

User-Centered Development of a Digital Platform to Provide Neurorehabilitation Patients With Feedback on Health Data

Nadine Domnik, Katarzyna Krasnopolska, Ramona Sylvester, Jens Bansi, Jaeyong Song, Roman Gonzenbach, Olivier Lambercy

Submitted to: Journal of Medical Internet Research
on: September 30, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 42

 Multimedia Appendixes 43

 Multimedia Appendix 1..... 43



User-Centered Development of a Digital Platform to Provide Neurorehabilitation Patients With Feedback on Health Data

Nadine Domnik¹ MSc; Katarzyna Krasnopolska¹; Ramona Sylvester²; Jens Bansi^{2, 3}; Jaeyong Song¹; Roman Gonzenbach²; Olivier Lambercy^{1, 4}

¹ Rehabilitation Engineering Laboratory ETH Zurich Zurich CH

²Neurology Department Kliniken Valens Valens CH

³Department of Health OST - Eastern Switzerland University of Applied Sciences St. Gallen CH

⁴ Future Health Technologies Programme Singapore-ETH Centre Singapore SG

Corresponding Author:

Nadine Domnik MSc

Rehabilitation Engineering Laboratory
ETH Zurich
Gloriastrasse 37/39
Zurich
CH

Abstract

Background: An increasing amount of digital health data is being collected across rehabilitation settings - but its integration into routine clinical practice remains limited, despite its potential to motivate patients or inform clinical decision-making.

Objective: This study presents a mobile application that delivers structured, interpretable feedback from clinical and technology-based assessments to neurorehabilitation patients and healthcare practitioners (HCPs).

Methods: Developed through an iterative co-design process with 17 persons after a neurological injury and 15 HCPs, the app integrates common conventional assessment outcomes, and digital health metrics from the Virtual Peg Insertion Test - as a proof of concept for the integration of technology-based assessments. Stakeholders evaluated feedback types, visualization formats, and usability across three structured sessions.

Results: Progress bars and simple illustrations were consistently preferred for clarity and engagement. The resulting mobile application prototype demonstrated high usability (System Usability Scale: patients = 93.6; clinicians = 80.9) and enhanced the perceived value of feedback in supporting patient motivation and clinical decision-making, compared to perceptions reported before users were shown how such feedback could be presented.

Conclusions: These findings underline the potential of digital tools to bridge data interpretation gaps in neurorehabilitation and support value-based, personalized care delivery.

(JMIR Preprints 30/09/2025:85072)

DOI: <https://doi.org/10.2196/preprints.85072>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <https://preprints.jmir.org/preprint/85072>



Original Manuscript

User-Centered Development of a Digital Platform to Provide Neurorehabilitation Patients with Feedback on Health Data

Nadine Domnik^{1*}, Katarzyna Krasnopolska¹, Ramona Sylvester³, Jens Bansi^{3,4}, Jaeyong Song¹,
Roman Gonzenbach³,
Olivier Lambercy^{1,2}

^{1*}Rehabilitation Engineering Laboratory, ETH Zurich, Gloriastrasse 37/39, Zurich, 8092, Switzerland.

²Singapore-ETH Centre, Future Health Technologies Programme, CREATE Campus, Singapore, 138602, Singapore.

³Neurology Department, Kliniken Valens, Taminaplatz 1, Valens, 7317, Switzerland.

⁴Department of Health, OST - Eastern Switzerland University of Applied Sciences, Rosenbergstrasse 59, St. Gallen, 9000, Switzerland.

*Corresponding author(s). E-mail(s): nadine.domnik@hest.ethz.ch; Contributing authors: ramona.sylvester@kliniken-valens.ch; jens.bansi@kliniken-valens.ch; jay.song@hest.ethz.ch; roman.gonzenbach@kliniken-valens.ch; olivier.lambercy@hest.ethz.ch;

Abstract

Background: An increasing amount of digital health data is being collected across rehabilitation settings - but its integration into routine clinical practice remains limited, despite its potential to motivate patients or inform clinical decision-making.

Objective: This study presents a mobile application that delivers structured, interpretable feedback from clinical and technology-based assessments to neurorehabilitation patients and healthcare practitioners (HCPs).

Methods: Developed through an iterative co-design process with 17 persons after a neurological injury and 15 HCPs, the app integrates common conventional assessment outcomes, and digital health metrics from the Virtual Peg Insertion Test - as a proof of concept for the integration of technology-based assessment. Stakeholders evaluated feedback types, visualization formats, and usability across three structured sessions.

