [doi: 10.1145/3132525.3132542]
20. Favre B, Cheung K, Kazemian S, et al. Automatic human utility evaluation of ASR systems: does WER really predict performance? Presented at: Interspeech 2013; Aug 25-29, 2013 URL: https://www.isca-archive.org/interspeech_2013/favre13_interspeech.pdf [accessed 2025-12-15]
21. Yamamoto K, Suzuki I, Shitara A, Ochiai Y. See-through captions: real-time captioning on transparent display for deaf and hard-of-hearing people. Presented at: ASSETS ’21: Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibilit; Oct 18-22, 2021. [doi: 10.1145/3441852.3476551]
22. Suzuki I, Yamamoto K, Shitara A, Hyakuta R, Iijima R, Ochiai Y. See-through captions in a museum guided tour: exploring museum guided tour for deaf and hard-of-hearing people with real-time captioning on transparent display. Presented at: Computers Helping People with Special Needs: 18th International Conference, ICCHP-AAATE 2022; Jul 11-15, 2022. [doi: 10.1007/978-3-031-08648-9_64]

XSL•FO
RenderX

23. Olwal A, Balke K, Votintcev D. Wearable subtitles: augmenting spoken communication with lightweight eyewear for all-day captioning. Presented at: UIST ’20: Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology; Oct 20-23, 2020 URL: https://dl.acm.org/doi/proceedings/10.1145/3379337 [accessed 2025-12-15] [doi: 10.1145/3379337.3415817]
24. Speech-to-Text. Google Cloud. URL: https://cloud.google.com/speech-to-text [accessed 2025-05-27]
25. Morris AC, Maier V, Green P. From WER and RIL to MER and WIL: improved evaluation measures for connected speech recognition. Presented at: Interspeech 2004; Oct 4-8, 2004 URL: https://www.isca-archive.org/interspeech_2004/morris04_interspeech.pdf [accessed 2025-12-15] [doi: 10.21437/Interspeech.2004-668]
26. jitsi/jiwer. GitHub. URL: https://github.com/jitsi/jiwer [accessed 2025-05-27]
27. Ye-Yi W, Acero A, Chelba C. Is word error rate a good indicator for spoken language understanding accuracy. Presented at: 2003 IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat No03EX721); Nov 30 to Dec 4, 2003. [doi: 10.1109/ASRU.2003.1318504]
28. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors 2004;46(1):50-80. [doi: 10.1518/hfes.46.1.50_30392] [Medline: 15151155]
29. Sweller J. Cognitive load during problem solving: effects on learning. Cogn Sci 1988 Jun;12(2):257-285. [doi: 10.1016/0364-0213(88)90023-7]
30. Paas F, Renkl A, Sweller J. Cognitive load theory and instructional design: recent developments. Educ Psychol 2003 Jan 1;38(1):1-4. [doi: 10.1207/S15326985EP3801_1]
31. Bickmore TW, Paasche-Orlow MK. The role of information technology in health literacy research. J Health Commun 2012;17 Suppl 3(sup3):23-29. [doi: 10.1080/10810730.2012.712626] [Medline: 23030559]
32. Afonja T, Olatunji T, Ogun S, Etori NA, Owodunni A, Yekini M. Performant ASR models for medical entities in accented speech. Presented at: Interspeech 2024; Sep 1-5, 2024 URL: https://www.isca-archive.org/interspeech_2024/afonja24_interspeech.pdf [accessed 2025-12-15]
33. Michalopoulos G, Wang Y, Kaka H, Chen H, Wong A. UmlsBERT: clinical domain knowledge augmentation of contextual embeddings using the unified medical language system metathesaurus. Presented at: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics; Jun 6-11, 2021 URL: https://aclanthology.org/2021.naacl-main [accessed 2025-12-15] [doi: 10.18653/v1/2021.naacl-main.139]
34. Hsie C, Moreira C, Nobre IB, et al. DALL-m: context-aware clinical data augmentation with llms. arXiv. Preprint posted online on May 1, 2025. [doi: 10.48550/ARXIV.2407.08227]
35. Wu LY, Jain D. EvolveCaptions: empowering DHH users through real-time collaborative captioning. arXiv. Preprint posted online on Oct 2, 2025. [doi: 10.48550/ARXIV.2510.02181]
36. Spehar B, Tye-Murray N, Myerson J, Murray DJ. Real- time captioning for improving informed consent: patient and physician benefits. Reg Anesth Pain Med 2016;41(1):65-68. [doi: 10.1097/AAP.0000000000000347] [Medline: 26650432]

Abbreviations

ASR: automatic speech recognition
DHH: Deaf/deaf or hard of hearing
HIPAA: Health Insurance Portability and Accountability Act
HL: hearing loss
SRS: speech recognition system
WER: word error rate

Edited by S Munce; submitted 14.Jun.2025; peer-reviewed by A Roundtree, SD Ubur; revised version received 21.Nov.2025; accepted 03.Dec.2025; published 15.Jan.2026.

Please cite as:
Hughes SE, Wu LY, Ma LJ, Jain D, McKee MM
Assessing the Role of Medical Caption Technology to Support Physician-Patient Communication for Patients With Hearing Loss: Mixed Methods Pilot Study
JMIR Rehabil Assist Technol 2026;13:e79073
URL: https://rehab.jmir.org/2026/1/e79073
doi:10.2196/79073

© Sarah E Hughes, Liang-Yuan Wu, Lindsay J Ma, Dhruv Jain, Michael M McKee. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

XSL•FO RenderX

and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Therapeutic Conversational Agent (Solace) for Management of Chronic Pain: Acceptability and Usability Study

P Maxwell Slepian[1,2,3,4], MS, PhD; Stephanie Buryk-Iggers[1], PhD; Anna M Lomanowska[1], PhD; Binh Nguyen[4], PhD; Tahir Janmohamed[4], PEng, MBA; Hance Clarke[1,2,3,4], MD, PhD; Joel Katz[1,2,3,4,5], PhD; Nils G Niederstrasser[6], PhD

[1]Department of Anesthesia and Pain Management, University Health Network, Toronto, ON, Canada
[2]Department of Anesthesiology and Pain Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[3]University of Toronto Centre for the Study of Pain, University of Toronto, Toronto, ON, Canada
[4]ManagingLife, Inc, Toronto, ON, Canada
[5]Department of Psychology, Faculty of Health, York University, Toronto, ON, Canada
[6]School of Psychology, Sport, and Health Sciences, University of Portsmouth, Portsmouth, England, United Kingdom

Corresponding Author:
P Maxwell Slepian, MS, PhD
Department of Anesthesia and Pain Management
University Health Network
3 EN 317A
200 Elizabeth St.
Toronto, ON, M5G 2C4
Canada
Phone: 1 7242890586
Email: maxwell.slepian@uhn.ca

Abstract

Background: Chronic pain is a critical cause of personal suffering and societal concern. However, treatment options remain inadequate, and access to efficacious treatment is limited by geography, economics, and scale. Digital health interventions for chronic pain are easily scaled solutions to this problem, and autonomous conversational agents represent a new frontier in this treatment domain. Despite their potential impact, conversational agents powered by generative artificial intelligence (AI) have yet to be developed or examined for the treatment of chronic pain.

Objective: We aimed to develop and test Solace, a first-of-its-kind, expert-trained generative AI conversational agent designed to deliver support grounded in principles of evidence-based pain psychology.

Methods: We conducted an acceptability and usability study of Solace in a group of individuals with chronic pain. Participants (n=175) were recruited from Prolific, an online crowdsourcing platform, and interacted with Solace for at least 25 minutes. Self-report measures of system usability, treatment acceptability, and therapeutic alliance were completed after the interaction, and clinically relevant pain-related measures were completed before and after the interaction.

Results: Participants rated the usability of Solace to be excellent (System Usability Scale mean score 85.04, SD 13.6) and found it to be an acceptable intervention for chronic pain. The therapeutic alliance between participants and Solace was rated highly. Participants also demonstrated improvements in anxiety, pain interference, kinesiophobia, and pain resilience. Safety guardrails designed to identify and manage instances of suicidal ideation, injury, or requests for medication recommendations performed appropriately during the study. Solace is a usable and acceptable treatment for chronic pain that facilitates forming a strong therapeutic alliance with users. A single 30-minute conversation with Solace was also associated with improvements in several clinically relevant domains.

Conclusions: Solace is a usable and acceptable expert-trained generative AI conversational agent for pain management. Randomized clinical trials are needed to evaluate the efficacy of Solace as a strategy for the treatment of chronic pain.

Trial Registration: OSF Registries osf.io/9c7tv; https://osf.io/9c7tv

(JMIR Rehabil Assist Technol 2026;13:e87689) doi:10.2196/87689

XSL•FO
RenderX

KEYWORDS

artificial intelligence; chatbot; chronic pain; digital health; pain psychology

Introduction

Chronic pain affects approximately 20% of adults worldwide, and is an enormous source of personal suffering as well as a major public health challenge [1,2]. Psychological interventions are efficacious for chronic pain. Across numerous large-scale randomized controlled trials, cognitive behavioral therapy, in particular, has been robustly associated with improvements in pain intensity and pain-related disability [3]. There is also accumulating evidence that other psychological treatment approaches, including mindfulness-based interventions and acceptance and commitment therapy (ACT), are similarly effective [4-7].

Despite the wealth of evidence for pain psychology interventions, widespread implementation of these interventions has been limited. This may be due, in part, to systemic barriers to adoption, including insufficient finances, time constraints, poor accessibility, and social stigma [8,9]. However, one of the key limitations is simply the lack of competent providers. For instance, difficulty accessing competent providers is the most common barrier reported by individuals with chronic pain [8]. This is further complicated by the fact that most licensed health care providers practicing therapy (eg, psychiatrists, psychologists, counselors, and social workers) report little to no education or training in pain treatment [8].

Digital health interventions (DHIs) have emerged as one of the primary means to overcome these barriers and bring psychological interventions for pain to scale [10,11]. DHIs exist in many forms, ranging from distance-based replications of services via telephone or videoconferencing to web-based treatment programs and mobile applications. Across these services, content can be delivered in either a fully self-guided format or supported by a provider or “coach,” either synchronously or asynchronously [12]. These interventions have broadly demonstrated efficacy, but uptake remains limited, with studies reporting high dropout rates and low adherence [10,13].

More recently, researchers have begun to examine the provision of therapy by conversational agents, or chatbots, that enable patients to have interactive conversations in real time [13]. A recent scoping review identified beneficial effects of chatbots on pain intensity, but results were equivocal for pain interference and mental health outcomes [13]. These chatbots have generally been designed according to rule-based response algorithms rather than large language model (LLM)–based conversational support [13-16]. Such approaches have demonstrated that conversational agents can build a strong therapeutic alliance with users [15].

The use of LLM-based generative artificial intelligence (AI) conversational agents offers an alternative to algorithm-based chatbots. To date, not one study has examined fully generative AI conversational agents for the treatment of chronic pain. However, such generative AI chatbots have been studied for general mental health treatment. A clinical trial of a generative AI chatbot, Therabot, identified clinically significant improvement in symptoms of depression, anxiety, and feeding/eating disorder risk after 4 and 8 weeks, compared to waitlist controls [17]. Generative AI chatbots are also being applied to coping with medical concerns. In a recent study, health care providers identified a generative AI chatbot designed to provide coping support for adolescents with cancer as a feasible option for support provision, although the impact on clinical outcomes remains to be examined [18].

We sought to develop and test Solace, an expert-trained, generative AI chatbot built by ManagingLife, Inc. Solace was created by pairing a purpose-built knowledge base with a multiagent architecture leveraging a variety of commercially available foundational models. This study describes a prospective feasibility evaluation of the acceptability and usability of Solace in a group of participants reporting chronic pain (n=175). Participants engaged in an open-ended 30-minute interaction with Solace and completed measures of treatment acceptability, system usability, and therapeutic alliance after the conversation. Pain-related self-report measures were also completed before and after the participant’s interaction with Solace. We hypothesized that individuals would find Solace to be a usable and acceptable treatment for chronic pain.

Methods

Study Design

This study was a prospective feasibility evaluation [19,20] of the AI-driven conversational agent, Solace. The design was guided by the Developmental and Exploratory Clinical Investigation of Decision-Support Systems Driven by Artificial Intelligence (DECIDE-AI), providing an early-stage clinical evaluation checklist of AI-specific reporting items (Multimedia Appendix 1) [21].

AI Chatbot

Overview

Solace is a stand-alone, web-based digital health application developed by ManagingLife [22]. It interacts with users to gather information related to the user’s experience of pain and delivers appropriate, structured, interactive support grounded in evidence-based pain psychology techniques (eg, psychoeducation and ACT) to bolster emotional coping and self-management in individuals living with chronic pain. The development process focused on ensuring appropriate psychological support for adults with chronic pain through systematic programming and testing phases.

Technical Infrastructure

This study used the Solace v1.0 beta system version. The technical backbone of the chatbot system was built on top of a cloud-based digital health platform, Manage My Pain, developed by ManagingLife [22]. The main components of the Manage My Pain platform included user authentication, secure cloud hosting, and compliance with international data privacy

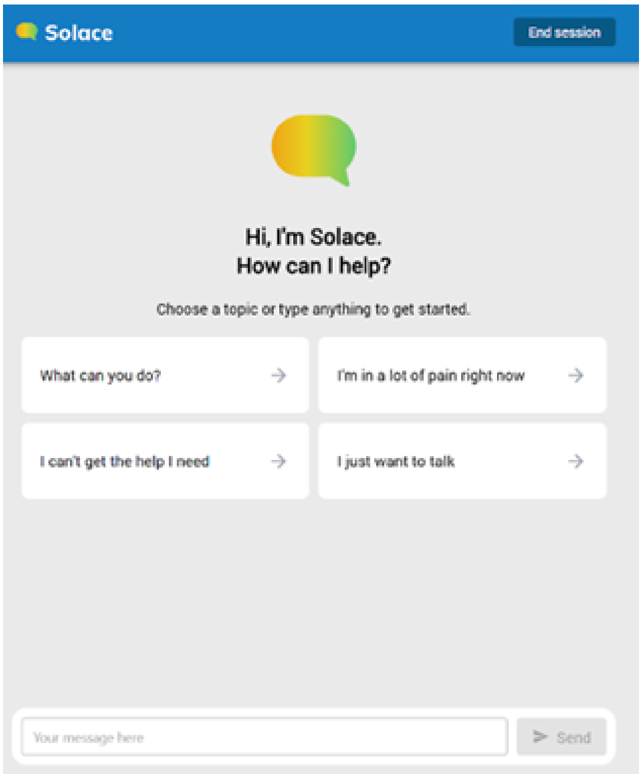
XSL•FO
RenderX

regulations, including end-to-end encryption and role-based access controls. Solace was designed to extend this platform by leveraging Amazon Web Services Bedrock to build a multiagent architecture using a library of LLMs. Each agent is responsible for a specialized task, and, together, these agents are coordinated by a centralized orchestration layer that manages conversational flow, memory retention, response timing, and system-level safety controls [23]. The foundational model versions, underlying platform, and all settings were fixed throughout the study period.

Conversational Design and Rules

Users began their experience with Solace by typing freely or selecting from starter prompts (Figure 1). User inputs were accepted without restrictions on length or content. Solace followed a set of core design principles developed through iterative collaboration with subject matter experts (SMEs) to ensure therapeutic integrity and user safety. These included the following:

Figure 1. The user options for starting a Solace conversation, including the use of the free-text field or starter prompts.



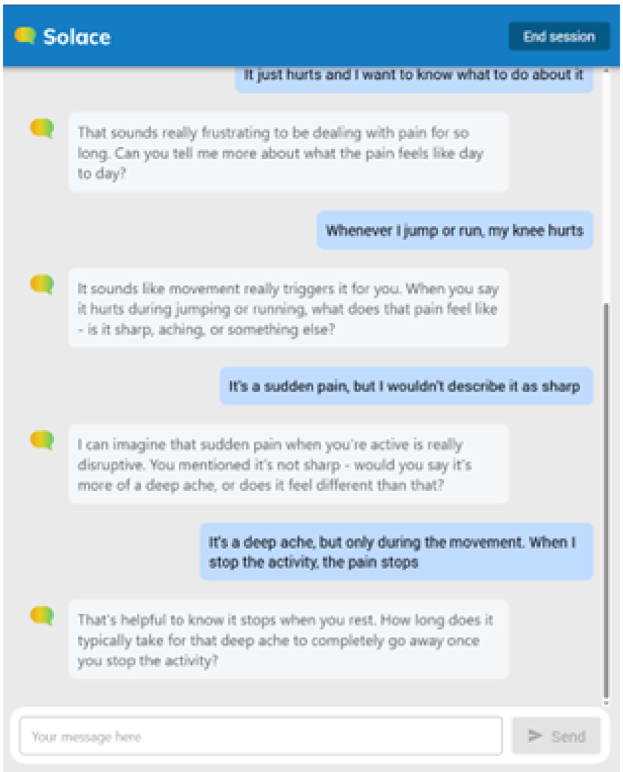
Hi, I'm Solace.
How can I help?

Choose a topic or type anything to get started.

1. Open-ended engagement, encouraging users to freely express their thoughts and emotions
2. Plain language to ensure accessibility and reduce cognitive burden
3. Emotional sensitivity, with empathic responses embedded throughout the conversation
4. Therapeutic guidance toward evidence-based strategies and self-understanding
5. Context continuity, using memory variables to support coherent, personalized dialogue
6. Multilevel safety guardrails, with screening for sensitive content (eg, mention of self-harm, harm to others, and opioid use) and fallback responses designed in collaboration with clinical experts.

Solace's outputs drew from knowledge bases composed from psychological frameworks, including but not limited to ACT, CBT, and dialectical behavior therapy and were presented to the user as (typically) 1-4 sentences (Figure 2).

Figure 2. Examples of Solace outputs during a user conversation.



It just hurts and I want to know what to do about it

That sounds really frustrating to be dealing with pain for so long. Can you tell me more about what the pain feels like day to day?

Whenever I jump or run, my knee hurts

It sounds like movement really triggers it for you. When you say it hurts during jumping or running, what does that pain feel like - is it sharp, aching, or something else?

It's a sudden pain, but I wouldn't describe it as sharp

I can imagine that sudden pain when you're active is really disruptive. You mentioned it's not sharp - would you say it's more of a deep ache, or does it feel different than that?

It's a deep ache, but only during the movement. When I stop the activity, the pain stops

That's helpful to know it stops when you rest. How long does it typically take for that deep ache to completely go away once you stop the activity?

Testing and Optimization

Initial development and testing were led by a clinical pain psychologist (PMS) and included a cohort of 8 SMEs, including other clinical psychologists, pain physicians, and pain researchers. SMEs conducted simulated therapy sessions with Solace and provided both real-time "think-aloud" feedback and postsession feedback to the development team, focusing on the conversational tone of Solace, therapeutic alignment, and emotional sensitivity. The development team worked with SMEs to translate their feedback into Solace's conversational logic, which included refining pacing, language, and safety mechanisms.

Within Solace, user interactions are processed as string-based text inputs and routed through an orchestration layer that assigns each request to an appropriate specialized agent, which generates a response. Each specialized agent represents a specific element within the structure and therapeutic style of a clinical psychology session (eg, 1 agent for managing transitions between stages within a session), and has been fine-tuned and reviewed during development by an SME.

All generated responses are subsequently evaluated by a guardrail validation component that performs safety and compliance checks. If the validator fails, the response is discarded, and a specialized guardrail agent generates an alternative response (see "Safety and Errors" below); this agent and its outputs are likewise reviewed and refined by the SME. If validation succeeds, the approved response is returned to the user.

Stress-testing scenarios were used to explore how the system responded to ambiguous or sensitive inputs. System-level performance metrics, including response latency, session duration, and output length, were recorded during internal testing. This series of tests was part of an iterative process, informing and optimizing prompt development and the design of response templates.

Safety and Errors

System-level errors were defined as failure to generate a response, excessive latency, or loss of conversational flow. Errors were identified using a series of validator checkpoints embedded within the system architecture. Checkpoints initiated fallback mechanisms to trigger predefined safe response templates. System anomalies and platform reliability were further monitored using session-level logs and performance data. Detection of risks was supported by multilevel safety guardrails built into Solace's conversation logic and prompt structures and relied on a combination of rule-based screening and automated classification techniques. The safety guardrails were designed to identify and mitigate risks in the following safety-sensitive domains: (1) disclosure of content indicating potential harm to self, suicidal ideation, or harm to others; (2) active psychological or medical emergencies; and (3) recommendations related to prescription medication use.

When high-risk content was detected, Solace generated predefined fallback responses designed to redirect the user to seek professional help, reinforce system boundaries, and discourage further engagement with unsafe topics.

Participant Eligibility and Enrollment

Participants were recruited online through Prolific [24], an online research participation platform, and assigned an anonymous Prolific identification number. The target sample size for this study was 145 participants based on sufficiently approximating the population means across multiple continuous

XSL•FO
RenderX

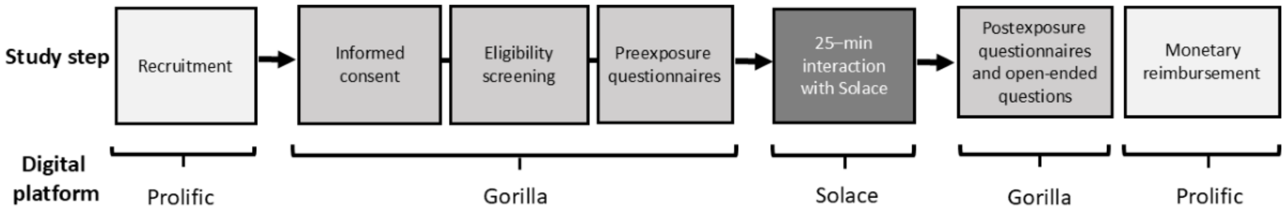
variables that are expected to have normal distributions. We aimed to recruit 195 participants to accommodate an anticipated 25% rate of attrition [25].

Those who consented were redirected to Gorilla [26], a cloud-based research platform hosting the questionnaires, to screen for eligibility and to access the survey. Participants were considered eligible based on the following criteria: (1) 18 years of age or older, (2) English speaking, and (3) living with chronic pain (ie, pain that has persisted for longer than 3 months). Chronic pain status and participant demographics were provided via the Prolific platform. Those eligible were asked to enter demographic information (age and sex) and to complete the baseline self-report measures.

Chatbot Engagement

Participants were informed that Solace is a nonclinical investigational tool under development and is not intended to provide medical advice or replace professional care. Training or an opportunity for familiarization with Solace was not provided to the participants prior to the study. Participants were prompted to engage with Solace, which was described as a "conversational agent, developed by ManagingLife. Solace provides pain management support and evidence-based strategies to help you understand and manage pain in your daily life." Participants were instructed to engage with Solace for at least 30 minutes from the start of the conversation, and then manually end the session. If users ended the session earlier than 30 minutes, they were informed how much time was left. Participants were then congratulated for completing their time with Solace and were provided with a link to return to Gorilla to complete the same self-report measures as at baseline, as well as a set of postexposure self-report measures. After completing the questionnaires, they were thanked for their participation, debriefed about the study hypotheses, and signposted to additional support if needed. The full participant flow through the study is depicted in Figure 3.

Figure 3. Participant flow through the study steps and associated digital platforms.



Measures

Participant Demographics

Participant demographics, including sex, ethnicity, employment status, and pain duration, were obtained from Prolific records.

System Usability

The System Usability Scale (SUS) is a 10-item questionnaire designed to measure the usability of a system, including digital products or digital systems [27]. Participants are asked to respond to items regarding usability (eg, "I thought the system was easy to use") on a scale from 1=strongly disagree to 5=strongly agree. The SUS is scored by subtracting the sum of the negatively worded items (eg, "I thought there was too much inconsistency in this system") from the sum of positively worded items (eg, "I think that I would use this system frequently"), adding 20, and multiplying by 2.5 to get a score out of 100. An SUS score of ≥85 has been identified as "excellent."

Treatment Acceptability

Two self-report inventories were included to examine treatment acceptability. The Treatment Acceptability/Adherence Scale (TAAS) is a 10-item self-report questionnaire designed to assess treatment acceptability and anticipated adherence to psychological interventions for anxiety and related disorders [28]. The TAAS has shown strong convergent and discriminant validity as well as excellent internal consistency [28]. In this study, the TAAS also demonstrated excellent internal consistency (α=.85). The Treatment Evaluation Inventory-Short Form (TEI-SF) is a 9-item measure of the perceived acceptability of behavioral treatments, and captures both positive (eg, "I would find this treatment to be an acceptable way of dealing with my chronic pain") and negative (eg, "I believe this treatment is likely to harm or injure me") perceptions of treatment [29]. The TEI-SF had excellent internal consistency in this study (α=.85). Both measures were adapted to the context of chronic pain treatment, which has been done in previous studies.

Therapeutic Alliance

The Working Alliance Inventory-Client Version (WAI-C) is a 36-item questionnaire used to measure therapeutic alliance in face-to-face therapy [30]. The WAI-C captures the client's perspective on 3 elements of therapeutic alliance, including the bond between therapist and client, agreement on the goals of therapy, and agreement on tasks to accomplish those goals. The WAI-C has excellent internal consistency (α=.87) and is strongly related to psychotherapy treatment satisfaction and outcomes. It has also been shown to predict treatment response to therapeutic interventions for pain. This instrument was adjusted for use within the context of an interaction with an AI-agent to align with the study's exploratory aims.

Pain-Related Measures

Patient Reported Outcomes Measurement Information System-29, version 2.0 (PROMIS-29) provides a profile of physical and mental health and assesses 7 health domains, including physical function, pain interference, pain intensity, anxiety, depression, fatigue, sleep disturbance, and social role integration. The PROMIS-29 has demonstrated strong validity in chronic pain populations. Internal consistency of subscales ranged from α=.88 to α=.96, indicating excellent consistency.

The questionnaire was administered using the static short forms of the domains [31].

The Tampa Scale of Kinesiophobia (TSK) was included as a measure of fear of movement (and reinjury) [32]. Participants rate items relevant to this fear (eg, "Pain always means I injured my body") on a scale from 0=strongly disagree to 4=strongly agree. The TSK has been extensively validated for individuals with chronic pain, and internal consistency in this study was adequate (α=.70)

The Pain Resilience Scale (PRS) is a 14-item measure of an individual's ability to maintain positive physical and emotional functioning despite pain. Participants are asked to rate agreement to items (eg, "I push through the pain") on a Likert scale from 0=not at all to 5=all the time [33,34]. The PRS has 2 subscales, behavioral perseverance and cognitive-affective positivity. The PRS has demonstrated psychometric validity among individuals with chronic pain [33], and internal consistency in this sample was excellent (α=.95).

The Chronic Pain Acceptance Questionnaire (CPAQ) is a 20-item measure with two subscales: (1) the degree to which one engages in life activities regardless of pain and (2) willingness to experience pain [35]. The CPAQ has been extensively validated for individuals with chronic pain [36,37], and the internal consistency in this study was adequate (α=.70).

System Metrics

User engagement and system performance were abstracted from transcripts of each participant session and associated data. User engagement characteristics included the number and length of prompts entered into Solace, the number and length of Solace's responses, and total session duration. Response readability was calculated using the Flesch Reading Ease Score and the Flesch-Kincaid Grade Level [38,39]. Safety was assessed by manual review of all conversations that contained content relevant to the guardrails.

Analysis

Analyses were primarily descriptive in nature. Participant responses on measures of usability, treatment acceptability, and therapeutic alliance were averaged (rather than summed) to give an indication of the overall level of agreement with measure items. Median (IQR) ratings for individual items were also reported. Averages with SD or counts, where applicable, were reported for system-level metrics. Only participants who completed at least 25 minutes of interaction with Solace were included for analyses of usability, treatment acceptability, and therapeutic alliance, whereas intent-to-treat analyses (eg, including all participants) were conducted using linear mixed effects models to analyze changes in clinically relevant self-report outcomes. Of note, due to a technical issue, 1 item was missed from the WAI-C (item 31) and 1 item from the CPAQ (item 8).

Ethical Considerations

The University of Portsmouth Ethics Board approved this study (REB SHFEC 2025-073), and the study has been registered with the Open Science Framework. Written informed consent was obtained from all participants prior to taking part in the study. Participants were presented with information about the study, made aware of their right to withdraw from the study at any time, and provided consent. Participants were compensated £18 (approximately US $24) for completing the study.

All identifying information was removed from the data before storage, and the data were further anonymized by assigning ID codes. Data pertaining to participants' responses to the structured and open-ended questionnaires and to their interactions with Solace were anonymized. Anonymized data were temporarily stored on the servers of Gorilla and ManagingLife, respectively. Data were extracted and transferred to the University of Portsmouth's secure server. The data will be retained for a minimum of 10 years in accordance with the University of Portsmouth Retention Schedule for Research Data.

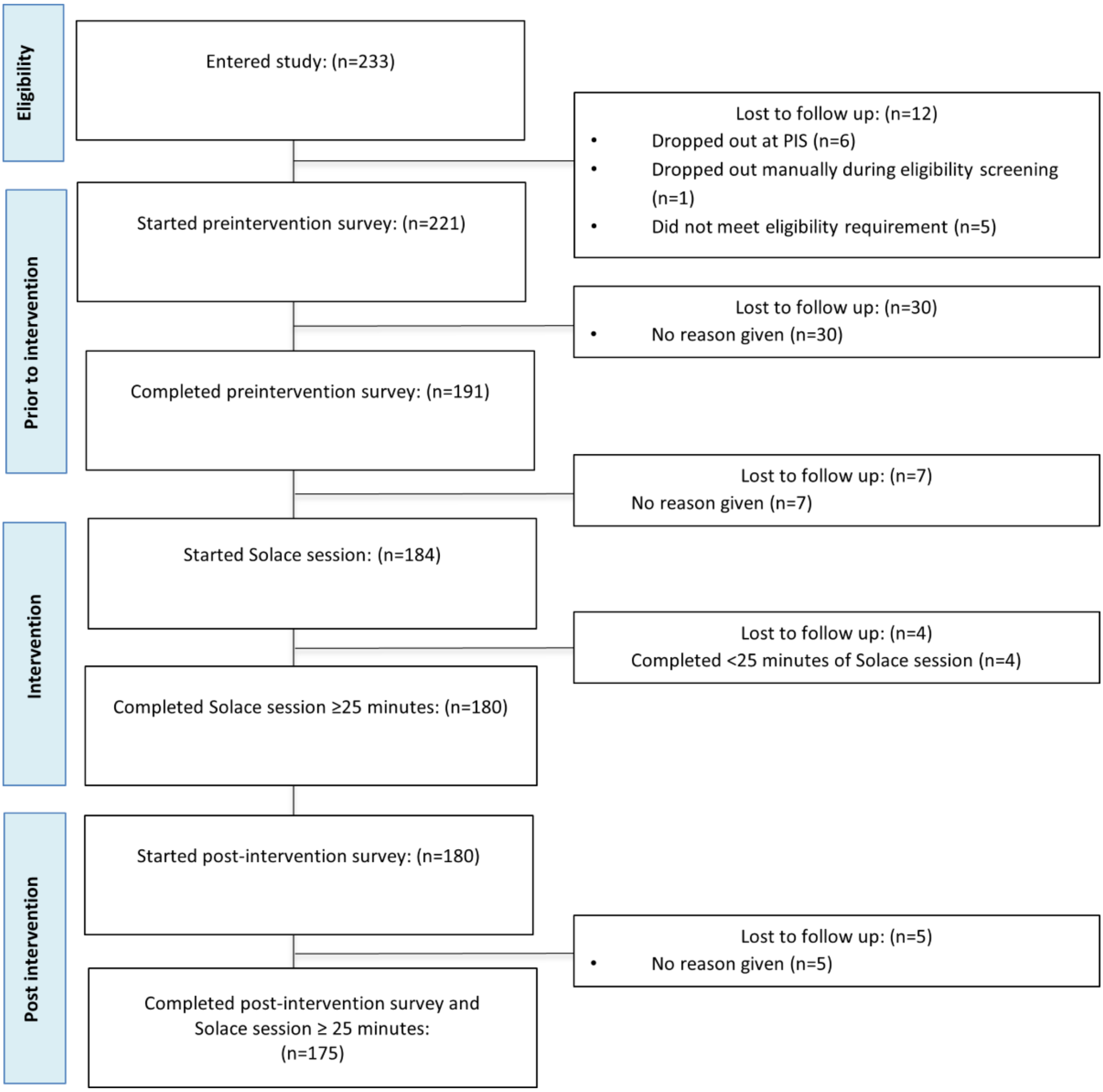
Results

Participant Characteristics and Flow

The flow of participants through the study is depicted in Figure 4. A total of 221 participants consented and completed the preinteraction survey. Of these, 184 began a session with Solace, and 175 engaged with Solace for at least 25 minutes. Participant demographics are listed in Table 1. All participants passed attention checks.

XSL•FO
RenderX

Figure 4. Flow diagram of participants who completed a ≥25-minute Solace session and self-report constructs. PIS: preintervention survey.



XSL•FO
RenderX

Table 1. Participant characteristics (n=175).

Characteristic	Value
Age (years), mean (SD)	42.54 (12.62)
Sex, n (%)	
Male	81 (46.3)
Female	93 (53.1)
Prefer not to say	1 (0.6)
Ethnicity, n (%)	
Asian	12 (6.9)
Black	13 (7.4)
Mixed	9 (5.1)
White	139 (80.8)
Other	2 (1.1)
Employment status, n (%)	
Full time	93 (53.1)
Part time	28 (16.0)
Unemployed and job seeking	9 (5.1)
Not in paid work (eg, homemaker, retired, or disabled)	23 (13.1)
Other	8 (4.6)
No data	14 (9.7)
Pain duration, n (%)	
3-6 months	9 (5.1)
6-12 months	10 (5.7)
1-2 years	16 (9.1)
2-5 years	52 (29.7)
5-10 years	32 (18.3)
10-20 years	43 (24.7)
>20 years	12 (6.9)
No data	1 (0.6)

Usage Characteristics and Safety Performance

The interaction sessions with Solace lasted an average of 38.74 (SD 11.7) minutes. On average, participants provided 29.2 (SD 10.1) input prompts to the system with an average length of 19.8 (SD 12.4) words, and Solace gave responses averaging 49.73 (SD 9.8) words on each occasion. On average, the readability of Solace’s responses was at the early high school level (Flesch-Kincaid Grade Level mean 9.56, SD 2.8; Flesch Reading Ease mean 56.44, SD 15.6). The distribution of Flesch-Kincaid Grade Levels and Flesch Reading Ease scores is depicted in Figures 5 and 6, respectively.

Figure 5. Distribution of Flesch-Kincaid Grade Levels across all responses generated by Solace.

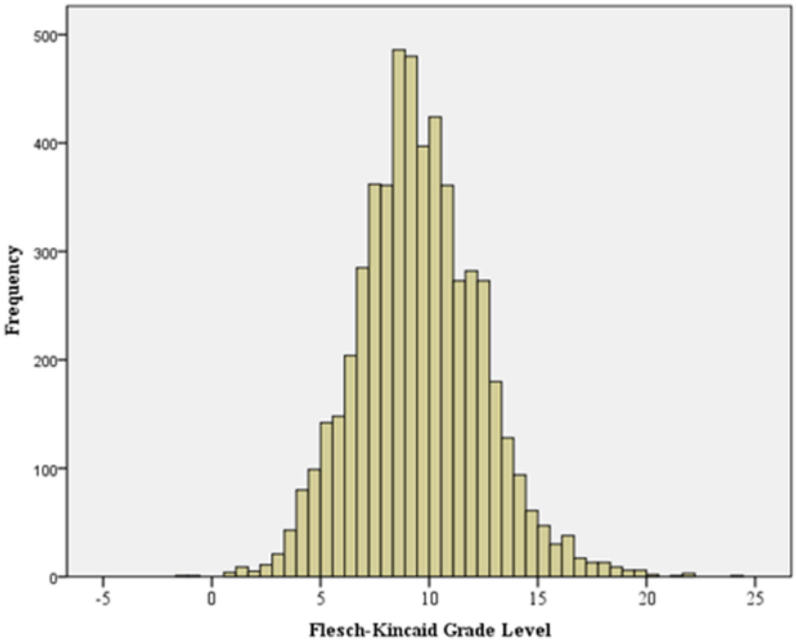
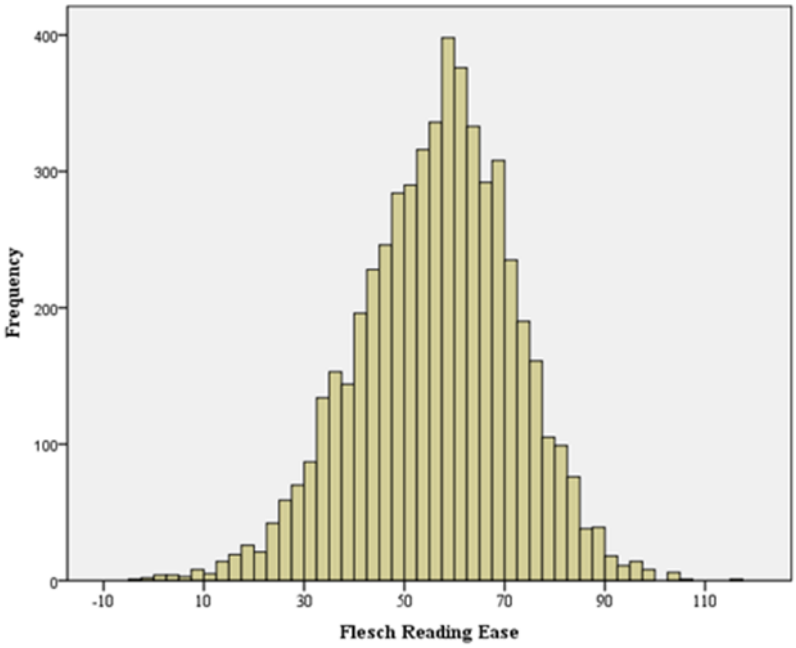


Figure 6. Distribution of Flesch Reading Ease scores across all responses generated by Solace.



Notably, 2 transcripts featured discussions of death, and 89 transcripts featured discussions of medication. Upon review by a clinical health psychologist (PMS), neither instance where death was discussed constituted suicidal ideation, self-harm, or harm to others, and Solace's responses were appropriate. Likewise, none of the conversations regarding medication included elicitation of recommendations, and, as such, guardrails were not activated.

System Usability

The SUS was completed as a measure of how easy participants found Solace to use. Participants rated the usability of Solace as excellent (mean 85.04, SD 13.6). On a scale from 1=strongly disagree to 5=strongly agree, participants rated all items as "strongly agree" except "I would like to use this system frequently," and "I found the various functions in this system were well-integrated," which were both rated as "agree." Item-level results are presented in Table 2.

XSL•FO
RenderX

Table 2. System Usability Scale item-level results (n=175). Scale ranges from 1=strongly disagree to 5=strongly agree.