Results: Progress bars and simple illustrations were consistently preferred for clarity and engagement. The resulting mobile application prototype demonstrated high usability (System Usability Scale: patients = 93.6; clinicians = 80.9) and enhanced the perceived value of feedback in supporting patient motivation and clinical decision-making, compared to perceptions reported before users were shown how such feedback could be presented.

Conclusions: These findings underline the potential of digital tools to bridge data interpretation gaps in neurorehabilitation and support value-based, personalized care delivery.

Keywords: Neurorehabilitation; digital health; user-centered design; clinical decision-making; health data visualization; mobile health (mHealth); value-based health care (VBHC); digital feedback application

1 Introduction

During neurorehabilitation, both conventional and technology-based assessments are essential for diagnosing impairments, tracking progress, and tailoring treatments¹⁻⁵. Standardized tools such as the Nine Hole Peg Test⁶, the Timed Up and Go⁷, and the Functional Independence Measure⁸ are widely used⁹. Although they capture different domains of the International Classification of Functioning, Disability and Health (ICF) framework¹⁰, they are

subjective, time-consuming, and lack sensitivity^{11–13}.

Technology-based assessments address these limitations through automated data collection, improved objectivity, and greater sensitivity. Examples include (markerless) motion capture^{14–16}, wearable sensors¹⁷, and robotic tools such as the Kinesio- logical Instrument for Normal and Altered Reaching Movement (KINARM)^{18,19} or the Virtual Peg Insertion Test (VPIT)^{20,21}. However, their outputs can be abstract, poorly validated, and difficult to integrate into clinical workflows²².

Despite their potential to support the implementation of precision medicine, integrating both conventional and digital assessments systematically into clinical routine decision-making remains challenging. In many cases, assessment outcomes are difficult to synthesize and communicate effectively during patient consultations, which can limit their ability to enhance patient motivation, satisfaction, and adherence to treatment. This limited adoption is largely driven by the high effort required to collect, summarize and interpret the data, which can outweigh the perceived benefit in many clinical contexts. Moreover, while the level of detail offered by technology-based assessments is valuable in research^{23–25}, its clinical relevance remains uncertain, and there is a lack of clear guidelines for selecting or applying metrics in practice²⁶.

Regular and well-communicated feedback on assessment outcomes has been shown to enhance patient motivation^{27,28}, support engagement in shared decision-making²⁹, and improve rehabilitation outcomes^{30,31}. Likewise, transparent and interpretable data can support clinical decisions that are often otherwise driven by prior experience and/or subjective judgment³². To leverage the full potential of both conventional and technology-based assessments, their results must be made accessible, interpretable, and actionable for both patients and healthcare practitioners (HCPs). Achieving this requires a systematic and collaborative approach to identify clinically meaningful outcome measures and establish health communication strategies that enhance their practical use in clinical routine^{5,33,34}.

Digital applications have the potential to bridge the gap between researchers, HCPs, and patients by enabling access to individualized health data and translating complex information into intuitive, user-friendly formats. By streamlining data collection and interpretation, these tools have the potential to support more frequent assessments and promote continuous communication of outcomes throughout the rehabilitation process. In recent years, the number of released health apps in areas such as mental health^{35,36},

consumer health³⁷, and fitness^{38,39}, increased rapidly. While many of these apps lack sufficient validation or fail to demonstrate clear benefits, others show strong potential to increase patient motivation and adherence^{35,36,38,40}. A key factor in making digital health tools effective and engaging for both patients and HCPs is the active involvement of end-users in the development process⁴¹⁻⁴⁶. Applying user-centered design principles helps ensure that these tools are better aligned with the real needs of all stakeholders^{3-5,40,41,46-50}. However, a key limitation is that many existing digital health applications are developed primarily for the consumer market rather than for integration into clinical workflows. As a result, their design often overlooks the specific requirements of clinical environments, thereby limiting their adoption and utility in routine care. This gap is particularly evident in the neurorehabilitation field, where existing digital health records still lack effective features for summarizing and visualizing health-related data. Consequently, both HCPs and patients are often left without accessible, user-friendly tools to monitor and understand rehabilitation progress.

In this paper we describe a three-step user-centered design process for developing a digital feedback application that displays digital health metrics from conventional and technology-based assessments to patients and HCPs along the neurorehabilitation journey. As first step, we defined the need and requirements and evaluated expectations for such feedback together with HCPs. Through iterative rounds of discussions and structured interviews with both neurological patients and HCPs, we identified an optimal type and amount of information, along with the appropriate visualizations to display health data in an intuitive and engaging way. Our evaluation focused on three key aspects: informativeness, understandability, and visual appeal. Building on this, we developed a proof-of-concept application that visualizes typical rehabilitation data collected during clinical routine assessments, including digital health metrics from the VPIT, and evaluated its usability. This work is important as it provides insights from end-users - patients and HCPs - on how to meaningfully present assessment outcomes and include digital health metrics into neurorehabilitation settings, which could open new avenues for empowering patients with their health data, and support HCPs in decision-making towards precision medicine.

2 Methods

2.1 Conceptual approach

We employed a structured co-design process to develop a digital application aimed at enhancing patient motivation and supporting clinical decision-making in neurorehabilitation. The initial requirements were identified through a multidisciplinary roundtable involving researchers from ETH Zurich and clinical staff from Kliniken Valens, including the heads of neurology, research, and therapy, as well as therapists with different backgrounds. The discussions focused on defining clinically relevant assessments, selecting the appropriate platform, and ensuring integration into routine care. The development process included three iterative feedback sessions with HCPs, designed to evaluate: (i) the content and visualization of patient information; (ii) the integration of digital health metrics; and (iii) overall usability. Between sessions, targeted follow-up meetings with researchers and HCPs enabled systematic incorporation of feedback into subsequent development stages (Figure 1Error: Reference source not found).

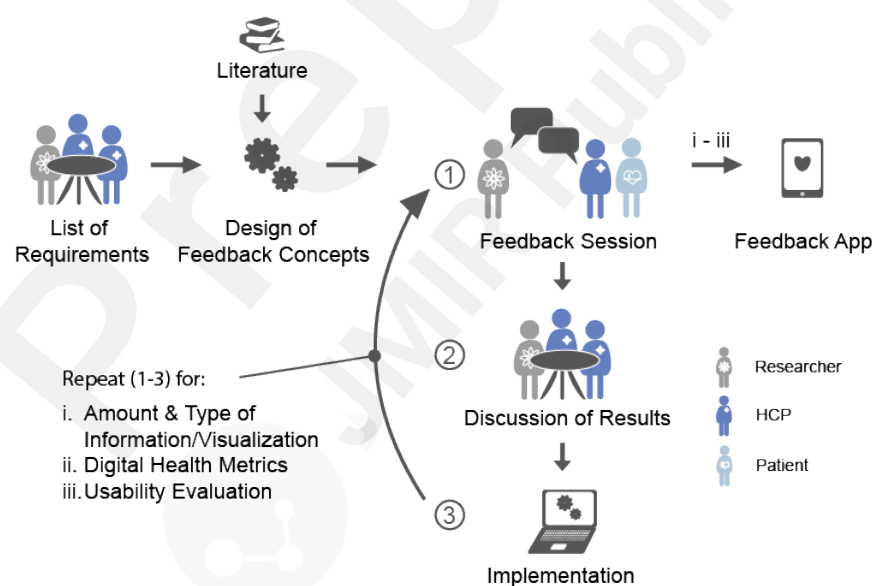


Figure 1: Overview of the co-design process for a digital feedback application for patients and healthcare practitioners (HCPs). The illustration shows a three-step iterative approach guided by co-design principles. Abbreviations: HCP, healthcare practitioner.

To ensure clinical relevance and align with current clinical practice, the application was designed to include the full battery of conventional assessments routinely conducted at admission and discharge at the Kliniken Valens. To explore integration of digital health data, we included the VPIT, a well validated tool for assessing upper limb function. The

metrics derived from the VPIT are representative of the digital health data typically obtained from wearable sensors or robotic devices used in neurorehabilitation assessments. We implemented the application as a mobile solution for iOS, with an optional future compatibility across additional operating systems. Currently, the app is intended for use on tablets, as this offers greater practicality during patient consultations compared to stationary terminals.

2.2 Participants

A total of 17 neurological inpatients diagnosed with Multiple Sclerosis (MS) and 15 distinct HCPs from the Kliniken Valens took part in one of three structured feedback sessions. Some HCPs participated in more than one session, resulting in a total of 26 HCP participations. An overview of participant demographics and roles is provided in Table 1. HCPs included physiotherapists (PT), occupational therapists (OT), sports therapists/scientists (ST), and one participant with a Health Science and Technology (HST) background. Participant demographics, including age and gender, varied across sessions.

Table 1: Participant demographics across the three structured feedback sessions. The table summarizes the number of participants, their roles, and basic demographic characteristics for both patients and HCPs involved in each of the three structured feedback sessions. Abbreviations: PT, physiotherapist; ST, sports therapist/scientist; OT, occupational therapist; HST Health Science and Technology degree; SD, standard deviation; F, female; M, male.