Items	Median (IQR)
1. I think that I would like to use this system frequently	4.00 (2.00)
2. I found the system unnecessarily complex	1.00 (1.00)
3. I thought the system was easy to use	5.00 (1.00)
4. I think that I would need the support of a technical person to be able to use this system	1.00 (0.00)
5. I found the various functions in this system were well integrated	4.00 (2.00)
6. I thought there was too much inconsistency in this system	1.00 (1.00)
7. I would imagine that most people would learn to use this system very quickly	5.00 (1.00)
8. I found the system very cumbersome to use	1.00 (1.00)
9. I felt very confident using the system	5.00 (1.00)
10. I needed to learn a lot of things before I could get going with this system	1.00 (0.00)

Treatment Acceptability

Participants completed 2 measures of treatment acceptability, the TEI-SF and the TAAS. Average scores for the TEI-SF and TAAS, as well as individual item characteristics (median, IQR), are presented in Tables 3 and 4, respectively. On a scale from 1=strongly disagree to 5=strongly agree, participants reported an average rating of 3.77 (SD 0.7) on the TEI-SF. On the TAAS, participants, on average, rated treatment acceptability as 5.47 (SD 0.92) on a scale from 1=disagree strongly to 7=agree strongly.

Table 3. Treatment Evaluation Inventory-Short Form item-level results (n=175). Scale ranges from 1=strongly disagree to 5=strongly agree.

Items	Median (IQR)
1. I would find this treatment to be an acceptable way of dealing with my chronic pain	4.00 (1.00)
2. I would be willing to use this procedure if prescribed or recommended by a physician	4.00 (1.00)
3. I like the procedures that may be used in this treatment	4.00 (1.00)
4. I believe this treatment is likely to be effective	4.00 (2.00)
5. I believe that I will experience discomfort during the treatment	3.00 (1.00)
6. I believe this treatment is likely to result in permanent improvement	3.00 (1.00)
7. I believe this treatment is likely to harm or injure me	1.00 (1.00)
8. I believe it would be acceptable to use this treatment with individuals who cannot choose treatments for themselves	4.00 (1.00)
9. Overall, I have a positive reaction to this treatment	4.00 (1.00)

Table 4. Treatment Adherence/Acceptability Scale item-level results (n=175). Scale ranges from 1=disagree strongly to 7=agree strongly.

Items	Median (IQR)
1. If I began this treatment, I would be able to complete it	6.00 (2.00)
2. If I participated in this treatment, I would adhere to its requirements	6.00 (2.00)
3. I would find this treatment exhausting	3.00 (3.00)
4. It would be distressing for me to participate in this treatment	2.00 (2.00)
5. Overall, I would find this treatment intrusive	2.00 (2.00)
6. This treatment would provide effective ways to help me cope with my chronic pain	5.00 (1.00)
7. I would prefer to try another type of psychological treatment instead of this one	3.00 (3.00)
8. I would prefer to receive medication for my chronic pain instead of this treatment	3.00 (2.00)
9. I would recommend this treatment to a friend with chronic pain	5.00 (2.00)
10. If I began this treatment, I would likely drop out	2.00 (3.00)

XSL•FO
RenderX

Therapeutic Alliance

The WAI-C was completed to gauge the therapeutic alliance between participants and Solace. Overall, participants reported a strong working alliance with Solace (mean 188.03, SD 36.1), with a shared sense of goals rated as the strongest component of the therapeutic alliance (mean 67.86, SD 13.3), followed by agreement on tasks (mean 61.6, SD 12.7), and the therapeutic bond (mean 58.6, SD 12.6). WAI-C individual item scores are presented in Table 5.

Table 5. Working Alliance Inventory-Client Version item-level results (n=175). Scale ranges from 1=never to 7=always.

Items	Median (IQR)
1. I felt uncomfortable with Solace	1.00 (1.00)
2. Solace and I agreed about the things I will need to do in therapy to help improve my situation	6.00 (2.00)
3. I was worried about the outcome of the session	1.00 (1.00)
4. What I was doing in therapy gave me new ways of looking at my problem	5.00 (2.00)
5. Solace and I understood each other	6.00 (2.00)
6. Solace perceived accurately what my goals were	6.00 (2.00)
7. I find what I was doing in therapy confusing	1.00 (1.00)
8. I believe that Solace liked me	5.00 (2.00)
9. I wish Solace and I could have clarified the purpose of our session	2.00 (4.00)
10. I disagreed with Solace about what I ought to get out of therapy	1.00 (1.00)
11. I believe the time Solace and I were spending together was not spent efficiently	1.00 (2.00)
12. Solace did not understand what I was trying to accomplish in therapy	1.00 (1.00)
13. I was clear on what my responsibilities were in therapy	6.00 (2.00)
14. The goals of the session were important for me	6.00 (2.00)
15. I find what Solace and I were doing in therapy were unrelated to my concerns	1.00 (1.00)
16. I feel that the things I did in therapy helped me to accomplish the changes I wanted	5.00 (2.00)
17. I believe Solace was genuinely concerned for my welfare	5.00 (3.00)
18. I was clear as to what Solace wanted me to do in that session	6.00 (2.00)
19. Solace and I respected each other	7.00 (2.00)
20. I feel that Solace was not totally honest about its feelings toward me	1.00 (0.00)
21. I was confident in Solace’s ability to help me	5.00 (3.00)
22. Solace and I were working toward mutually agreed upon goals	6.00 (2.00)
23. I feel that Solace appreciated me	5.00 (3.00)
24. We agreed on what was important for me to work on	6.00 (2.00)
25. As a result of therapy, I became clearer as to how I might be able to change	6.00 (3.00)
26. Solace and I trusted each other	6.00 (3.00)
27. Solace and I had different ideas about what my problems were	1.00 (1.00)
28. My relationship with Solace was very important to me	4.00 (4.00)
29. I had the feeling that if I said or did the wrong thing Solace would stop working with me	1.00 (0.00)
30. Solace and I collaborated on setting goals for my therapy	6.00 (2.00)
31. I was frustrated by the things I was doing in therapy	—[a]
32. We had a good understanding of the kind of changes that would be good for me	6.00 (2.00)
33. The things that Solace was asking me to do did not make sense to me	1.00 (1.00)
34. I did not know what to expect as the result of my therapy	2.00 (3.00)
35. I believe the way were working with my problem was correct	6.00 (2.00)
36. I feel Solace cared about me even when I did things it did not approve of	5.00 (3.00)

[a] Not available.

XSL•FO
RenderX

Clinical Outcomes

Clinically relevant outcomes were completed before and after the conversation with Solace and were analyzed using linear mixed effects models. Descriptive statistics and test results are presented in Table 6. Following Bonferroni correction to account for multiple comparisons, participants reported significant improvement in kinesiophobia, pain resilience, and pain willingness. On the PROMIS-29, participants reported improvement in anxiety and pain interference.

Table 6. Changes in clinically relevant outcomes.

Outcomes	Baseline, mean (SD)	Posttest, mean (SD)	*F* test (*df*)	*P* value
PRS[a]	28.07 (10.6)	30.57 (10.6)	41.62 (1, 193.6)	<.001
PRS-Behavioral Perseverance	11.64 (4.2)	12.1 (3.9)	6.79 (1, 196.8)	.01
PRS-Cognitive-Affective Positivity	16.43 (7.7)	18.48 (7.6)	51.95 (1, 194.1)	<.001
Kinesiophobia	43.4 (7.1)	41.79 (8.0)	24.73 (1, 191.5)	<.001
CPAQ[b]	61.66 (12.6)	62.5 (12.2)	1.71 (1, 201.1)	.19
CPAQ-Activity Engagement	31.71 (8.0)	33.02 (8.1)	16.46 (1,194.05)	<.001
CPAQ-Pain Willingness	29.95 (8.5)	29.48 (9.1)	0.37 (1, 194.05)	.54
PROMIS-29[c]				
Physical Function	15.15 (3.5)	15.0 (3.6)	2.38 (1, 191.7)	.13
Anxiety	10.46 (3.9)	10.0 (4.1)	17.04 (1, 191.2)	<.001
Depressive Symptoms	9.58 (4.6)	9.42 (4.6)	2.82 (1, 192.2)	.09
Sleep	12.87 (4.0)	12.86 (3.9)	0.01 (1, 193.6)	.93
Social Function	13.0 (3.9)	13.08 (3.8)	0.37 (1, 195.1)	.55
Pain Interference	11.60 (4.0)	11.15 (3.9)	10.89 (1, 195.1)	.001
Fatigue	12.95 (4.1)	12.7 (4.4)	3.78 (1, 191.2)	.05

[a] PRS: Pain Resilience Scale.

[b] CPAQ: Chronic Pain Acceptance Questionnaire.

[c] PROMIS-29: Patient Reported Outcome Measurement Information System-29.

Discussion

This study conducted a feasibility evaluation of Solace, an expert-trained generative AI conversational agent designed to use pain psychology–informed self-management strategies for chronic pain, following the DECIDE-AI framework [40]. Participants reporting chronic pain completed self-report measures before and after at least 25 minutes of conversation with Solace. Broadly, participants rated Solace to have excellent usability, good treatment acceptability, and reported a strong therapeutic alliance. Moreover, participants reported improvement in several clinically relevant domains from before to after the interaction with Solace.

The reported system usability (mean 85.04, SD 13.6) was rated well above the benchmark established for digital health applications (SUS score >65), indicating excellent usability [41,42]. Notably, SUS scores were also higher than those reported for other generative AI tools, such as ChatGPT, when rated by either general users (mean 67.44) or health care professionals (mean 64.52, SD 13.91) [43,44]. Although the SUS has been frequently used to evaluate AI applications [45], the SUS may not capture the nuance of using a generative AI conversational agent such as Solace [41]. Future studies should include more specific measures of usability designed to analyze generative AI conversational agents, such as the Bot Usability Scale [46].

Solace was deemed an acceptable health application for chronic pain and was well-received by participants. Participant responses on the TEI-SF and the TAAS were largely on par with other DHIs for chronic pain [47,48]. Participants responded with agreement to almost all items. This includes statements such as “I believe this treatment is likely to be effective,” “I would find this treatment to be an acceptable way of dealing with my chronic pain,” and “this treatment would provide effective ways to help me cope with my chronic pain.” Participants also endorsed items related to the belief that they could adhere to treatment with Solace, such as “If I began this treatment, I would be able to complete it,” and beliefs that Solace is a safe intervention (ie, strongly disagreeing with “I believe this

XSL•FO
RenderX

treatment is likely to harm or injure me"). It is worth noting that participants rarely expressed the strongest agreement with the TEI-SF or TAAS items. This may be due to the noninterventional nature of the evaluation and the lack of set treatment expectations. It will be important to further evaluate acceptability in the context of a treatment trial or clinical practice.

Participants reported developing a strong therapeutic alliance with Solace over the course of the 25-minute conversation. In particular, participants reported a strong shared sense of treatment goals (ie, strongly disagreeing with the statement "I find what Solace and I were doing in therapy was unrelated to my concerns"). Several previous studies have examined the experience of therapeutic alliance with a conversational agent, and have found similarly strong working alliance [15,17]. It is perhaps understandable that participants reported the therapeutic bond as the weakest aspect of the alliance. Items designed to capture this on the WAI-C reflect a personal relationship that might not be relevant for interactions with a conversational agent, such as items asking if participants felt Solace liked or trusted them. New tools might be needed in this area to thoroughly assess therapeutic alliance with nonhuman conversational agents [49]. It is also notable that this study was not presented to participants as a treatment designed to help them with their pain, despite the conversational focus on treatment-related topics. It will be important to replicate these promising indications of therapeutic alliance with Solace in a treatment setting, and to examine if alliance improves with longer treatment duration, an important predictor of treatment success [50,51].

Safety of Solace and other generative AI conversational agents is a critical consideration in their development as DHIs. However, this could not be fully assessed as guardrails designed to manage suicidality, self-harm, medical emergencies, or medication recommendations were developed, but user prompts did not include safety concerns, and therefore, guardrails were not activated in this study. It will be necessary to analyze a much larger volume of interactions to fully test guardrail performance.

After the 25-minute conversation with Solace, participants demonstrated improvement in several meaningful clinical domains, including kinesiophobia, resilience, activity engagement, as well as anxiety and pain interference on the PROMIS-29 measure. Across these factors, there was 4%-9% improvement, with the greatest improvement seen in pain resilience. There are several difficulties in interpreting these changes. Notably, although participants were told Solace was a treatment in development and self-report constructs inherently included reference to treatment, they were also told that this was not a treatment study. There was also a mismatch between the wording on several of the measures and the time frame of the study. In particular, instructions for the subscales of the PROMIS-29 ask respondents to reference how they felt "in the last 7 days," whereas the study might have lasted as little as 25 minutes. Despite these challenges in interpretation, improvements were all in the beneficial direction, held up to statistical correction, and likely reflect the participants' feelings at the time of responding. Although the change was modest in magnitude, this is a promising indication for further treatment-focused research on Solace.

Future research on Solace should take several directions. As described above, a program of clinical research should be undertaken to evaluate the efficacy and effectiveness of Solace, either as a stand-alone program or integrated into the Manage My Pain application, a digital pain self-management application developed by ManagingLife. This will include a fuller test of the effectiveness of the safety guardrails in identifying discussions involving recommendations for medication use, suicidal ideation, and harm to self or others. Further evaluation of the nature of users' interactions with Solace will also be critical. These should focus on both the therapeutic relationship as well as evaluation of safety features of the application. Moreover, algorithmic fairness was not assessed in this study. Subsequent studies reporting on Solace will aim to compare model performance (sensitivity and specificity) across demographic subgroups, including sex, age, and ethnicity, to ensure equitable accuracy and lack of bias. Patient engagement is also needed to more fully address the lived experience of individuals with chronic pain.

There is room to further improve Solace's functionality. In particular, Solace's responses required a reading comprehension level considered roughly equivalent to early-middle high school. Future iterations of Solace should focus on the accessibility of language to improve utility by the general population. More detailed analysis of conversations with Solace is also warranted to ensure that conversational content is adherent to prompting.

There are several notable limitations of this study. First, the study was conducted via an online crowdsourcing platform. Although the sample recruited in this study is similar in several respects (ie, age, sex, and duration of pain) to individuals with chronic pain recruited from community samples [52], there were several notable distinctions, and also limited demographic data collection. In particular, participants in this study were primarily White and working full-time. As such, it is possible that community or treatment-seeking individuals with chronic pain would have a different experience of Solace. Although extensive research on chronic pain has been conducted via the Prolific platform [53,54], future research needs to be conducted in clinical samples with verified diagnoses. Second, the study was not designed as a test of Solace as a treatment. As such, participants did not necessarily have a treatment-related interaction with Solace, which could impact perceptions of acceptability and therapeutic alliance. Third, the sensitivity and specificity of the guardrail system were not fully tested, given the few instances that required guardrail activation. Finally, the short time frame of the study and the lack of a control condition limit the interpretation of the clinical impact of Solace. It is possible that alternative explanations, such as regression to the mean or social desirability, could account for the change associated with Solace use. Subsequent studies should examine prolonged or repeated interactions and feature comparisons to a control condition to enhance internal validity.

In conclusion, participants found Solace to be a usable system and an acceptable treatment for chronic pain after a 25-minute conversation. Participants reported a strong therapeutic alliance

XSL•FO
RenderX

and demonstrated improvement in several clinically relevant domains. Furthermore, examination of safety guardrails for Solace indicated appropriate performance. Solace is a promising therapeutic tool, and clinical trials are needed to fully examine its clinical efficacy as a DHI for chronic pain.

Data Availability

The data are available from the University of Portsmouth Research Portal, which functions as a data repository.

Funding

This research was funded by the University of Portsmouth's Higher Education Innovation Fund.

Conflicts of Interest

TJ is the founder and CEO of ManagingLife. BN is an employee of ManagingLife. PMS is an independent consultant for ManagingLife. JK and HC are unpaid members of the ManagingLife Advisory Board, serving as research director and medical director, respectively, and providing guidance on the product and the company's research initiatives. All other authors declare no conflicts of interest.

Multimedia Appendix 1

Developmental and Exploratory Clinical Investigation of Decision-Support Systems Driven by Artificial Intelligence (DECIDE-AI) checklist.

[DOCX File , 2338 KB - rehab_v13i1e87689_app1.docx]

References

1. Reitsma ML, Tranmer JE, Buchanan DM, Vandenkerkhof EG. The prevalence of chronic pain and pain-related interference in the Canadian population from 1994 to 2008. Chronic Dis Inj Can 2011;31(4):157-164. [doi: 10.24095/hpcdp.31.4.04]
2. Hogan ME, Taddio A, Katz J, Shah V, Krahn M. Incremental health care costs for chronic pain in Ontario, Canada: a population-based matched cohort study of adolescents and adults using administrative data. Pain 2016;157(8):1626-1633. [doi: 10.1097/j.pain.0000000000000561] [Medline: 26989805]
3. Williams ACDC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 2020;8:CD007407. [doi: 10.1002/14651858.CD007407.pub4] [Medline: 32794606]
4. Zhu M, Wong SY, Zhong CC, Zeng Y, Xie L, Lee EK, et al. Which type and dosage of mindfulness-based interventions are most effective for chronic pain? A systematic review and network meta-analysis. J Psychosom Res 2025;191:112061 [FREE Full text] [doi: 10.1016/j.jpsychores.2025.112061] [Medline: 40010103]
5. Martinez-Calderon J, García-Muñoz C, Rufo-Barbero C, Matias-Soto J, Cano-García FJ. Acceptance and commitment therapy for chronic pain: an overview of systematic reviews with meta-analysis of randomized clinical trials. J Pain 2024;25(3):595-617. [doi: 10.1016/j.jpain.2023.09.013]
6. Ma TW, Yuen ASK, Yang Z. The efficacy of acceptance and commitment therapy for chronic pain: a systematic review and meta-analysis. Clin J Pain 2023;39(3):147. [doi: 10.1097/ajp.0000000000001096]
7. Hilton L, Hempel S, Ewing BA, Apaydin E, Xenakis L, Newberry S, et al. Mindfulness meditation for chronic pain: systematic review and meta-analysis. Ann Behav Med 2017;51(2):199-213 [FREE Full text] [doi: 10.1007/s12160-016-9844-2] [Medline: 27658913]
8. Darnall BD, Scheman J, Davin S, Burns JW, Murphy JL, Wilson AC, et al. Pain psychology: a global needs assessment and national call to action. Pain Med 2016;17(2):250-263 [FREE Full text] [doi: 10.1093/pm/pnv095] [Medline: 26803844]
9. Driscoll MA, Edwards RR, Becker WC, Kaptchuk TJ, Kerns RD. Psychological interventions for the treatment of chronic pain in adults. Psychol Sci Public Interest 2021;22(2):52-95. [doi: 10.1177/15291006211008157] [Medline: 34541967]
10. Valentijn PP, Tymchenko L, Jacobson T, Kromann J, Biermann CW, AlMoslemany MA, et al. Digital health interventions for musculoskeletal pain conditions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022;24(9):e37869 [FREE Full text] [doi: 10.2196/37869] [Medline: 36066943]
11. Kazdin AE. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav Res Ther 2017;88:7-18. [doi: 10.1016/j.brat.2016.06.004] [Medline: 28110678]
12. El-Tallawy SN, Pergolizzi JV, Vasiliu-Feltes I, Ahmed RS, LeQuang JK, Alzahrani T, et al. Innovative applications of telemedicine and other digital health solutions in pain management: a literature review. Pain Ther 2024;13(4):791-812. [doi: 10.1007/s40122-024-00620-7] [Medline: 38869690]
13. Souza FL, Bowman H, Yang F, Hesam-Shariati N, Linke J, Gilanyi YL, et al. Conversational agents to support pain management: a scoping review. Eur J Pain 2025;29(5):e70016. [doi: 10.1002/ejp.70016] [Medline: 40170347]

XSL•FO
RenderX

14. Sinha C, Cheng AL, Kadaba M. Adherence and engagement with a cognitive behavioral therapy-based conversational agent (Wysa for Chronic Pain) among adults with chronic pain: survival analysis. JMIR Form Res 2022;6(5):e37302 [FREE Full text] [doi: 10.2196/37302] [Medline: 35526201]
15. Beatty C, Malik T, Meheli S, Sinha C. Evaluating the therapeutic alliance with a free-text CBT conversational agent (Wysa): a mixed-methods study. Front Digit Health 2022;4:847991 [FREE Full text] [doi: 10.3389/fdgth.2022.847991] [Medline: 35480848]
16. Meheli S, Sinha C, Kadaba M. Understanding people with chronic pain who use a cognitive behavioral therapy-based artificial intelligence mental health app (Wysa): mixed methods retrospective observational study. JMIR Hum Factors 2022;9(2):e35671 [FREE Full text] [doi: 10.2196/35671] [Medline: 35314422]
17. Heinz MV, Mackin DM, Trudeau BM, Bhattacharya S, Wang Y, Banta HA, et al. Randomized trial of a generative AI Chatbot for mental health treatment. NEJM AI 2025;2(4):AIoa2400802. [doi: 10.1056/aioa2400802]
18. Hasei J, Hanzawa M, Nagano A, Maeda N, Yoshida S, Endo M, et al. Empowering pediatric, adolescent, and young adult patients with cancer utilizing generative AI chatbots to reduce psychological burden and enhance treatment engagement: a pilot study. Front Digit Health 2025;7:1543543 [FREE Full text] [doi: 10.3389/fdgth.2025.1543543] [Medline: 40070545]
19. McCradden MD, Stephenson EA, Anderson JA. Clinical research underlies ethical integration of healthcare artificial intelligence. Nat Med 2020;26(9):1325-1326. [doi: 10.1038/s41591-020-1035-9] [Medline: 32908273]
20. Van der Vegt AH, Scott I, Dermawan K, Schnetler R, Kalke V, Lane P. Implementation frameworks for end-to-end clinical AI: derivation of the SALIENT framework. J Am Med Inform Assoc 2023;30(9):1503-1515 [FREE Full text] [doi: 10.1093/jamia/ocad088] [Medline: 37208863]
21. The DECIDE-AI Steering Group. DECIDE-AI: new reporting guidelines to bridge the development-to-implementation gap in clinical artificial intelligence. Nat Med 2021;27(2):186-187. [doi: 10.1038/s41591-021-01229-5] [Medline: 33526932]
22. ManagingLife Inc. 2025. URL: https://www.managinglife.com/ [accessed 2026-02-21]
23. IEEE Computer Society, Standards Committee Working Group. Standard for large language model agents for AI-powered education. AI Standards Hub. 2023. URL: https://aistandardshub.org/forums/topic/standard-for-large-language-model-agents-for-ai-powered-education-3/ [accessed 2026-02-21]
24. Prolific. URL: https://www.prolific.com [accessed 2026-02-28]
25. Macea DD, Gajos K, Daglia Calil YA, Fregni F. The efficacy of web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010;11(10):917-929 [FREE Full text] [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
26. GORILLA. URL: https://www.gorilla.sc/ [accessed 2026-02-28]
27. Brooke J. SUS: a retrospective. J Usability Stud 2013;2:29-40 [FREE Full text]
28. Milosevic I, Levy HC, Alcolado GM, Radomsky AS. The treatment acceptability/adherence scale: moving beyond the assessment of treatment effectiveness. Cogn Behav Ther 2015;44(6):456-469. [doi: 10.1080/16506073.2015.1053407] [Medline: 26091250]
29. Kelley ML, Heffer RW, Gresham FM, Elliott SN. Treatment Evaluation Inventory—Short Form. APA PsycNet 2016. [doi: 10.1037/t49940-000]
30. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. Psychother Theory Res Pract 1979;16(3):252-260. [doi: 10.1037/h0085885]
31. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS-29 v2.0 profile physical and mental health summary scores. Qual Life Res 2018;27(7):1885-1891 [FREE Full text] [doi: 10.1007/s11136-018-1842-3] [Medline: 29569016]
32. Swinkels-Meewisse EJCM, Zuyd H, Vlaeyen JW, Leuven KU, Oostendorp RAB. Psychometric properties of the Tampa Scale for Kinesiophobia and the Fear-Avoidance Beliefs Questionnaire in acute low back pain. Man Ther 2003;8(1):29-36. [doi: 10.1054/math.2002.0484] [Medline: 12586559]
33. Ankawi B, Slepian PM, Himawan LK, France CR. Validation of the pain resilience scale in a chronic pain sample. J Pain 2017;18(8):984-993 [FREE Full text] [doi: 10.1016/j.jpain.2017.03.013] [Medline: 28428092]
34. Slepian PM, Ankawi B, Himawan LK, France CR. Development and initial validation of the pain resilience scale. J Pain 2016;17(4):462-472 [FREE Full text] [doi: 10.1016/j.jpain.2015.12.010] [Medline: 26748044]
35. McCracken LM, Vowles KE, Eccleston C. Acceptance of chronic pain: Component analysis and a revised assessment method. Pain 2004;107(1-2):159-166. [doi: 10.1016/j.pain.2003.10.012] [Medline: 14715402]
36. Fish RA, McGuire B, Hogan M, Morrison T, Stewart I. Validation of the chronic pain acceptance questionnaire (CPAQ) in an internet sample and development and preliminary validation of the CPAQ-8. Pain 2010;149(3):435-443. [doi: 10.1016/j.pain.2009.12.016] [Medline: 20188472]
37. Vowles KE, McCracken LM, McLeod C, Eccleston C. The Chronic Pain Acceptance Questionnaire: confirmatory factor analysis and identification of patient subgroups. Pain 2008;140(2):284-291. [doi: 10.1016/j.pain.2008.08.012] [Medline: 18824301]
38. Flesch R. A new readability yardstick. J Appl Psychol 1948;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
39. Eleyan D, Othman A, Eleyan A. Enhancing software comments readability using Flesch Reading Ease Score. Information 2020;11(9):430. [doi: 10.3390/info11090430]

XSL•FO
RenderX

40. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, DECIDE-AI expert group. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ 2022;377:e070904 [FREE Full text] [doi: 10.1136/bmj-2022-070904] [Medline: 35584845]
41. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
42. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
43. Aljamaan F, Malki KH, Alhasan K, Jamal A, Altamimi I, Khayat A, et al. ChatGPT-3.5 system usability scale early assessment among healthcare workers: horizons of adoption in medical practice. Heliyon 2024;10(7):e28962 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e28962] [Medline: 38623218]
44. Mulia AP, Piri PR, Tho C. Usability analysis of text generation by ChatGPT OpenAI using system usability scale method. Procedia Comput Sci 2023;227:381-388. [doi: 10.1016/j.procs.2023.10.537]
45. Nikou SA, Guliya A, Van VS, Chang M. A generative artificial intelligence empowered chatbot: system usability and student teachers' experience. In: Sifaleras A, Lin F, editors. Generative Intelligence and Intelligent Tutoring Systems. Gener Intell Intell Tutoring Syst Cham: Springer Nature Switzerland; 2024:330-340.
46. Borsci S, Malizia A, Schmettow M, van der Velde F, Tariverdiyeva G, Balaji D, et al. The chatbot usability scale: the design and pilot of a usability scale for interaction with AI-based conversational agents. Pers Ubiquit Comput 2021;26(1):95-119. [doi: 10.1007/s00779-021-01582-9]
47. Scheutzow J, Attoe C, Harwood J. Acceptability of web-based mental health interventions in the workplace: systematic review. JMIR Ment Health 2022;9(5):e34655 [FREE Full text] [doi: 10.2196/34655] [Medline: 35544305]
48. Hennessy RW, Rumble D, Christian M, Brown DA, Trost Z. A graded exposure, locomotion-enabled virtual reality app during walking and reaching for individuals with chronic low back pain: cohort gaming design. JMIR Serious Games 2020;8(3):e17799 [FREE Full text] [doi: 10.2196/17799] [Medline: 32773381]
49. Malouin-Lachance A, Capolupo J, Laplante C, Hudon A. Does the digital therapeutic alliance exist? Integrative review. JMIR Ment Health 2025;12:e69294-e69294. [doi: 10.2196/69294]
50. Holmes MB, Scott A, Camarinos J, Marinko L, George SZ. Working Alliance Inventory (WAI) and its relationship to patient-reported outcomes in painful musculoskeletal conditions. Disabil Rehabil 2023;45(8):1363-1369. [doi: 10.1080/09638288.2022.2060337] [Medline: 35416110]
51. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. J Couns Psychol 1989;36(2):223-233. [doi: 10.1037/0022-0167.36.2.223]
52. Mills SEE, Nicolson KP, Smith BH. Chronic pain: a review of its epidemiology and associated factors in population-based studies. Br J Anaesth 2019;123(2):e273-e283 [FREE Full text] [doi: 10.1016/j.bja.2019.03.023] [Medline: 31079836]
53. Niederstrasser NG, Attridge N. Associations between pain and physical activity among older adults. PLoS One 2022;17(1):e0263356 [FREE Full text] [doi: 10.1371/journal.pone.0263356] [Medline: 35089966]
54. Todd J, Pickup B, Sharpe L. Interpretation bias and the transition from acute to chronic pain. Pain 2024;165(2):357-364. [doi: 10.1097/j.pain.0000000000003016] [Medline: 37624880]

Abbreviations

ACT: acceptance and commitment therapy
AI: artificial intelligence
CPAQ: Chronic Pain Acceptance Questionnaire
DECIDE-AI: Developmental and Exploratory Clinical Investigation of Decision-Support Systems Driven by Artificial Intelligence
DHI: digital health intervention
LLM: large language model
PROMIS-29: Patient Reported Outcomes Measurement Information System-29, version 2.0
PRS: Pain Resilience Scale
SME: subject matter expert
SUS: System Usability Scale
TAAS: Treatment Acceptability/Adherence Scale
TEI-SF: Treatment Evaluation Inventory-Short Form
WAI-C: Working Alliance Inventory-Client Version
TSK: Tampa Scale of Kinesiophobia

XSL•FO
RenderX

Edited by S Munce; submitted 12.Nov.2025; peer-reviewed by J Greenberg, KH Lin; comments to author 19.Dec.2025; revised version received 27.Jan.2026; accepted 10.Feb.2026; published 20.Mar.2026.

Please cite as:
Slepian PM, Buryk-Iggers S, Lomanowska AM, Nguyen B, Janmohamed T, Clarke H, Katz J, Niederstrasser NG
A Therapeutic Conversational Agent (Solace) for Management of Chronic Pain: Acceptability and Usability Study
JMIR Rehabil Assist Technol 2026;13:e87689
URL: https://rehab.jmir.org/2026/1/e87689
doi:10.2196/87689
PMID:41711568

©P Maxwell Slepian, Stephanie Buryk-Iggers, Anna M Lomanowska, Binh Nguyen, Tahir Janmohamed, Hance Clarke, Joel Katz, Nils G Niederstrasser. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 20.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Barriers to Adoption of Electronic Low Vision Aids Among Eye Care Professionals in Jordan: Descriptive Cross-Sectional Study

Areej Okasheh-Otoom, PhD

Faculty of Applied Medical Sciences, Jordan University of Science and Technology (JUST), PO Box 3030, Irbid, Jordan

Corresponding Author:
Areej Okasheh-Otoom, PhD
Faculty of Applied Medical Sciences, Jordan University of Science and Technology (JUST), PO Box 3030, Irbid, Jordan

Abstract

Background: Digital, smart, and electronic low vision aids (LVAs) have expanded options for visual rehabilitation and functional independence among people with visual impairment. However, adoption of these technologies remains limited, particularly in low- and middle-income countries such as Jordan, where access, affordability, and training resources may be constrained.

Objective: To examine barriers to the adoption of electronic LVAs and identify factors associated with their use among eye care professionals in Jordan.

Methods: A descriptive cross-sectional survey was conducted among 270 eye care professionals working in hospitals, rehabilitation centers, and private clinics across Jordan. The questionnaire assessed awareness, training exposure, institutional support, and perceived barriers related to electronic LVAs. Descriptive statistics and inferential analyses were used to examine adoption patterns and predictors, with statistical significance set at *P*<.05.

Results: Of the 270 participants, 156 (57.8%) were optometrists, 78 (28.9%) were ophthalmologists, and 36 (13.3%) were low vision specialists. The mean age was 36 (SD 8) years, and the mean professional experience was 12 (SD 6) years. Overall, 117 of 270 (42.2%) participants reported current use or recommendation of electronic LVAs. The most frequently reported barriers were high device cost (n=213, 79%), lack of training (n=184, 68.1%), limited institutional support (n=173, 64%), and low patient awareness (n=154, 57%). In multivariable analysis, greater training exposure (odds ratio [OR] 1.82, 95% CI 1.31 - 2.53; *P*<.001), stronger institutional support (OR 1.48, 95% CI 1.12 - 1.96; *P*=.008), and higher awareness scores (OR 1.35, 95% CI 1.05 - 1.72; *P*=.02) were positively associated with aid adoption, whereas high device cost was negatively associated with aid adoption (OR 0.41, 95% CI 0.27 - 0.62; *P*<.001).

Conclusions: Adoption of electronic LVAs among eye care professionals in Jordan remains limited. Cost, training exposure, and institutional support are key factors influencing uptake. These findings suggest that strengthening professional training and institutional support may facilitate broader integration of electronic LVAs into low vision rehabilitation services.

(JMIR Rehabil Assist Technol 2026;13:e87685) doi:10.2196/87685

KEYWORDS

low vision; electronic low vision aids; assistive technology; digital rehabilitation; adoption barriers; Jordan; optometry; ophthalmology

Introduction

Visual impairment affects more than 2.2 billion people worldwide, and at least 1 billion cases could have been prevented or remain unaddressed [1]. Low vision rehabilitation aims to enhance residual vision and functional independence using optical and electronic assistive technologies [2]. Recent advances in portable, digital, and smart low vision aids (LVAs), including magnifiers, wearable displays, and smartphone-based systems, have significantly improved navigation and access to information [2-5]. These devices, collectively termed electronic LVAs, complement traditional optical aids by offering adjustable magnification, contrast enhancement, and speech feedback [6,7].

Despite growing evidence supporting their functional and psychosocial benefits [8-11], adoption of electronic LVAs remains limited globally [12]. Reported barriers include high device cost, inadequate professional training, limited awareness among clinicians and patients, and insufficient institutional and policy support [13-15]. These challenges are particularly pronounced in low- and middle-income countries, where limited service infrastructure, fragmented referral pathways, and absence of sustainable funding mechanisms further restrict access to electronic low vision technologies [2].

In Jordan, visual impairment represents the most prevalent form of disability, affecting approximately 6% to 7% of the population aged 5 years or older, according to national statistics [16]. Although optometric education and eye care infrastructure

XSL•FO
RenderX

have expanded in recent years, integration of structured low vision rehabilitation services within the health care system remains inconsistent [7,15,17]. Available local evidence indicates that assistive and digital technologies are used predominantly in educational or professional contexts, with limited incorporation into formal low vision rehabilitation pathways. Regional studies further suggest that fewer than half of eye care professionals routinely prescribe or recommend electronic LVAs, underscoring persistent gaps in training, access, and service delivery [3,18,19]. Understanding context-specific barriers in Jordan is therefore essential for developing targeted professional training and service-level interventions.

Technology adoption frameworks, including the technology acceptance model and the unified theory of acceptance and use of technology, provide a useful lens for examining these challenges [20]. These models emphasize perceived usefulness, ease of use, and facilitating conditions as key determinants of adoption [21]. In low vision rehabilitation, these constructs translate into clinician awareness, perceived patient benefit, affordability, and availability of institutional support. While prior studies have examined assistive technology use in various contexts, limited evidence exists regarding how these adoption factors interact among eye care professionals involved in low vision care in Jordan. The aim of this study was to examine awareness, training exposure, institutional support, perceived barriers, and factors associated with the adoption and recommendation of electronic low vision aids among eye care professionals in Jordan.

Methods

Study Design

A descriptive cross-sectional study was conducted among eye care professionals working in public and private hospitals, university eye clinics, and vision rehabilitation centers across Jordan. This study was implemented as an online cross-sectional survey and is reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guidelines (Checklist 1). Invitations were distributed electronically through established professional networks, including national professional associations, institutional mailing lists, and closed professional communication groups (eg, email lists and secure messaging platforms) commonly used by ophthalmologists, optometrists, and low-vision specialists in Jordan. These networks included clinicians working in public, private, and academic settings across Jordan. Distribution was not stratified by geographic region or sector; therefore, participation depended on individual access to and engagement with these professional channels. Survey items were presented in a fixed order. Completion of all mandatory closed-ended items was required for submission, whereas open-ended questions were optional. Only fully completed questionnaires were included in the analysis.

Eligible participants included optometrists, ophthalmologists, and low-vision specialists currently practicing in Jordan. Inclusion criteria were (1) possession of a valid professional license from the Ministry of Health, (2) at least 1 year of clinical experience, and (3) direct involvement in patient care. Professionals in administrative roles or not directly engaged in patient management were excluded. A total of 300 invitations were distributed electronically, of which 270 completed responses were received (response rate 90%). The response rate was calculated at the overall survey level based on completed questionnaires relative to the total number of invitations distributed. The survey was distributed through professional online groups and networks of eye care clinicians. Because the questionnaire was not sent to a fixed list of individuals, accurate denominators for each professional group were not available; therefore, group-specific response rates could not be calculated. An a priori sample size estimation was conducted using G*Power (version 3.1) [22] for multiple linear regression analysis. Assuming a medium effect size (f^2=0.15), an α level of .05, statistical power of 0.80, and up to 10 predictors, the minimum required sample size was 118 participants. The final sample of 270 respondents exceeded this requirement, providing adequate statistical power for the planned analyses.

A self-administered electronic questionnaire was developed following a review of existing validated surveys on assistive technology adoption [10-12,14,18,23]. The instrument comprised five domains: (1) awareness, (2) training exposure, (3) institutional support, (4) perceived barriers, and (5) adoption of electronic low vision aids (Table 1). Items were rated on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). Optional open-ended questions were included at the end of the survey to allow participants to provide additional comments or perspectives not captured by the closed-ended items. Selected anonymized comments are presented in the Results section to illustrate key quantitative findings; no formal qualitative coding or thematic analysis was performed.

XSL•FO
RenderX

Table . Survey domains, representative items, and mean Likert scores (5-point scale; 1=strongly disagree, 5=strongly agree; N=270 respondents).

Domain and representative item	Score, mean (SD)
Awareness	
I am familiar with the main types of electronic LVAs.[a]	3.6 (0.7)
I understand the difference between optical and electronic devices.	3.5 (0.8)
I know about digital magnifiers and smart glasses.	3.8 (0.8)
I am aware of the benefits of electronic LVAs in improving reading and mobility.	3.7 (0.7)
I have read or heard about electronic LVAs through professional sources.	3.4 (0.9)
I know where to obtain information about new low vision technologies.	3.5 (0.8)
Training exposure	
I attended workshops or training sessions on electronic LVAs.	2.7 (0.9)
I received hands-on demonstration during my education.	2.8 (0.8)
I participated in online webinars related to electronic LVAs.	2.9 (0.9)
I practiced using electronic magnifiers with patients.	2.6 (0.9)
I have self-trained using online tutorials or videos.	2.7 (0.8)
Institutional support	
My workplace provides access to low vision devices for demonstration.	2.9 (0.8)
My department encourages use of electronic aids in patient management.	3.0 (0.9)
Institutional funding is available for electronic LVA training.	2.7 (0.9)
My clinic has a designated low vision service or room.	2.8 (0.8)
Technical support is available for device maintenance.	2.8 (0.9)
Perceived barriers	
The high cost of devices limits their use in clinical practice.	4.2 (0.7)
Lack of professional training prevents effective use of electronic LVAs.	3.9 (0.8)
Limited institutional support restricts adoption.	4.1 (0.8)
Low patient awareness	3.8 (0.8)
There is a lack of Arabic-language software support.	3.7 (0.9)
Devices are difficult to use for older adults or individuals with limited formal education.	3.9 (0.7)
There is limited technical support for device maintenance.	3.6 (0.8)
Import regulations make devices unavailable.	3.5 (0.9)

XSL•FO
RenderX

Domain and representative item	Score, mean (SD)
Adoption behavior	
I currently prescribe or recommend electronic LVAs to my patients.	3.5 (0.8)
I routinely demonstrate electronic LVAs in my clinic.	3.4 (0.8)
I feel confident teaching patients how to use electronic LVAs.	3.6 (0.7)
I intend to integrate electronic LVAs more frequently in future practice.	3.8 (0.8)
I believe electronic LVAs improve my patients' quality of life.	4.1 (0.7)

[a]LVA: low vision aid.

The survey was administered in English, as English is the primary language of instruction and clinical documentation among eye care professionals in Jordan. Participants were able to request clarification for any items if needed. The questionnaire underwent expert review by 3 senior vision rehabilitation specialists to establish content validity, followed by pilot testing with 25 clinicians. Pilot data were excluded from the final analysis. Minor linguistic refinements were made to improve clarity and internal consistency. Internal consistency was assessed using Cronbach α for each domain (Table 2). Values ≥0.70 were considered acceptable.

Table . Internal consistency (Cronbach α) of questionnaire domains based on the final survey sample (N=270).

Domain	Cronbach α
Awareness	0.86
Training exposure	0.81
Institutional support	0.83
Perceived barriers	0.88
Adoption behavior	0.79

Data were analyzed using SPSS software (version 29; IBM Corp). Descriptive statistics summarized participant characteristics and mean domain scores. Inferential analyses included chi-square tests to compare aid adoption rates across professional groups, independent-sample *t*-tests to compare training hours between adopters and nonadopters, and multivariable binary logistic regression to identify predictors of electronic low vision aid adoption. Multicollinearity was assessed using variance inflation factors (<2). Model fit was evaluated using the Hosmer-Lemeshow test (*P*>.05). Statistical significance was set at *P*<.05 (2-tailed).

Professional experience was assessed using a single self-reported item asking participants to indicate their years of clinical practice in eye care. Awareness of electronic and digital low vision aids was measured using a composite awareness score derived from multiple Likert scale items assessing familiarity with device types, indications, and clinical applications. Training exposure was assessed by asking participants whether they had received formal or informal training related to electronic LVAs (yes or no), as well as the total number of training hours completed.

Internal consistency was calculated using Cronbach α for each domain (Table 2). Values ≥0.70 were considered acceptable. The α coefficients were calculated using responses from the final survey participants. Expert review (n=3) was used solely for content validity and was not included in the final internal consistency analysis.

Ethical Considerations

This study was approved by the Institutional Review Board of Jordan University of Science and Technology (IRB/642) and was conducted in accordance with the principles of the Declaration of Helsinki. Participation was voluntary, and electronic informed consent was obtained from all participants prior to survey initiation. The survey was administered anonymously through a secure online platform. No personally identifiable information (such as names, identification numbers, or contact details) was collected. Data were stored on password-protected devices accessible only to the principal investigator and were analyzed in aggregate form to ensure participant confidentiality. Measures were implemented to minimize duplicate submissions.

Results

A total of 270 eye care professionals completed the survey, representing a 90% response rate. Participants included 156 optometrists (57.8%), 78 ophthalmologists (28.9%), and 36 low vision specialists (13.3%) working across multiple practice settings in Jordan. The mean age was 36 (SD 8) years, and 149 of 270 (55.2%) participants were female. Respondents reported

XSL•FO
RenderX

a mean of 12 (SD 6) years of professional experience, indicating a clinically experienced cohort.

Across all participants, the mean awareness score regarding electronic LVAs was 3.2 (SD 0.7) on a 5-point Likert scale, suggesting moderate familiarity. The mean training exposure time was 18.4 (SD 12.5) hours. However, substantial variability was observed, with many clinicians reporting fewer than 10 hours of formal training. These data are summarized in Table 3.

Table . Participant characteristics (N=270).

Characteristic	Value
Profession, n (%)	
Optometrist	156 (57.8)
Ophthalmologist	78 (28.9)
Low vision specialist	36 (13.3)
Gender, n (%)	
Female	149 (55.2)
Male	121 (44.8)
Age (y), mean (SD)	36 (8)
Professional experience (y), mean (SD)	12 (6)
Awareness score (1-5), mean (SD)	3.2 (0.7)
Training hours, mean (SD)	18.4 (12.5)

Profession-specific analyses revealed that low vision specialists had the highest training exposure (mean 28.5, SD 9.2 hours) and awareness score (mean 3.8, SD 0.6), followed by optometrists (mean 17.6, SD 11.3 training hours; mean 3.3, SD 0.7 awareness score) and ophthalmologists (mean 12.1, SD 8.5 training hours; mean 3.0, SD 0.6 awareness score). Gender and years of experience were not significantly associated with awareness or training exposure (*P*>.05).

Among all respondents, 117 of 270 professionals (42.2%) reported that they currently prescribe or recommend electronic LVAs to patients. Adoption was significantly higher among low vision specialists (22/36, 61%) than optometrists (69/156, 44.2%) and ophthalmologists (26/78, 33%) (χ^2=9.64; *P*=.008). Eye care professionals who reported active aid adoption also demonstrated substantially greater training exposure (mean 23.6, SD 11.2 hours) than nonadopters (mean 14.9, SD 9.8 hours), a difference that was statistically significant (*t*=5.74; *P*<.001). Additionally, adopters also had higher awareness scores than nonadopters (mean 3.6, SD 0.6 vs mean 3.0, SD 0.7; *P*<.001) and higher institutional support scores (mean 3.4, SD 0.8 vs mean 2.9, SD 0.7; *P*=.002).

Overall, these results suggest that professional experience alone does not drive aid adoption; rather, the presence of structured training and institutional support are key determinants of electronic LVA use. The perceived barriers domain comprised 8 Likert scale items representing 8 distinct barriers to electronic LVA adoption (Table 1). Respondents most frequently cited high device cost (213/270, 79%), followed by lack of training opportunities (184/270, 68.1%), limited institutional support (173/270, 64%), and low patient awareness (154/270, 57%; Table 4) as barriers to electronic LVA adoption.

Table . Barriers to electronic low vision aid adoption among Jordanian eye care professionals (N=270).

Perceived barriers	Respondents, n (%)
Import regulations make devices unavailable	111 (41.1)
There is limited technical support for device maintenance	124 (45.9)
Devices are difficult to use for older adults or individuals with limited formal education	130 (48.1)
There is a lack of Arabic-language software support	140 (51.9)
Low patient awareness	154 (57)
Limited institutional support restricts adoption	173 (64)
Lack of professional training prevents effective use of electronic low vision aids	184 (68.1)
The high cost of devices limits their use in clinical practice	213 (79)

Subgroup analysis revealed that cost concerns were particularly pronounced among optometrists (131/156, 84.0%), whereas low vision specialists reported greater concern regarding limited patient awareness (22/36, 61%) and lack of Arabic-language

XSL•FO
RenderX

device interfaces (20/36, 56%). In contrast, ophthalmologists more frequently emphasized institutional limitations (56/78, 72%) and unclear referral mechanisms (50/78, 64%) as major obstacles.

One-way ANOVA demonstrated significant between-group differences for lack of professional training ($F_{2,267}$=6.21; *P*=.002) and limited maintenance or technical support ($F_{2,267}$=4.89; *P*=.008). Post hoc analyses indicated that optometrists perceived these barriers more strongly than ophthalmologists and low vision specialists. Although cost- and patient-related barriers were commonly reported across all professional groups, no statistically significant between-group differences were observed for these domains (Multimedia Appendix 1). Comments from participants reinforced these findings, as illustrated by the following representative quotes:

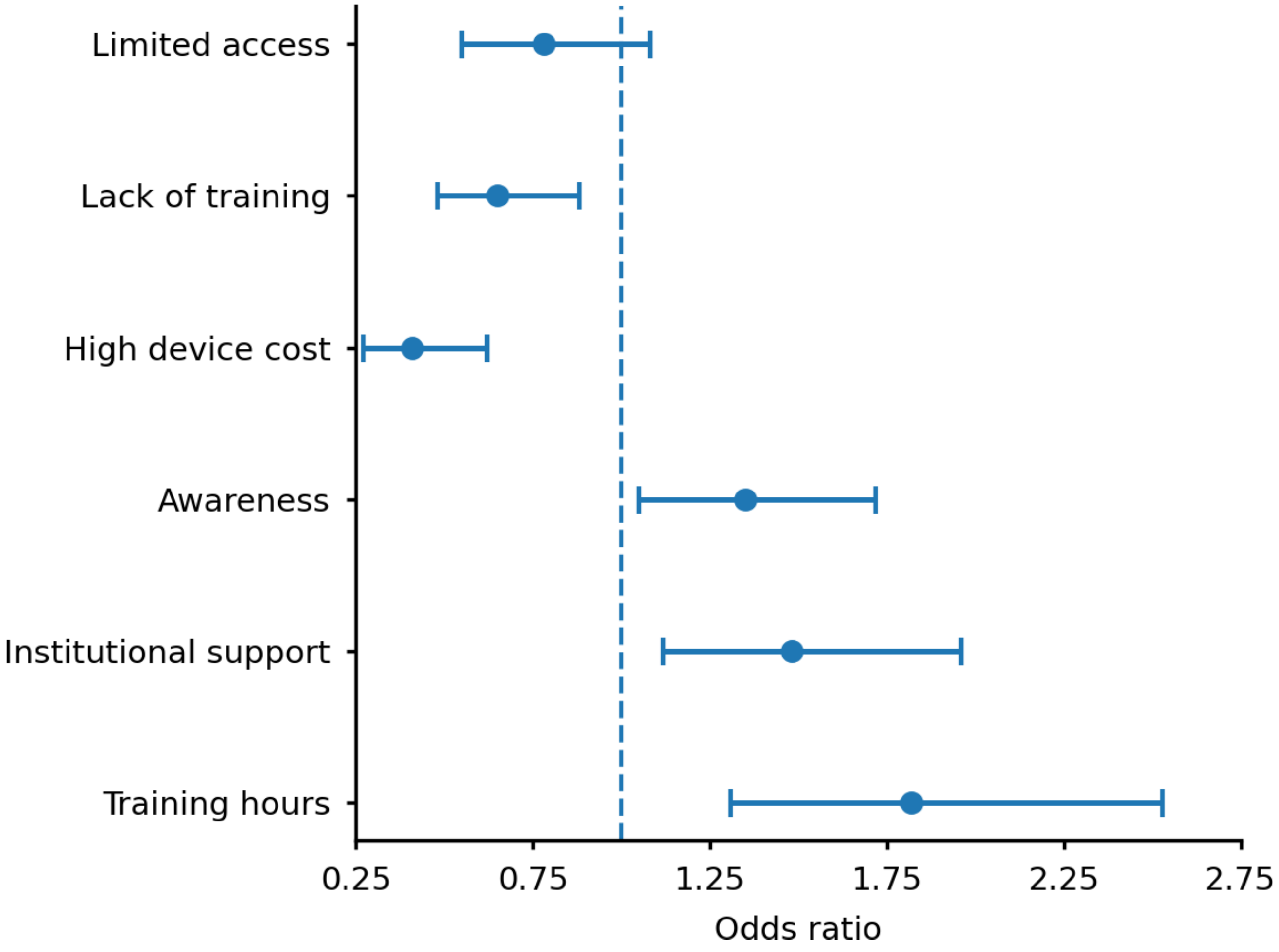
> *Even when devices are available, few clinicians are trained to demonstrate them effectively.* [Participant 47, optometrist]
>
> *Cost remains prohibitive for many families, especially without Ministry of Health support.* [Participant 112, ophthalmologist]

These results point to a multifactorial set of barriers (financial, educational, and systemic) that collectively hinder widespread electronic LVA integration into clinical practice.

A binary logistic regression model was fitted to identify predictors of electronic LVA adoption (Multimedia Appendix 2). The model explained 42% of the variance in adoption behavior (Nagelkerke R^2=0.42) and correctly classified 79% of cases. Training hours emerged as the strongest positive predictor of electronic low vision aid adoption (odds ratio [OR] 1.82, 95% CI 1.31 - 2.53; *P*<.001). Higher institutional support (OR 1.48, 95% CI 1.12 - 1.96; *P*=.008) and greater awareness (OR 1.35, 95% CI 1.05 - 1.72; *P*=.02) were also significantly associated with increased adoption. In contrast, high device cost was a strong negative predictor (OR 0.41, 95% CI 0.27 - 0.62; *P*<.001), while lack of training was associated with a reduced likelihood of adoption (OR 0.65, 95% CI 0.48 - 0.88; *P*=.006). Limited access was not a statistically significant predictor (OR 0.78, 95% CI 0.55 - 1.08; *P*=.14).

Figure 1 illustrates the direction and strength of associations, showing that training exposure and institutional support exerted the most substantial positive influence, while high device cost exerted the strongest negative effect.

Figure 1. Forest plot showing adjusted odds ratios (with 95% CIs) for predictors of electronic low vision aid adoption.



Discussion

This study examined barriers to the adoption and recommendation of electronic LVAs among eye care professionals in Jordan. The main findings indicate that high device cost and lack of professional training were significantly associated with a lower likelihood of recommending electronic LVAs, whereas higher awareness and stronger institutional

XSL•FO
RenderX

support were positively associated with adoption. Although overall awareness scores were moderate, fewer than half of respondents reported current use or recommendation of these devices, indicating a persistent gap between awareness and routine clinical implementation. These findings are consistent with reports from other low- and middle-income countries, where cost and training constraints are repeatedly identified as dominant barriers to electronic assistive technology uptake [10-12,14,18,23].

Multivariable regression analysis further demonstrated that training exposure, institutional support, and awareness independently predicted adoption of electronic LVAs, whereas perceived device cost was negatively associated with adoption. Together, these findings suggest that adoption is influenced by both individual-level factors (eg, professional knowledge and perceived benefit) and system-level factors (eg, organizational support and resource availability). This dual influence is consistent with prior research indicating that isolated improvements in clinician awareness are insufficient to promote sustained technology adoption without parallel investment in training and institutional infrastructure.

International studies report similar patterns of underuse. For example, surveys of eye care professionals in high-income settings have shown that fewer than half routinely recommend electronic or digital magnification devices despite recognizing their potential clinical benefit [13]. Studies from the Asia Pacific region and the Middle East likewise highlight limited institutional resources, high device cost, and insufficient professional training as persistent barriers to adoption [2,20]. These parallels suggest that the challenges identified in this study reflect broader implementation issues rather than context-specific resistance to technology.

Interpretation of these findings should also consider broader health system context in Jordan. Reports and policy analyses indicate that low vision rehabilitation services are often fragmented, with limited availability of designated low vision clinics, variable referral pathways, and uneven access to continuing professional education in rehabilitation technologies [1,12,24,25]. Although these system-level characteristics were not directly measured in this survey, they provide important context for understanding why institutional support and training emerged as prominent determinants of adoption. Global guidance from the World Health Organization emphasizes the need for integrated national rehabilitation frameworks to support equitable access to assistive and digital technologies, particularly in resource-constrained settings [3,20]. In this study, clinicians who perceived stronger institutional facilitation and greater patient benefit were more likely to recommend electronic LVAs, consistent with core constructs of the technology acceptance model and the unified theory of acceptance and use of technology framework [21].

The findings underscore the need for coordinated strategies to support electronic LVA adoption. Structured professional training in digital rehabilitation should be incorporated into undergraduate curricula and continuing professional development programs. Health care institutions may benefit from establishing demonstration or referral pathways that allow clinicians to trial devices with patients. In addition, funding mechanisms—such as subsidy schemes or public-private partnerships—may help mitigate cost barriers, particularly for older adults and children with visual impairment. These implications are aligned with international calls to strengthen rehabilitation systems and data collection under the World Health Organization Rehabilitation 2030 initiative [3].

Several limitations warrant consideration. The cross-sectional design precludes causal inference, and adoption behavior was assessed using self-reported measures rather than objective indicators such as prescription records or device use data. In addition, although the survey instrument was informed by established technology adoption frameworks, formal construct validity testing was not undertaken. This study also did not explicitly examine potential regional or socioeconomic variation within the eye care workforce, which may influence access to training, institutional support, and technology availability. These factors may limit generalizability and should be addressed in future research.

Future studies should use longitudinal or mixed methods designs to examine how training interventions, institutional policies, and funding mechanisms influence adoption over time. Incorporating validated measurement instruments, objective adoption indicators, and stratified analyses by region or practice setting would further strengthen the evidence base and inform targeted implementation strategies.

In conclusion, adoption of digital, smart, and electronic LVAs among eye care professionals in Jordan remains limited. Cost, training exposure, and institutional support were the primary factors associated with uptake in this study. These findings suggest that strengthening professional training opportunities, improving institutional facilitation, and exploring sustainable funding mechanisms may support broader integration of digital assistive technologies within low vision rehabilitation services. Greater coordination among clinicians, policymakers, and industry stakeholders may help address persistent implementation barriers and improve access to low vision care.

Acknowledgments

The author thanks the participating optometrists, ophthalmologists, and low vision specialists across Jordan for their valuable contributions. Appreciation is also extended to the statistical support team and research assistants who facilitated data entry and analysis.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

XSL•FO
RenderX

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Subgroup analysis of mean barrier scores by professional category (N=270).
[DOCX File, 14 KB - rehab_v13i1e87685_app1.docx]

Multimedia Appendix 2
Logistic regression analysis for predictors of electronic LVA adoption.
[DOCX File, 14 KB - rehab_v13i1e87685_app2.docx]

Checklist 1
CHERRIES checklist.
[DOCX File, 15 KB - rehab_v13i1e87685_app3.docx]

References

1. World report on vision. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/world-report-on-vision [accessed 2026-02-04]
2. Burton MJ, Ramke J, Marques AP, et al. The Lancet Global Health Commission on Global Eye Health: vision beyond 2020. Lancet Glob Health 2021 Apr;9(4):e489-e551. [doi: 10.1016/S2214-109X(20)30488-5] [Medline: 33607016]
3. Rehabilitation 2030: a call for action. World Health Organization. 2017. URL: https://www.who.int/publications/m/item/rehabilitation-2030-a-call-for-action [accessed 2026-02-04]
4. Miller A, Crossland MD, Macnaughton J, Latham K. The usefulness of a wearable electronic vision enhancement system for people with age-related macular degeneration: a randomized crossover trial. Transl Vis Sci Technol 2025 Sep 2;14(9):8. [doi: 10.1167/tvst.14.9.8] [Medline: 40905748]
5. Miller A, Macnaughton J, Crossland MD, Latham K. “Such a lot of bother”: qualitative results of a home trial of a wearable electronic vision enhancement system for people with age-related macular degeneration. Ophthalmic Physiol Opt 2025 May;45(3):699-712. [doi: 10.1111/opo.13453] [Medline: 39865322]
6. Crossland MD, Silva RS, Macedo AF. Smartphone, tablet computer and e-reader use by people with vision impairment. Ophthalmic Physiol Opt 2014 Sep;34(5):552-557. [doi: 10.1111/opo.12136] [Medline: 25070703]
7. Bakkar MM, Alzghoul EA, Haddad MF. Clinical characteristics and causes of visual impairment in a low vision clinic in northern Jordan. Clin Ophthalmol 2018;12:631-637. [doi: 10.2147/OPTH.S153754] [Medline: 29662299]
8. Virgili G, Acosta R, Bentley SA, Giacomelli G, Allcock C, Evans JR. Reading aids for adults with low vision. Cochrane Database Syst Rev 2018 Apr 17;4(4):CD003303. [doi: 10.1002/14651858.CD003303.pub4] [Medline: 29664159]
9. van Nispen RM, Virgili G, Hoeben M, et al. Low vision rehabilitation for better quality of life in visually impaired adults. Cochrane Database Syst Rev 2020 Jan 27;1(1):CD006543. [doi: 10.1002/14651858.CD006543.pub2] [Medline: 31985055]
10. Sivakumar P, Vedachalam R, Kannusamy V, et al. Barriers in utilisation of low vision assistive products. Eye (Lond) 2020 Feb;34(2):344-351. [doi: 10.1038/s41433-019-0545-5] [Medline: 31388131]
11. Sarika G, Venugopal D, Sailaja MVS, Evangeline S, Krishna Kumar R. Barriers and enablers to low vision care services in a tertiary eye care hospital: a mixed method study. Indian J Ophthalmol 2019 Apr;67(4):536-540. [doi: 10.4103/ijo.IJO_1215_18] [Medline: 30900589]
12. Stolwijk ML, van Nispen RMA, van der Ham AJ, Veenman E, van Rens G. Barriers and facilitators in the referral pathways to low vision services from the perspective of patients and professionals: a qualitative study. BMC Health Serv Res 2023 Jan 21;23(1):64. [doi: 10.1186/s12913-022-09003-0] [Medline: 36681848]
13. Codina CJ, Rhodes M. Low vision services provision throughout NHS trusts in the UK. Br Ir Orthopt J 2023;19(1):64-70. [doi: 10.22599/bioj.293] [Medline: 37577067]
14. Arshad M, Younas A, Amin N. Assessment of level of awareness and barriers to low vision rehabilitation among optometrists in Pakistan. J Vis Impair Blind 2024 Mar;118(2):110-117. [doi: 10.1177/0145482X241248461]
15. Qutishat Y, Shublaq S, Masoud M, Alnuman N. Low vision profile in Jordan: a vision rehabilitation center-based study. Healthcare (Basel) 2021;9(1):20. [doi: 10.3390/healthcare9010020]
16. The reality of disability: “functional difficulties” in jordan, based on the data of the General Population and Housing Census 2015. : Department of Statistics (Jordan); 2021 URL: https://dosweb.dos.gov.jo/DataBank/Analytical_Reports/Disability_2021.pdf [accessed 2026-02-24]

XSL•FO
RenderX

17. Gammoh Y, Moodley V. Situation analysis of optometric education in the Eastern Mediterranean region. East Mediterr Health J 2023 Mar 26;29(3):170-176. [doi: 10.26719/emhj.23.019] [Medline: 36987622]
18. Ovenseri-Ogbomo GO, Alghamdi W. Knowledge, attitudes, and practices of optometrists regarding low vision services in Saudi Arabia. TOOPHTJ 2021 Nov 10;15(1):217-288. [doi: 10.2174/1874364102115010217]
19. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
20. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
21. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
22. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
23. Alsaqr AM. Barriers to low vision services among optometrists in Saudi Arabia. TOOPHTJ 2021 Oct 5;15(1):178-188. [doi: 10.2174/1874364102115010178]
24. Global report on assistive technology. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240049451 [accessed 2026-02-24]
25. IAPB vision atlas. International Agency for the Prevention of Blindness. 2023. URL: https://visionatlas.iapb.org/ [accessed 2026-02-06]

Abbreviations

CHERRIES: Checklist for the Reporting of Results of Internet E-Surveys
LVA: low vision aid
OR: odds ratio

Edited by S Munce; submitted 12.Nov.2025; peer-reviewed by A Anderson, MM Hatamleh; revised version received 06.Feb.2026; accepted 10.Feb.2026; published 02.Mar.2026.

Please cite as:
Okasheh-Otoom A
Barriers to Adoption of Electronic Low Vision Aids Among Eye Care Professionals in Jordan: Descriptive Cross-Sectional Study
JMIR Rehabil Assist Technol 2026;13:e87685
URL: https://rehab.jmir.org/2026/1/e87685
doi:10.2196/87685

© Areej Okasheh-Otoom. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Smart Textile Biofeedback Training System for Upper Limb Rehabilitation After Stroke: Co-Design Development and Evaluation Study

Maria Munoz-Novoa[1,2], PhD; Li Guo[3], PhD; Anna Björkquist[3], BSc; Morten B Kristoffersen[4,5], PhD; Peiman Khorramshahi[6], MSc; Leif Sandsjö[2,7], PhD; Margit Alt Murphy[1,8], PhD

[1]Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden

[2]Department of Work Life and Social Welfare, Faculty of Caring Science, Work Life and Social Welfare, University of Borås, Allégatan 1, Borås, Sweden

[3]Department of Textile Technology, Swedish School of Textiles, University of Borås, Borås, Sweden

[4]Chalmers Industriteknik, Gothenburg, Sweden

[5]Department of Engineering Technology, Technical University of Denmark, Ballerup, Denmark

[6]Daralabs AB, Gothenburg, Sweden

[7]Design & Human Factors, Department of Industrial and Materials Science, Chalmers University of Technology, Gothenburg, Sweden

[8]Department of Health and Rehabilitation, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden

Corresponding Author:
Leif Sandsjö, PhD
Department of Work Life and Social Welfare, Faculty of Caring Science, Work Life and Social Welfare, University of Borås, Allégatan 1, Borås, Sweden

Abstract

Background: An increasing number of rehabilitation technologies are being developed to support upper limb rehabilitation after stroke, with smart textile solutions for surface electromyography (sEMG) emerging as a promising approach. Early end-user involvement is crucial for developing user-friendly and clinically valid rehabilitation tools.

Objective: This study aims to refine and evaluate the prototype design and usability of a smart textile biofeedback system for self-administered upper limb training after stroke.

Methods: The training system includes a knitted smart textile sleeve with integrated electrodes over the forearm muscles, an sEMG unit, and tablet-based biofeedback software. An iterative co-design process was followed, including initial testing, demonstration sessions with end users (9 clinicians and 10 individuals with stroke), and a final evaluation of the co-design process. Participants' experiences were gathered through semistructured interviews, analyzed using content analysis, and the User Experience Questionnaire. The co-design team included experts in stroke rehabilitation, textile engineering, biomedical engineering, software development, and human factors, as well as a research partner with lived experience after stroke.

Results: The perspectives of the end users and the expert team were collectively integrated into prototype refinements of the sleeve and training software to meet the needs of the intended target group. The experiences of end users formed 2 main categories: "This could be an exciting new training tool for stroke rehabilitation" and "The tool works well, but some changes could enhance independent training." End users found the smart textile sleeve and biofeedback system easy to use and saw potential for integrating it into their training routines. Both end-user groups rated the system as attractive, stimulating, and novel.

Conclusions: The results of this study establish a necessary ground toward the development of a smart textile sEMG biofeedback system for self-administered upper limb training after stroke. Findings from the co-design process support the continued development and evaluation of the system as a self-administered upper limb training tool for individuals living with stroke.

(JMIR Rehabil Assist Technol 2026;13:e77999) doi:10.2196/77999

KEYWORDS

stroke; surface electromyography; sEMG; surface electrodes; textile electrodes; smart textiles; upper limb function; self-administered; rehabilitation; co-design

XSL•FO
RenderX

Introduction

Regaining upper limb function is a primary goal for individuals after a stroke [1]. In the acute phase, about 50% to 70% of individuals experience impaired arm and hand function, and in the long term, only a smaller percentage fully recover their previous level of function [2-5]. Upper limb impairment often leads to difficulties in performing daily activities and limitations in participation [6,7]. Most improvements occur early in recovery, with slower progress continuing into the chronic phase [5,8]. To follow this recovery pattern, rehabilitation programs are mostly concentrated in the acute and subacute phases, while in the chronic phase, the responsibility for continued training often falls on the individuals themselves [9,10]. To support long-term rehabilitation [1], various rehabilitation technologies have been developed to facilitate independent home-based training, ultimately aiming to improve functioning [11,12].

Surface electromyography (sEMG) biofeedback is a rehabilitation technology that provides real-time visual feedback on muscle activity, thereby supporting motor learning mechanisms for more effective movement execution [13,14]. Previous research in stroke populations indicates that sEMG biofeedback interventions can increase upper limb muscle strength, active range of motion, and awareness of the paretic arm [15-18]. sEMG biofeedback systems commonly use disposable pre-gelled Ag/AgCl electrodes [19]. As an alternative to these traditional electrodes, smart textile solutions with integrated textile-based electrodes offer improved comfort, easier application, and a reduced risk of skin irritation [20-24]. Textile solutions enable prolonged, repeated use, which makes them suitable for self-administered rehabilitation in both clinical and home settings. However, smart textile solutions for sEMG applications in upper limb stroke rehabilitation are in the early stages of development, with existing studies limited to controlled research environments and/or small sample sizes [25-27].

A challenge in developing new rehabilitation technologies is their implementation in real-world settings [12,28,29]. A failure to meet the needs of end users (patients and clinicians) and underestimating the complexity and costs of implementation in clinical practice are common reasons for limited clinical uptake [30-32]. Adopting collaborative co-design methodologies that involve end users in the development process can help identify and address specific user needs early on. This ensures better alignment among the technology, its intended use, and context and has been shown to improve the applicability and acceptance of rehabilitation technologies [30,33,34].

This study aims to refine and evaluate the prototype design and usability of a smart textile biofeedback system for self-administered upper limb training after a stroke, using a co-design approach.

Methods

Ethical Considerations

The study was approved by the Ethics Review Authority (reference 2019-00450/1074-18 and amendment 2024-01703-02) and follows the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Checklist 1) [35]. Before entering the study, all participants completed an approved informed consent form, as authorized by the Swedish Ethics Review Authority, including information about the possibility to withdraw from the study at any time without providing a reason. All collected data were anonymized and safely stored accessible only to the research team. No compensation was given for participation.

Co-Design Process

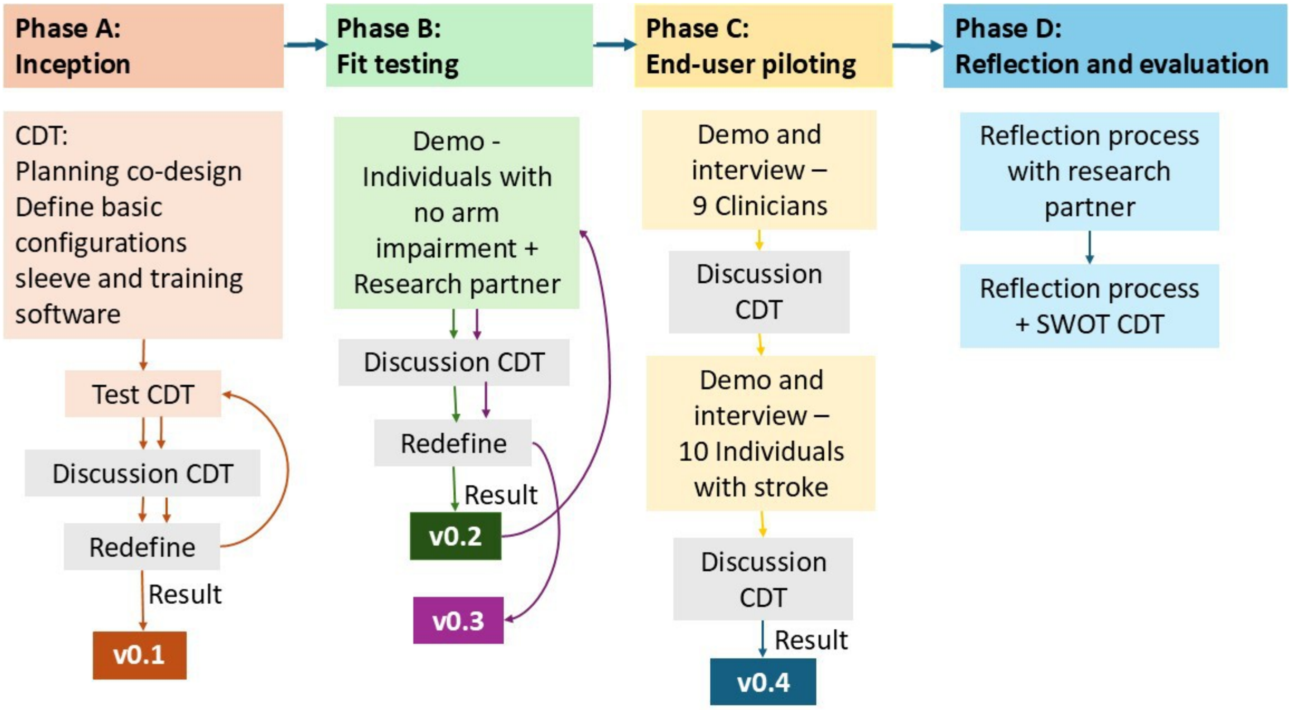
The participatory action research model [36] and the Medical Research Council framework for complex interventions [37] were used to guide the iterative and reflective co-design process, which included 4 phases (Figure 1). Each phase involved iterative cycles of problem definition, planning, acting, observing, reflecting, and redefining the problem [36]. At the end of the codesign process, a strengths, weaknesses, opportunities, and threats analysis [38] was conducted to evaluate the co-design process. The aims, procedures, and actions of each co-design phase are detailed in Table 1.

The co-design team included 7 experts from various fields, including stroke rehabilitation, biomedical engineering, textile engineering, signal processing, software development, and human factors. Most team members had over 15 years of experience in their respective areas, including 1 industry representative with a background in commercial product development. The team also included 2 junior researchers, each with approximately 5 years of experience.

A research partner with a university degree and over 5 years of lived experience with stroke was also included in the co-design team to provide end-user insights during the co-design process. Following the iterative approach of the co-design process, regular physical and digital meetings were held within the co-design team between December 2023 and February 2025.

XSL•FO
RenderX

Figure 1. The 4 phases of the co-design process, guided by the participatory action research model and the Medical Research Council framework, showing the iterations leading to gradually evolving prototype versions (v0.1-v0.4). CDT: co-design team; SWOT: strengths, weaknesses, opportunities, and threats.



XSL•FO
RenderX

Table . Co-design process and refinements of the training system.

Phase	Aim	Procedures	Actions
A: Inception	To set priorities for the co-design process and discuss basic configurations for the training system.	• The CDT[a] met several times to set priorities and define the initial configuration of the training system. • The CDT also tested the first version of the sleeve and software.	• Version 0.1: a conically shaped, circularly knitted sleeve with 2 electrode pairs (wrist flexors and extensors) and a reference electrode at the elbow was developed. The electrodes were connected via snap-button connectors to a wireless sEMG[b] acquisition unit (Shimmer3 EMG, Shimmer Sensing) with an external band. To enhance signal quality, water was applied by 3 finger strokes to each electrode prior to use. Biofeedback was delivered through a tablet using a custom-made application (iOS and Android compatible). Real-time sEMG signals from 2 channels (wrist flexors and extensors) were used during a calibration period to set training thresholds. For training, a vertical bar (presented next to the sEMG signals) turned green when the user reached a target threshold.
B: Fit testing	To test the training system with individuals without upper limb impairments and refine its design for improved fit and usability.	• Individual demonstration sessions with the research partner and with 11 participants without arm impairments. • Refinements and production of the sleeve.	• Version 0.2: sEMG signals were obtained from all participants, although the comfort of the sleeve varied across participants. Based on their feedback, medium and large sizes were produced, and a reference line following the middle finger was added to facilitate electrode alignment during donning. • Version 0.3: a second feedback round led to further refinements, including a knob for easier elbow alignment, a Velcro strap to attach the sEMG unit to the sleeve (replacing the external band), and the addition of a small sleeve size. Software updates included adding a separate window for visualizing the bar activity, tracking repetitions exceeding the threshold, and implementing calibration with a manual stop.

XSL•FO
RenderX

Phase	Aim	Procedures	Actions
C: End-user piloting	To assess usability and investigate perceptions and experiences of the training system from clinicians and individuals with stroke.	• Individual demonstration sessions with 9 stroke rehabilitation clinicians and 10 individuals with stroke. • Semistructured interviews and the UEQ[c] with end users at the end of the demonstration sessions.	• Version 0.4: on the basis of the collected feedback by end users, an additional Velcro strap was added to the sleeve to secure cables and simplify donning and doffing. The medium-sized sleeve was also downsized for a better fit. Software improvements included: fewer and larger buttons, naming of channels, adding a calibration countdown, larger size of the bars with repetition counts, higher color contrast, and a performance report.
D: Reflection and evaluation	To reflect on and to evaluate the co-design process.	• SWOT[d] analysis and reflection workshop with CDT. • Individual reflection session with research partner.	• Reflections from the CDT highlighted strengths such as strong collaboration, research partner involvement, and interdisciplinary expertise. Weaknesses included lack of an external facilitator and limited access to in-house software developers. Opportunities and threats related to future implementation were discussed. The research partner provided additional insights on her experience and valued her active role throughout the co-design process.

[a] CDT: co-design team.

[b] sEMG: surface electromyography.

[c] UEQ: User Experience Questionnaire.

[d] SWOT: strengths, weaknesses, opportunities, and threats.

Participants and Data Collection

In phase A, the co-design team defined the basic configuration of the training system (version 0.1). A textile sleeve with 2 pairs of integrated textile electrodes, placed over the wrist extensors and flexors, and a reference electrode placed on the elbow was produced. The sleeve was connected to a 2-channel wireless sEMG system (Shimmer3 EMG, Shimmer Sensing). The biofeedback software provided visual feedback of the sEMG signals displayed on a tablet from the wrist extensor muscles (eg, extensor carpi radialis brevis) and flexor muscles (eg, flexor carpi ulnaris). The sEMG signals from specific muscle activation were used to calibrate muscle activity thresholds, which were then visualized during training through a vertical bar display. This provided real-time feedback and allowed for adjustments to individualized muscle activity threshold settings.

In phase B, an initial test demonstration session was conducted with the research partner, followed by individual sessions with 11 participants without arm impairments. Convenience sampling was used, and age, gender, and anthropometric data were collected from this sample (Multimedia Appendix 1).