Sessio n	Group	N	Age (mean $\pm$ SD)	Gender (F/M)	Role
1	Patient	5	40.8 $\pm$ 15.6 years	2/3	—
1	s HCPs	8	35.8 $\pm$ 10.4 years	5/3	PT (n=3), ST (n=2), OT (n=2), HST (n=1)
2	Patient	6	56.5 $\pm$ 7.8 years	1/5	—
2	s HCPs	9	35.2 $\pm$ 9.3 years	3/6	PT (n=5), ST (n=2), OT (n=2)
3	Patient	6	46.1 $\pm$ 14.8 years	5/1	—
3	s HCPs	8	33.1 $\pm$ 10.3 years	1/7	PT (n=3), ST (n=2), OT (n=2), HST (n=1)

2.3 Conventional assessments

The conventional assessments at Kliniken Valens were grouped into six categories - endurance, mobility, strength, coordination, speed, and global health/function - to support clearer interpretation in the app, Supplementary Table S1. These categories were based on the ICF framework and refined with HCP input.

2.4 Virtual Peg Insertion Test (VPIT)

The VPIT is a technology-aided assessment of upper limb function that integrates a haptic device with a virtual environment ²⁰. The task involves goal-directed reaching, in which users are instructed to insert nine virtual pegs into nine holes displayed on a board as quickly and precisely as possible. During task execution, kinetic and kinematic data is continuously recorded. From this data, ten digital health metrics are computed to quantify various aspects of movement quality, including smoothness, efficiency, accuracy, speed, and grip force control ²¹. These metrics enable sensitive and objective characterization of upper

limb impairments, and have been validated in diverse patient populations, including e.g. individuals with MS^{51,52}, stroke⁵³, and ataxia^{54,55}. Importantly, the VPIT served as a model for integrating digital health metrics into clinical feedback systems, offering a template for incorporating validated, objective assessments into user-facing tools. As such, it offers a clinically relevant and generalizable use case for exploring how complex sensor-based outputs can be effectively visualized and communicated to both patients and HCPs. Its integration into the co-designed digital feedback application supports our broader goal of making digital health data more accessible and actionable, fostering patient engagement and shared decision-making in neurorehabilitation. In collaboration with HCPs from Kliniken Valens, the VPIT metrics were grouped into the same five functional categories as the conventional assessments. An overview of the ten core VPIT metrics, including their descriptions and associated categories, is provided in Supplementary Table S2.

2.5 Structured feedback sessions

We conducted three structured feedback sessions to iteratively evaluate key aspects of the app design, including information presentation, metric representation, and overall usability. Each session followed a standardized procedure with two participants - either HCPs or patients - participating simultaneously. To minimize bias, all participants first received a neutral written introduction outlining the session's purpose. In sessions two and three, this was supplemented by a short video and an explanation of the VPIT.

Participants were then shown a series of feedback concepts and completed structured questionnaires. These included Likert-scale items (1 = strongly disagree to 5 = strongly agree)⁵⁶, open-ended questions, and preference rankings of the concepts. Concepts were rated on motivational value, visual appeal, informativeness, and understandability. Additional items assessed the perceived usefulness of performance feedback - either for enhancing therapy motivation (patients) or supporting clinical decision-making (HCPs) - as well as attitudes toward technology-based metrics, using the VPIT as an example. Participants were encouraged to ask both technical and content-related questions throughout, ensuring a full understanding of the material.

After completing the questionnaire, participants could share additional thoughts in a brief open discussion. All reflections were manually transcribed, and quotes included in this

manuscript were translated from German into English.

The feedback concepts were developed according to established user-centered design principles^{41,45,48}, balancing clarity and representativeness while minimizing bias during evaluation.

2.5.1 Feedback session 1 – type and amount of information

The first structured feedback session focused on evaluating how different types and levels of information, as well as their visual presentation, influenced user perceptions of usefulness, clarity, and engagement. To reflect a realistic interaction flow, we structured the session around three conceptual screens: (i) a start screen summarizing overall therapy progress, (ii) an assessment overview screen displaying all assessment outcomes, and (iii) an assessment detail screen providing more in-depth information on a selected assessment outcome. The three-level screen structure was designed to balance conciseness and depth, allowing users to quickly access key insights or explore detailed results as needed. Limiting the structure to three screens aimed to maintain a shallow navigation hierarchy, thereby enhancing usability in rehabilitation contexts, where cognitive and physical demands on users may be heightened⁵⁷.