In phase C, individual demonstration sessions were conducted with 2 end-user groups: 9 clinicians (6 physiotherapists, 2 occupational therapists, and 1 rehabilitation doctor) and 10 individuals with stroke (Tables S1 and S2 in Multimedia Appendix 2). Participants were recruited through purposive sampling. Clinicians were recruited from stroke rehabilitation units in hospitals and primary care centers, while individuals with stroke were recruited through advertisements at rehabilitation centers in the Gothenburg urban area. Inclusion criteria for individuals with stroke were as follows: aged over 18 years, upper limb impairment after stroke, and detectable sEMG signal from the paretic arm. Exclusion criteria for all participants included open wounds, other conditions affecting arm function, uncorrected visual impairment, and inability to understand or follow instructions in Swedish or English. Demographic data, prior sEMG experience, and familiarity with using technologies (eg, smartphones, tablets, and computers) were collected from all end users. For individuals with stroke, additional information was gathered on stroke type and time since stroke, while for clinicians, data on years of experience and work setting were collected.

All demonstration sessions were conducted at Gothenburg University and led by a physiotherapist experienced with the training system (MM-N). In these sessions, the participants were first instructed on how to use the sleeve and software and

XSL•FO
RenderX

then asked to try it independently, with additional support provided as needed. At the end of the sessions, the participants were interviewed for about 30 minutes and asked to fill in the User Experience Questionnaire (UEQ) [39,40]. Additionally, in the demonstration sessions with individuals with stroke, upper limb function and spasticity were assessed using the Fugl-Meyer Assessment of Upper Extremity [41] and the Modified Ashworth Scale [42], respectively. From the demonstration sessions with end users, 1 clinician tested only the sleeve due to technical issues with the software and 1 individual with stroke provided limited feedback due to aphasia.

Evaluations and Data Analysis

Semistructured Interviews

The semistructured interviews included questions about the end users' perceptions and experiences using the sleeve and training software during the demonstration sessions, as well as their opinions on the training tool's potential for independent use in clinics or home settings. For clinicians, questions were adapted to explore how they would foresee using the training system in stroke rehabilitation (Multimedia Appendix 1). All interviews were led by the same physiotherapist who conducted the demonstration sessions (MM-N, with more than 5 years of research experience in stroke rehabilitation), unless a participant was not fluent in English, in which case the interview was conducted in Swedish by another physiotherapist (MAM, with more than 20 years of experience in stroke rehabilitation). All interviews were audio-recorded, transcribed verbatim by 2 researchers from the co-design team (MM-N and AB), and stored securely, with access restricted to the authors. The field notes and spontaneous feedback collected from participants during the demonstration sessions were also used to complement the interview data.

Content analysis using the approach described by Graneheim and Lundman [43] was used to analyze the interviews [44]. The analysis involved extracting meaning units, condensing them, coding, and sorting codes into categories and subcategories (Multimedia Appendix 1). Since the interviews were conducted following a single demonstration session, the analysis focused on manifest content, emphasizing explicit and observable elements [43,44]. Interviews conducted in Swedish were translated into English and checked for accuracy by a native Swedish speaker (AB). The analyses were conducted in English. Initial coding was carried out by a member of the co-design team (MM-N), while a second researcher (AB), a research engineer with over 5 years of experience in textile technology who was not involved in data collection but was familiar with the data, participated in the review and refinement of the codes. Interviews with clinicians and individuals with stroke were analyzed together, but coding was grouped separately for each participant type using different colors to facilitate category development while distinguishing differences between the groups. Two researchers (AB and MM-N) independently developed initial categories, followed by an iterative process of discussions with a third researcher (MAM) to refine and resolve conflicts until final categories were agreed upon. The research partner was also included in the final stages of the analysis, providing additional input on the selected categories and subcategories.

User Experience Questionnaire

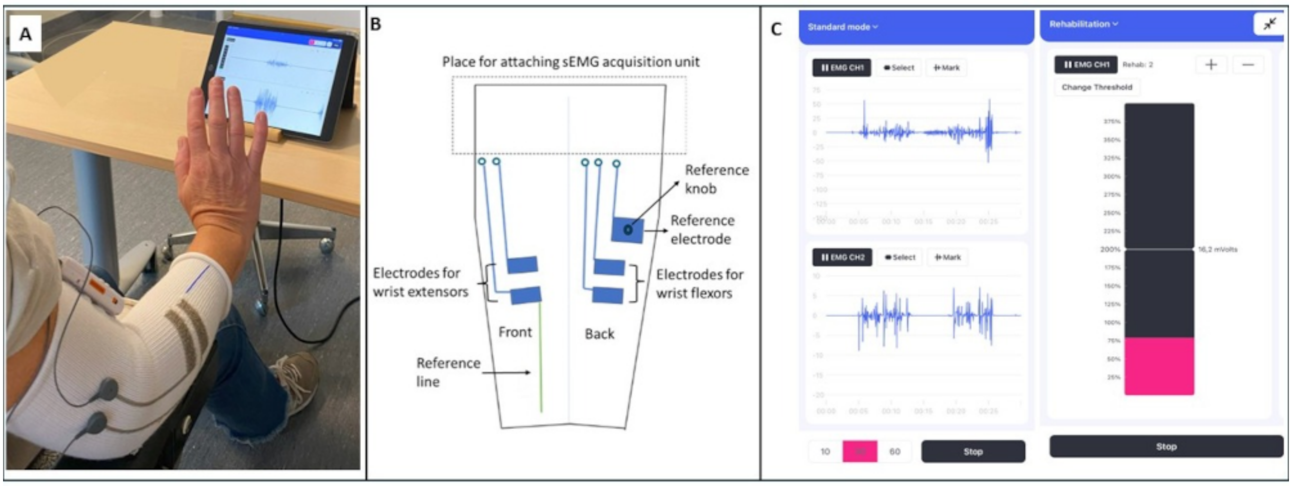
The UEQ captures a broad range of user experience aspects, helping to identify both strengths and areas for improvement [39,40]. It consists of 26 items rated on a 7-point scale, with each item presented as a pair of opposite terms. The questionnaire evaluates 6 dimensions: attractiveness (overall impression), perspicuity (ease of use), efficiency (task completion without unnecessary effort), dependability (feeling in control), stimulation (excitement and motivation), and novelty (innovation and creativity) [39,40]. The UEQ benchmark [45] was used to interpret the results on 5 levels: excellent, good, above average, below average, and bad.

Results

The specific refinements of the training system during the iterative phases (A-D) of the study are detailed in Table 1. The setup of the training system, the smart textile sleeve configuration, and the biofeedback software used in the demonstration sessions (version 0.3) are shown in Figure 2.

XSL•FO
RenderX

Figure 2. The smart textile sleeve biofeedback training system version 0.3. (A) Demonstration session setup, (B) smart textile sleeve configuration, and (C) biofeedback software (left panel: real-time sEMG signals from extensor [CH1] and flexor [CH2] channels; right panel: muscle activity bar visualization with adjustable threshold settings for training). CH1: channel 1; CH2: channel 2; sEMG: surface electromyography.



End Users' Experiences and Perceptions

From the interviews with clinicians and individuals with stroke, 2 main categories and 5 subcategories emerged, as illustrated in Textbox 1. Quotes are presented below each subcategory, where "S" indicates individuals with stroke and "C" refers to clinicians.

Textbox 1. Categories and subcategories emerging from content analysis.

Categories and subcategories

- This could be a new exciting training tool for stroke rehabilitation.
    - Rewarding and motivating to see muscle activity.
    - Use will be determined by individual needs.
    - Instructions and guidance will help independent use.
- The training tool works fine, but some changes could enhance independent training.
    - It can be difficult at first, but it is possible to find a way to make it work.
    - More is needed for independent long-term use.

This Could Be a New Exciting Training Tool for Stroke Rehabilitation

Participants with stroke and clinicians found the training system to be novel, exciting, and potentially useful for upper limb rehabilitation after stroke. Both groups expressed interest in incorporating the training system into their routines.

Rewarding and Motivating to See Muscle Activity

The real-time visual feedback on muscle activity was described as intuitive, useful [2], and rewarding by most participants. Individuals with stroke liked seeing their muscle activity, especially during training with the bar visualization, as many had thought they had little or no muscle activity due to their impairment. Some participants reported feeling more "connected" to their affected arm when seeing the activity visualized on the screen [3,4]. They described the feedback as not only giving them a sense of control over their arm but also possibly making the training more engaging, interesting, and motivating. Clinicians also highlighted the benefits of visual feedback, particularly for patients with limited muscle activity, as it could help them recognize their remaining muscle function and visualize their progress over time. Participants with stroke and clinicians also noted that visual feedback could enhance motivation and engagement and increase awareness of the affected arm, ultimately supporting and empowering independent training.

> ***S5:*** *That's, for me, [seeing the muscle activity] is something good to see, because then I see that something actually happens…I mean, it's boring to just sit like that, but if you see something happening, then I know my muscles are generating that kind of activity.*
>
> ***C3:*** *I like the visual feedback you get, even though we just do like a slight movement, you see the graph is changing. And if I put it on stroke patients, I think it's really good because sometimes when you tell a stroke patient, "Oh, can you do this movement?" They are quite sad, because they can't do the movement, but [with this training] you can see, "Oh, they are initiating the movement."*

XSL•FO
RenderX

Use Will Be Determined by Individual Needs

Most participants considered the training system more suitable for the chronic or subacute phases after a stroke. In the acute and subacute phases, individuals with stroke felt it could be overwhelming due to the many other challenges they faced at that stage, a concern shared by the clinicians. Some clinicians also noted that cost, time, and technical constraints could limit its use in clinical settings, while others described the tool as a complement to other training in the subacute phase. Both groups saw it as feasible and beneficial for home training in the chronic phase, especially considering that the available training opportunities are fewer compared to the subacute phase. However, they noted that some level of upper limb function in the paretic arm would be necessary, as high spasticity or limited elbow and finger control could make it difficult to don and doff the sleeve independently. Clinicians also emphasized that a certain level of cognition is required to handle and manage the training without therapist or carer support. Both groups mentioned that limited prior experience with technology could be a potential barrier to using the software.

> ***S6:*** *I think it’s a very good idea [the training tool]. Eh, it’s a problem for people like me [individuals with stroke], in the winter, to get out to training places when it’s snowy, so it’s very nice to make it possible to train home as well.*
>
> ***C7:*** *I think if you have one arm [functional arm], you could do it. You still need good cognitive abilities. Maybe not being too old. Some patients who are old are afraid of technological things because they feel uncertain. Even though they can manage… they don’t think they can. It is very different from person to person.*

Instructions and Guidance Will Help Independent Use

Individuals with stroke, as well as clinicians, viewed independent home training positively, with individuals with stroke especially finding it exciting and convenient. However, both groups emphasized that clear instructions and support are essential for successful independent use. While they appreciated the simple and clear guidance provided during the demonstration session, they felt additional help with software navigation and sleeve placement would be required. Individuals with stroke generally believed they could train independently or with partial caregiver assistance at home after initial guidance. However, they desired more structured support, including written training protocols, visual and written instructions, in-software reminders, and clinician follow-ups (in person or online) to ensure success. Clinicians also stressed the need for clear, repeated instructions in multiple formats to facilitate adherence. Additionally, clinicians highlighted the importance of hands-on training for patients and caregivers, along with in-person or online follow-ups to monitor tool usage and progress.

> ***S2:*** *It was difficult [to place the sleeve independently]. I think it can be even more difficult if you, kind of, get it [the training system] home and don't get to try it on first.*
>
> ***C1:*** *...They [individuals with stroke] need guidance, as it is right now [the tested version], I think...but otherwise, I think it’s quite easy to use with guidance.*

The Training Tool Works Fine, but Some Changes Could Enhance Independent Training

Overall, individuals with stroke and clinicians described the sleeve as functional and easy to use, noting that further improvements to the software were needed to promote independent training.

It Can Be Difficult at First, but It Is Possible to Find a Way to Make It Work

Most participants found the sleeve easy to use and comfortable, and they appreciated the provided instructions as well as the reference line and elbow knob for guidance. Individuals with stroke and more severe arm impairment experienced some difficulties with getting the hand inside the sleeve, rotating the sleeve into the correct position, and/or pulling the sleeve up on the arm. However, many found strategies by themselves during the demonstration session and believed that, with practice, it would not be a major problem. Another challenge was locating the reference electrode at the elbow and applying water to the reference and flexor electrodes due to a limited range of arm movement and/or sensory impairment. In the end, all participants with stroke were able to place the sleeve correctly, although some needed more time and minor assistance.

Perceptions of the software varied. Younger participants in both groups found the software easier to use, while others described it as a bit complicated and not intuitive. The buttons on the screen were too small and difficult to see for many. In general, most individuals with stroke and clinicians liked that a tablet was used for feedback, as it is larger than a phone but also portable compared to a computer screen. They emphasized that initial guidance and clear instructions on the software would be essential to ensure successful use.

> ***S7:*** *Yes, it was difficult to begin with [to put on the sleeve], but I think it’s something you get used to.*
>
> ***C6:*** *To me, it was easy and fine. I don’t know if a patient with a paresis would have difficulties having it drawn up on the middle of the arm …and also, placing it correctly on the elbow might be difficult.*

More Is Needed for Independent Long-Term Use

Independent home use was a possibility that many described as feasible, but they suggested improvements for better usability and long-term adherence. To simplify sleeve placement, participants recommended using a mirror or tablet webcam to better view the electrodes on the dorsal side of the arm and the reference electrode. Additional reference points, shorter electrode cables, and a pull strap for easier placement were also suggested. Regarding the software, both participants with stroke and clinicians desired a more user-friendly, engaging, and intuitive interface, with suggestions including games, videos, and motivational elements to make training more engaging. Clinicians emphasized adding functional, goal-oriented tasks and applying resistance to enhance engagement and facilitate training. Fewer and larger clickable buttons, simpler calibration

XSL•FO
RenderX

procedures, and adding more colors and contrast were other proposals for software improvements. Clinicians also suggested incorporating audio feedback to guide training and augment visual feedback. Additionally, both groups highlighted the need for customizable training parameters to facilitate sessions, such as calibration time and the number of repetitions, as well as the ability to track progress, which is particularly important for independent training. When considering home training, participants thought that a short daily training session of about 15 minutes would be feasible—short enough to avoid mental fatigue. A clinician also suggested using a bag to keep the sleeve clean between training sessions and noted that multiple sleeves might be needed if one requires cleaning or washing.

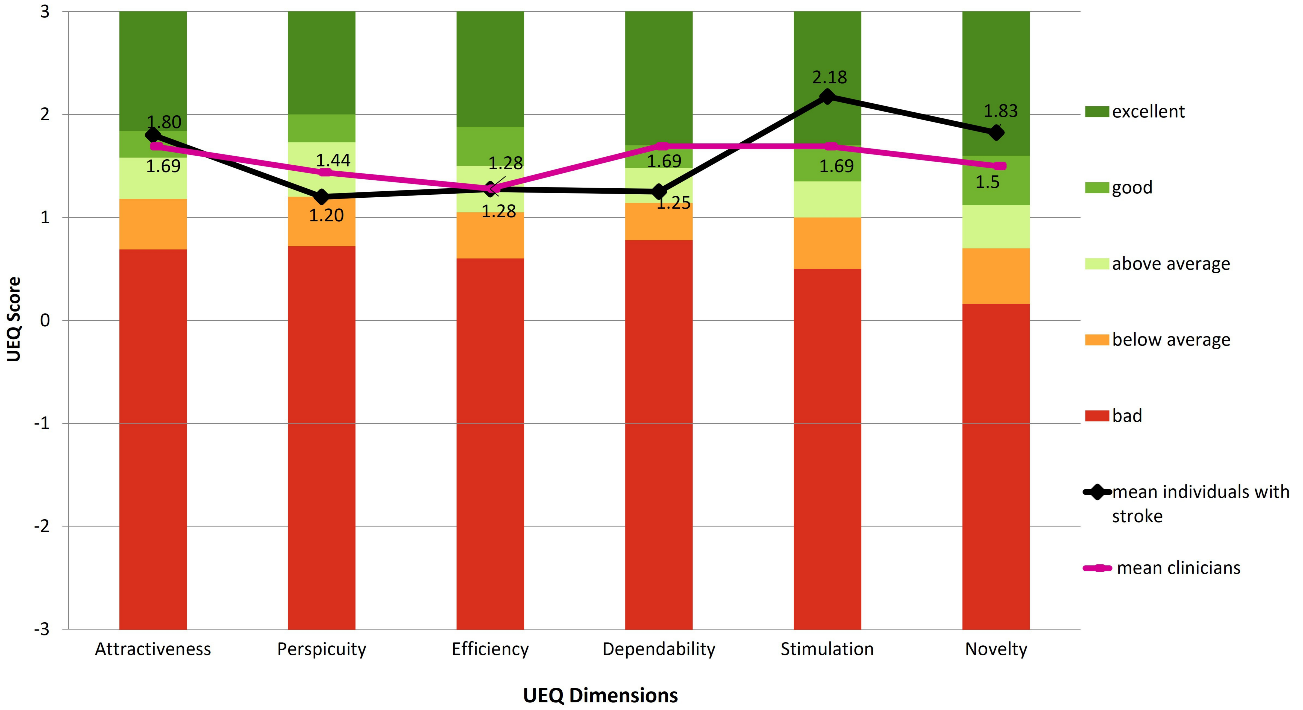
> ***S10:*** *On the tablet, I think it's too many knobs [buttons] to press on…It should be easier. One or two presses.*

> ***C2:*** *I'm thinking this [the training tool] was easy to use. But maybe adding some extra, like instructions... instructions like how fast you will do the exercise, how many repetitions, how long will I keep it. Maybe adding…like an audio, like a voice, [for example] if you have bad vision.*

User Experience Questionnaire

The UEQ was completed by all individuals with stroke and all but 1 clinician, who did not have the opportunity to try the software. Both groups rated *attractiveness* as "good" and *efficiency* as "above average." Clinicians rated *perspicuity* as "above average," while participants with stroke rated it lower, between "above average" and "below average." *Dependability* was rated "good" by clinicians and "above average" by individuals with stroke. Participants with stroke rated *stimulation* and *novelty* as "excellent," while clinicians rated them as "good" (Figure 3).

Figure 3. The results from the User Experience Questionnaire (UEQ) in relation to the standard UEQ benchmark levels. Scores range from −3 (very negative evaluation) to +3 (very positive evaluation), with 0 representing a neutral response. Line plots indicate the mean responses from individuals with stroke (black line; n=10) and clinicians (pink line; n=8).



Reflection and Evaluation of the Co-Design Process

The strengths of the co-design process identified by the co-design team were strong collaboration, effective communication, and research partner involvement, ensuring an end-user perspective. Flexibility, team commitment, and optimal facilities were also highlighted. Not having access to an external facilitator and software developers within the co-design team was identified as a weakness, with the latter contributing to delays in communication and iterative development. Opportunities included creating visual simulations before user involvement, direct contact with clinicians and patients, and a positive view of the training tool's potential value. Threats such as overdevelopment, uncertainty about real-world applicability, and challenges with commercialization and regulatory approval were also identified.

At the end of the co-design process, a reflection session with the research partner revealed that she enjoyed the experience and was pleasantly surprised by how much she liked it, as it was her first time in this role. She appreciated being involved not only in testing the training system but also in the reflective process. Overall, she found the experience fun and felt good about contributing to the research.

Discussion

Principal Findings

Supported by collaboration between experts from different fields and end users, the iterative co-design process provided valuable

XSL•FO
RenderX

insights into the technical, clinical, and user-centered aspects of the training system, highlighting its potential usability, acceptability, and areas for improvement based on end users' needs. The participants with stroke were able to independently, or with minor assistance, don and doff the sleeve and operate the training software on the tablet. Both clinicians and individuals with stroke recognized the training tool's potential for self-administered training. This study provides a starting point for the ongoing development and improvement of the training system.

The structured co-design process in this study, grounded in the participatory action research model [36] and the Medical Research Council framework [37], provided a strong foundation for the development of the training system and the shared understanding necessary for its continued refinement. This approach supported a systematic and iterative development of the training system and may serve as a useful model for future rehabilitation technology development [46,47]. The interdisciplinary nature of the co-design team enriched the process through diverse perspectives. However, differences in disciplinary knowledge and workflows, as reported in previous studies [48-50], may have prolonged the decision-making process. The co-design team suggested that adding a neutral facilitator to the team could have helped to advance the process further [33,51]. Integrating end-user feedback on design improvements with development constraints and limited available resources is a complex but well-recognized challenge in rehabilitation technology development [36,52]. This challenge was addressed through regular co-design meetings and systematic documentation, which enabled balanced prioritization of modifications based on both end-user feedback and practical constraints. Even though not all suggestions from the end users were integrated, they offered valuable insights for future improvements to the training system.

The overall positive perceptions of the smart textile biofeedback training system among individuals with stroke and clinicians underscore its potential as a possible solution for self-administered rehabilitation at home and in clinical contexts. Initial difficulties with donning and doffing the sleeve, particularly among participants with more severe upper limb impairments, highlighted the importance of user-centered design and individualized support during the development of the training system [30,33,34]. All participants were ultimately able to use the system independently or with minimal assistance, suggesting its accessibility across varying levels of motor function. The integration of textile electrodes simplified the donning process, possibly reducing barriers associated with conventional electrode placement and supporting independent use [20-24], especially for individuals with limited fine motor control. Although the system was evaluated in single sessions, the positive perceptions and willingness to incorporate it into daily training routines indicate promising initial usability and acceptability. Further longitudinal studies are needed to assess sustained engagement and rehabilitation outcomes over time.

The visual feedback, consistent with previous evidence [53-56], was perceived as intuitive and helpful in increasing awareness of muscle activity. The participants also found that being able to see and control their muscle activity increased their engagement and motivation. Future iterations, as suggested by the end users, will include more engaging interactive games and/or virtual environments, to facilitate long-term use [57-60]. Additionally, more advanced signal processing techniques, such as myoelectric pattern recognition, could be valuable for detecting specific movement patterns and enabling more intuitive control [14,61]. Ease of use and engaging designs have been identified as key factors in increasing treatment adherence [50,62,63] and will therefore be prioritized in the continued development of the training system for self-administered rehabilitation.

The participants also emphasized the need for structured support from the therapist, including regular follow-ups and adjustments to the training program. They highlighted the importance of clear instructions, in written or video formats, to support independent use and long-term adherence. Clear guidance, personalized feedback, and ongoing clinician support are essential for the successful use and sustained engagement with home-based rehabilitation technologies [50,63-65]. Both participants with stroke and clinicians found the training tool suitable for individuals with limited upper limb function but raised some concerns for individuals with limited cognitive capacity or experience with digital tools. On the basis of these insights, future iterations of the training system will focus on incorporating therapist support, clear instructional materials, the potential integration of a virtual coach, and appropriate user targeting to promote effectiveness and broader adoption in home-based rehabilitation.

Strength and Limitations

The ongoing input from the research partner was a strength of this study, offering valuable feedback that kept the co-design team focused on end users' needs. The iterative co-design process, supported by the triangulation of interviews, field notes, and questionnaires, as well as ongoing reflexive discussions within the team, strengthened the study's credibility, transparency, and overall trustworthiness [46,47]. Additionally, interviews, field notes, and a UEQ provided a broader understanding of participants' experiences.

A strength that also posed a limitation was that some participants were recruited through professional networks and, in many cases, the first author conducted both the sessions and the interviews. While this facilitated access and trust, it may have introduced sampling bias and relationship bias, as prior familiarity could have influenced participants' responses and the researcher's interpretation, potentially limiting the diversity and representativeness of perspectives.

The end users' perceptions were primarily used to inform the design and perceived usefulness of the training system, rather than to evaluate its practical application in clinical settings. The brief, single-session exposure further limited feedback to initial impressions rather than long-term use. Future studies are needed to evaluate the usability, feasibility, and potential effectiveness of the system in people with stroke after a longer training period.

Conclusions

This study developed and evaluated the usability of a smart textile sEMG biofeedback system designed for self-administered

XSL•FO
RenderX

upper limb rehabilitation after stroke. The co-design process, involving a multidisciplinary team of experts and end users, played an important role in identifying users' needs and preferences, contributing to a more targeted and user-centered design. Individuals with stroke and clinicians found the training tool useful and relevant for rehabilitation. The findings will guide ongoing development and clinical evaluation, but further research is needed to determine the system's effectiveness as a home-based rehabilitation tool for people with stroke.

Acknowledgments

The authors thank all the participants, including individuals with stroke, clinicians, researchers, and individuals with no arm impairment, for their time and effort in joining this study. The authors give special thanks to their research partner, Johanna Persson, for her invaluable guidance in the co-design process and to Elin Gustafsson for assisting with participant recruitment. The authors also acknowledge Bola AB for producing the textile prototypes used in this study and Katharina S Sunnerhagen for her support throughout the process.

Funding

This study was funded by the Promobilia Foundation (A23169); KK-Stiftelsen (20220222); Strokeförbundets Stiftelser och Fonder (S-10119402024-09-15); Norrbacka-Eugeniastiftelsen (837/24); Stiftelsen Hjalmar Svenssons (HJSV2024058); Swedish Society for Medical Research (S19-0074); and the Swedish state under the agreement between the Swedish government and the county councils, the ALF-agreement (ALF-GBG-932697, ALF-GBG 965508, and ALF-GBG 1006050).

Authors' Contributions

Conceptualization – LS (lead), LG (equal), MAM (equal), AB (supporting), MBK (supporting), MM-N (supporting), PK (supporting)
Data curation – MM-N (lead), LG (supporting), LS (supporting), MAM (supporting), PK (supporting)
Formal analysis – MM-N (lead), MAM (equal), AB (supporting)
Funding acquisition - MAM (lead), LG (equal), LS (equal), MM-N (supporting)
Investigation – MM-N (lead), MAM (equal), AB (supporting), LG (supporting), LS (supporting), MBK (supporting), PK (supporting)
Methodology – MM-N (lead), LG (equal), LS (equal), MAM (equal), AB (supporting), MBK (supporting), PK (supporting)
Project administration – MN-M (lead), LG (equal), MAM (equal), LS (supporting)
Resources – MAM (lead), LG (equal), LS (equal), MM-N (equal), PK (equal), AB (supporting), MBK (supporting)
Software – PK (lead), LG (equal), LS (equal), AB (supporting), MAM (supporting), MBK (supporting), MM-N (supporting)
Supervision – MAM (lead), LG (supporting), LS (supporting), MBK (supporting)
Validation – MAM (lead), LG (supporting), LS (supporting), MBK (supporting)
Visualization – MM-N (lead), LG (supporting), LS (supporting), MAM (supporting)
Writing – original draft – MM-N (Lead)
Writing – review & editing – MM-N (lead), LG (equal), LS (equal), MAM (equal), MBK (equal) AB (supporting), PK (supporting)

Conflicts of Interest

AB, LG, LS, and PK are founders and stakeholders of ZiniaCare, a startup company aiming to make the described biofeedback system commercially available, thereby presenting a potential conflict of interest. The company did not provide funding or support for the study.

Multimedia Appendix 1
Characteristics of individuals without arm impairment, interview templates, and an example of content analysis from end-user interviews.
[DOCX File, 55 KB - rehab_v13i1e77999_app1.docx]

Multimedia Appendix 2
Characteristics of individuals with stroke and clinicians.
[DOCX File, 47 KB - rehab_v13i1e77999_app2.docx]

Checklist 1
COREQ checklist.
[PDF File, 447 KB - rehab_v13i1e77999_app3.pdf]

XSL•FO
RenderX

References

1. Pollock A, St George B, Fenton M, Firkins L. Top 10 research priorities relating to life after stroke--consensus from stroke survivors, caregivers, and health professionals. Int J Stroke 2014 Apr;9(3):313-320. [doi: 10.1111/j.1747-4949.2012.00942.x] [Medline: 23227818]
2. Persson HC, Parziali M, Danielsson A, Sunnerhagen KS. Outcome and upper extremity function within 72 hours after first occasion of stroke in an unselected population at a stroke unit. A part of the SALGOT study. BMC Neurol 2012 Dec 29;12:162. [doi: 10.1186/1471-2377-12-162] [Medline: 23273107]
3. Nijland RHM, van Wegen EEH, Harmeling-van der Wel BC, Kwakkel G, EPOS Investigators. Presence of finger extension and shoulder abduction within 72 hours after stroke predicts functional recovery: early prediction of functional outcome after stroke: the EPOS cohort study. Stroke 2010 Apr;41(4):745-750. [doi: 10.1161/STROKEAHA.109.572065] [Medline: 20167916]
4. Harris JE, Eng JJ. Paretic upper-limb strength best explains arm activity in people with stroke. Phys Ther 2007 Jan;87(1):88-97. [doi: 10.2522/ptj.20060065] [Medline: 17179441]
5. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: a systematic review. Lancet Neurol 2009 Aug;8(8):741-754. [doi: 10.1016/S1474-4422(09)70150-4] [Medline: 19608100]
6. Raghavan P. Upper limb motor impairment after stroke. Phys Med Rehabil Clin N Am 2015 Nov;26(4):599-610. [doi: 10.1016/j.pmr.2015.06.008] [Medline: 26522900]
7. Ekstrand E, Brogårdh C. Life satisfaction after stroke and the association with upper extremity disability, sociodemographics, and participation. PM R 2022 Aug;14(8):922-930. [doi: 10.1002/pmrj.12712] [Medline: 34541828]
8. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
9. Teasell R, Mehta S, Pereira S, et al. Time to rethink long-term rehabilitation management of stroke patients. Top Stroke Rehabil 2012;19(6):457-462. [doi: 10.1310/tsr1906-457] [Medline: 23192711]
10. Boehme C, Toell T, Lang W, Knoflach M, Kiechl S. Longer term patient management following stroke: a systematic review. Int J Stroke 2021 Oct;16(8):917-926. [doi: 10.1177/17474930211016963] [Medline: 33949269]
11. Arntz A, Weber F, Handgraaf M, et al. Technologies in home-based digital rehabilitation: scoping review. JMIR Rehabil Assist Technol 2023 Jul 27;10:e43615. [doi: 10.2196/43615] [Medline: 37253381]
12. Chen Y, Abel KT, Janecek JT, Chen Y, Zheng K, Cramer SC. Home-based technologies for stroke rehabilitation: a systematic review. Int J Med Inform 2019 Mar;123:11-22. [doi: 10.1016/j.ijmedinf.2018.12.001] [Medline: 30654899]
13. Giggins OM, Persson UM, Caulfield B. Biofeedback in rehabilitation. J Neuroeng Rehabil 2013 Jun 18;10:60. [doi: 10.1186/1743-0003-10-60] [Medline: 23777436]
14. Munoz-Novoa M, Kristoffersen MB, Sunnerhagen KS, Naber A, Ortiz-Catalan M, Alt Murphy M. Myoelectric pattern recognition with virtual reality and serious gaming improves upper limb function in chronic stroke: a single case experimental design study. J Neuroeng Rehabil 2025 Jan 17;22(1):6. [doi: 10.1186/s12984-025-01541-y] [Medline: 39825410]
15. Munoz-Novoa M, Kristoffersen MB, Sunnerhagen KS, Naber A, Alt Murphy M, Ortiz-Catalan M. Upper limb stroke rehabilitation using surface electromyography: a systematic review and meta-analysis. Front Hum Neurosci 2022;16:897870. [doi: 10.3389/fnhum.2022.897870] [Medline: 35669202]
16. Nelson LA. The role of biofeedback in stroke rehabilitation: past and future directions. Top Stroke Rehabil 2007;14(4):59-66. [doi: 10.1310/tsr1404-59] [Medline: 17698458]
17. Woodford H, Price C. Emg biofeedback for the recovery of motor function after stroke. Cochrane Database Syst Rev 2007 Apr 18;2007(2):CD004585. [doi: 10.1002/14651858.CD004585.pub2] [Medline: 17443550]
18. Kim JH. The effects of training using EMG biofeedback on stroke patients upper extremity functions. J Phys Ther Sci 2017 Jun;29(6):1085-1088. [doi: 10.1589/jpts.29.1085] [Medline: 28626331]
19. Hermens H, Freriks B, Merletti R, et al. SENIAM 8: European recommendations for surface electromyography. : Roessingh Research and Development; 1999 URL: http://www.seniam.org/pdf/contents8.PDF [accessed 2026-03-13]
20. Guo L, Sandsjö L, Ortiz-Catalan M, Skrifvars M. Systematic review of textile-based electrodes for long-term and continuous surface electromyography recording. Text Res J 2020 Jan;90(2):227-244. [doi: 10.1177/0040517519858768]
21. Searle A, Kirkup L. A direct comparison of wet, dry and insulating bioelectric recording electrodes. Physiol Meas 2000 May;21(2):271-283. [doi: 10.1088/0967-3334/21/2/307] [Medline: 10847194]
22. Fleury A, Sugar M, Chau T. E-textiles in clinical rehabilitation: a scoping review. Electronics (Basel) 2015;4(1):173-203. [doi: 10.3390/electronics4010173]
23. Khan MA, Saibene M, Das R, Brunner I, Puthusserypady S. Emergence of flexible technology in developing advanced systems for post-stroke rehabilitation: a comprehensive review. J Neural Eng 2021 Dec 2;18(6). [doi: 10.1088/1741-2552/ac36aa] [Medline: 34736239]
24. Korzeniewska E, Krawczyk A, Mróz J, Wyszyńska E, Zawiślak R. Applications of smart textiles in post-stroke rehabilitation. Sensors (Basel) 2020 Apr 22;20(8):2370. [doi: 10.3390/s20082370] [Medline: 32331218]
25. Yang G, Deng J, Pang G, et al. An IoT-enabled stroke rehabilitation system based on smart wearable armband and machine learning. IEEE J Transl Eng Health Med 2018;6:2100510. [doi: 10.1109/JTEHM.2018.2822681] [Medline: 29805919]

XSL•FO
RenderX

26. De Marchis C, Santos Monteiro T, Simon-Martinez C, Conforto S, Gharabaghi A. Multi-contact functional electrical stimulation for hand opening: electrophysiologically driven identification of the optimal stimulation site. J Neuroeng Rehabil 2016 Mar 8;13:22. [doi: 10.1186/s12984-016-0129-6] [Medline: 26955873]
27. Lorussi F, Carbonaro N, De Rossi D, Paradiso R, Veltink P, Tognetti A. Wearable textile platform for assessing stroke patient treatment in daily life conditions. Front Bioeng Biotechnol 2016;4:28. [doi: 10.3389/fbioe.2016.00028] [Medline: 27047939]
28. Langan J, Subryan H, Nwogu I, Cavuoto L. Reported use of technology in stroke rehabilitation by physical and occupational therapists. Disabil Rehabil Assist Technol 2018 Oct;13(7):641-647. [doi: 10.1080/17483107.2017.1362043] [Medline: 28812386]
29. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251. [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
30. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17. [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
31. Redfern J, McKevitt C, Wolfe CDA. Development of complex interventions in stroke care: a systematic review. Stroke 2006 Sep;37(9):2410-2419. [doi: 10.1161/01.STR.0000237097.00342.a9] [Medline: 16902171]
32. Alt Murphy M, Pradhan S, Levin MF, Hancock NJ. Uptake of technology for neurorehabilitation in clinical practice: a scoping review. Phys Ther 2024 Feb 1;104(2):pzad140. [doi: 10.1093/ptj/pzad140] [Medline: 37856528]
33. Singh H, Benn N, Fung A, et al. Co-design for stroke intervention development: results of a scoping review. PLoS ONE 2024;19(2):e0297162. [doi: 10.1371/journal.pone.0297162] [Medline: 38354160]
34. Shah SGS, Robinson I. Benefits of and barriers to involving users in medical device technology development and evaluation. Int J Technol Assess Health Care 2007;23(1):131-137. [doi: 10.1017/S0266462307051677] [Medline: 17234027]
35. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): A 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
36. Cornish F, Breton N, Moreno-Tabarez U, et al. Participatory action research. Nat Rev Methods Primers 2023;3(1). [doi: 10.1038/s43586-023-00214-1]
37. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
38. Pickton DW, Wright S. What's SWOT in strategic analysis? Strategic Change 1998 Mar;7(2):101-109. [doi: 10.1002/(SICI)1099-1697(199803/04)7:2<101::AID-JSC332>3.0.CO;2-6]
39. Schrepp M, Hinderks A, Thomaschewski J. Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. 2014 Presented at: Design, User Experience, and Usability: Theories, Methods, and Tools for Designing the User Experience: 3rd International Conference, DUXU 2014; Jun 22-27, 2014; Heraklion, Crete, Greece p. 383-392. [doi: 10.1007/978-3-319-07668-3_37]
40. Schrepp M, Hinderks A, Thomaschewski J. Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). Int J Interact Multimed Artif Intell 2017;4(6):103-108. [doi: 10.9781/ijimai.2017.09.001]
41. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. Scand J Rehabil Med 1975;7(1):13-31. [Medline: 1135616]
42. Gregson JM, Leathley M, Moore AP, Sharma AK, Smith TL, Watkins CL. Reliability of the Tone Assessment Scale and the Modified Ashworth Scale as clinical tools for assessing poststroke spasticity. Arch Phys Med Rehabil 1999 Sep;80(9):1013-1016. [doi: 10.1016/s0003-9993(99)90053-9] [Medline: 10489001]
43. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
44. Lindgren BM, Lundman B, Graneheim UH. Abstraction and interpretation during the qualitative content analysis process. Int J Nurs Stud 2020 Aug;108:103632. [doi: 10.1016/j.ijnurstu.2020.103632] [Medline: 32505813]
45. Schrepp M, Hinderks A, Thomaschewski J. Construction of a benchmark for the User Experience Questionnaire (UEQ). Int J Interact Multimed Artif Intell 2017;4(4):40-44. [doi: 10.9781/ijimai.2017.445]
46. Malterud K. Qualitative research: Standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]
47. Ahmed SK. The pillars of trustworthiness in qualitative research. J Med Surg Public Health 2024 Apr;2:100051. [doi: 10.1016/j.glmedi.2024.100051]
48. Williams J, Craig TJ, Robson D. Barriers and facilitators of clinician and researcher collaborations: a qualitative study. BMC Health Serv Res 2020 Dec 5;20(1):1126. [doi: 10.1186/s12913-020-05978-w] [Medline: 33278896]
49. Domecq JP, Prutsky G, Elraiyah T, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014 Feb 26;14:89. [doi: 10.1186/1472-6963-14-89] [Medline: 24568690]
50. Mitchell J, Shirota C, Clanchy K. Factors that influence the adoption of rehabilitation technologies: a multi-disciplinary qualitative exploration. J Neuroeng Rehabil 2023 Jun 20;20(1):80. [doi: 10.1186/s12984-023-01194-9] [Medline: 37340496]

XSL•FO
RenderX

51. Eriksson S, Wallgren P, Sandsjö L, Karlsson M. Genuine co-design: An activity theory analysis involving emergency nurses in an interdisciplinary new product development project of a novel medical device. Int J Hum Factors Ergon 2021;8(4):331-369. [doi: 10.1504/IJHFE.2021.119054]
52. Barbosa IM, Alves PR, Silveira ZC. Upper limbs' assistive devices for stroke rehabilitation: a systematic review on design engineering solutions. J Braz Soc Mech Sci Eng 2021 May;43(5). [doi: 10.1007/s40430-021-02919-4]
53. Feldner HA, Papazian C, Peters K, Steele KM. "It's all sort of cool and interesting…but what do i do with it?" A qualitative study of stroke survivors' perceptions of surface electromyography. Front Neurol 2020;11:1037. [doi: 10.3389/fneur.2020.01037] [Medline: 33041981]
54. Feldner HA, Howell D, Kelly VE, McCoy SW, Steele KM. "Look, your muscles are firing!": A qualitative study of clinician perspectives on the use of surface electromyography in neurorehabilitation. Arch Phys Med Rehabil 2019 Apr;100(4):663-675. [doi: 10.1016/j.apmr.2018.09.120] [Medline: 30392855]
55. Cardoso VF, Valencia N, Loterio FA, et al. Towards an upper limb rehabilitation tool after stroke based on surface electromyography biofeedback and virtual reality. Res Biomed Eng 2022;38(3):1017-1025. [doi: 10.1007/s42600-022-00218-y]
56. Munoz-Novoa M, Andersson C, Sunnerhagen KS, Alt Murphy M. A novel intervention for upper limb rehabilitation in people with stroke combining myoelectric pattern recognition, virtual reality, and serious gaming: a qualitative study. Disabil Rehabil 2025 Jul;47(15):3930-3937. [doi: 10.1080/09638288.2024.2434643] [Medline: 39625112]
57. Lohse KR, Hilderman CGE, Cheung KL, Tatla S, Van der Loos HFM. Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. PLoS ONE 2014;9(3):e93318. [doi: 10.1371/journal.pone.0093318] [Medline: 24681826]
58. Burke JW, McNeill MDJ, Charles DK, Morrow PJ, Crosbie JH, McDonough SM. Optimising engagement for stroke rehabilitation using serious games. Vis Comput 2009 Dec;25(12):1085-1099. [doi: 10.1007/s00371-009-0387-4]
59. Bargeri S, Scalea S, Agosta F, et al. Effectiveness and safety of virtual reality rehabilitation after stroke: an overview of systematic reviews. EClinicalMedicine 2023 Oct;64:102220. [doi: 10.1016/j.eclinm.2023.102220] [Medline: 37745019]
60. Doumas I, Everard G, Dehem S, Lejeune T. Serious games for upper limb rehabilitation after stroke: a meta-analysis. J Neuroeng Rehabil 2021 Jun 15;18(1):100. [doi: 10.1186/s12984-021-00889-1] [Medline: 34130713]
61. Pregnolato G, Rimini D, Baldan F, et al. Clinical features to predict the use of a sEMG Wearable Device (REMO®) for hand motor training of stroke patients: a cross-sectional cohort study. Int J Environ Res Public Health 2023 Mar 14;20(6):5082. [doi: 10.3390/ijerph20065082] [Medline: 36981992]
62. Hochstenbach-Waelen A, Seelen HAM. Embracing change: Practical and theoretical considerations for successful implementation of technology assisting upper limb training in stroke. J Neuroeng Rehabil 2012 Aug 2;9:52. [doi: 10.1186/1743-0003-9-52] [Medline: 22856548]
63. Vourganas I, Stankovic V, Stankovic L, Kerr A. Factors that contribute to the use of stroke self-rehabilitation technologies: a review. JMIR Biomed Eng 2019;4(1):e13732. [doi: 10.2196/13732]
64. Mahmood A, Deshmukh A, Natarajan M, et al. Development of strategies to support home-based exercise adherence after stroke: a Delphi consensus. BMJ Open 2022 Jan 6;12(1):e055946. [doi: 10.1136/bmjopen-2021-055946] [Medline: 34992120]
65. Odetunde MO, Olaoye OA, Obajobi OA, Yusuf AA, Awotidebe TO. Adherence to home-based exercise programs among stroke survivors and perspectives of informal caregivers: a mixed method study. Bull Fac Phys Ther 2024;29(1). [doi: 10.1186/s43161-024-00224-4]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
sEMG: surface electromyography
UEQ: User Experience Questionnaire

Edited by S Munce; submitted 23.May.2025; peer-reviewed by A Kerr, D Hobbs; accepted 19.Feb.2026; published 13.Apr.2026.