For each screen, participants were shown a set of alternative feedback concepts that varied in content complexity and visual format. These included combinations of textual summaries, color-coded timelines, bar and line charts, spider plots, emojis, and progress bars. A text- and emoji-only version was also included in the first set to serve as a low-complexity baseline. These types of visualizations were selected as for their common use in digital health or consumer applications, with bar and line graphs being the most prevalent in digital health contexts^{58,59}.

An overview of the three screens and their respective feedback concepts is presented in Figure 2. This layered approach was designed to assess preferences across varying levels of information granularity and visual representation.

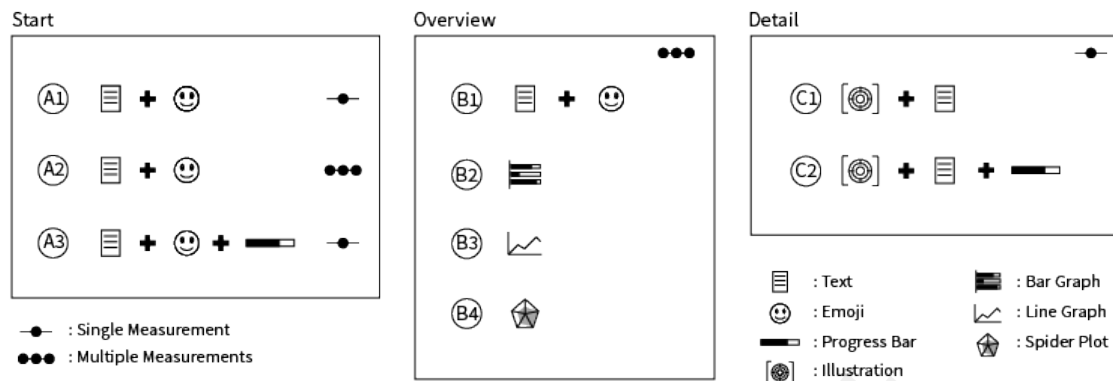


Figure 2: Illustration of feedback concepts shown during the first structured feedback session. The figure presents the three conceptual screens used to evaluate feedback design: (i) the start screen with overall performance summaries, (ii) the metric overview screen comparing visualization types for multiple assessments, and (iii) the assessment detail screen showing in-depth explanations of single assessment outcomes. Each screen included multiple concepts (labeled with 1-4) varying in complexity, visual format, and content presentation.

In addition to the concept-specific ratings, patients were asked whether feedback on assessments could enhance their motivation for therapy, and which types of feedback they would prefer to receive. HCPs were asked whether feedback could support clinical decision-making and help motivate patients. Further, HCPs were asked if they think that the VPIT could be helpful for them to evaluate upper limb function of the patients. The questions whether patients think that structured feedback could enhance their motivation and whether HCPs think that it could support in decision-making was asked before providing feedback concepts and afterwards, so that we get an unbiased opinion first. Both groups were also invited to indicate the kinds of technology-based assessment feedback they considered most meaningful.

2.5.2 Feedback session 2 – presentation of digital health metrics

The second structured feedback session focused on how digital health metrics could be effectively communicated to patients and HCPs through intuitive, engaging visual designs. To enhance understandability, the ten VPIT metrics (Supplementary Table S2) along with overall task completion time were grouped into five lay-language metrics: *total time & speed*, *movement smoothness*, *accuracy*, *efficiency*, and *grabbing & holding*. These categories were derived from the original research-based classifications but adapted in collaboration with HCPs to ensure that terminology was more intuitive and meaningful for patients and HCPs. To support both aggregation of metrics with different units and easier interpretation by non-expert users, each VPIT metric was first normalized to a 0–100 scale, where 0 represents the

most impaired neurological subject in our database (excluding outliers), and 100 corresponds to the median performance of age-matched healthy controls. Without normalization, the raw values would be too abstract or technical to convey meaningful feedback to patients or HCPs. The normalized metrics within each category were then averaged, without weighting, to create composite scores representing broader functional domains (e.g., dimensionless log jerk and spectral arc length were combined under movement smoothness). This approach preserved metric validity while enhancing interpretability for end-users. Details of the metric groupings are provided in the description column of Supplementary Table S2.

Participants were presented with three sets of feedback concepts (Figure 3), each designed to explore a different aspect of digital health metric feedback delivery.