Please cite as:
Munoz-Novoa M, Guo L, Björkquist A, Kristoffersen MB, Khorramshahi P, Sandsjö L, Alt Murphy M
A Smart Textile Biofeedback Training System for Upper Limb Rehabilitation After Stroke: Co-Design Development and Evaluation Study
JMIR Rehabil Assist Technol 2026;13:e77999
URL: https://rehab.jmir.org/2026/1/e77999
doi:10.2196/77999

XSL•FO
RenderX

© Maria Munoz-Novoa, Li Guo, Anna Björkquist, Morten B Kristoffersen, Peiman Khorramshahi, Leif Sandsjö, Margit Alt Murphy. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Voice-Assisted Technology for People With Parkinson's Disease Experiencing Speech and Voice Difficulties: Co-Designing Solutions Using Design Thinking

Jodie Mills[1], BSc, MSc; George Kernohan[2], BSc, PhD; Katy Pedlow[1], BSc, PhD; Orla Duffy[1], BSc, PhD

[1]School of Health Sciences, Faculty of Life and Health Sciences, University of Ulster, Belfast, Northern Ireland, United Kingdom

[2]School of Nursing, Faculty of Life and Health Sciences, University of Ulster, Belfast, United Kingdom

Corresponding Author:
Jodie Mills, BSc, MSc
School of Health Sciences
Faculty of Life and Health Sciences
University of Ulster
2-24 York Street
Belfast, Northern Ireland, BT15 1AP
United Kingdom
Phone: 44 2895365308
Email: mills-j12@ulster.ac.uk

Abstract

Background: While smart speakers are emerging as a novel health care technology, people with Parkinson's Disease (PwPD) and speech and language therapists (SaLTs) have reported difficulties using smart speakers with speech and voice impairments in research. To date, PwPD have identified frustration with having to repeat themselves to be understood, devices timing out before they had finished speaking, and being unable to have a conversation with smart speakers. SaLTs have reported technical and practical challenges in implementing voice-assisted technology tools. Both PwPD and SaLTs indicated a lack of knowledge about what smart speakers could do, as well as concerns about privacy and the listening nature of the devices.

Objective: This study aims to co-design solutions that support the use of smart speakers for speech and voice difficulties experienced by PwPD.

Methods: Based on the Design Thinking framework, a multistage design process was conducted, involving a lay steering group and 2 online co-design workshops. Twenty participants, including PwPD, carers, SaLTs, design and technology experts, and third-sector staff, collaborated during the co-design workshops. The ideate phase included brainstorming and ranking, and conventional content analysis was used to specify prototypes.

Results: Two main prototypes were created: (1) education and guidance, including privacy and therapeutic usage guides for PwPD and SaLTs to address troubleshooting and delivery considerations; and (2) new speech and language therapy (SLT)–specific features for smart speakers. Participants provided feedback on their experiences of co-design, highlighting feeling valued, the balance of perspectives, and making improvement suggestions. Feedback aligned with the UK standards for public involvement.

Conclusions: Smart speakers could enhance accessibility, therapy engagement, and long-term speech outcomes, offering scalable, cost-effective solutions to support SLT services, patient independence, and reduced service demand. Smart speaker solutions with a SLT focus enable PwPD to self-manage speech and voice difficulties at home and reinforce therapy gains between clinic visits. Co-designed with users, these prototypes are intended to address health disparities and relieve pressure on SLT services, offering a scalable and sustainable solution that enhances efficiency and supports ongoing rehabilitation within health care systems.

(JMIR Rehabil Assist Technol 2026;13:e84364) doi:10.2196/84364

KEYWORDS

voice-assisted technology; speech and language therapy; co-design; participatory methods; dysarthria; speech and voice; Parkinson's Disease

XSL•FO
RenderX

Introduction

Background

Voice-assisted technology (VAT) is defined as a device that uses natural language processing or automatic speech recognition (ASR) to interpret spoken language and translate it into actionable requests.

Smart speakers are commercially available VAT devices that are controlled using voice commands and are usually connected to the internet (current examples include Amazon Alexa and Google Assistant). They can feature built-in control systems for tasks on demand, including smart home automation, providing general information (not limited to weather, recipes, or health information), person-to-person calls, sending and receiving messages, and playing music. New models with screens can also support audio and video streaming. Smart speakers are readily available for purchase and use by the general public [1].

It has been reported that VAT prompts some participants with speech difficulties to modify their speech to enable interaction with VAT [2-6]. People with Parkinson's disease (PwPD) have reported adapting their speech by speaking more slowly, loudly, and clearly when interacting with a smart speaker [4]. Considering that 90% of PwPD present with reduced speech intelligibility and limited vocal loudness [7], VAT may hold potential as a therapeutic adjunct in speech and language therapy (SLT). This prior evidence indicates that VAT may enhance access to therapy [8].

Some therapists have reported using VAT to promote improved volume, clarity, and intelligibility of speech [9]. In addition to offering biofeedback on speech clarity, these tools have provided structured opportunities for home-based practice, fostering self-awareness and supporting the self-management of dysarthria and other speech difficulties [9,10]. PwPD have reported increased clarity of speech and volume when using VAT, and have used VAT as a communication partner to practice their speech and rebuild confidence in using their voice [11]. Both speech and language therapists (SaLTs) and PwPD agree that the objective nature of VAT is key to promoting interaction and providing feedback on speech. Textbox 1 presents a hypothetical vignette illustrating how a person with speech or voice difficulties may interact with smart speakers. This vignette is informed by the understanding and findings of previous research [11].

Textbox 1. Case study

1. Case

John is 65 and has had Parkinson's disease for 5 years. His phonation is impacted by poor breath support, resulting in a breathy, hoarse voice with low volume. His articulation is reduced, resulting in imprecise speech production, which reduces speech clarity and intelligibility. His speech is also hypernasal, and nasal emissions are noted. He has hypokinetic dysarthria. At home, his family can understand him, but he is frequently told that they "can't hear him" and that he "needs to speak up." This is frustrating for John, as he reports that "he feels like he is shouting," which suggests impaired self-awareness of his speech.

2. Use

John uses his smart speaker daily. When he speaks to the smart speaker, it replies with "Sorry, I didn't get that" approximately 50% of the time. As a result, John raises his volume and repeats his request. Often, he uses a loud voice, overarticulates his words, slows down, and speaks as soon as he takes a breath. The smart speaker responds when he uses these strategies. This demonstrates that smart speakers provide feedback on volume and clarity of speech in the form of an external cue: "Sorry, I didn't get that." This can encourage increased self-awareness of speech volume and intelligibility, and result in the use of LOUD, clear speech strategies. As John's smart speaker can time out before he has finished speaking, he uses *adaptive listening mode* (available on Amazon devices), which is found in the accessibility settings and gives him longer to speak.

John also plays the game "Word Tennis" on his smart speaker. He has to think of words within a semantic category quickly and remember to use a LOUD, clear voice when answering. This task focuses on a word-level activity within the speech hierarchy and adds a cognitive load to increase difficulty, which aligns with Lee Silverman Voice Treatment (LSVT) LOUD principles [12]. He also enjoys sport and cooking and uses his smart speaker to search for recipes. Common functional requests include "Add cheese and potatoes to my shopping list," "Show me my cooking library," and "What was the Man United score today?" Sometimes, he even uses his smart speaker like a diary: "Leave a sticky note for...", where he records a voice note on his smart speaker to remind someone to feed their dog.

3. Summary

Overall, interacting with his smart speaker allows John to practice a LOUD, clear voice at home, with external feedback on speech volume and clarity that may help improve his self-awareness.

Despite the facilitators discussed in Textbox 1, several barriers to the effective use of VAT among PwPD and SaLTs remain [9,11]. For example, PwPD have reported feeling frustrated by needing to repeat themselves to be understood, by devices timing out before they had finished speaking, and by being unable to have a conversation with their smart speaker [11]. SaLTs also indicated that they faced technical and practical challenges in implementing VAT tools [9]. Both PwPD and SaLTs reported a lack of knowledge about smart speaker capabilities and concerns surrounding privacy and data security.

Addressing these challenges is essential to enable the integration of VAT into SLT practice. Design Thinking is a user-centered innovation framework used to guide the development of new health care technologies, often utilizing co-design approaches [13-15]. It offers a structured approach to identifying problems and generating solutions through empathy, collaboration, and iterative prototyping and testing. This study is informed by the define, ideate, and prototyping phases of the Design Thinking process. Figure 1 outlines the Design Thinking process, and Table 1 shows the connections between the phases of the Design

XSL•FO
RenderX

Thinking framework, the specific research questions to be addressed, and the methods used.

Co-design has been used to foster collaboration that stimulates new ideas, clarifies concepts, and creates solutions that prioritize the needs and lived experiences of end users [16]. Co-design workshops have been used in SLT, health technology research, and with older adult populations [17-19], with improved outcomes for technology adoption compared with noncollaborative design processes [17,19]. Co-design is critical when developing technologies for SLT [18] and has value in engaging people with communication difficulties [20-22]. We set out to follow the co-design cycle and principles, meeting the criteria for true co-design under the ladder of co-production [23-25], through the identification and development of recommendations from participants with communication difficulties [26,27].

Figure 1. The Design Thinking stages [13] from empathize to prototyping, which can be used during the development of new health care technologies [15].

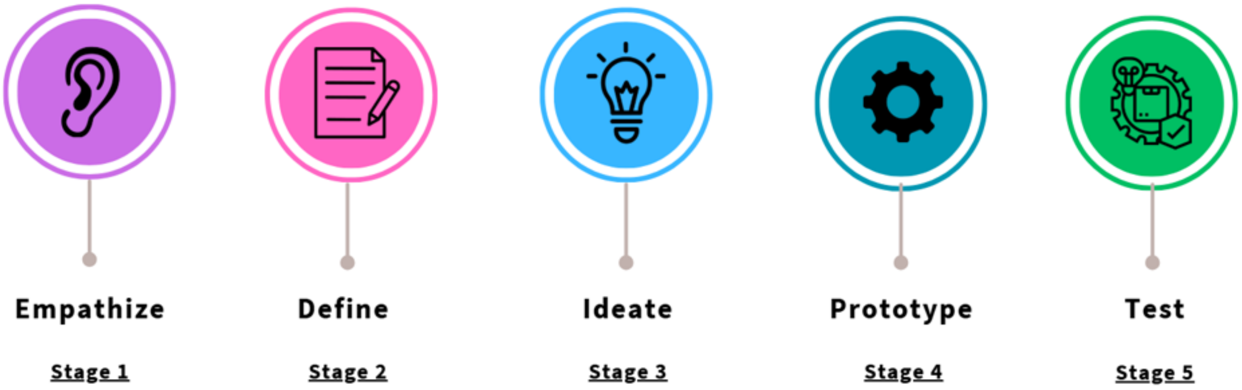


Table 1. Ways in which existing barriers to the therapeutic use of VAT[a] by people with Parkinson’s can be solved.

Objective	Objective 1: To consider and select problem statements from the perspectives of experts by experience.	Objective 2: To create solutions to problem statements associated with VAT usage, alongside people with Parkinson’s, carers, SaLTs[b], technology and design experts, and third-sector representatives.	Objective 3: To prioritize co-designed solutions to inform prototype VAT interventions.	Objective 3: To prioritize co-designed solutions to inform prototype VAT interventions.
Design Thinking stage	Stage 2: Define	Stage 3: Ideate	Stage 3: Ideate	Stage 4: Prototype
Method	Patient, public involvement workshop	Co-design workshop A	Co-design workshop B	Inductive content analysis

[a] VAT: voice-assisted technology.

[b] SaLT: speech and language therapist.

This research is intended to address previously noted barriers to VAT use [9,11] and aims to co-design solutions to previously identified challenges with VAT by working with PwPD, carers, SaLTs, charity representatives, and technology and design experts. This approach was taken to ensure that new technologies can be used in ways that meet end user needs. We sought to create solutions by using commercial technology, without coding or modifying VAT devices. This ensures that solutions are low cost and accessible, enhancing the potential for wider adoption of VAT in SLT contexts.

Aim

We set out to co-design solutions to support the use of smart speakers in SLT to improve volume and intelligibility for PwPD, using a Design Thinking framework. The research addressed the following question: “How can we facilitate the therapeutic use of VAT by people with Parkinson's Disease?”

Our study objectives (mapped onto Design Thinking stages) are as follows:

- To consider and select problem statements from the perspectives of experts by experience (Define).
- To co-create solutions to problem statements associated with VAT usage, alongside PwPD, carers, SaLTs, technology and design experts, and third-sector representatives (Ideate).
- To prioritize co-designed solutions to inform prototype VAT interventions (Prototype).

Methods

Participatory Co-Design Approach

Participatory methods such as co-design allow interventions to be designed around end user needs. This study co-designed solutions to previously identified barriers regarding the use of VAT when speech and voice difficulties were present. This results in technology that more readily meets user needs [28] and helps to avoid digital exclusion [29]. Workshops were held online, removing geographical and physical barriers and

XSL•FO
RenderX

enabling SaLTs from throughout the United Kingdom to share their experiences.

Ethical Considerations

Ethical approval was granted by the Ulster University Research Ethics Committee in January 2025 (approval number FCNUR-24-078-A). This study is part of a larger PhD project using Design Thinking. Previous phases of work aligned with the empathize stage, and the current co-design phase aligns with the define, ideation, and prototyping stages. All participants provided informed consent before the workshops. Participant outputs were anonymous, and ground rules were agreed upon to maintain confidentiality. Participants did not receive payment or financial incentives.

Patient and Public Involvement

A patient and public involvement (PPI) steering group was established to provide a voice for key stakeholders and ensure their active role in shaping the research. This group included a SaLT with firsthand experience using VAT in clinical settings, a person living with Parkinson’s, and a caregiver. These 3 experts by experience coassessed the barriers to VAT usage identified in previous research [9,11] and co-decided the top 5 problems that reflected their experiences, in keeping with the cycle of coproduction [24].

Study Recruitment

PwPD and carers were recruited via a third-sector organization (Parkinson’s UK) using advertisements on the Parkinson’s UK research portal, Research Support Network monthly emails, and flyers at local Parkinson’s support groups in Northern Ireland. SaLTs were recruited through the Royal College of Speech and Language Therapists, including the Parkinson’s Clinical Excellence Network. In addition, Parkinson’s UK staff and technology or design experts were recruited through the lead author’s (JM) professional network. Some participants had established a relationship with the lead researcher through work with the local branch of Parkinson’s UK in Northern Ireland.

Previous research was used to determine the number of participant collaborators invited to share their experiences during the workshops (n=20) [30,31]. Participants were asked to contact the research team to express interest in the study and were screened according to predefined criteria (Table 2). Potential participants were sent study information and consent forms by email or post, depending on preference, and were asked to indicate their availability. Once consent was obtained, participants completed a demographic survey and received links for the online workshops. This enabled interaction between diverse experiences. Participants were placed into smaller, experience-diverse groups of 4-5 participants to encourage idea generation in a safe and supportive environment.

XSL•FO
RenderX

Table 2. Inclusion and exclusion criteria for people with Parkinson's, carers, speech and language therapists, third-sector representatives, and technology or design experts.

Participant group	Inclusion	Exclusion
People with Parkinson's	• Adults over 18 years old • Mild to moderate dysarthria/voice difficulties (to include users of augmentative, alternative communication) • Diagnosis of Parkinson's disease • Current or previous use of VAT[a] • Have access to a laptop, with a camera, that facilitates videoconferencing software	• Moderate or severe cognitive impairment • History of other neurological disorders
Carers	• Adults over 18 years old • Live with or care for a PwPD[c] or both • Experience of facilitating the use of VAT with a PwPD • Have access to a laptop, with a camera, that facilitates videoconferencing software	N/A[b]
Speech and language therapists	• Adults over 18 years old • Who currently have/have had a clinical caseload of PwPD in the past 5 years • Who have used VAT in practice and have basic knowledge of the devices • Have a laptop, with a camera, that facilitates videoconferencing software	N/A
Third-sector staff	• Adults over 18 years old • Currently working in a third-sector organization for PwPD • Involvement and relationships with the local Parkinson's community • Basic knowledge of speech and voice difficulties in Parkinson's disease • Have a laptop, with a camera, that facilitates videoconferencing software	N/A
Technology/design experts	• Adults over 18 years old • Experience of VAT and detailed knowledge of its capabilities or relevant experience in designing or developing health care technologies	N/A

[a] VAT: voice-assisted technology.

[b] N/A: not applicable.

[c] PwPD: people with Parkinson's disease.

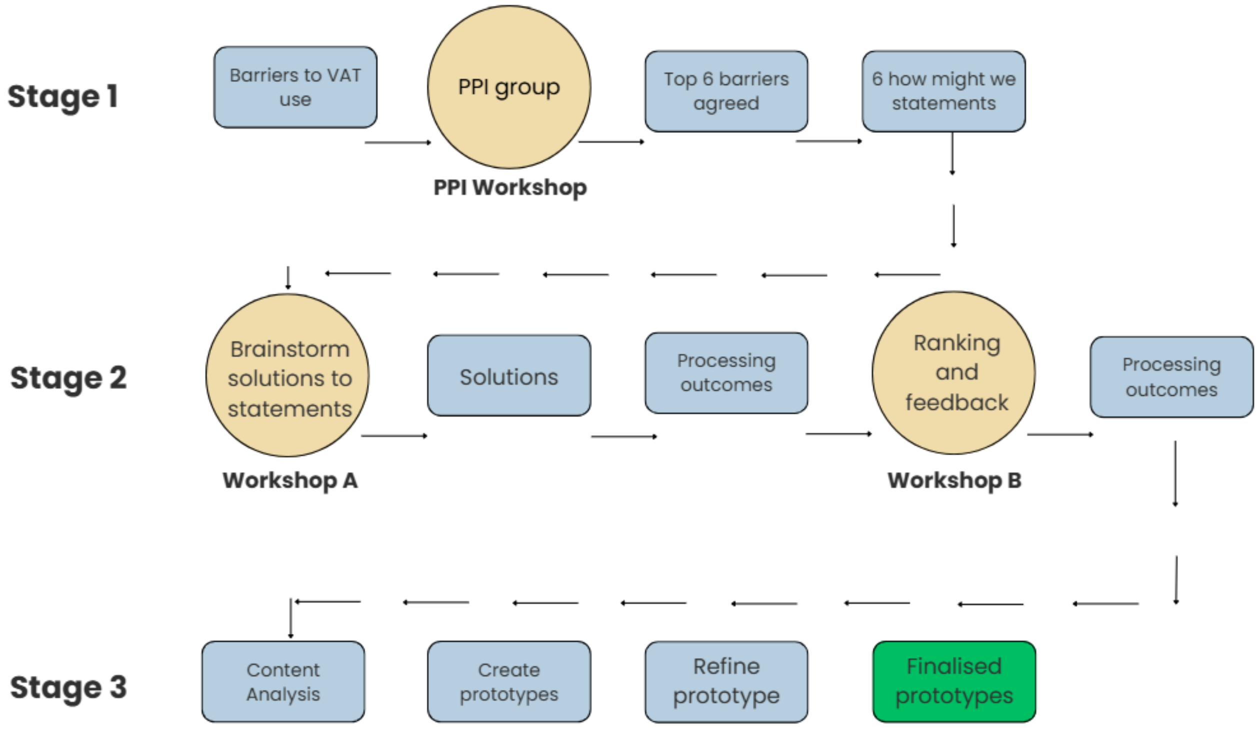
Procedure

Overview

Using principles underpinned by Design Thinking and Participatory methodology, a series of 2 co-design workshops were undertaken, informed by insights gained from previous research [30-32]. The 3-stage co-design procedure is outlined below (see Figure 2) and aligns with the define, ideate, and prototype phases of the Design Thinking framework, as presented in Figure 1.

XSL•FO
RenderX

Figure 2. The co-design process from the development of problem statements, through workshop completion, analysis, prototype creation, and refinement. PPI: patient and public involvement; VAT: voice-assisted technology.



Stage 1: Define Phase

Barriers identified by PwPD, their carers, and SaLTs during an earlier stage of the research [9,11] were reviewed by the PPI group, ensuring that the research was shaped by lived experience. Researchers JM and OD were present during the workshop. The group identified and agreed on the top 6 problems to be brought forward to workshop A, and, as per Design Thinking guidance, these were reframed into “How Might We” statements. These “How Might We” statements are used in Design Thinking to help people reframe a problem-focused perspective into solution-focused thinking [33], and previous research highlights the value of shaping research based on lived experience, as this ensures that the problems being solved are meaningful to end users [34]. The final 6 “How Might We” statements are presented in Table 3.

XSL•FO
RenderX

Table 3. Procedures and outcomes for stages 1, 2, and 3 of the co-design process.

Stage	Design Thinking framework stage	Procedure	Outcomes
1	Define	**PPI[a] workshop**	
		• The PPI group was presented with all barriers from previous research [9,11]. • JM shared her screen via videoconferencing, using Canva PowerPoint software to present the barriers identified in previous research and to assist with live decision-making. This acted as contemporaneous notes. OD also took field notes. • PPI group identified 6 top problems—felt to be a reasonable number of barriers to brainstorm solutions to within an hour. • 6 problems reframed into "How Might We" statements.	• The following "How Might We" statements were used in workshop A: • How might we help people understand smart speaker privacy and reduce their fears? • How might we help people when smart speakers do not work? • How might we help people to have a conversation with a smart speaker? • How might we help people to know what smart speakers can do? • How could we deliver this information? • How could smart speaker technology be adapted?
2a	Ideate	**Co-design workshop A**	
		• A document with problem statements and examples of the problems was emailed to the participants before the workshop. • Welcome and introduction PowerPoint, which presented an overview of previous research, problems experienced by people using smart speakers, and an overview of the co-design process. • Two example brainstorming activities completed. • Smaller groups discussed 6 problem statements, rotating after 10 minutes on each problem statement. All groups completed different problems at the same time. • Facilitators shaped discussions and noted contributions on a live Word document. Each problem statement had a separate Word document. • Participants were invited to share solutions to the problem statements around using VAT[b] with a speech or voice difficulty. • Participants were thanked for their time and the workshop ended.	• Solutions to each problem statement were reviewed. Similar solutions for each problem statement were grouped together and combined. No content was removed, while ensuring there was a feasible number of ideas to rank. • Ideas are presented fully in Multimedia Appendix 1, as they appeared during the workshop. • A reduced number of solutions were placed into tables for each group to rank in workshop B.
2b	Ideate	**Co-design workshop B**	
		• Welcome and recap of solutions generated in workshop A. Participants received a document of the solutions ahead of the workshop to aid their preparation. • Each facilitator worked through 3 solutions documents with their group and ranked their top 5 priorities for each solution. • A similar procedure to workshop A was followed: each solution was captured in a separate Word document, and groups rotated after 10 minutes on each problem statement, inputting to the same solution document. • Each participant was asked a short series of questions about their experience of the co-design workshops. Questions were based on the UK standards for public involvement. • Participants were thanked for time, next steps were explained, and the workshop ended.	• Lead researcher (JM) collated solutions and rankings and removed any solutions that were not ranked. • Solutions ranked and their rankings taken forward to stage 3 (prototyping).

XSL•FO
RenderX

Stage	Design Thinking framework stage	Procedure	Outcomes
3	Prototyping	**5 stages of inductive content analysis conducted for solutions and rankings from workshop B**	
		• Initial and final codes were discussed for refinement and agreement with the research team and broadly grouped into themes. • Themes were mapped into 2 prototypes.	• Two prototypes were created, namely, prototype 1 (education and guidance) and prototype 2 (developing new SLTc-specific features for smart speakers). • An overview of the link between workshop 2 outputs and final prototypes is available in Multimedia Appendix 2. • Participant feedback was collated.

[a] PPI: patient and public involvement.

[b] VAT: voice-assisted technology.

[c] SLT: speech and language therapy.

Stage 2: Ideate Phase

Co-design workshops were facilitated by 2 qualified SaLTs (JM and OD), 2 health care professionals (KP and RB), and an academic (GK), with participants working in smaller groups during the workshops. Facilitator guidance and training before the workshops ensured methodological consistency when participants worked within the groups (Multimedia Appendix 3). Co-design principles of valuing lived experience, sharing power, and respect were presented at the beginning of each workshop, aiming to reduce power imbalances between researchers and participants. Both workshops lasted approximately 1 hour and were conducted via videoconferencing to enable data collection across a wider geographical area [35], avoid travel, and facilitate workshops in the evenings.

Data collection was recorded as notes in live Word (Microsoft Corporation) documents. By recording content-only contributions, participant anonymity was ensured from the outset, as no identifiers were associated with the contributions. Although participants may have been known to other group members, they were asked to respect everyone’s right to confidentiality by not sharing contributions outside the group setting. Additionally, workshops were not audio- or video-recorded, in keeping with co-design principles. Although this may have contributed to some data loss, the live recording of workshop contributions helped to mitigate this risk.

In workshop A, participants brainstormed solutions to the “How Might We” statements shown above. The process for this workshop is shown in Table 3. Following workshop A, these solutions were refined by combining similar ideas and removing duplicates (Multimedia Appendix 1) in preparation for workshop B. In workshop B, participants reviewed the solutions to the 6 problem statements created during workshop A and were asked to rank the top solutions for each problem statement from 1 (top priority) to 5 (lower priority). Ranking is regarded as a way to prioritize and reach an agreement [36]. To ensure priorities accurately reflected participants’ lived experience, each problem statement was ranked by 2 groups; however, this meant that not every group ranked every problem statement. The workshop procedure is outlined in Table 3, and the solutions are presented in the “Results” section. At the conclusion of the workshop, participants were asked by facilitators to provide feedback on their experiences of the co-design process. Following workshop B, the lead researcher (JM) collated the rankings and removed any solutions that were not ranked.

Stage 3: Prototyping

Outcomes from the workshops were analyzed using content analysis and used to create prototypes, in line with the Design Thinking process. Conventional inductive content analysis, following 5 stages, was used to allow categories to emerge directly from the workshop outputs and to reduce the volume of information [37].

The lead researcher (JM) read through the workshop outputs and any associated field notes several times to become immersed in the data. This supported note-making on initial ideas in a reflexive journal, allowing consideration of connections, similarities, and differences within the data. This process also highlighted that participants generally lacked knowledge about smart speakers, were fearful of hackers, and wanted speech-accessible smart speakers. These insights challenged the lead researcher’s confirmation bias, encouraging empathy with the experiences of PwPD and allowing the research to be shaped by user needs. This highlighted the importance of creating new features for smart speakers that better meet users’ needs, as well as utilizing existing features. Reflexivity also allowed the lead researcher to reflect on her multiple roles as a SaLT, facilitator, and analyst, and the potential for these roles to introduce interpretation bias toward a clinical perspective. As a result, an audit trail was developed to demonstrate the analysis process and enhance trust in decision-making during analysis [38].

Following data immersion, the lead author created a mind map to inductively group ideas and develop initial codes for analysis. Initial codes were shared with OD, KP, and GK for discussion and refinement, enhancing credibility through investigator triangulation and peer debriefing. The final codes were both descriptive (eg, “privacy concern”) and interpretative (eg, “need to increase motivation for speech practice”), and the meaning of each code was documented to ensure reliability during coding.

XSL•FO
RenderX

Subsequently, the workshop outputs were coded. This was conducted by hand, using graph paper and colored pens to assign meaning to each output.

Coding was conducted 3 times on separate days by 1 coder (JM) and was presented to the research team for discussion, redrafting, and agreement. It is acknowledged that coding by a single researcher may introduce bias, and, upon reflection, the involvement of 2 coders may have enabled data triangulation and enhanced data credibility. Despite this, peer debriefing helped to minimize potential impacts on the analysis of results. Similar codes were grouped into broader themes to capture meaning across the outputs. This process was also completed by hand, using colored pens to illustrate relationships between codes. Reflexive notes were recorded, discussed with coauthors (OD, KP, and GK), and refined accordingly. Finally, themes were conceptually mapped into 2 prototypes, in keeping with the Design Thinking Framework.

The content analysis process described above, from ranked ideas to the creation of prototypes, is available in Multimedia Appendix 2 as an audit trail, enhancing the credibility of the outputs [39]. All solutions that were ranked in workshop B were included in the prototypes to ensure that the prototypes reflected the wants and needs of participants. Furthermore, direct quotations, where available, from participant feedback are presented to provide a direct voice and to link outputs with interpretations [32]. Findings were sent to all participants for member checking to ensure that the written findings reflected their lived experiences and to enhance the rigor of the research.

Results

Study Participants

A total of 20 participants were recruited; 19 participated in co-design workshop A, and 16 in co-design workshop B. Overall, 15 participants took part in both workshops (Tables 4 and 5).

Table 4. Makeup of breakout rooms in workshop A.

Group number	Number of participants, n	Participants				
		People with Parkinson’s disease	Carer	Speech and language therapist	Third sector	Technology/design
1	4	✓	✓	✓	N/A[a]	✓
2	3	✓	✓	✓	N/A	N/A
3	4	✓	✓	✓	✓	N/A
4	4	✓	✓	✓	N/A	✓
5	4	✓	N/A	✓	N/A	✓

[a] N/A: not applicable.

Table 5. Makeup of breakout rooms in workshop B.

Group number	Facilitator	Number of participants	Participants				
			People with Parkinson’s disease	Carer	Speech and language therapist	Third sector	Technology/design
1	GK	4	✓	✓	✓	N/A[a]	✓
2	OD	4	✓	✓	✓	N/A	N/A
3	JM	4	✓	✓	✓	✓	N/A
4	KP	4	✓	N/A	✓	✓	✓

[a] N/A: not applicable.

The ranking of solutions aligned with the 6 problem statements is outlined in Multimedia Appendix 4. For problem statements 1, 2, 4, and 6, 3 solutions were ranked by both groups, and 4 solutions were ranked by 1 group. For problem statement 3, 2 solutions were ranked by both groups, and 5 solutions were ranked by 1 group. For problem statement 5, all 4 solutions were ranked by both groups, as only 4 options were presented. Solutions that were not given a rank by any group were removed from the results presented below. The full list of ideas available for ranking during workshop B is provided in Multimedia Appendix 1.

Stage 1: Prototyping Results

Rankings were collated into 2 main prototypes by the primary researcher and agreed upon by the team: (1) educational guidance on the therapeutic use of smart speakers, and (2) developing new SLT-specific features for smart speakers (Multimedia Appendices 5 and 6). These prototypes present the results outlined above, emphasizing cross-cutting themes. Participants’ experiences of the co-design process are also presented. The process from solutions to prototypes is fully detailed in Multimedia Appendix 2.

XSL•FO
RenderX

Prototype 1: Education and Guidance

Guides for PwPD and SaLTs, detailing how to use smart speakers to improve volume, intelligibility, and clarity of speech, were unanimously agreed upon by participants. The contents of these guides are described in Multimedia Appendix 5. Participants highlighted a gap in knowledge between the traditionally available features of smart speakers and an understanding of how these features could be repurposed to benefit speech and voice in Parkinson's disease. The suggested skills catalog for therapy would create a repository of standard smart speaker features and skills that could be utilized with therapeutic intent by SaLTs and PwPD. Suggestions included integrating prompts and positive reinforcement by building routines, for example: "Could you speak louder?" or "Well done, great practice today." PwPD felt that verbal prompts to speak louder or clearer, along with positive reinforcement from smart speakers, would replicate cuing provided by SaLTs during direct therapy and motivate home practice.

Participants indicated that routines could be used to practice scripted conversations, and that these should be personalized, include prompts to help sustain conversations, and contain only personal information that users felt comfortable sharing with their smart speaker.

It was evident that not all participants were aware of these accessibility features, which are designed to maximize engagement with smart speakers, and that education in this area may help to encourage more natural conversational reciprocity. For example, the conversation mode available on Amazon Alexa devices.

Participants also indicated that education about privacy relating to smart speaker use was required. It was reported that education for both PwPD and SaLTs would help to alleviate fears regarding personal data storage and General Data Protection Regulation (GDPR) concerns.

Prototype 2: Developing New SLT-Specific Features for Smart Speakers

Participants indicated that an Alexa skill could be created to support speech therapy, as shown in Multimedia Appendix 6. Suggestions included delivering LSVT through a smart speaker or developing a speech therapy game to support speech and voice practice. Participants suggested that this could include increased feedback, such as visual cues on a screen for volume and speech clarity and live transcription of speech that repeats back what was heard, to support self-awareness in PwPD. It is acknowledged that newer Amazon Alexa models, such as the Echo Show 10, already offer subtitling features within the settings, which provide live captioning of speech or video calls.

Additionally, participants were excited about the potential for artificial intelligence (AI) integration within smart speakers and suggested that this could be used to enable more intelligent conversations with the device. Many participants indicated that current smart speakers lacked this capability. Although intelligent conversation has not yet been integrated as a core functionality across Amazon Alexa devices, skills such as ChatGPT were perceived to facilitate live, functional conversation. Furthermore, Alexa Plus, a paid feature for Amazon devices, uses generative AI to remember previous interactions and continue conversations over time. It also offers 5 personalities, which may help users feel as though they are conversing with a person rather than a device. However, Alexa Plus is not yet available in Northern Ireland, where this research was conducted. Additionally, the *follow-up* mode within Alexa accessibility settings prevents users from having to repeat the device wake word, which Amazon suggests supports a more conversational interaction with smart speakers.

Participants also indicated that extended listening time for smart speakers would prevent mid-sentence interruptions. It is acknowledged that Amazon Alexa devices currently offer an *adaptive listening* feature in the accessibility settings, which extends input time and accommodates speech differences. Although a few participants were aware of this feature, they did not indicate its impact on their smart speaker interactions.

Furthermore, enhanced privacy features were suggested. Again, it is understood that, under Alexa privacy settings, voice commands can be enabled to clear Alexa voice history; for example, "Alexa, delete everything I've ever said." Additionally, although Alexa cannot be trained to respond only to certain voices, there is an option to set up a voice profile to receive more personalized content and prevent unauthorized voice purchases.

Although participants acknowledged that adapting smart speakers to better recognize dysarthric speech could hamper their therapeutic value, they felt that this would improve accessibility for the devices more generally. They sought devices that could gradually learn their speech patterns over time, as well as deal effectively with regional accents. Notably, there is currently no research exploring the impact of improved speech recognition in smart speakers on therapy outcomes in SLT.

In addition to ranking solutions, participants were asked about their experience of co-design using questions based on the UK standards for public involvement. Overall, participants valued the online workshop format, which facilitated engagement for those with limited mobility. They felt the workshops were informal yet professional and found the tasks interesting, positively challenging them to think of solutions. Small groups were reportedly the right size for supported discussions, and participants felt this was an effective way to gather substantial information. Participants discussed their expectations and involvement in co-design, describing feeling included and respected:

> *I had some experience of delivering co-design, so I had an idea of how it should be done...I felt valued, and felt everyone has been really equally valued, no matter how you're coming at it; person with Parkinson's, speech therapist, whatever. We've all been treated equally, with respect.* [Person from Parkinson's UK]
>
> *I felt heard and respected throughout and you did a good job of facilitating conversations for us to feel heard.* [SaLT]
>
> *The carer and patient are heard. So often in NHS setting they are the last ones to be heard y'know, what*

XSL•FO
RenderX

> *would they know. But here, they were put front and centre.* [Carer]

Participants also felt that the right people were involved in the co-design process and that there was a good balance between perspectives:

> *It's involved so many stakeholders that come from that same place of making improvements for people living with Parkinson's. It was great to see various individuals are spoken to and included.* [PwPD]
>
> *There was a really good balance of people from different backgrounds...It absolutely worked and its so important to get everyone's view; it's mostly important to hear people with Parkinson's, carers you work alongside. You get a really holistic picture of what is the most important thing from different perspectives.* [SaLT]
>
> *It was useful to be able to discuss together in a group and helpful to consider all views: SaLTs, patients and tech experts.* [PwPD]

Participants provided feedback on engagement challenges and future improvements. Some PwPD or carers felt that a bridging workshop between creating and ranking solutions would be helpful. This could have included a session to discuss all brainstormed solutions and integrate them with real-world examples. Although elements of this were included in the workshops, they felt that a third workshop would have given them time to digest the large number of solutions and some more complex ideas before ranking them. One clinician who was unable to attend the first brainstorming workshop felt that this would have helped orient her more fully before the ranking task. Others suggested that color grouping or collapsing solutions for each problem statement by themes may have made it easier to rank statements. Participants also indicated that more prompts were required to remind them to think creatively and that "anything was possible."

Participants also discussed the project's focus on smart speakers, as well as their advantages and disadvantages. A few participants felt that it would be easier to create an app, as many are available for smartphones, and most people use these devices. However, most participants felt that the voice interaction of smart speakers offered advantages over smartphones, particularly for people with a tremor. Additionally, participants felt that smart speakers could remind and motivate users to practice, whereas with an app, users often have to self-motivate or remind themselves.

Discussion

Principal Findings

This study aimed to co-produce solutions to support smart speaker use for speech and voice difficulties and to inform a future intervention. PwPD, carers, SaLTs, Parkinson's UK staff, and technology and design experts collaborated during 2 online co-design workshops to brainstorm and prioritize solutions to problems identified in prior research [9,11]. Two prototypes were developed: (1) education and guidance on the therapeutic use of smart speakers and (2) the development of new speech therapy–specific features for smart speakers. By incorporating collaborators' priorities and needs, the study offers a foundation for a future smart speaker–based intervention for speech and voice therapy in PwPD.

Impact of Co-Design

This project recognizes the need to involve end users early and meaningfully when designing health care interventions [40,41], contributing to the quality and relevance of co-designed outcomes [42]. This aligns with the Design Thinking framework, specifically the ideate and prototyping phases. While co-production with people with aphasia is increasing, there is limited evidence on co-design in SLT, especially for motor speech disorders [21,32,43]. Therefore, this research continues to contribute to and develop the evidence base regarding co-design in SLT, particularly for people with dysarthria. This study is unique, as it is believed to be the first co-design study with PwPD who have speech and voice difficulties that co-designs solutions to problems experienced when using commercial VAT technology.

Participants described personal benefits of co-design, including gaining knowledge, social interaction, and feeling heard and validated, echoing previous co-production findings [32,44] and aligning with public involvement standards [45]. These benefits are particularly relevant for PwPD, who often experience reduced participation due to speech and voice issues [46], highlighting how co-design can empower participants. Power sharing and partnership can enhance engagement and lead to more patient-centered outcomes [47], and involving SaLTs may also improve future implementation of such tools into clinical practice [48]. In wider co-design research in SLT, participants with communication difficulties report improved confidence, motivation, and sense of well-being [27], and their involvement can lead to more and better-quality outcomes [43]. Overall, this demonstrates how co-production can allow participants, such as PwPD, to feel in control, empowered, and validated. For PwPD, this co-design study both physically and metaphorically provided them with a voice, building on current evidence. Despite this, wider research also acknowledges that relinquishing power in research can be challenging for researchers, requiring an active effort to make the co-design process truly collaborative [18].

Additionally, collaborator feedback highlighted the importance of skilled facilitation in enabling communication during workshops. Although evidence on co-design facilitation strategies for people with speech and voice difficulties is limited, facilitators used clinical experience and evidence-based strategies [32] to support PwPD. These included allowing preparation time before workshops, building rapport, giving extra time to speak, screen-sharing key points, regularly checking understanding, and summarizing discussions [43,49,50]. Such approaches are crucial for inclusive and accessible co-design. Some collaborators suggested improvements, such as offering more workshops and using multimedia formats to make tasks easier, which extends the evidence base on co-design with PwPD who have speech and voice difficulties. This balance of positive experiences and suggested improvements reflects the range of participants, lived experiences, and heterogeneous needs. Advantages and

XSL•FO
RenderX

disadvantages of co-design methods should be evaluated from a range of perspectives to achieve a balance between the needs of a diverse group of PwPD.

Participants indicated that training for SaLTs in the therapeutic use of smart speakers for speech and voice difficulties was a priority. Wider research supports this finding, showing that education and guidance are required to support therapeutic adoption by SaLTs and PwPD [4,9-11], and that digital health interventions for older adults should include education in effective device use, digital literacy skills, and technical support throughout [51,52]. Tailored education and guidance may contribute to PwPD and SaLTs successfully adopting and using smart speakers to support speech and voice difficulties. As such, this study begins to advance understanding of how to support VAT adoption into clinical SLT practice. While smart speaker features make them valuable tools for chronic health management among older adults [53,54], older people in particular can struggle to comprehend the full range of smart speaker functions [11,54].

Guidance should clearly link device features to SLT goals to promote understanding and demonstrate how devices can help people achieve their SLT practice and related goals [55,56]. This may positively impact digital literacy for PwPD, supporting device adoption and regular use [56], again contributing to advances in knowledge regarding the clinical adoption of VAT. Similarly, SaLTs in our earlier research made several content suggestions for guidance to empower them to use VAT [9], including sample therapy plans, scripts, goal-setting frameworks, and evidence-based practice. However, this is the first study to collate these elements into an education and guidance prototype for SaLTs. Simplified guidance is particularly important, as clinicians often discontinue technologies they perceive as overly complex for clients [57]. Similar requirements for implementation guidance have been reported in SLT research using commercial technologies, such as virtual reality (VR) [19,58,59]. These studies highlight that therapeutic usage guides should promote the ease of use and usefulness of commercial technologies to support clinical adoption and provide opportunities to trial the devices. However, it is important to note that although commercial VR technology was used, the VR program itself was specifically created by researchers. This suggests that guidance must explain how smart speakers' out-of-the-box "Alexa skills" are relevant to SLT, given that the commercial use of the technology is not intended to be therapeutic. Unlike custom VR programs, smart speakers are off-the-shelf products not originally designed for health care. Therefore, guidance must explicitly link commercial features to therapeutic aims and support clinicians in adapting features to individual client needs, ultimately contributing to the adoption of VAT into clinical practice.

Furthermore, privacy and data protection are significant barriers to the adoption of smart speakers [60]. Common concerns include the recording of conversations and data misuse, which can deter both clinicians and clients [51,61-63]. To address this, usage guides for PwPD should include clear, accessible privacy information, support informed consent, and clearly explain how devices handle user data [64]. Given SaLTs' responsibility for safeguarding client data, guidance should map VAT's GDPR compliance and potential risks to SLT governance policies, such as Data Protection Impact Assessments. This study, therefore, begins to answer questions posed by previous research [10] regarding how VAT may be implemented in accordance with clinical governance requirements. Previous findings highlight that many SaLTs lack clarity on which technologies meet governance and GDPR standards [57]. Reassuring both clinicians and PwPD about privacy may improve confidence and facilitate adoption [65]. Future evaluation of guidance acceptability and usability could apply frameworks such as the Technology Acceptance Model or Unified Theory of Acceptance and Use of Technology 2.

Delivery

Participants suggested delivering training through Royal College of Speech and Language Therapy–led webinars, live demonstrations, and group sessions led by trained SaLTs. While previous research has not identified optimal delivery formats [4,10], this study provides new insights into practical implementation and contributes to the evidence base regarding the therapeutic use of VAT in SLT clinical practice. The literature indicates that older adults often prefer hands-on, task-based learning supported by written instructions [66,67]. A training program using VAT as a tool for activities of daily living with adults with cognitive communication disorders indicated a need for written, easy-to-follow instructions, with hands-on support to overcome low technological literacy [67]. Group-based workshops can offer a supportive, low-pressure environment for exploration and skill-building with in-person support [52]. This is particularly important for users with limited experience or confidence in digital tools.

Findings indicate that SaLTs are central to introducing and supporting smart speaker use in therapy. When clinicians demonstrate relevance and ease of use, PwPD may be more likely to adopt the technology [67]. By increasing perceived usefulness and reducing concerns, training can enhance performance expectancy and digital engagement. Additionally, previous research on integrating commercial technologies in SLT has highlighted the importance of multifaceted training approaches, including device trials, workshops, clinical manuals, and information technology support [58,68]. Additional methods, such as guided observation and co-delivered interventions, may be necessary to bridge the gap between knowledge and practice [58,69]. As such, this study begins to address gaps in knowledge regarding the implementation of VAT as a therapeutic tool for speech and voice difficulties associated with Parkinson's disease.

Participants highlighted the need to develop SLT-specific features for smart speakers, designed for therapeutic use. For example, Cassano et al [70] described a SaLT building a custom skill. At the time of publication, at least three speech therapy Alexa skills existed: Speech Therapy Practice, Speech Device Practice, and Let's Talk. Additionally, 2 further speech therapy skills were identified but are no longer publicly available on the Amazon Skills store: Speech Doctor, as discussed by Makin et al [71], and Speech Therapy by Cathal Killeen. Notably, Speech Therapy Practice is a live Alexa skill developed by a SaLT that enables people with aphasia to practice very basic

XSL•FO
RenderX

words and phrases, such as colors, opposites, who/what questions, and yes/no questions. While this may potentially act as a starting point for SaLTs, the skill lacks applicability to practicing phrases and sentences and, in its current state, is unlikely to meet the speech practice needs of PwPD. To date, there are no specific Alexa skills for adults with Parkinson's or targeting dysarthria, and our research highlights the potential for future development. Future research may seek to work with developers to create an Alexa skill for this population that can be used to support home practice of speech therapy exercises. Features may include prompts for loud, clear speech; increased feedback on volume and intelligibility with suggestions for improvement; the ability to monitor progress; visual displays and biofeedback; reminders to complete therapy tasks; and LSVT-style exercises with gamification [9,11]. Such features align with wider studies integrating technology into SLT and related areas, including apps using Google Glass [72], smart speaker–based physical activity interventions [73,74], and social engagement tools for people with disabilities [75]. Development platforms like Alexa Skills Kit and Alexa Blueprint may offer scalable, cost-effective options, enabling a focus on increasing motivation, engagement, and potential adherence to intervention programs. A curated hub of Alexa skills that can be used for SLT goals may also support clinical implementation. For example, Esquivel et al [76] developed a repository of Alexa skills and recommendations for people with disabilities, by people with disabilities. Future research may explore the acceptability of a speech therapy–specific Alexa skill and its implementation within clinical practice.

Given the commercial nature of smart speakers, it may be beneficial to first assess their current therapeutic value before creating bespoke skills. As our study focused on co-design processes and did not include a formal evaluation of intervention usability, effectiveness, or acceptability, future research may consider testing the current prototypes to determine real-world clinical impact and user outcomes. This study establishes the rationale for a future feasibility study to examine the effectiveness of VAT as a therapeutic tool for speech and voice difficulties in Parkinson's disease. At the time of writing, no studies have been conducted in this area using commercial VAT. Emerging SLT research shows benefits for speech clarity in populations with intellectual disabilities and speech sound disorders [5,71], citing immediate rewards, spaced practice, enhanced autonomy, intrinsic motivation, and reduced social barriers as mechanisms of change in speech. However, these interventions do not follow established SLT intervention protocols [5,71]. Therefore, future studies should evaluate the effectiveness and usability of smart speakers for PwPD using principles of neuroplasticity and motor learning from SLT protocols, such as LSVT LOUD or Speak Out!

Despite this, the challenges surrounding the therapeutic use of smart speakers cannot be ignored. Smart speakers rely on evolving ASR models, a type of AI, which are continually being improved. ASR models can change without warning, presenting a risk to the reliability of baseline measurements and the measurement of therapy goals [77]. Furthermore, ASR errors are often higher than expected for dysarthric speech, speakers of minority languages, and those with regional accents [78-80]. Without clear and specific feedback on device or speech errors, both PwPD and SaLTs are left without information about where the "error" lies, whether it is speech- or device-related. These risks may reinforce maladaptive speech behaviors if speech practice is based on inconsistent or misleading responses from the device. It may also damage client motivation and confidence, with PwPD blaming themselves for technological errors. Research demonstrates that speakers can attribute ASR errors to themselves and link this to their sense of identity, including racial, regional, and locational identity [77]. To mitigate this lack of transparency, it is essential that SaLTs educate potential VAT users on strategies for adapting speech, managing frustration, and correctly interpreting VAT errors, as well as raising awareness of the limited ASR training on dysarthric speech and some minority or foreign languages [9,11]. This highlights the importance of a therapeutic usage manual for smart speakers for people with speech and voice difficulties and for SaLTs, as indicated in the current findings.

However, it should be noted that projects such as Voiceitt, Google Euphonia, and Project Relate aim to improve ASR accuracy in recognizing dysarthric speech, and Accessible Voice Interaction Technology for Aphasia (AVITA) aims to improve the accessibility of smart speakers [81], which may have the unintended consequence of limiting certain therapeutic applications of smart speakers in SLT. When smart speaker recognition is improved, speech difficulties no longer affect recognition, meaning all speech is easily recognized. This can be problematic, as speech that may not be intelligible in real life is recognized by devices. Consequently, this hampers therapeutic applications, because positive biofeedback provided by smart speakers does not reflect the speaker's intelligibility to unfamiliar listeners in everyday contexts. Indeed, participants in this research indicated that future adaptations of smart speakers, outside of therapeutic contexts, should aim to better recognize dysarthric speech and regional accents. Future smart speaker designs may bridge the gap between standard out-of-the-box devices and fully customized skills. For example, smart speakers could allow users to set recognition thresholds, enabling both increased accessibility for users and therapeutic usage for clinicians. Given the rapid pace of innovation, continued review of emerging literature and technologies is recommended throughout the development and implementation stages.

Limitations

This co-design study offered valuable insights into developing VAT tools for PwPD with speech and voice difficulties; however, limitations are evident.

Participants suggested an additional workshop between the ideation and prioritization phases, that is, between workshops A and B. A bridging session could have allowed more reflection and improved understanding, potentially leading to rankings that more accurately reflected lived experience. Furthermore, although recruitment was successful, there was some participant dropout between workshops A and B. This necessitated merging groups in workshop B, which may have influenced group dynamics and limited continuity of discussion.

XSL•FO
RenderX

Despite efforts to recruit a diverse group, the sample was small (n=20), and certain perspectives, such as those of people with advanced Parkinson's or severe dysarthria, were underrepresented. This may limit the generalizability of the findings.

Finally, given the rapidly evolving technology landscape in AI and ASR, some recommendations may become outdated by the time of implementation. This includes changes in smart speaker capabilities, privacy policies, and integration with large language models (eg, AI conversational agents).

Conclusions

This study highlights the value of co-designing smart speaker interventions with PwPD, carers, SaLTs, third sector representatives and technology and design experts to address challenges in using VAT for speech therapy. Using a participatory Design Thinking approach, user-centered solutions were generated to improve the accessibility, usability, and therapeutic potential of smart speakers.

Two prototypes were developed: (1) education and guidance for PwPD and SaLTs, and (2) speech therapy–specific smart speaker features.

The outputs balance commercial technology with clinical needs, focusing on privacy, troubleshooting, and feedback for home use, while reinforcing co-design as a powerful method for developing digital health tools. Co-design also ensured that interventions reflected lived experience and clinical insight, enhancing the likelihood of adoption and sustained use. This research strengthens the evidence for co-design in SLT and supports smart speakers as tools to enhance therapy access, promote self-management, and reduce pressure on SLT services. Future work should develop and evaluate these prototypes to assess their real-world impact and scalability.

Acknowledgments

We thank Rory Bradley for his contributions to the co-design workshops in this research. The authors would also like to thank Donna McGuckin, Emmet Leyden, and Dr Marc Parker for their contribution to this research as members of the patient and public involvement (PPI) group.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results condensed from workshop A.
[DOCX File , 22 KB - rehab_v13i1e84364_app1.docx]

Multimedia Appendix 2
Ranking to prototype creation.
[DOCX File , 28 KB - rehab_v13i1e84364_app2.docx]

Multimedia Appendix 3
Facilitator guidance.
[DOCX File , 21 KB - rehab_v13i1e84364_app3.docx]

Multimedia Appendix 4
Ranking results.
[DOCX File , 21 KB - rehab_v13i1e84364_app4.docx]

Multimedia Appendix 5
Prototype 1: education and guidance on the therapeutic use of smart speakers.
[DOCX File , 17 KB - rehab_v13i1e84364_app5.docx]

Multimedia Appendix 6
Prototype 2: developing new speech therapy–specific features for smart speakers.
[DOCX File , 16 KB - rehab_v13i1e84364_app6.docx]

References

1. The Infinite Dial UK. Edison Research. 2025. URL: https://www.edisonresearch.com/wp-content/uploads/2025/05/Infinite-Dial-UK-2025-webinar-presentation-for-download.pdf [accessed 2025-09-01]
2. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput-Hum Interact 2020 Sep 16;27(4):1-27. [doi: 10.1145/3373759]

XSL•FO
RenderX

3. Cave R, Bloch S. The use of speech recognition technology by people living with amyotrophic lateral sclerosis: a scoping review. Disabil Rehabil Assist Technol 2023 Oct;18(7):1043-1055 [FREE Full text] [doi: 10.1080/17483107.2021.1974961] [Medline: 34511007]
4. Duffy O, Synnott J, McNaney R, Brito Zambrano P, Kernohan WG. Attitudes toward the use of voice-assisted technologies among people with Parkinson disease: findings from a web-based survey. JMIR Rehabil Assist Technol 2021 Mar 11;8(1):e23006 [FREE Full text] [doi: 10.2196/23006] [Medline: 33704072]
5. Smith E, Sumner P, Hedge C, Powell G. Smart speaker devices can improve speech intelligibility in adults with intellectual disability. Int J Lang Commun Disord 2021 May;56(3):583-593. [doi: 10.1111/1460-6984.12615] [Medline: 33772998]
6. Bleakley A, Rough D, Roper A, Lindsay S, Porcheron M, Lee M, et al. Exploring smart speaker user experience for people who stammer. 2022 Presented at: ASSETS '22: The 24th International ACM SIGACCESS Conference on Computers and Accessibility; October 23-26, 2022; Athens, Greece p. 23-26. [doi: 10.1145/3517428.3544823]
7. Miller N. Communication changes in Parkinson's disease. Pract Neurol 2017 Aug;17(4):266-274. [doi: 10.1136/practneurol-2017-001635] [Medline: 28687681]
8. Miller N, Deane KHO, Jones D, Noble E, Gibb C. National survey of speech and language therapy provision for people with Parkinson's disease in the United Kingdom: therapists' practices. Int J Lang Commun Disord 2011;46(2):189-201. [doi: 10.3109/13682822.2010.484849] [Medline: 21401817]
9. Mills J, Duffy O, Pedlow K, Kernohan G. Exploring speech and language therapists' perspectives of voice-assisted technology as a tool for dysarthria: qualitative study. JMIR Rehabil Assist Technol 2025 Sep 02;12:e75044 [FREE Full text] [doi: 10.2196/75044] [Medline: 40896823]
10. Kulkarni P, Duffy O, Synnott J, Kernohan WG, McNaney R. Speech and language practitioners' experiences of commercially available voice-assisted technology: web-based survey study. JMIR Rehabil Assist Technol 2022 Jan 05;9(1):e29249 [FREE Full text] [doi: 10.2196/29249] [Medline: 34989694]
11. Mills J, Duffy O, Pedlow K, Kernohan G. Exploring the perceptions of voice-assisted technology as a tool for speech and voice difficulties: focus group study among people with Parkinson disease and their carers. JMIR Rehabil Assist Technol 2025 Jul 16;12:e75316 [FREE Full text] [doi: 10.2196/75316] [Medline: 40669073]
12. Ramig L, Halpern A, Spielman J, Fox C, Freeman K. Speech treatment in Parkinson's disease: randomized controlled trial (RCT). Mov Disord 2018 Nov;33(11):1777-1791 [FREE Full text] [doi: 10.1002/mds.27460] [Medline: 30264896]
13. Altman M, Huang TT, Breland JY. Design thinking in health care. Prev Chronic Dis 2018 Sep 27;15:E117 [FREE Full text] [doi: 10.5888/pcd15.180128] [Medline: 30264690]
14. Sumner J, Chong LS, Bundele A, Wei Lim Y. Co-designing technology for aging in place: a systematic review. Gerontologist 2021 Sep 13;61(7):e395-e409 [FREE Full text] [doi: 10.1093/geront/gnaa064] [Medline: 32506136]
15. Brown T. Design and thinking. Harvard Educational Review 2013 Apr 01;86(1):84-92. [doi: 10.17763/1943-5045-83.1.243]
16. Vargas C, Whelan J, Brimblecombe J, Allender S. Co-creation, co-design, co-production for public health - a perspective on definition and distinctions. Public Health Res Pract 2022 Jun 15;32(2):3222211 [FREE Full text] [doi: 10.17061/phrp3222211] [Medline: 35702744]
17. Benjamin-Thomas TE, Corrado AM, McGrath C, Rudman DL, Hand C. Working towards the promise of participatory action research: learning from ageing research exemplars. International Journal of Qualitative Methods 2018 Dec 26;17(1):e-1609406918817953. [doi: 10.1177/1609406918817953]
18. Roper A, Skeat J. Innovation through participatory design: collaborative qualitative methods in the development of speech-language pathology technology. Int J Speech Lang Pathol 2022 Oct;24(5):527-532. [doi: 10.1080/17549507.2022.2050943] [Medline: 35506478]
19. Brassel S, Brunner M, Power E, Campbell A, Togher L. Speech-language pathologists' views of using virtual reality for managing cognitive-communication disorders following traumatic brain injury. Am J Speech Lang Pathol 2023 Mar 23;32(2S):907-923. [doi: 10.1044/2022_ajslp-22-00077]
20. Wilson S, Roper A, Marshall J, Galliers J, Devane N, Booth T, et al. Codesign for people with aphasia through tangible design languages. CoDesign 2015 Jan 09;11(1):21-34. [doi: 10.1080/15710882.2014.997744]
21. Cruice M, Aujla S, Bannister J, Botting N, Boyle M, Charles N, et al. Creating a novel approach to discourse treatment through coproduction with people with aphasia and speech and language therapists. Aphasiology 2021 Jul 03;36(10):1159-1181. [doi: 10.1080/02687038.2021.1942775]
22. Shiggins C, Coe D, Gilbert L, Research Collaboration A, Mares K. Development of an “aphasia-accessible participant in research experience survey” through co-production. Aphasiology 2022 Jan 31;39(12):1659-1692. [doi: 10.1080/02687038.2021.1996532]
23. Arnstein S. A Ladder Of Citizen Participation. Journal of the American Institute of Planners 1969 Jul;35(4):216-224 [FREE Full text] [doi: 10.1080/01944366908977225]
24. Elwyn G, Nelson E, Hager A, Price A. Coproduction: when users define quality. BMJ Qual Saf 2020 Sep;29(9):711-716 [FREE Full text] [doi: 10.1136/bmjqs-2019-009830] [Medline: 31488570]
25. Penny J, Slay J. Commissioning for outcomes and co-production: a practical guide for local authorities. New Economics Foundation. London, UK: New Economics Foundation; 2014. URL: https://neweconomics.org/2014/06/commissioning-outcomes-co-production [accessed 2026-01-25]

XSL•FO
RenderX

26. Connery A, Salsberg J. Exploring participatory health research and its application to speech and language therapy research practices. Int J Lang Commun Disord 2024;59(4):1257-1268. [doi: 10.1111/1460-6984.12994] [Medline: 38130139]
27. McMenamin R, Griffin M, Grzybowska B, Pound C. Working together: experiences of people with aphasia as co-researchers in participatory health research studies. Aphasiology 2021 Jun 25;39(12):1601-1622. [doi: 10.1080/02687038.2021.1923948]
28. Cook T, Boote J, Buckley N, Vougioukalou S, Wright M. Accessing participatory research impact and legacy: developing the evidence base for participatory approaches in health research. Educational Action Research 2017 Jul 31;25(4):473-488. [doi: 10.1080/09650792.2017.1326964]
29. Roper A, Wilson S, Neate T, Marshall J. Speech and language. In: Yesilada Y, Harper S, editors. A Foundation for Research. London, UK: Springer; 2019:121-131.
30. McNaney R, Tsekleves E, Synnott J. Future opportunities for IoT to support people with Parkinson's. 2020 Presented at: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; April 25-30, 2020; Honolulu, HI p. 1-15 URL: https://dl.acm.org/doi/10.1145/3313831.3376871 [doi: 10.1145/3313831.3376871]
31. Mc Menamin R, Isaksen J, Manning M, Tierney E. Distinctions and blurred boundaries between qualitative approaches and public and patient involvement (PPI) in research. Int J Speech Lang Pathol 2022 Oct 28;24(5):515-526 [FREE Full text] [doi: 10.1080/17549507.2022.2075465] [Medline: 35762365]
32. Clay P, Walton T, Malin E, Hutchinson M, Levitt K, Williams C, et al. Better conversations with Parkinson’s: co-production of a novel speech and language therapy intervention with people living with Parkinson’s. Research for All 2024;8(1):1-17. [doi: 10.14324/RFA.08.1.07]
33. Beckman SL. To frame or reframe: where might design thinking research go next? California Management Review 2020 Mar 02;62(2):144-162. [doi: 10.1177/0008125620906620]
34. Mah H, Dobson R, Thomson A. The importance of lived experience: a scoping review on the value of patient and public involvement in health research. Health Expect 2025 Apr 26;28(2):e70205-e70205 [FREE Full text] [doi: 10.1111/hex.70205] [Medline: 40135803]
35. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. International Journal of Qualitative Methods 2019 Sep 11;18(1):1-8. [doi: 10.1177/1609406919874596]
36. Larkman CS, Lanyon L, Rose ML. Co-designing solutions to the challenges speech pathologists and interpreters encounter when they collaborate to deliver aphasia therapy. Disabil Rehabil 2025 Dec;47(24):6468-6482 [FREE Full text] [doi: 10.1080/09638288.2025.2486459] [Medline: 40188384]
37. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
38. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
39. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
40. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
41. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
42. National Institute for Health and Care Research (NIHR). Briefing notes for researchers in public involvement in NHS, health and social care research. NIHR. 2021. URL: https://www.nihr.ac.uk/documents/briefing-notes-for [accessed 2025-05-26]
43. Jayes M, Moulam L, Meredith S, Whittle H, Lynch Y, Goldbart J, et al. Making public involvement in research more inclusive of people with complex speech and motor disorders: the I-ASC project. Qual Health Res 2021 Jun 28;31(7):1260-1274 [FREE Full text] [doi: 10.1177/1049732321994791] [Medline: 33645331]
44. Bird M, Ouellette C, Whitmore C, Li L, Nair K, McGillion MH, et al. Preparing for patient partnership: a scoping review of patient partner engagement and evaluation in research. Health Expect 2020 Jun 10;23(3):523-539 [FREE Full text] [doi: 10.1111/hex.13040] [Medline: 32157777]
45. UK standards for public involvement. UK Public Involvement Standards Development Partnership/National Institute for Health and Care Research (NIHR). URL: https://sites.google.com/nihr.ac.uk/pi-standards/standards [accessed 2025-05-26]
46. Gillivan-Murphy P, Miller N, Carding P. Voice treatment in Parkinson’s disease: patient perspectives. JPRLS 2019 Jul;Volume 9(1):29-42. [doi: 10.2147/jprls.s180183]
47. Carman KL, Dardess P, Maurer M, Sofaer S, Adams K, Bechtel C, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. Health Aff (Millwood) 2013 Feb;32(2):223-231. [doi: 10.1377/hlthaff.2012.1133] [Medline: 23381514]

XSL•FO
RenderX

48. Singer I, Klatte IS, de Vries R, van der Lugt R, Gerrits E. Using co-design to develop a tool for shared goal-setting with parents in speech and language therapy. Int J Lang Commun Disord 2022 Nov 20;57(6):1281-1303 [FREE Full text] [doi: 10.1111/1460-6984.12753] [Medline: 35859264]
49. Burton A, Ogden M, Cooper C. Planning and enabling meaningful patient and public involvement in dementia research. Curr Opin Psychiatry 2019 Nov;32(6):557-562. [doi: 10.1097/YCO.0000000000000548] [Medline: 31306247]
50. Swinburn K. Working alongside people with stroke and aphasia. In: Volkmer A, Broomfield K, editors. Seldom Heard Voices in Service User Involvement: The How and Why of Meaningful Collaboration. Havant, UK: J&R Press; 2022:129-150.
51. McCloud R, Perez C, Bekalu MA, Viswanath K. Using smart speaker technology for health and well-being in an older adult population: pre-post feasibility study. JMIR Aging 2022 May 09;5(2):e33498 [FREE Full text] [doi: 10.2196/33498] [Medline: 35532979]
52. Masters K, Correia R, Nemethy K, Benjamin J, Carver T, MacNeill H. Online learning in health professions education. Part 2: Tools and practical application: AMEE Guide No. 163. Medical Teacher 2023 Sep 23;46(1):18-33. [doi: 10.1080/0142159x.2023.2259069]
53. Kim S, Choudhury A. Exploring older adults’ perception and use of smart speaker-based voice assistants: a longitudinal study. Computers in Human Behavior 2021 Nov 18;124(3):106914-106924. [doi: 10.1016/j.chb.2021.106914]
54. Chang F, Sheng L, Gu Z. Investigating the integration and the long-term use of smart speakers in older adults’ daily practices: qualitative study. JMIR Mhealth Uhealth 2024 Feb 12;12:e47472 [FREE Full text] [doi: 10.2196/47472] [Medline: 38345844]
55. Arnold A, Kolody S, Comeau A, Miguel Cruz A. What does the literature say about the use of personal voice assistants in older adults? A scoping review. Disabil Rehabil Assist Technol 2024 Jan;19(1):100-111. [doi: 10.1080/17483107.2022.2065369] [Medline: 35459429]
56. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019 Apr 08;21(4):e11604 [FREE Full text] [doi: 10.2196/11604] [Medline: 30958272]
57. Tazi F, Nandakumar A, Dykstra J, Rajivan P, Das S. Privacy, security, and usability tradeoffs of telehealth from practitioners’ perspectives. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2023 Oct 25;67(1):1862-1867. [doi: 10.1177/21695067231199689]
58. Vaezipour A, Aldridge D, Koenig S, Theodoros D, Russell T. A mixed-method investigation of clinician acceptance, barriers and enablers of virtual reality technology in communication rehabilitation. Disabil Rehabil 2022 Jul 14;44(15):3946-3958. [doi: 10.1080/09638288.2021.1895333] [Medline: 33715566]
59. Mills J, Duffy O. Speech and language therapists' perspectives of virtual reality as a clinical tool for autism: cross-sectional survey. JMIR Rehabil Assist Technol 2025 Feb 27;12:e63235-e63235 [FREE Full text] [doi: 10.2196/63235] [Medline: 40014826]
60. Maccario G, Naldi M. Alexa, is my data safe? The (ir)relevance of privacy in smart speakers reviews. International Journal of Human–Computer Interaction 2022 Apr 20;39(6):1244-1256. [doi: 10.1080/10447318.2022.2058780]
61. Lustgarten SD, Garrison YL, Sinnard MT, Flynn AW. Digital privacy in mental healthcare: current issues and recommendations for technology use. Curr Opin Psychol 2020 Dec;36:25-31 [FREE Full text] [doi: 10.1016/j.copsyc.2020.03.012] [Medline: 32361651]
62. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220 [FREE Full text] [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
63. Quinn K, Leiser Ransom S, O'Connell C, Muramatsu N, Marquez DX, Chin J. Assessing the feasibility and acceptability of smart speakers in behavioral intervention research with older adults: mixed methods study. J Med Internet Res 2024 Aug 30;26:e54800 [FREE Full text] [doi: 10.2196/54800] [Medline: 39213034]
64. Zhang R, Li H, Chen A, Liu Z, Lee Y. AI privacy in context: a comparative study of public and institutional discourse on conversational AI privacy in the US and Chinese social media. Social Media + Society 2024 Oct 28;10(4):1-19. [doi: 10.1177/20563051241290845]
65. Saripalle R, Patel R. From command to care: a scoping review on utilization of smart speakers by patients and providers. Mayo Clin Proc Digit Health 2024 Jun;2(2):207-220 [FREE Full text] [doi: 10.1016/j.mcpdig.2024.03.002] [Medline: 40207172]
66. Mitzner TL, Fausset CB, Boron JB, Adams AE, Dijkstra K, Lee CC, et al. Older adults' training preferences for learning to use technology. Proc Hum Factors Ergon Soc Annu Meet 2008 Sep;52(26):2047-2051 [FREE Full text] [doi: 10.1177/154193120805202603] [Medline: 25309139]
67. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. Front Public Health 2024;12:1447747 [FREE Full text] [doi: 10.3389/fpubh.2024.1447747] [Medline: 39555039]
68. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251.e1 [FREE Full text] [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]

XSL•FO
RenderX

69. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
70. Cassano F, Pagano A, Piccinno A. Supporting speech therapies at (smart) home through voice assistance. 2021 Presented at: International Symposium on Ambient Intelligence; October 6-8, 2021; Salamanca, Spain. [doi: 10.1007/978-3-031-06894-2_10]
71. Makin L, Smith E, Hedge C, Sumner P, Powell G. Smart speakers are an acceptable and feasible speech practice tool for children with speech difficulties. Disabil Rehabil Assist Technol 2025 Oct 17;20(7):2510-2521 [FREE Full text] [doi: 10.1080/17483107.2025.2491636] [Medline: 40244154]
72. McNaney R, Vines J, Roggen D, Balaam M, Zhang P, Poliakov I, et al. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Apr 26 Presented at: CHI '14: The SIGCHI Conference on Human Factors in Computing Systems; May 1, 2014; Toronto, Canada p. 2551-2554. [doi: 10.1145/2556288.2557092]
73. Jansons P, Fyfe J, Via JD, Daly RM, Gvozdenko E, Scott D. Barriers and enablers for older adults participating in a home-based pragmatic exercise program delivered and monitored by Amazon Alexa: a qualitative study. BMC Geriatr 2022 Mar 25;22(1):248-258 [FREE Full text] [doi: 10.1186/s12877-022-02963-2] [Medline: 35337284]
74. Carlin A, Logue C, Flynn J, Murphy MH, Gallagher AM. Development and feasibility of a family-based health behavior intervention using intelligent personal assistants: randomized controlled trial. JMIR Form Res 2021 Jan 28;5(1):e17501 [FREE Full text] [doi: 10.2196/17501] [Medline: 33507155]
75. Greuter S, Balandin S, Watson J. Social games are fun: exploring social interactions on smart speaker platforms for people with disabilities. In: CHI PLAY '19 Extended Abstracts: Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2019 Presented at: CHI PLAY '19: The Annual Symposium on Computer-Human Interaction in Play; October 22-25, 2019; Barcelona, Spain p. 429-435. [doi: 10.1145/3341215.3356308]
76. Esquivel P, Gill K, Chung C, Ding D, Faieta J. Smart speakers and skill use: what do we know? Disabil Rehabil Assist Technol 2025 Feb 08;20(2):488-497. [doi: 10.1080/17483107.2024.2387801] [Medline: 39115270]
77. Cave R. How people living with amyotrophic lateral sclerosis use personalized automatic speech recognition technology to support communication. J Speech Lang Hear Res 2024 Nov 07;67(11):4186-4202. [doi: 10.1044/2024_jslhr-24-00097]
78. Mengesha Z, Heldreth C, Lahav M, Sublewski J, Tuennerman E. "I don't think these devices are very culturally sensitive."—Impact of automated speech recognition errors on African Americans. Front Artif Intell 2021;4:725911 [FREE Full text] [doi: 10.3389/frai.2021.725911] [Medline: 34901836]
79. Green J, MacDonald RL, Jiang PP, Cattiau J, Heywood R, Cave R. Automatic speech recognition of disordered speech: personalized models outperforming human listeners on short phrases. 2021 Presented at: The 22nd Annual Conference of the International Speech Communication Association; August 30 to September 3, 2021; Brno, Czechia p. 4778-4782 URL: https://www.isca-archive.org/interspeech_2021/green21_interspeech.pdf [doi: 10.21437/interspeech.2021-1384]
80. Tobin J, Nelson P, MacDonald B, Heywood R, Cave R, Seaver K, et al. Automatic speech recognition of conversational speech in individuals with disordered speech. J Speech Lang Hear Res 2024 Nov 07;67(11):4176-4185. [doi: 10.1044/2024_jslhr-24-00045]
81. Improving access to Alexa for people with aphasia. City of St George's, University of London. 2025. URL: https://tinyurl.com/45fcp8k6 [accessed 2025-09-10]

Abbreviations

AI: artificial intelligence
ASR: automatic speech recognition
GDPR: General Data Protection Regulation
LSVT: Lee Silverman Voice Treatment
PPI: patient and public involvement
PwPD: people with Parkinson's disease
SaLT: speech and language therapist
SLT: speech and language therapy
VAT: voice-assisted technology
VR: virtual reality

XSL•FO
RenderX

Edited by S Munce; submitted 22.Sep.2025; peer-reviewed by D Singhal, Y Lin; comments to author 04.Nov.2025; revised version received 18.Dec.2025; accepted 13.Jan.2026; published 04.Feb.2026.

Please cite as:
Mills J, Kernohan G, Pedlow K, Duffy O
Voice-Assisted Technology for People With Parkinson's Disease Experiencing Speech and Voice Difficulties: Co-Designing Solutions Using Design Thinking
JMIR Rehabil Assist Technol 2026;13:e84364
URL: https://rehab.jmir.org/2026/1/e84364
doi:10.2196/84364
PMID:

©Jodie Mills, George Kernohan, Katy Pedlow, Orla Duffy. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 04.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Influence of Digitalization on Multidisciplinary Poststroke Rehabilitation Practice: Qualitative Study

Ann Marie Hestetun-Mandrup[1,2], MHS; Charlotta Hamre[2], MHS, PhD; Anne Lund[1], Prof Dr; Anne Catrine Trægde Martinsen[1,2], Prof Dr; Hong-Gu He[3,4], Prof Dr; Minna Pikkarainen[1], Prof Dr

[1]Department of Rehabilitation Science and Health Technology, Faculty of Health, OsloMet – Oslo Metropolitan University, Pb. 4 St. Olavs Plass, Oslo, Norway

[2]Department for Research and Education, Sunnaas Rehabilitation Hospital, Oslo, Norway

[3]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[4]National University Health System, Singapore, Singapore

Corresponding Author:
Ann Marie Hestetun-Mandrup, MHS
Department of Rehabilitation Science and Health Technology, Faculty of Health, OsloMet – Oslo Metropolitan University, Pb. 4 St. Olavs Plass, Oslo, Norway

Abstract

Background: Leveraging digital technologies in health care is recognized as essential for effective and efficient services. However, significant challenges remain in implementing these technologies in stroke rehabilitation practice, and research on their influence is limited.