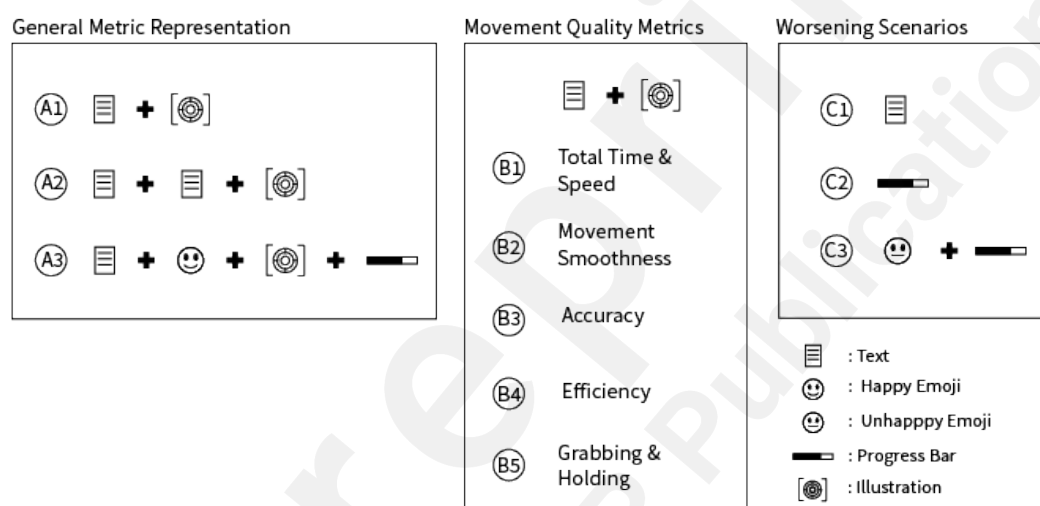


Figure 3: Overview of feedback concepts presented in the second structured feedback session. The figure summarizes three sets of visual feedback concepts shown to participants: Set 1 presents different designs for a single digital health metric (movement smoothness); Set 2 shows all five aggregated metrics with metric name, icons, visual illustrations, and explanations; and Set 3 explores how to communicate worsening outcomes, ranging from neutral to emotionally framed representations.

Set 1 explored the general presentation of a single metric, using movement smoothness as an example. The three concepts varied in their level of explanation and visual complexity. Each included the metric name, an icon, and different combinations of visual elements such as color-coded bars (green for healthy, red for impaired), progress bars, emojis, and explanatory text.

Set 2 provided an overview of all five aggregated VPIT metrics. Each metric was shown with its name, icon, short descriptive text, and a corresponding illustration. This set aimed to assess whether the metrics were perceived as clear, useful, and engaging.

Set 3 addressed how to communicate worsening outcomes. It included three

visualizations with varying emotional tone and visual emphasis. The first concept displayed a neutral message (“No Improvement”) without emojis or colored bars. The second used a green progress bar to indicate decline but excluded emojis and text. The third showed worsening results with an orange progress bar and an unhappy orange emoji. Each concept included an information button linking to an HCP-validated disclaimer explaining that performance changes may reflect normal day-to-day variation and should not be interpreted as clinically significant.

Participants rated each concept based on visual appeal, informativeness, and understandability. Additional open-ended questions captured qualitative impressions. Patients and HCPs were also asked whether they found the disclaimer clear and reassuring.

2.5.3 Feedback session 3 – usability evaluation of the app prototype

In the third feedback session, participants evaluated the usability of a high-fidelity app prototype created using Figma⁶⁰. The prototype featured randomly generated data and allowed participants to freely explore its interactive components, including screen navigation and button functionalities. Further, instead of just illustrations for the VPIT metrics, we included animations.

Participants were instructed to interact with the app independently. HCPs were asked to imagine they were preparing for a consultation and tasked with reviewing a patient’s most recent assessment results, identifying any worsening trends, and checking the summary screen. Patients were instructed to imagine they had just completed a new assessment and were asked to explore their feedback, compare it with previous results, and interpret the meaning of their scores. Participants were encouraged to ask questions during the interaction if clarification was needed.

Following the exploration phase, all participants completed the System Usability Scale (SUS)⁶¹ to evaluate overall usability. In addition, they responded to several targeted questions using 5-point Likert scales and open-ended questions. These assessed whether participants perceived structured feedback as valuable, and if there is something to improve in the app design and implementation. Items also addressed the desire for comparative feedback (e.g., performance relative to healthy individuals or other patients) and whether participants would like to celebrate personal milestones. Participants were also asked to suggest changes

or propose features to add or remove.

2.6 Data analysis

2.6.1 Quantitative analysis

Likert-scale responses collected across the three structured feedback sessions were analyzed using descriptive statistics in Python (version 3.13.2). For each item, the median and IQR were calculated separately for patients and HCPs. Median and IQR were chosen due to the ordinal nature of the data and their robustness to outliers, particularly in the context of small sample sizes. Accordingly, no inferential statistical testing was performed. Participants also ranked the feedback concepts within each concept set. Median ranks were computed to identify overall user preferences for different content types and visual formats. Visualizations such as heatmaps and summary tables were created in Python to support interpretation.

2.6.2 Usability assessment

In the third feedback session, the System Usability Scale (SUS) was used to evaluate the prototype's usability. Raw SUS responses were converted into standardized scores (0–100) following established scoring procedures. Group-level means, and standard deviations were calculated separately for patients and HCPs using Python. The scores were interpreted using standard usability benchmarks to contextualize the results.

2.6.3 Qualitative analysis

Open-ended responses from all three feedback sessions were analyzed using qualitative content analysis. A single coder manually reviewed all responses and identified recurring themes related to clarity, motivational impact, or visual design. Codes were grouped into categories to extract higher-level patterns. Representative quotes from both patients and HCPs were selected to illustrate these insights. The qualitative findings were used to complement the quantitative results and guide iterative refinements of the app design.

3 Results

3.1 Participant overview

In this study, a total of 17 persons after neurological injury and 15 HCPs participated across three feedback sessions (Table 1). Several of those 15 HCPs took part in multiple sessions. In feedback session 1 five patients and eight HCPs participated, in feedback session 2 six patients and nine HCPs participated and in feedback session 3 six patients and eight HCPs participated.

3.2 Feedback session 1 – type and amount of information

The primary aim of this feedback session was to assess the preferred type and quantity of information to be presented to patients and HCPs within a digital feedback application. To reduce potential bias, participants were asked general questions about structured feedback both prior (pre) and then following their exposure to the different feedback concepts (post). The feedback concepts included various mockup screens designed to give participants an idea of how feedback might be presented - for example, different data visualization formats such as bar charts, line graphs, spider plots, text, or even emojis. These concepts were created for a range of topics. In addition to presenting and evaluating feedback concepts, HCPs were asked whether feedback on completed assessments could aid clinical decision-making (pre: median 4.0 [IQR 3.5–4.0], post: 4.0 [4.0–5.0]), while patients were asked whether such feedback might increase their motivation to engage more consistently in therapy (pre: 4.0 [4.0–5.0], post: 4.0 [4.0–5.0]). Median responses therefore remained consistent before and after viewing the feedback concepts. Additionally, HCPs were asked whether providing feedback to patients could enhance their motivation (4.5 [4.0–5.0]) and whether digital health metrics such as VPIT-derived feedback would be useful in evaluating upper limb function (4.0 [4.0–5.0]). Participants evaluated nine feedback concepts across three screen types: the *start screen* which summarized overall performance, the *overview screen* which presented multiple assessment outcomes, and the *detail screen* which displayed results for a single metric. All concepts were rated on visual appeal, informativeness, and understandability using Likert scales (Figure 4, Supplementary Table S3), and were ranked by preference (Table 2).