Objective: This study aimed to explore the current influence of digital technologies on stroke rehabilitation practices and consider how these technologies could shape the future landscape of rehabilitation for multidisciplinary health care professionals in poststroke rehabilitation.

Methods: A qualitative, exploratory design was used. Data were collected from 12 experienced multidisciplinary health care professionals at 2 Norwegian rehabilitation settings via semistructured interviews, and the data were analyzed using reflexive thematic analysis. Data analysis was guided by social practice theory.

Results: The 12 participants included experienced physiotherapists, occupational therapists, speech therapists, nurses, physicians, and social workers. The following three main themes were generated: (1) Outsourcing information about and to stroke survivors: coordination and continuity within and across services (subthemes on follow-up and interservice collaboration, and user-centered approaches); (2) Navigating the ambivalence of remaining human relations in digital psychosocial support conversations (highlighting multidisciplinary challenges in building relational depth and addressing sensitive topics); and (3) Enhancing digital supplements for assessment and engagement in motor rehabilitation (subthemes on progress monitoring and motor skills exercises). Overall, the use of digital technologies in specialized stroke rehabilitation practices was seen as an adjunct to practices. While digital technologies influenced rehabilitation practices, ambivalence and challenges were noted, particularly in digitalizing multidisciplinary psychological support and exercise programs. Systems for sharing medical records and goal-setting apps, which enhance coordination and involve stroke survivors, were emphasized as future digital technologies that can shape stroke rehabilitation.

Conclusions: Health care professionals used various technologies in their daily specialist practices, as well as for the coordination and follow-up of stroke survivors after referral to community services. This study identified several organizational processes, roles, standards, and rules that can act as barriers or drivers to implementing digital technologies in practice. Viewing familiar digital technology as a supplement to existing practices, rather than as a singular solution for all areas of specialized stroke rehabilitation, offers significant potential for quality improvement. These findings provide valuable insights for technology developers, health care personnel, and user groups in specialized neurological rehabilitation settings.

(JMIR Rehabil Assist Technol 2026;13:e77753) doi:10.2196/77753

KEYWORDS

digital therapeutics; digital technology; goals; practice theory; rehabilitation; stroke

XSL•FO
RenderX

Introduction

The World Health Organization (WHO) has promoted digital health to enhance the quality and accessibility of health care services [1]. Leveraging digital technologies in health care is recognized as essential for effective and efficient services [2]. Limited resources have led to earlier hospital discharge and outsourced rehabilitation services aided by digital technologies, which are areas where politicians, clinicians, stroke survivors, and researchers have turned their attention to [3]. WHO's initiative of Digital and Assistive Technologies for Aging (DATA) [4] further underscores the importance of using accessible digital technologies in developing and sustaining new stroke care pathways [5]. The term "digital technologies" refers to electronically driven devices or applications [6,7]. Digital technologies encompass various information and communication technologies (ICTs) used in medicine and health care for disease management and wellness enhancement [6]. These technologies include devices, such as computers, smartphones, and tablets, along with their associated applications and internet connectivity [7]. Studies have shown that digital technologies improve activity and function [8-11], social participation [12], and medical adherence among individuals with stroke [13-15]. Technologies, such as gaming, videoconferencing, online health websites, apps, and wearables, have been especially explored for their potential to enhance exercise programs in both clinic and home settings in the stroke population [16]. Despite the rapid adaptation of digital technologies after the COVID-19 pandemic, scaling up remains slow [17].

Effective neurorehabilitation and tailored digital interventions are crucial for future stroke home rehabilitation [18], as they are cost-effective, portable, and motivational solutions to support daily activities [19]. However, challenges persist in home implementation [19], which might explain the lack of superior effects on supervised digital motor rehabilitation at home over conventional training for stroke survivors [18]. Although studies have found overall positive attitudes toward digital technologies in poststroke rehabilitation from the perspective of individuals with stroke, due to increased accessibility and convenience, there is an advocated need for personal contact, shared responsibility, and collaboration with health care professionals (HCPs) [20].

Building on these findings, rehabilitation specialists have introduced many digital solutions for home-based rehabilitation [21], yet significant challenges remain in adapting these technologies in practice, and research on their influence is limited. Despite guidelines and research recommending technology use in stroke rehabilitation, actual implementation and clinical validation are inconsistent [16,22-24]. A systematic review by Borg et al [25] focused on older people and people with disabilities and highlighted the need for research on technology adoption in practice, particularly from a theoretical, multidisciplinary approach. A longitudinal study across inpatient, outpatient, and community rehabilitation settings found that the type and amount of technology use varied between neurological populations, with the most frequent and varying use in stroke rehabilitation [22]. Contrary to this, Langan et al [16] found that physical and occupational therapists do not widely incorporate digital technologies into their practices, often prescribing written, paper-based, and nontechnological home exercises after hospital discharge. The primary barriers often include financial constraints, lack of technological experience, and limited time [26,27]. Despite inconsistencies in technological uptake, researchers assert that organizational readiness for eHealth reveals a dichotomy, with rehabilitation staff feeling ready, but practice preparedness being slower [2].

A broader and multidisciplinary perspective is essential to fully capture the complexity of rehabilitation practices [2,27,28]. This study addresses these opportunities for further inquiry. Existing research argues that the challenge in technology application often lies in social behavior around technology use, rather than the technology itself [29,30]. A second critical challenge lies in understanding organizational readiness for the adaptation of digital technologies in rehabilitation settings [2]. This includes the preparedness of both the organization and its members to adapt and integrate these technologies in their daily practices [2]. A third challenge is that existing studies often overlook the broader social context in which these technologies are used in practice [31]. Therefore, there is a need to consider the interconnectedness of health and social care when evaluating how digital technologies shape practices, where digital solutions should aim to lower barriers and facilitate drivers for access to rehabilitation [32].

We applied social practice theory as a theoretical framework to understand how digitalization affects stroke rehabilitation practices at the individual and social levels. Social practice theory, a behavioral theory, explores how people engage with the social world and how it impacts human behavior [33,34]. This applied practice theory is based on work by Reckwitz and incorporates elements from Bourdieu, Giddens, Foucault, and Schatzki [35]. Instead of focusing only on concepts, such as intersubjectivity, embodiment, language, and power within organizations, it incorporates all these concepts and takes a broader perspective by examining what people do in practice and the interconnected elements of individual actions and institutional structures [33]. Practice entails routine behavior involving linked key concepts, such as agents' skills and meanings needed for practices, the things that shape and enable practices, the rules and social norms guiding practices, and the process structure [35]. This theory, which has gained recognition in various fields, provides a useful lens to address the challenges and solutions related to digital inclusivity [36]. It has been used in environmental research [37], education [33], digital information [38], and public and clinical health [31,34,39,40]. The application of this theory in poststroke settings is novel, and this study adds to the limited qualitative research on the changes in practice through digitalization [40]. This study applies the concept of digitalized stroke rehabilitation as a social practice to explore how digitalization influences rehabilitation within a local context. Therefore, this study aims to examine the current influence of digital technologies on stroke rehabilitation practices and consider how these technologies could shape the future landscape of rehabilitation. This is performed through the following research question: How do digital technologies shape the current and perceived future everyday practices of multidisciplinary HCPs in poststroke rehabilitation?

Methods

Design

In this descriptive qualitative study, a theoretical framework of social practice theory (henceforth "practice theory") [35] was applied as a lens to explore and interpret the structures shaping the use of digital technologies in 2 Norwegian stroke rehabilitation settings. Inductively generated thematic themes were mapped onto practice theory concepts. This design allowed for an in-depth exploration and assessment of the relationship between various theoretical key components of the challenges and opportunities of digitalization in stroke rehabilitation, providing insights that could shape the future landscape of rehabilitation.

Settings and Participants

Multidisciplinary HCPs from 2 rehabilitation wards in Norway were invited to participate via email and staff meetings. Recruitment at the first hospital (June 2022 to January 2023) and the second hospital (April to May 2024) involved multidisciplinary HCPs, facilitated by team coordinators. Purposeful sampling was used to recruit appropriate and information-rich data for the in-depth qualitative study. Participants were selected for their direct experience with multidisciplinary specialized stroke rehabilitation. There were no a priori guidelines for the amount or the specific data required, such as regular engagement with new technologies. The study started broadly and narrowed naturally as we became interested in what qualifies as specialized stroke rehabilitation and how technologies shape practices, justifying the theoretical delineation of the sample [41]. Both settings specialize in poststroke rehabilitation, regularly engaging with new technologies. Some technologies have been implemented in practice, while others are still being tested or are under consideration.

The first rehabilitation hospital is a large specialist facility with subacute and poststroke wards, allowing for both short- and long-term rehabilitation. Recruitment occurred in the poststroke ward, with an average of 10 days poststroke at admission and an annual average of 279 stroke survivors admitted. The second ward is a community-based clinic within a hospital and focuses mainly on subacute stroke rehabilitation, with an average of 32 days poststroke at admission and an annual range of 150‐200 stroke survivors admitted. Despite differences in organization, both settings were chosen for their high staffing levels for specialized and intensive multidisciplinary stroke rehabilitation. The first hospital qualifies as performing complex rehabilitation due to the involvement of at least six different professions. Maximum variation purposive sampling [42] was used to recruit 12 experienced HCPs, ensuring diversity in age, sex, and work experience (Table 1). Participants had at least 2 years of experience in poststroke rehabilitation, and both female and male professionals were included. Interviews were conducted on-site at the hospital during appropriate times to minimize interference with daily practices.

XSL•FO
RenderX

Table . Demographic information of the health care professionals who participated in this study (N=12).

ID	Profession	Age (years)	Time working in stroke rehabilitation (years)	Workplace and rehabilitation practice	Current digital technologies used in rehabilitation
1.1	Speech therapist	31	2 - 5	Specialized stroke rehabilitation hospital	Videoconference, digital speech training, video fluoroscopy, and electronic medical records
1.2	Occupational therapist	35	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, pilot project on stroke survivor access to medical records, QR codes to health-related webpages, robotic technology training, digital group motor rehabilitation, and electronic medical records
1.3	Physiotherapist	30	2 - 5	Specialized stroke rehabilitation hospital	Videoconference, virtual reality lab, and electronic medical records
1.4	Occupational therapist	38	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, webinars, virtual reality, and electronic medical records
1.5	Nurse/head of the ward	42	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, virtual reality lab, robotic technology training, and electronic medical records
1.6	Nurse	53	2 - 5	Specialized stroke rehabilitation hospital	Videoconference and electronic medical records
1.7	Physiotherapist	52	25 - 30	Specialized stroke rehabilitation hospital	Videoconference, web-based exercise programs, and electronic medical records
1.8	Social worker	40	5 - 10	Specialized stroke rehabilitation hospital	Videoconference, national health webpages, electronic individual plan, and electronic medical records
1.9	Physician/head physician	57	15 - 20	Specialized stroke rehabilitation hospital	Videoconference, webinars, and electronic medical records
2.1	Physiotherapist	54	<30	Rehabilitation clinic within a hospital	Videoconference, pulse monitoring, digital blood pressure equipment, digital exercise program, and electronic medical records
2.2	Physiotherapist	35	5 - 10	Rehabilitation clinic within a hospital	Virtual reality, videoconference, pulse monitoring, step counter, iPad, dynamometer, and electronic medical records

XSL•FO
RenderX

ID	Profession	Age (years)	Time working in stroke rehabilitation (years)	Workplace and rehabilitation practice	Current digital technologies used in rehabilitation
2.3	Physiotherapist	36	5 - 10	Rehabilitation clinic within a hospital	Web-based exercises, pulse monitoring, dynamometer, biometric equipment, Nintendo Wii, step counter, and electronic medical records

Theory

In this study, practice theory offered a lens through which to examine how different HCPs talked about their actions. Grounded in constructionism, the theory posits that learning and knowledge are collectively produced through shared understandings, practices, and language [33]. Practice theory situated digitalization in poststroke rehabilitation within 2 practice settings, and its terminology served as an analytical advancement for understanding how digital technologies influence poststroke practices. Its key concepts are mapped in Table 2 to clarify the description of each concept as it relates to specialized poststroke rehabilitation settings.

XSL•FO
RenderX

Table . Glossary of practice theory concepts related to poststroke rehabilitation practice.

Concept	Practice theory concepts	Example related to poststroke rehabilitation practice
Agent (including skills)	An agent is a carrier of actions with the necessary skills, and this encompasses both bodily and mental routines. It involves ways of understanding the social world and implicit and explicit know-how [35].	Health care professional agents are connected through specialized hospital processes [43] and are norm-following individuals shaped by professional know-how, who adhere to stroke guidelines and hospital policies relative to their hierarchy at the hospital [40]. They use their minds and bodies to perform routine actions during stroke rehabilitation, such as instructing stroke survivors on walking while encouraging progress. In both contexts, they interact with medically stable stroke survivors past the acute phase, when they are cleared for intense rehabilitation. Occupational therapists, physiotherapists, nurses, medical social workers, and speech and language therapists are central to practices when motor rehabilitation is the main focus [43].
Things	To carry out a practice, agents use things in a certain way, such as physical entities or material conditions, where actions take place. However, writing, printing, and electronic media are also a part of things [35].	Things represent all therapeutic devices, equipment, and health solutions that are necessary to carry out the practice of stroke rehabilitation. Physical elements, such as conversation rooms and rooms for exercise or written material, encompass the things that the agents use or interact with [35].
Rules	Structural rules construct a practice around communication chains and signs. Agents endorse cultural norms and meanings to objects, determining actions through routinized nonsubjective understanding [35].	Health care professionals in poststroke rehabilitation follow recommended interventions from recent rehabilitation guidelines and national clinical practice guidelines for stroke rehabilitation [43]. Communication between professionals often occurs via health information systems, electronic medical records, and IT platforms where rehabilitation plans, progress notes, and patient outcomes are documented and coded according to standardized protocols. Communication with patients or their families typically happens through face-to-face consultations, telephone follow-ups, or digital platforms for remote monitoring, ensuring adherence to evidence-based practices [40].
Process structure	Routinized social practices and processes occur in a sequence over time and through repetition, forming the core of the social structure. Changes in practices arise internally, as practitioners challenge or resist established routines, and externally, as intersecting practices influence each other. Practitioners improvise new actions in novel situations, prompting internal shifts, while external interactions among diverse practices foster adaptation and transformation [35].	In poststroke rehabilitation, health care professionals facilitate a process-centered rehabilitation model for stroke survivors. Although structured, it starts flexibly with initial consultations, followed by planning, a discharge meeting, and ultimately self-care. The process includes multidisciplinary subprocesses like therapeutic examinations and collaboration with other institutions [44].

Data Collection

Individual semistructured interviews were conducted over 2 years, from 2022 to 2024, each lasting 40 to 60 minutes. Information power guided the inclusion of 12 interviews, based on data relevance and richness [45] and when sufficient data “told a coherent, convincing, and comprehensive story” [41]. The interview guide, collaboratively developed by the study’s authors (Multimedia Appendix 1), explored participants’ stroke rehabilitation practices, shared decision-making, and the use of technology and future digital innovations. The inclusion of decision-making as a part of the stroke rehabilitation process was inspired by its person-centered foundation, drawing on literature by scholars, such as Rose et al [46] and Bomhof-Roordink et al [47], in formulating interview guide questions. Interviews were conducted in Norwegian, allowing for probing questions and guided conversations. Audio recordings were securely stored using services for sensitive data. A transcription software assisted with transcription to Norwegian, which was later manually edited for accuracy. Deidentified selected quotations were then translated into English and checked for accuracy by the research team and a secure university-based artificial intelligence (AI) tool

XSL•FO
RenderX

(ChatGPT-4, OpenAI). The research team retained responsibility for final wording and meaning.

Data Analysis

Thematic analysis was chosen to identify, analyze, and interpret patterns across varied meanings [48]. This approach is well-suited as a method for exploring complex phenomena, as it allows for a rich and nuanced understanding of the HCPs' perspectives on how they construct and navigate their practices. The flexibility of thematic analysis enables it to be effectively applied across diverse theoretical frameworks and research questions, making it appropriate for this study's aim of understanding multidisciplinary HCPs' practices. We used reflexive thematic analysis (RTA), specifically developed by Braun and Clarke [48], because it aligns with the interpretative paradigm, emphasizing the active role between researchers and interviewees in generating knowledge and the subjective, situated nature of data interpretation. The in-depth engagement with the data enabled us to explore the social, cultural, and contextual complexities of digitalization in stroke rehabilitation practices, as well as the iterative and reflective process of capturing the underlying meanings and nuances of HCPs' perspectives [48]. The 6 phases of RTA—familiarization, coding, theme generation, review, refinement, and write up—were applied in a back-and-forth manner to accommodate curiosity and interpretation. Familiarization and coding were performed in NVivo (Lumivero), while theme generation, refinement, and write up were conducted in Word (Microsoft Corp). We conducted 2 analyses. The first analysis followed RTA, and the second aligned the findings according to practice theory concepts using Excel (Microsoft Corp). Both descriptive ("semantic," according to Braun and Clarke [48]) and interpretative ("latent") codes were utilized [49] with emphasis on capturing the empirical data during the coding stage. After several rounds of coding, the codes were manually organized into main themes, with practice theory guiding further analysis. The themes from the RTA did not change in meaning, but further descriptions and interpretations related to themes and subthemes were inspired by practice theory as a lens. During theme refinement, an additional analysis mapping qualitative data according to practice theory's 4 concepts was conducted [35] (Multimedia Appendix 2). Practice tasks, forming the final subthemes, were defined. An example of the six phases of RTA, including analyses from both RTA and practice theory, is presented in Multimedia Appendix 3. Key concepts are written in italics in the Results section. Each concept was examined from the points of both the current and future states to understand how digital technologies shape HCPs' practices.

Reflexivity

The researchers' backgrounds acknowledge how the authors' positionality influenced data analysis [48], with several having clinical experience with stroke survivors (AMH-M, CH, AL, and HGH). All authors have academic backgrounds, experience in digital solutions, and multidisciplinary expertise in the fields of rehabilitation within occupational therapy, physiotherapy, nursing, computer science, and medical physics and technology. The first author, a doctoral student and experienced physiotherapy specialist, conducted all interviews, considering subjectivity to be a resource in conducting and analyzing data; maintained a listener's role; and emphasized her role as a researcher rather than an HCP. Still, her clinical background allowed her own experiences to resonate with participants, enabling her to listen and repeat participants' recognizable stories. Reflexivity was considered early in the analysis process by recording initial assumptions after each interview and writing memos and annotations in NVivo during coding to ensure the quality of analytical points and the possibility of reviewing interpretations. A description related to all codes also supported analytical interpretation. To ensure rigor, the authors collaborated in regular meetings and workshops to discuss and enhance understanding, develop comprehensive descriptions, and refine the analysis, strengthening confirmability, coherence, and clear theme boundaries [48,50].

Ethical Considerations

As the study's focus on health service research fell outside the scope of the Health Research Act § 2, ethical approval from The Norwegian National Research Ethics Committee for Medical and Health Research was not required (reference number: 279418). Approval was granted by the Norwegian Agency for Shared Services in Education and Research (reference number: 668899) to ensure privacy protection and adherence to ethical standards. Participants received oral and written information about the study, and written informed consent was obtained prior to the interviews. The first author reiterated consent form details, emphasizing the voluntary nature of participation, confidentiality, and the participants' right to withdraw from the study. Participants did not receive any compensation for their involvement in the study.

Results

Themes and Subthemes

The results are presented in themes and subthemes and exemplified with selected quotes extracted from the analysis based on practice theory.

Theme 1: Outsourcing Information About and to Stroke Survivors: Coordination and Continuity Within and Across Services

Overview

HCPs actively leveraged digital technologies to enable better coordination and continuity of care across services. Digital tools facilitated collaboration not only between HCPs in cross-service collaboration but also with stroke survivors, relatives, and interpreters, forming the first subtheme. Through their practices, they sought to harness digital health to streamline communication and improve rehabilitation outcomes, despite the challenges of sustaining structured coordination. Maintaining continuity after referrals to community services was challenging, requiring constant goal adjustments and balancing person-centered approaches. The rule of involving stroke survivors in their rehabilitation was viewed as a necessary but complex procedure, reflected in the second subtheme "navigating user-centered approaches to rehabilitation for stroke survivors in the digital era."

XSL•FO
RenderX

Subtheme 1.1: Follow-Up and Continuity in Rehabilitation in Collaboration With Stroke Survivors and Between HCPs in Cross-Service Collaboration

In their current practices, HCPs use digital technologies, primarily videoconferencing, for stroke follow-up after discharge, involving additional *agents* such as stroke survivors, community HCPs, interpreters, Norwegian Labor and Welfare Administration (NAV) advisors, and relatives. The importance of HCPs with ICT knowledge as an added skill was highlighted. Coordination and continuity were perceived as crucial owing to the challenges in maintaining structure after referrals to community services, but they often relied on the agency of stroke survivors.

> *Keeping track of who to contact and when is challenging. We've faced logistical issues regarding when the reablement team is supposed to come, and when it conflicts with orthopedic appointments. We lack an overview of these things when the stroke survivors aren't hospitalized and rely on information from stroke survivors or other responsible persons. We do have good contact with some teams, but the organization varies greatly between different districts, making it very inconsistent.* [Informant 2.2]

Increased coordination and continuity aligned with the explicit *rules* of increasing outpatient activity and outsourcing information about stroke survivors. For future practices, an HCP suggested creating a unified individual digital plan between specialist services and communities to address uncertainty about digital progress. This system would provide an overview and ensure that individuals needing long-term, coordinated services receive appropriate care and benefit from effective collaboration to meet their goals. The following comment was made:

> *It would have been quite easy to get started with creating the document and then pass it on to the community. I know that there is no common solution, because it is somewhat up to the community whether they have a digital individual plan, and at least it was a few years back, so I don't know if there is a common solution that all communities can use, but if there is, then there is great potential to look into it.* [Informant 1.8]

Others proposed sharing medical records reporting standardized test results across services and using apps with push-warnings for goal-setting visible to both the HCP and the stroke survivor. Future solutions could involve extended digital therapy skills to include other digital platforms, such as NAV's screening tools, forming a new type of *agent* bridging HCPs and NAV advisors. They also proposed a system between the stroke survivor and health care provider to navigate in-hospital appointments and communicate with staff.

> *Maybe an overview for the patient, in an app, or some kind of guide, showing you this way, you can follow this path to get there. Your physiotherapist is sick today, so your appointment is cancelled. This is somewhat similar to what has already started in other wards. We will also implement it here eventually. It just needs to be adapted to our ward.* [Informant 1.6]

There were societal encouragements for increasing the outpatient activity and outsourcing information to the stroke survivor, enabled by more videoconferencing. Closer follow-up and a more standardized program under the direction of either a general practitioner (GP) or hospitals were mentioned. Enhanced outpatient activity demanded restructuring practices, creating better access to stroke-related information, and introducing a joint communication system between the stroke survivor and the HCP. Digitalizing content on stroke exercise programs, fatigue, rehabilitation experience forums, and cognitive or physical constraints was suggested for improved coordination. However, the complexity arising from multiple medical record systems was highlighted, showing that these systems were used exclusively between HCPs and GPs, leaving out stroke survivors. This suggests the need for integration through a unified platform for visibility and collaboration.

Even though collaboration between HCPs across health services is essential, they expressed a lack of knowledge about available resources for stroke survivors outside their own settings, underscoring the role of stroke survivors in informing their own GPs about rehabilitation opportunities. Paper-based reports were still the main communication tool. Clarifying realistic expectations and outcomes for the stroke survivor's condition helped in establishing routines and maintaining standards and rehabilitation plans. To enable this, HCPs routinely used SMART (specific, measurable, achievable, relevant, and time-bound) goals and educated stroke survivors in SMART goals [51] based on the understanding of being *specific or measurable, achievable or attractive, realistic, and time-bound* (informant 1.5), which gave direction for the future rehabilitation process. However, HCPs expressed that stroke survivors' goals, as revealed in goal-setting conversations and documented in written reports, were often forgotten or changed. The flow of information heavily depended on the stroke survivor's effort to communicate with other HCPs. As a future extension of implementing SMART goal principles in home-based rehabilitation, some HCPs perceived that goals could shift to being more focused on daily routines rather than solely the therapeutic understanding of goals for stroke survivors. For future development, HCPs valued apps with push notifications for goal-setting visible to both HCPs and stroke survivors, and indicated that future practice could consist of increased health-related data and goals. The following comment was made:

> *I wish we could have a bit more follow-up along the way, mostly for the stroke survivor's sake, because it's quite easy to put that binder in a drawer when you get home. Then you have, in a sense, "free time" until they come back. Since we say that goal setting is so important, I wish we could have communication along the way....If they could almost get a push notification through an app.* [Informant 1.2]

Collaboration faced time constraints in cross-service communication, and HCPs highlighted the importance of "setting boundaries for oneself" (informant 1.4) to balance their workload and reduce personal-professional responsibilities.

XSL•FO
RenderX

This perspective was contrasted by another participant, who highlighted differences between specialist health care services and communities in care provision and time allocation, noting the challenges in aligning social expectations.

> *I imagine that in the communities, you might get a couple of assessment visits and then wait a few months for some assistive devices, and you might not get them. You certainly don't get anything weekly and can only address one thing at a time. If we propose ten measures, it might be that you can either spend ten minutes per measure spread over a few weeks, or half of the measures are left out, is the impression I have.* [Informant 1.2]

Conversely, an HCP made the following comment: “We can’t really demand that the community follow up on everything we do either” (informant 1.5). Some HCPs desired more follow-up time in their practices due to a known hesitancy in working with stroke survivors in communities, suggesting more knowledge on stroke rehabilitation and restructuring practices for better access to information and logistics. The potential benefits of a joint communication system include a less fragmented goal process and user involvement in rehabilitation goals for stroke survivors and their relatives.

Subtheme 1.2: Navigating User-Centered Approaches to Rehabilitation for Stroke Survivors in the Digital Era

Stroke survivors were routinely involved in their rehabilitation through discussions at several rehabilitation meetings, aiming to foster independence and proactive behavior. Digital preassessment conversations encouraged participation of stroke survivors in decision-making. The *rule* of involving stroke survivors in rehabilitation was to address their existential needs and was echoed by HCPs as fundamental to the nature of rehabilitation. They emphasized constant goal adjustments and the balance of person-centered care as a common but complex procedure. For instance, cognitive (eg, cognitive fatigue) or cultural determinants could undermine their accountability.

> *But of course, there are quite a few cognitive difficulties that cause significant disruption, which also means that if there are psychiatric and other conditions, stroke survivors may resist some of the services offered or do not take the initiative. It is often said that they are not motivated, as many lose due to cognitive impairment or other conditions, the ability to motivate themselves or make decisions and take initiative. And they often fall behind.* [Informant 1.9]

Relatives were involved, and they were invited to meetings either in person or remotely. HCPs supported stroke survivors’ relatives by providing written material for support, owing to their role of being both a close relative and a caregiver. The challenge of maintaining this balance was highlighted as follows:

> *It’s a challenge for relatives to live with someone who has had a severe stroke and to be both the closest relative, personal trainer, and everything else. They may need to be a home helper, assist with meals, and serve. It’s a significant burden. It is beneficial for them to have someone come in who can take on some of the burden.* [Informant 2.1]

HCPs initiated local opportunities to continue rehabilitation, often referring stroke survivors to other services, acknowledging the complexity of stroke rehabilitation where “most need more rehabilitation” (informant 1.7). Many stroke survivors returned to the same rehabilitation hospital owing to gaps in service availability, such as occupational therapy lacking in some communities. Future digital rehabilitation would require stroke survivors to be more involved in discussions about rehabilitation opportunities throughout the process and to take ownership, which are often omitted in current practices. The following comment was made:

> *[In digital communication with stroke survivors it’s important to] provide alternatives because there’s not just one option; there are often several factors to consider. So, it’s also about involving the stroke survivor more actively in the discussions about their future as well.* [Informant 1.6]

Concerns were expressed about current routines that excluded stroke survivors from discussions regarding alternatives in rehabilitation. HCPs viewed passivity after stroke and a lack of insight and comprehension as barriers to increased agency. However, HCPs observed a growing trend of stroke survivors taking an active role in their treatment. Although this was problematic for most owing to a lack of comprehensive medical and rehabilitation knowledge, there was a clear need for enhanced stroke survivor education. The potential for increased use of digitalized materials for relatives was also suggested. However, some HCPs criticized the assumption that most people are digitally proficient, noting that this responsibility often falls on not only stroke survivors but also their relatives. In situations where stroke survivors struggle to learn new skills, the assistance of relatives was seen as essential. The following comment was made:

> *Not all stroke survivors can use digital solutions, and then [the responsibility] might fall to their relatives. However, not all stroke survivors have resourceful relatives either, creating ambivalence about whether this user group can manage it. I believe that a large proportion can, whether it’s the stroke survivors or their relatives who take responsibility. As shown in videoconference meetings here, where relatives have participated, most manage if they get a brief introduction to what they need to do, how to do it, and when. Many will be able to be part of such an arrangement in the long term, but it is probably a long way to get there.* [Informant 1.3]

Education was needed for stroke survivors with cognitive difficulties or limited digital skills. High demands were placed on those facilitating digital interactions. There were divided opinions on whether the rehabilitation hospital or the community held the main responsibility for digital home exercises and follow-ups. Some claimed that districts were responsible for digital home rehabilitation and follow-up, while others believed that the hospital’s main future task was to spread and utilize their specialist expertise. Nonetheless, the user’s capabilities to become digitally self-sufficient were questioned, and an individualized approach was the top priority. The following comment was made:

> *So, when we say that we aim to provide equitable services to our users and stroke survivors, we must also remember that equitable can mean that some need in-person appointments while others can manage with digital solutions. There are many ahead of us in digital communication, including users and partners, and we are somewhat behind. But many are also far behind us and will never reach where we want to be, so we must remember to cater to the full range of our users. We cannot demand that everyone go digital. If we do, it must be very much based on collaboration. Maybe they can manage an app, but not video. We need to find the right levels.* [Informant 1.2]

Theme 2: Navigating the Ambivalence of Remaining Human Relations in Digital Psychosocial Support Conversations

Multidisciplinary HCPs played a crucial role in providing psychological support through communication, but grappled with the dual promise and limitation of digital technologies in providing support. Psychosocial support remains a multidisciplinary task as a person-centered rehabilitation approach often frames these practices. They engaged in both structured and unstructured conversations to assess existential struggles, self-perception changes following stroke, and quality of life, especially for stroke survivors with multiple cognitive and psychiatric challenges. These conversations often focused on enhancing well-being, and the rehabilitation *process* was about gaining knowledge about the difficulties and the approaches to tackle them.

While digital communication enabled flexibility and accessibility, providing effective support still required building trust and relational depth. HCPs emphasized therapeutic communication skills involving active listening, openness, and sharing, capturing sensitive information in both digital and in-person meetings. Conversations sometimes included topics like depression or suicidality, but often, these were viewed as expressions of an unwanted situation rather than an active plan. The significance of building a relationship with stroke survivors digitally was highlighted through practical strategies to achieve it.

> *Just giving them the time and space to express themselves, joking a lot. I use a lot of humor with the stroke survivors and try to show some understanding of their situation. But I find that if you can mention relevant life experiences, I think it’s a good way to build relationships, to dare to open up a bit, so we're not just sitting there as rigid authority figures, making it a one-sided conversation.* [Informant 1.1]

Some HCPs felt that building relations is harder to achieve digitally and raised concerns about digital health potentially diminishing their therapeutic skills, emphasizing the importance of human relationships over relying solely on technology.

> *I think one should be a bit cautious about saying that all therapists are now redundant. Because now they can just train with an app. I don't believe in that. It’s about human relationships.* [Informant 2.1]

This ambivalence was further shaped by the need to balance technological efficiency with the irreplaceable value of human connection. As professionals sought to navigate the tension between digital and in-person care, building relationships digitally was seen as a prerequisite for further effective support, particularly for follow-up in the later rehabilitation stages.

> *These conversations, right, which many experience as being more difficult- yes, life can become challenging when you return home. The challenges where perhaps there could also have been more follow-up conversations via video, but also something in a group setting. We currently have groups for those who come here, but the number of participants is somewhat limited Many live far away, and it’s difficult for them to travel from home. So, perhaps there should be a type of service specifically directed at stroke survivors and their relatives at a later stage.* [Informant 1.9]

Theme 3: Enhancing Digital Supplements for Assessment and Engagement in Motor Rehabilitation

Overview

HCPs embraced digital tools to enhance assessment, treatment, and patient engagement in motor rehabilitation. These tools provided detailed insights into physical progress, fostering participation and curiosity about their health metrics, as shown in the first subtheme “assessment and monitoring of progress.” HCPs emphasized individualized approaches and explored creative solutions, such as using virtual reality (VR) and augmented reality (AR), to simulate real-life scenarios or design exercises based on videos in home environments, as illustrated in the second subtheme “exercises and rehabilitation programs targeting motor skills.” However, digital tools also presented challenges, such as balancing safety concerns for patients with poor balance and ensuring engagement across generations. While digital tools offered significant potential for self-managed and supported rehabilitation, they were best suited for less complex cases.

Subtheme 3.1: Assessment and Monitoring of Progress

HCPs used various digital tools to enhance the assessment of stroke survivors’ progress and engagement in therapeutic activities. *Things* included pulse monitors, dynamometers with apps to measure and track physical strength and progress, video-based assessments like video fluoroscopy and endoscopic swallowing evaluations, VR tools like “Job Simulator” for cognitive assessments involving kitchen tasks, and step counters that display progress graphs. This fostered curiosity and provided detailed insights into treatment intensity. For instance, pulse monitors, which have been valued as indispensable over the past years, ensured that the pulse rate of stroke survivors remained within a target range.

XSL•FO
RenderX

> *We measure the pulse and check that we are in the zone as much as possible....And after using it for a few years now, I kind of don't understand how I managed without it. It says a lot about the patient's intensity.* [Informant 2.1]

Multidisciplinary assessment was integral to the *processes* of specialized stroke rehabilitation, focusing on holistic understanding and ongoing evaluation, which directly influenced HCPs' replicated behavior and their ability to act on it.

HCPs proposed using digital tools for preadmission consultations to evaluate engagement and proposed involving more community therapists, with local resources aiding preassessment. They suggested video-based visualizations of stroke survivors' home environments and the storage and processing of stroke survivor assessments. Monitoring was perceived as identifying stroke survivors in need of more interventions, but motivation and performance remained challenges. The following comment was made:

> *Monitoring by itself will not be enough, because when you look at the studies that have been done, at least in my experience, stroke survivors, even when they improve in arm function and actually have a higher capacity, don't use their arm as much when they get home. They revert to their usual patterns. Even though capacity increases, performance does not necessarily increase. This is a challenge that the monitor can detect, but the monitor itself will not address the stroke survivor's motivation, incentives, or desire to use the arm, or their understanding that they probably can walk more or use the arm more than they currently do.* [Informant 2.3]

Integrated feedback systems within the monitoring process were seen as potential solutions to motivation challenges. Nonetheless, the question of whether the encouragement of therapists alone is sufficient to enhance arm use or gait remains unresolved, and HCPs are yet to find effective solutions.

Subtheme 3.2: Exercises and Rehabilitation Programs Targeting Motor Skills

Both hospitals implemented gamified rehabilitation. In one hospital, VR was used for intensive balance training or as part of constraint-induced movement therapy for arm training, in addition to being utilized during digital group sessions. The other hospital used biometric equipment with force plates or the Nintendo Wii device in practice and VR for projects. One HCP had experience with asynchronous virtual individual exercises for stroke survivors living at home, when working in the community.

HCPs provided assistance either in person or remotely during the digital programs, but VR sessions required their presence. Language training, in which stroke survivors underwent naming therapy and practiced word recall, was more effective with an HCP present. *Rules* on therapy dose varied, but quantity over quality was mainly emphasized, with the intention of promoting cerebral blood flow and brain reorganization.

Despite efforts to fully utilize each stroke survivor's rehabilitation potential, challenges in maximizing their potential persisted. The following comment was made:

> *You really want that person to have the interventions implemented so they can achieve good functionality. It's very, very sad if there's potential that isn't fulfilled. Then our efforts here feel somewhat wasted. We can figure out and write down what is needed, but if it's not followed up, it feels a bit pointless. I just want people to get as well as they possibly can.* [Informant 1.7]

To enable greater potential for rehabilitation, HCPs proposed future practices involving goal-oriented digital rehabilitation opportunities and motivating exercises. However, they highlighted variations in the types of stroke survivors suitable for undergoing digital motor rehabilitation. As a result, the tasks of future HCPs would include designing motivating exercises while ensuring environmental safety and space. HCPs noted that poor balance and fall risk complicated independent exercise, making approaches to individualized adaptation and treatment effectiveness critical. Simple, self-managed technologies were seen as valuable for catering to different generations. For example, sitting exercises for arm rehabilitation were favored among HCPs owing to safety concerns, especially given that arm rehabilitation tends to receive less focus in stroke rehabilitation across services. Rehabilitation programs could be self-managed or supported by HCPs, depending on the complexity of the stroke survivor's condition. Moreover, they suggested the use of digital apps for low-threshold language training, incorporating features like animations, facial expressions, and tongue position guides. Other HCPs saw potential in creating videos of stroke survivors' homes to design VR or AR rehabilitation programs that simulate real-life challenges and create virtual training apartments, enhancing the transferability of rehabilitation practices to the home setting. Despite some doubts, many saw the potential of digital technologies to improve rehabilitation practices. The following comment was made:

> *Stroke survivors tend to walk very little. Many of them move very little. When they return home, they might train or have follow-up sessions with a physiotherapist maybe two, three days a week, at most. Some might attend a day center, but their overall movement remains limited. In that sense, if there were a low-threshold alternative to encourage them to get up and engage in physical activity would be beneficial.* [Informant 2.2]

Discussion

Principal Findings

This study explored how digital technologies shape the current and future practices of multidisciplinary HCPs in poststroke rehabilitation. According to the results and as linked to practice theory concepts illustrated in Figure 1, the primary agents (HCPs) collaborated with various other agents, including stroke survivors, relatives, interpreters, community HCPs, and NAV advisors, during their practices.