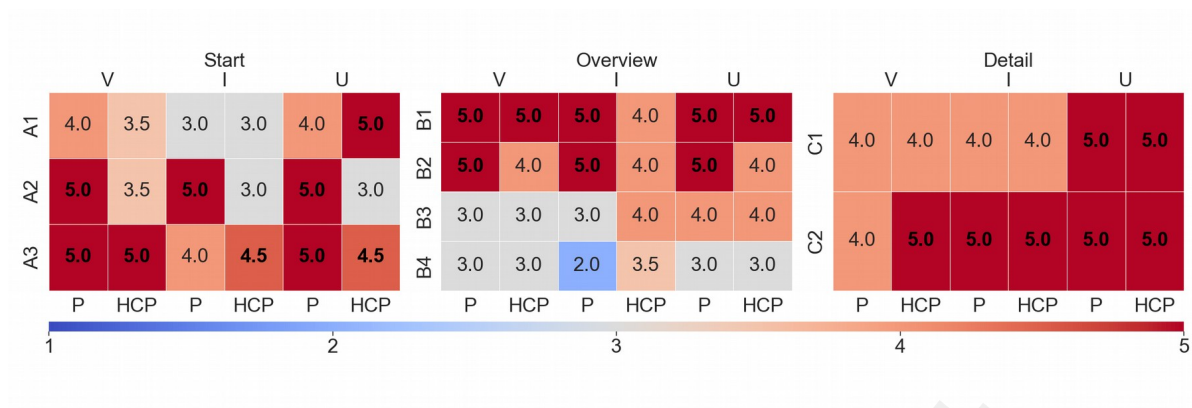


Figure 4: Median Likert-scale ratings for feedback concepts in feedback session 1. Participants rated nine feedback concepts on three dimensions, visual appeal (V), informativeness (I), and understandability (U), using a 5-point Likert scale. Concepts are grouped by screen: start screen, overview screen, and metric detail screen. For the start screen and the detail screen, concepts are ordered by increasing informational complexity (A1-A3; C1-C2); for the overview screen, by increasing visual complexity (B1-B4). Ratings are shown separately for patients (P) and healthcare practitioners (HCP). Color intensity reflects median score (red = favorable, blue = less favorable), with inter quartile ranges reported in Supplementary Table S3. Numbers in **bold** mark ratings above 4.5.

Table 2: Median concept rankings across screens for patients (P) and healthcare practitioners (HCP). The table summarizes rank preferences (median [IQR]) for each feedback concept shown during the three-screen app flow in feedback session 1. Concepts are grouped by screen and ordered as shown to reflect increasing informational (start and detail screen) or visual (overview screen) complexity. Lower ranks indicate higher preference (1 = most preferred). Numbers in **bold** mark the best ranked concepts.

Screen	Concept	Rank (P, n=5)	Rank (HCP, n=8)
Start	A1 (text + emoji)	3.0 [1.0–3.0]	2.0 [2.0–3.0]
	A2 (text + emoji + timeline)	2.0 [2.0–3.0]	2.5 [2.0–3.0]
	A3 (text + emoji + progress bar)	2.0 [1.0–2.5]	1.0 [1.0–1.5]
Overview	B1 (text + emoji)	3.0 [1.0–3.0]	2.5 [2.0–3.0]
	B2 (bar plot)	2.0 [2.0–2.0]	2.5 [2.0–4.0]
	B3 (line plot)	2.0 [2.0–4.0]	2.0 [1.0–2.5]
	B4 (spider plot)	3.0 [3.0–3.0]	3.0 [3.0–4.0]
Detail	C1 (illustration + text)	2.0 [2.0–2.0]	2.0 [1.0–2.0]
	C2 (illustration + text + progress bar)	1.0 [1.0–1.0]	1.0 [1.0–2.0]

Start screen: The concepts evolved from a basic emoji-based design (concept A1: text + emoji), to a timeline-enhanced version (concept A2), and finally to a version featuring a progress bar (concept A3). Both Likert scale ratings and rankings showed that concept A3 (progress bar) received the highest overall scores. Patients ranked concepts A3 and A2 (timeline) the highest (2.0 [1.0–2.5] and 2.0 [2.0–3.0], respectively), while HCPs rated concept A3 as their top choice (1.0 [1.0–1.5]). Concept A3 also scored highest among both groups in terms of visual appeal, informativeness, and understandability. One patient remarked, “The progress bar makes it clear at a glance if I got better or worse,” and an HCP commented, “It is easily understandable for patients.”

Overview screen: Concepts in this screen increased in visual complexity from text + emoji (B1) to bar plots (B2), line plots (B3), and spider plots (B4). HCPs ranked concept B3 slightly higher (2.0 [1.0–2.5]), while patients gave similar rankings to concepts B2 (2.0 [2.0–2.0]) and B3 (2.0 [2.0–4.0]). Both groups ranked concept B4 the lowest. Several participants commented on the need for absolute values, stating, “We want to see how much improvement or worsening there is in absolute values or percentages.” An HCP added, “If patients see norm values, they can compare what is good and where they need to improve. The same for us, we can then better compare where the patient is standing.”

Detail screen: Concept C2 (illustration + text + progress bar) was ranked highest by both patients and HCPs (1.0 [1.0–1.0] and 1.0 [1.0–2.0], respectively). Likert scores similarly favored this concept across all three dimensions. Participants made several comments on workflow, such as, “During clinical routine we don’t have time to study results, we want to directly be able to extract relevant information” and “Too many screens make the

system not usable, it is too much effort.”

Open feedback themes: Participants provided additional comments regarding feedback preferences. Several HCPs mentioned the importance of wording, e.g., “The language you use is important, it can be motivating for patients or not.” Concerns were raised about potential negative feedback effects, e.g., “What about if their results get worse? Then it is not motivating anymore.” Patients commented on the importance of defining goals, e.g., “What is my goal? I want it defined.” HCPs also suggested linking feedback to therapeutic recommendations, such as, “Add which therapy we can do