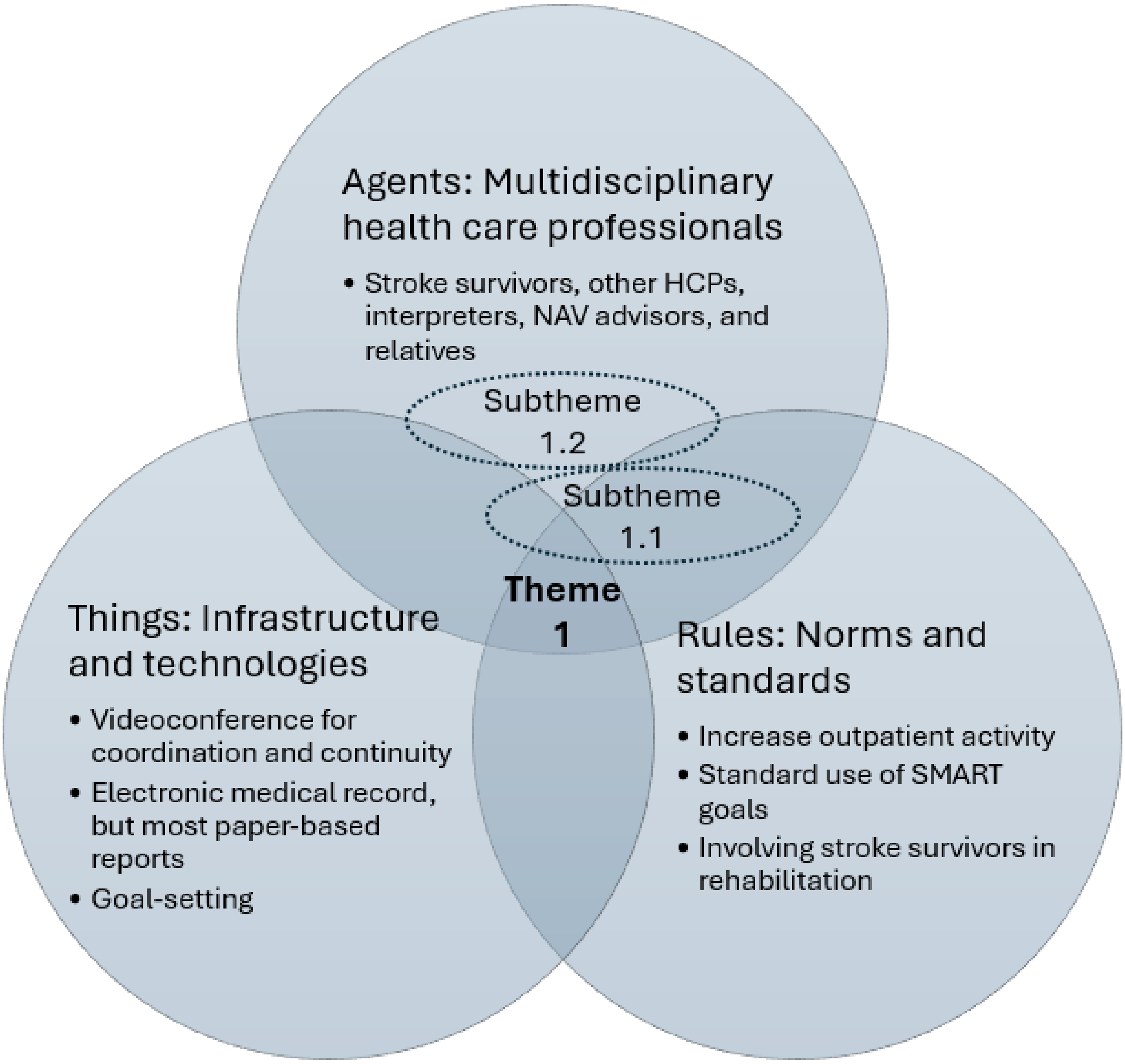
HCPs were influenced by a range of digital technologies, particularly videoconferencing, for rehabilitation coordination and continuity, although paper-based reports remained the primary communication tool across services and with stroke survivors. They used digital instruments for assessment and engagement, including gamified exercises. However, there was ambivalence in using digital technologies for multidisciplinary psychosocial support owing to concerns about their impact on human relationships.

Researchers have highlighted that contextual factors shape technology uptake, including professional discipline, rehabilitation settings, and device features [52]. The use of social practice theory helped us realize and identify how digitalization intersects with particular contextual factors that have not been emphasized in earlier stroke research, which can influence the digitalization of specialized poststroke rehabilitation practices. For example, this theory illuminated the relational dynamics between HCPs and stroke survivors by directing our attention to the required skills and know-how, as outlined in Table 2. Specifically, with regard to theme 2, this perspective highlighted how trust and relational depth are integral to the integration of digital tools. Additionally, this lens revealed how institutional structures, such as differing medical record systems and the digitalization of support conversations, altered established practice routines and were sometimes challenged or resisted by HCPs. Drawing from these core concepts and the results, this study highlights that the primary agents need digital therapy skills, potentially expanding to include knowledge of other digital platforms in cross-service collaboration. Previous research has shown that digital tools, such as apps, can give HCPs a better overview of stroke survivors' rehabilitation processes and can help in follow-up after discharge [53], particularly by using existing digital technologies in transitions between services. However, transitions between specialist health care and community-based services were found to be challenging in this study, requiring familiarization with different practices. Extended use of digital technologies could bridge this gap by enabling therapists to familiarize themselves with different practices during the coordination and continuity of rehabilitation. HCPs highlighted future unified digital systems for sharing medical records and test results across health care services, with goal-setting apps visible to both HCPs and stroke survivors. These could enhance coordination, user involvement in rehabilitation goals, and more active stroke survivor participation and agency in digital follow-up. Despite claims of low technological uptake among therapists [16], other longitudinal studies have found that in a multidisciplinary team, occupational therapists and physiotherapists used digital technologies more frequently across inpatient, outpatient, and community rehabilitation settings. However, studies agree that advancements are needed to match the rehabilitation context, including specific technology training and guidance of implementation [22]. Related responsibilities regarding whether the rehabilitation hospital or the community should lead this practice reorganization were found to be confusing among HCPs in this study. Additionally, they proposed future digital rehabilitation innovations visualizing stroke survivors' home environments and involving more community therapists while addressing safety concerns in stroke rehabilitation exercises. The rehabilitation process usually comprises the following 4 steps: assessment, goal-setting, intervention, and evaluation [54-56]. The results of this study relate to all aspects of rehabilitation and illustrate the complexity of how digitalization is shaping specialized stroke rehabilitation practices.

Consistent with our findings, other qualitative research has indicated that using digital technologies adds value in coordination and continuity, especially where user involvement is lacking. Comprehensive care coordination programs, including home telehealth, can aid stroke survivors and their relatives in managing stroke recovery through weekly real-time video calls and in-home messaging [57], enhancing accountability and self-management [58]. Moreover, tablet-based therapeutic programs can provide alternative access to rehabilitation services and follow-up. Studies have found positive attitudes among other HCPs toward digital apps for sharing stroke rehabilitation information with stroke survivors and their families [53]. However, researchers have highlighted the importance of infrastructure, equipment, space, and support in adapting telerehabilitation [17]. While ICT platforms facilitate communication and effective goal-setting among stakeholders in the rehabilitation process [32], they should also provide feedback to support the interest and progress of individuals with stroke [58,59], which is often limited or not adequately provided in current practices, as observed in this study. However, a range of digital technologies have been found to be integral to self-management for stroke survivors in the rehabilitation process, suggesting that these technologies can prevent therapists from taking over before users are able to do things themselves [20]. Nevertheless, both our findings and the findings of existing research have revealed a distinction between the therapeutic understanding of goal effectiveness and actual digital rehabilitation activities. Resistance to trivial digital technologies offering only engagement has been noted, and there are opportunities to target stroke survivor–specific goals and functional outcomes [53].

XSL•FO
RenderX

Figure 1. Diagram illustrating an example of the first overarching theme (with subthemes) and its relationship with social practice theory concepts. HCP: health care professional; NAV: Norwegian Labor and Welfare Administration; SMART: specific, measurable, achievable, relevant, and time-bound.



The exploration of how digital practices manifest within written documents and in community practices could be valuable, as discharge summaries are important for communication between health care levels and for ensuring continuity of rehabilitation [60]. There is potential for digitalizing written reports, and future research should be performed in this area. Although written reports are widely used, which is consistent with the findings of other studies, they lack the ability to facilitate both adherence and feedback [53]. While this was highlighted by the HCPs, it remains unclear whether therapists' decision-making regarding new technological infrastructure is constrained by institutional requirements of productivity, and if so, to what extent [61]. In this study, paper-based reports contrasted with the incentives of outsourcing digital information. Studies have found that stroke discharge summaries often report on activities of daily living and sensorimotor and general cognitive functions, but they frequently omit stroke survivors' needs and goals [60], and this is likely to increase with further digitalization and user feedback.

Our results contribute to the debate on whether technologies conflict with care and relationship building, emphasizing the need to transfer traditional communication skills, such as empathy, to digital platforms [62]. Claims about jeopardizing interpersonal connections due to being lost in data or a lack of in-person interactions have been made [63-66]. One reason technology may conflict with care is the limited sensory input it provides, highlighting the distinction between the critical social dimensions of humans as social sensors and some technologies as environmental sensors [67]. Despite this limitation, it has been found that using technologies can balance the roles toward a more symmetric relationship compared with the classic asymmetric stroke survivor-caregiver and HCP relationship often found in hospitals [62]. Digital platforms for collaboration, shared decision-making, and data storage can add value to a person-centered rehabilitation process [32], which, despite technology uptake, still drives rehabilitation and remains the rule of practice [52]. Video-based technologies enhance the ability to strengthen confidential and honest communication

XSL•FO
RenderX

[62], but they present ethical [24] and privacy considerations for family members, involve sharing sensitive information, and may invade stroke survivors' homes [68]. However, it is crucial not to lose or diminish face-to-face contact due to increased use of ICT, especially during the initial stage of establishing a relationship with a stroke survivor [12]. In our study, digital technologies interfered with some therapeutic communication skills, especially during communication with individuals having depressive or cognitive impairments, such as aphasia. Nonetheless, being reflexive about the technological environment in which therapists work and its impact on clinical practice has been advocated [69], despite disagreements in rating the importance of factors related to therapists' decision-making in implementing new technologies [61].

This study presented examples of therapists being unable to practice without certain technologies and concerns about losing therapeutic skills or positions owing to technology takeover. Concepts like technological determinism and disruption indicate that technologies shape societal structures and practices [70,71]. Similarly, the idea that new technologies disrupt social practices, norms, and even moral concepts has been revisited in the context of 21st-century technologies, particularly AI [71]. These concepts imply a reduced role for agents in changing practices, which we do not acknowledge, and this rather points back to the distribution of responsibility. In this study, we identified examples of technology that could shift responsibility and change routines. Therefore, new therapeutic communication skills are essential to adapt technology to people [70,72].

Assessments of stroke survivors' functional levels are needed not just in the initial phase but throughout rehabilitation to evaluate necessary adjustments. However, consideration of assessment accuracy must be taken into account [53]. Research supports that assessment is an important feature from the perspective of HCPs when using digital technologies in stroke rehabilitation [57,73]. In this study, digital monitoring enabled closer follow-up on progress and increased stroke survivors' self-management through collaboration with therapists. Shared understanding and priorities can promote technology implementation, including understanding its role relative to conventional therapy, training models, and evidence-based practice [52].

In motor rehabilitation, digital technologies appeared more complicated owing to the principles of intensity and repetition, which often require in-person assistance. While stroke rehabilitation may be too complex for full digitalization, partial implementation is feasible, especially combined with in-person support. Studies have found significant effectiveness of VR for varied outcomes [74], particularly for improving cognitive function compared with conventional care in a hospital setting [75]. Digital technologies for motor rehabilitation might be effective in a hospital setting but might face limitations in home environments [74]. In a previous study, participants continued the same types of exercises at home, transferring skills and preserving performance [75], which HCPs in our study noted as challenging. Additionally, studies have cautioned against the misfit between participants' needs and available technology, emphasizing the importance of coherent content and use to avoid misunderstandings and misuse in practice [32,53]. This approach depends on the stroke trajectory and the practice setting, as individuals may experience different recovery paths and require tailored interventions based on the stroke onset and the environment in which they receive rehabilitation [53]. The diverse perceptions between the 2 practices in this study appear to be related to engaging with people at different stages of stroke onset in the cross-section of hospitalization and home rehabilitation. The consideration of a continuous digital transition that follows stroke trajectories and becomes more self-managed in the chronic phase seems appropriate.

Researching rehabilitation practices raises questions about what is involved in stroke rehabilitation. Motor rehabilitation, as described by Kwakkel et al [43], involves a process that enhances the motor function, activity capacity, and daily performance of stroke survivors, aligning with the International Classification of Functioning, Disability and Health (ICF) framework [76]. Rehabilitation extends beyond restoring function, aiming to maximize independence and facilitate participation in home and community life [77]. Our study highlights varied perceptions of rehabilitation, underlining its multidisciplinary nature and the importance of coordinated and cross-professional contribution [32]. A digitalized rehabilitation service will require changes in the workflows of professionals, fostering cross-disciplinary collaboration to build capacity and ensure greater availability across time and settings. This shift will challenge traditional professional boundaries while breaking down silos between researchers and clinicians [78].

Strengths and Limitations

The data and findings in this study are unique as we applied social practice theory, which allowed us to comprehensively analyze the current and potential future influence of digital technologies for HCPs in poststroke rehabilitation and increased our awareness and reflexive interpretation of the contextual factors influencing these practices. It is acknowledged that the successful implementation of digital technologies depends on not only their inherent qualities but also how they are adopted and integrated into existing practices by HCPs. Maintaining an explorative approach, the findings are broad and consistent with the complexity and comprehensiveness of rehabilitation. A potential limitation of this study is the extended data collection period, which spanned 2 years but was divided into specific recruitment timepoints. While this allowed for a broader capture of experiences and perspectives, shifts in practice or external factors during this period may have influenced the findings. Some digital technologies, such as videoconference systems, VR tools, and webinar tools, were more frequently integrated into HCPs' daily practices at the first site during an earlier recruitment period (2022 - 2023), while step counters were more regularly used at the second site. These differences may reflect evolving practices influenced by project-based initiatives and specific time periods. These temporal variations should be considered when interpreting the results. Representatives of most multidisciplinary HCPs in specialized stroke rehabilitation participated in this study, with the majority recruited from a hospital rather than a community-based clinic. The inclusion of specialist stroke rehabilitation hospital wards provides a more specific context. As the participants emphasized cross-service collaboration and community collaboration, reflections from

XSL•FO
RenderX

community therapists could have added nuances to the findings. No psychologists participated as individual professionals, and only 1 HCP with a background in psychology took part. Future research could investigate the perspectives of psychologists and mental health specialists to expand the theme of psychosocial support and better tailor digital technologies to the psychological and emotional needs of stroke survivors. This study focused on the statements of HCPs regarding how they performed practices, while an ethnographic study would have provided detailed information about the implicit social structures in a specialized stroke rehabilitation practice. The use of RTA allowed for deep engagement with the data, enhancing transparency in the research process. Moreover, the collaboration between experienced researchers and clinicians ensured a robust and informed analysis, strengthening the study’s credibility. Although this study covered current and proposed future technologies, it might have missed emerging innovations or future novel digital tools that the participants might not have been aware of, such as generative AI and robotics, which could potentially revolutionize practices. The study acknowledged challenges in stroke survivor follow-up continuity after referral, suggesting that future research is needed to explore detailed solutions and develop strategies for effective digital integration in rehabilitation.

Conclusion

This study found that HCPs used different technologies during their daily specialist care practices, as well as for the coordination and follow-up of stroke survivors after referral to community services. While all practice themes influenced rehabilitation practices, ambivalence and challenges were noted, particularly in digitalizing multidisciplinary psychosocial support and exercise programs. Despite ambivalence, digital technologies can serve as valuable supplements for stroke recovery, enhancing accessibility, competence, information flow, and quality care. This study identified several factors, such as organizational processes, roles, standards, and rules, that can act as barriers or drivers to implementing digital technologies in practices, emphasizing coherence among content (thing), use (agent), and aim (rules). Systems for sharing medical records and goal-setting apps, which are aimed at enhancing coordination and stroke survivor involvement, were emphasized as future digital technologies that can shape stroke rehabilitation. Viewing technology as a supplement to existing practices, rather than a singular solution for all areas of specialized stroke rehabilitation, offers significant potential for quality improvement. By focusing on leveraging already familiar technologies, specialized stroke rehabilitation integrates agent skills, physical tools, and practice guidelines, forming strategies to enhance the use and effectiveness of digital technologies in the field. Our findings may inform technology developers, health care personnel, and user groups in specialized neurological rehabilitation settings.

Acknowledgments

We would like to thank all the health care professionals who participated in the individual interviews in this study.

Funding

No funding was received for this study.

Data Availability

Complete data sets are not publicly available, as this was not outlined in the ethical approval and informed consent form. The data are available from the corresponding author upon reasonable request.

Authors' Contributions

All authors contributed to this study, gave final approval for the version submitted for publication, and agreed to be accountable for the work as presented.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 17 KB - rehab_v13i1e77753_app1.docx]

Multimedia Appendix 2
Mapping of qualitative data.
[PDF File, 75 KB - rehab_v13i1e77753_app2.pdf]

Multimedia Appendix 3

XSL•FO
RenderX

Six phases of reflexive thematic analysis.
[DOCX File, 21 KB - rehab_v13i1e77753_app3.docx]

References

1. Digital Health Platform Handbook: Building a Digital Information Infrastructure (Infostructure) for Health: World Health Organization; 2020.
2. Touré M, Poissant L, Swaine BR. Assessment of organizational readiness for e-health in a rehabilitation centre. Disabil Rehabil 2012;34(2):167-173. [doi: 10.3109/09638288.2011.591885] [Medline: 21936712]
3. DiCarlo JA, Gheihman G, Lin DJ, 2019 Northeast Cerebrovascular Consortium Conference Stroke Recovery Workshop Participants. Reimagining stroke rehabilitation and recovery across the care continuum: results from a design-thinking workshop to identify challenges and propose solutions. Arch Phys Med Rehabil 2021 Aug;102(8):1645-1657. [doi: 10.1016/j.apmr.2021.01.074] [Medline: 33556351]
4. Khasnabis C, Holloway C, MacLachlan M. The Digital and Assistive Technologies for Ageing initiative: learning from the GATE initiative. Lancet Healthy Longev 2020 Dec;1(3):e94-e95. [doi: 10.1016/S2666-7568(20)30049-0] [Medline: 36094191]
5. Haji Mukhti MI, Ibrahim MI, Tengku Ismail TA, et al. Exploring the need for mobile application in stroke management by informal caregivers: a qualitative study. Int J Environ Res Public Health 2022 Oct 10;19(19):12959. [doi: 10.3390/ijerph191912959] [Medline: 36232257]
6. Yew SQ, Trivedi D, Adanan NIH, Chew BH. Facilitators and barriers to the implementation of digital health technologies in hospital settings in lower- and middle-income countries since the onset of the COVID-19 pandemic: scoping review. J Med Internet Res 2025 Mar 6;27:e63482. [doi: 10.2196/63482] [Medline: 40053793]
7. Fischl C, Asaba E, Nilsson I. Exploring potential in participation mediated by digital technology among older adults. J Occup Sci 2017 Jul 3;24(3):314-326. [doi: 10.1080/14427591.2017.1340905]
8. Veerbeek JM, Langbroek-Amersfoort AC, van Wegen EEH, Meskers CGM, Kwakkel G. Effects of robot-assisted therapy for the upper limb after stroke: a systematic review and meta-analysis. Neurorehabil Neural Repair 2017;31(2):107-121. [doi: 10.1177/1545968316666957] [Medline: 27597165]
9. Lin JJ, Mamykina L, Lindtner S, Delajoux G, Strub HB. Fish’n’Steps: encouraging physical activity with an interactive computer game. Presented at: UbiComp’06: Proceedings of the 8th International Conference on Ubiquitous Computing; Sep 17-21, 2006. [doi: 10.1007/11853565_16]
10. Paul L, Wyke S, Brewster S, et al. Increasing physical activity in stroke survivors using STARFISH, an interactive mobile phone application: a pilot study. Top Stroke Rehabil 2016 Jun;23(3):170-177. [doi: 10.1080/10749357.2015.1122266] [Medline: 27077973]
11. English C, Ceravolo MG, Dorsch S, et al. Telehealth for rehabilitation and recovery after stroke: state of the evidence and future directions. Int J Stroke 2022 Jun;17(5):487-493. [doi: 10.1177/17474930211062480] [Medline: 34983266]
12. Gustavsson M, Ytterberg C, Guidetti S. Exploring future possibilities of using information and communication technology in multidisciplinary rehabilitation after stroke - a grounded theory study. Scand J Occup Ther 2020 Apr;27(3):223-230. [doi: 10.1080/11038128.2019.1666918] [Medline: 31545665]
13. White J, Janssen H, Jordan L, Pollack M. Tablet technology during stroke recovery: a survivor’s perspective. Disabil Rehabil 2015;37(13):1186-1192. [doi: 10.3109/09638288.2014.958620] [Medline: 25212736]
14. Fruhwirth V, Berger L, Gattringer T, et al. Evaluation of a newly developed smartphone app for risk factor management in young patients with ischemic stroke: a pilot study. Front Neurol 2021;12:791545. [doi: 10.3389/fneur.2021.791545] [Medline: 35069420]
15. Taylor E, Fusari G, Darzi A, Jones F. Is a novel digital system for arm and hand rehabilitation suitable for stroke survivors? A qualitative process evaluation of OnTrack. BMJ Open 2023 Aug 18;13(8):e062119. [doi: 10.1136/bmjopen-2022-062119] [Medline: 37597873]
16. Langan J, Subryan H, Nwogu I, Cavuoto L. Reported use of technology in stroke rehabilitation by physical and occupational therapists. Disabil Rehabil Assist Technol 2018 Oct;13(7):641-647. [doi: 10.1080/17483107.2017.1362043] [Medline: 28812386]
17. Munce S, Andreoli A, Bayley M, et al. Clinicians’ experiences of implementing a telerehabilitation toolkit during the COVID-19 pandemic: qualitative descriptive study. JMIR Rehabil Assist Technol 2023 Mar 10;10:e44591. [doi: 10.2196/44591] [Medline: 36897634]
18. Hestetun-Mandrup AM, Toh ZA, Oh HX, He HG, Martinsen ACT, Pikkarainen M. Effectiveness of digital home rehabilitation and supervision for stroke survivors: a systematic review and meta-analysis. Digit Health 2024;10:20552076241256861. [doi: 10.1177/20552076241256861] [Medline: 38832099]
19. Vourganas I, Stankovic V, Stankovic L, Kerr A. Factors that contribute to the use of stroke self-rehabilitation technologies: a review. JMIR Biomed Eng 2019;4(1):e13732. [doi: 10.2196/13732]
20. Hestetun-Mandrup AM, Hamre C, Lund A, Martinsen ACT, He HG, Pikkarainen M. Exploring people with stroke’s perceptions of digital technologies in post-stroke rehabilitation - a qualitative study. Disabil Rehabil 2026 Jan;48(2):359-368. [doi: 10.1080/09638288.2025.2504615] [Medline: 40382688]

XSL•FO
RenderX

21. Willems EMG, Vermeulen J, van Haastregt JCM, Zijlstra GAR. Technologies to improve the participation of stroke patients in their home environment. Disabil Rehabil 2022 Nov;44(23):7116-7126. [doi: 10.1080/09638288.2021.1983041] [Medline: 34607474]
22. Pearce LMN, Pryor J, Redhead J, Sherrington C, Hassett L. Advanced technology in a real-world rehabilitation setting: longitudinal observational study on clinician adoption and implementation. J Med Internet Res 2024 Dec 30;26:e60374. [doi: 10.2196/60374] [Medline: 39753210]
23. Jarvis K, Thetford C, Turck E, Ogley K, Stockley RC. Understanding the barriers and facilitators of digital health technology (DHT) implementation in neurological rehabilitation: an integrative systematic review. Health Serv Insights 2024;17:11786329241229917. [doi: 10.1177/11786329241229917] [Medline: 38690403]
24. Silva GS, Andrade JD. Digital health in stroke: a narrative review. Arq Neuropsiquiatr 2024 Aug;82(8):1-10. [doi: 10.1055/s-0044-1789201] [Medline: 39187259]
25. Borg J, Gustafsson C, Landerdahl Stridsberg S, Zander V. Implementation of welfare technology: a state-of-the-art review of knowledge gaps and research needs. Disabil Rehabil Assist Technol 2023 Feb;18(2):227-239. [doi: 10.1080/17483107.2022.2120104] [Medline: 36103349]
26. Jette DU, Halbert J, Iverson C, Miceli E, Shah P. Use of standardized outcome measures in physical therapist practice: perceptions and applications. Phys Ther 2009 Feb;89(2):125-135. [doi: 10.2522/ptj.20080234] [Medline: 19074618]
27. Brouns B, Meesters JJL, Wentink MM, et al. Why the uptake of eRehabilitation programs in stroke care is so difficult-a focus group study in the Netherlands. Implement Sci 2018 Oct 29;13(1):133. [doi: 10.1186/s13012-018-0827-5] [Medline: 30373611]
28. Weiner BJ, Amick H, Lee SYD. Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields. Med Care Res Rev 2008 Aug;65(4):379-436. [doi: 10.1177/1077558708317802] [Medline: 18511812]
29. Wears RL, Berg M. Computer technology and clinical work: still waiting for Godot. JAMA 2005 Mar 9;293(10):1261-1263. [doi: 10.1001/jama.293.10.1261] [Medline: 15755949]
30. Dyb K, Berntsen GR, Kvam L. Adopt, adapt, or abandon technology-supported person-centred care initiatives: healthcare providers' beliefs matter. BMC Health Serv Res 2021 Mar 17;21(1):240. [doi: 10.1186/s12913-021-06262-1] [Medline: 33731078]
31. Frost J, Wingham J, Britten N, et al. The value of social practice theory for implementation science: learning from a theory-based mixed methods process evaluation of a randomised controlled trial. BMC Med Res Methodol 2020 Jul 6;20(1):181. [doi: 10.1186/s12874-020-01060-5] [Medline: 32631324]
32. Gåsvær JI, Jepsen R, Heldal I, Sudmann T. Supporting collaboration in rehabilitation trajectories with information and communication technologies: scoping review. JMIR Rehabil Assist Technol 2023 Jul 11;10(1):e46408. [doi: 10.2196/46408] [Medline: 37432715]
33. Lloyd A. Information Literacy Landscapes: Information Literacy in Education, Workplace and Everyday Contexts: Elsevier; 2010.
34. Harries T, Rettie R, Gabe J. Shedding new light on the (in)compatibility of chronic disease management with everyday life - social practice theory, mobile technologies and the interwoven time-spaces of teenage life. Sociol Health Illn 2019 Sep;41(7):1396-1409. [doi: 10.1111/1467-9566.12952] [Medline: 31124176]
35. Reckwitz A. Toward a theory of social practices: a development in culturalist theorizing. Eur J Soc Theory 2002;5(2):243-263. [doi: 10.1177/13684310222225432]
36. Tan KSY, Chan CML. Unequal access: applying Bourdieu's practice theory to illuminate the challenges of ICT use among senior citizens in Singapore. J Aging Stud 2018 Dec;47:123-131. [doi: 10.1016/j.jaging.2018.04.002] [Medline: 30447865]
37. Galvin R. Humans and stuff: interweaving social and physical science in energy policy research. Energy Res Soc Sci 2017 Apr;26:98-102. [doi: 10.1016/j.erss.2017.01.012]
38. Halford S, Savage M. Reconceptualizing digital social inequality. Information, Communication & Society 2010 Oct;13(7):937-955. [doi: 10.1080/1369118X.2010.499956]
39. Guni JN, Wanjala SW, Manguro G, et al. Using social practice theory in measuring perceived stigma among female sex workers in Mombasa, Kenya. BMC Public Health 2023 May 26;23(1):972. [doi: 10.1186/s12889-023-15809-2] [Medline: 37237349]
40. Niemelä R, Pikkarainen M, Ervasti M, Reponen J. The change of pediatric surgery practice due to the emergence of connected health technologies. Technol Forecast Soc Change 2019 Sep;146:352-365. [doi: 10.1016/j.techfore.2019.06.001]
41. Morse JM. The significance of saturation. Qual Health Res 1995 May;5(2):147-149. [doi: 10.1177/104973239500500201]
42. Gill SL. Qualitative sampling methods. J Hum Lact 2020 Nov;36(4):579-581. [doi: 10.1177/0890334420949218] [Medline: 32813616]
43. Kwakkel G, Stinear C, Essers B, et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. Eur Stroke J 2023 Dec;8(4):880-894. [doi: 10.1177/23969873231191304] [Medline: 37548025]
44. Langhorne PF, Coupar FB, Pollock AP. Motor recovery after stroke: a systematic review. Lancet Neurol 2009 Aug;8(8):741-754. [doi: 10.1016/S1474-4422(09)70150-4] [Medline: 19608100]

XSL•FO
RenderX

45. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
46. Rose A, Rosewilliam S, Soundy A. Shared decision making within goal setting in rehabilitation settings: a systematic review. Patient Educ Couns 2017 Jan;100(1):65-75. [doi: 10.1016/j.pec.2016.07.030] [Medline: 27486052]
47. Bomhof-Roordink H, Gärtner FR, Stiggelbout AM, Pieterse AH. Key components of shared decision making models: a systematic review. BMJ Open 2019 Dec 17;9(12):e031763. [doi: 10.1136/bmjopen-2019-031763] [Medline: 31852700]
48. Braun V, Clarke V. Thematic analysis: a practical guide. QMiP Bulletin 2022;1(33):46-50. [doi: 10.53841/bpsqmip.2022.1.33.46]
49. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qualitative Psychology 2022;9(1):3-26. [doi: 10.1037/qup0000196]
50. Braun V, Clarke V. Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: a review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). Palliat Med 2024 Jun;38(6):608-616. [doi: 10.1177/02692163241234800] [Medline: 38469804]
51. Bovend’Eerdt TJH, Botell RE, Wade DT. Writing SMART rehabilitation goals and achieving goal attainment scaling: a practical guide. Clin Rehabil 2009 Apr;23(4):352-361. [doi: 10.1177/0269215508101741] [Medline: 19237435]
52. Pearce LMN, Hassett L, Sherrington C, Pryor J. Human interactions remain at the heart of rehabilitation with advanced technology: a practice-embedded longitudinal qualitative study with allied health clinicians. J Neuroeng Rehabil 2025 Mar 7;22(1):52. [doi: 10.1186/s12984-025-01576-1] [Medline: 40050924]
53. Epalte K, Grjadovojs A, Bērziņa G. Use of the digital assistant Vigo in the home environment for stroke recovery: focus group discussion with specialists working in neurorehabilitation. JMIR Rehabil Assist Technol 2023 Apr 14;10(1):e44285. [doi: 10.2196/44285] [Medline: 37058334]
54. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet 2011 May 14;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5] [Medline: 21571152]
55. Wade DT. Describing rehabilitation interventions. Clin Rehabil 2005 Dec;19(8):811-818. [doi: 10.1191/0269215505cr923ed] [Medline: 16323380]
56. Lexell J, Brogårdh C. The use of ICF in the neurorehabilitation process. NeuroRehabilitation 2015;36(1):5-9. [doi: 10.3233/NRE-141184] [Medline: 25547759]
57. Lutz BJ, Chumbler NR, Roland K. Care coordination/home-telehealth for veterans with stroke and their caregivers: addressing an unmet need. Top Stroke Rehabil 2007;14(2):32-42. [doi: 10.1310/tsr1402-32] [Medline: 17517572]
58. Jayasree-Krishnan V, Ghosh S, Palumbo A, Kapila V, Raghavan P. Developing a framework for designing and deploying technology-assisted rehabilitation after stroke: a qualitative study. Am J Phys Med Rehabil 2021 Aug 1;100(8):774-779. [doi: 10.1097/PHM.0000000000001634] [Medline: 33141773]
59. Peretti A, Amenta F, Tayebati SK, Nittari G, Mahdi SS. Telerehabilitation: review of the state-of-the-art and areas of application. JMIR Rehabil Assist Technol 2017 Jul 21;4(2):e7. [doi: 10.2196/rehab.7511] [Medline: 28733271]
60. Solbakken LM, Sundseth A, Langhammer B, Brovold T. Are physiotherapists and occupational therapists following the guidelines for discharge summary?-an analysis of the content of physiotherapists’ and occupational therapists’ discharge summaries and their adherence to stroke guideline recommendations. PLoS ONE 2024;19(9):e0308039. [doi: 10.1371/journal.pone.0308039] [Medline: 39226253]
61. Chen CC, Bode RK. Factors influencing therapists’ decision-making in the acceptance of new technology devices in stroke rehabilitation. Am J Phys Med Rehabil 2011 May;90(5):415-425. [doi: 10.1097/PHM.0b013e318214f5d8] [Medline: 21765257]
62. Solli H, Hvalvik S. Nurses striving to provide caregiver with excellent support and care at a distance: a qualitative study. BMC Health Serv Res 2019 Nov 27;19(1):893. [doi: 10.1186/s12913-019-4740-7] [Medline: 31771566]
63. Narbutaitienė J, Björklund Carlstedt A, Fischl C. Stroke survivors’ experiences and meaning of digital technology in daily life: a phenomenological study. Disabil Rehabil Assist Technol 2024 May;19(4):1334-1342. [doi: 10.1080/17483107.2023.2174605] [Medline: 36740734]
64. Walker RC, Tong A, Howard K, Palmer SC. Patient expectations and experiences of remote monitoring for chronic diseases: systematic review and thematic synthesis of qualitative studies. Int J Med Inform 2019 Apr;124:78-85. [doi: 10.1016/j.ijmedinf.2019.01.013] [Medline: 30784430]
65. Fiske A, Buyx A, Prainsack B. The double-edged sword of digital self-care: physician perspectives from Northern Germany. Soc Sci Med 2020 Sep;260:113174. [doi: 10.1016/j.socscimed.2020.113174] [Medline: 32659512]
66. Lucivero F. Lessons about so-called “difficult” patients from the UK controversy over patient access to electronic health records. AMA J Ethics 2017 Apr 1;19(4):374-380. [doi: 10.1001/journalofethics.2017.19.4.stas1-1704] [Medline: 28430572]
67. Nectoux S, Magee L, Soldatic K. Sensing technologies, digital inclusion, and disability diversity. J Comput Mediat Commun 2023 Aug 4;28(5). [doi: 10.1093/jcmc/zmad026]
68. Wherton J, Greenhalgh T, Hughes G, Shaw SE. The role of information infrastructures in scaling up video consultations during COVID-19: mixed methods case study into opportunity, disruption, and exposure. J Med Internet Res 2022 Nov 10;24(11):e42431. [doi: 10.2196/42431] [Medline: 36282978]

XSL•FO
RenderX

69. Killingback C, Naylor J. Emerging technologies in healthcare: interpersonal and client-based perspectives. In: Hayre CM, Muller D, Scherer M, Hackett PMW, Gordley-Smith A, editors. Emerging Technologies in Healthcare: CRC Press; 2024:1-14. [doi: 10.1201/9781003272786]
70. Hallström J. Embodying the past, designing the future: technological determinism reconsidered in technology education. Int J Technol Des Educ 2022 Mar;32(1):17-31. [doi: 10.1007/s10798-020-09600-2]
71. Löhr G. Conceptual disruption and 21st century technologies: a framework. Technol Soc 2023 Aug;74:102327. [doi: 10.1016/j.techsoc.2023.102327]
72. Lindberg J, Kvist E, Lindgren S. The ongoing and collective character of digital care for older people: moving beyond techno-determinism in government policy. J Technol Hum Serv 2022 Oct 2;40(4):357-378. [doi: 10.1080/15228835.2022.2144588]
73. Marwaa MN, Kristensen HK, Guidetti S, Ytterberg C. Physiotherapists' and occupational therapists' perspectives on information and communication technology in stroke rehabilitation. PLoS ONE 2020;15(8):e0236831. [doi: 10.1371/journal.pone.0236831] [Medline: 32857781]
74. Demeco A, Zola L, Frizziero A, et al. Immersive virtual reality in post-stroke rehabilitation: a systematic review. Sensors (Basel) 2023 Feb 3;23(3):1712. [doi: 10.3390/s23031712] [Medline: 36772757]
75. Torrisi M, Maresca G, De Cola MC, et al. Using telerehabilitation to improve cognitive function in post-stroke survivors: is this the time for the continuity of care? Int J Rehabil Res 2019 Dec;42(4):344-351. [doi: 10.1097/MRR.0000000000000369] [Medline: 31464812]
76. McDougall J, Wright V, Rosenbaum P. The ICF model of functioning and disability: incorporating quality of life and human development. Dev Neurorehabil 2010;13(3):204-211. [doi: 10.3109/17518421003620525] [Medline: 20450470]
77. Stinear CM, Lang CE, Zeiler S, Byblow WD. Advances and challenges in stroke rehabilitation. Lancet Neurol 2020 Apr;19(4):348-360. [doi: 10.1016/S1474-4422(19)30415-6] [Medline: 32004440]
78. Bernhardt J, Hayward KS, Kwakkel G, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the stroke recovery and rehabilitation roundtable taskforce. Neurorehabil Neural Repair 2017 Sep;31(9):793-799. [doi: 10.1177/1545968317732668] [Medline: 28934920]

Abbreviations

AI: artificial intelligence
AR: augmented reality
GP: general practitioner
HCP: health care professional
ICT: information and communication technology
NAV: Norwegian Labor and Welfare Administration
RTA: reflexive thematic analysis
SMART: specific, measurable, achievable, relevant, and time-bound
VR: virtual reality
WHO: World Health Organization

Edited by A Mastropietro; submitted 19.May.2025; peer-reviewed by H Sadeghsalehi, J Hao, NA Merriman; revised version received 16.Dec.2025; accepted 17.Dec.2025; published 17.Feb.2026.

Please cite as:
Hestetun-Mandrup AM, Hamre C, Lund A, Martinsen ACT, He HG, Pikkarainen M
Exploring the Influence of Digitalization on Multidisciplinary Poststroke Rehabilitation Practice: Qualitative Study
JMIR Rehabil Assist Technol 2026;13:e77753
URL: https://rehab.jmir.org/2026/1/e77753
doi:10.2196/77753

© Ann Marie Hestetun-Mandrup, Charlotta Hamre, Anne Lund, Anne Catrine Trægde Martinsen, Hong-Gu He, Minna Pikkarainen. Originally published in JMIR Rehabilitation and Assistive Technology (https://rehab.jmir.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Rehabilitation and Assistive Technology, is properly cited. The complete bibliographic information, a link to the original publication on https://rehab.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX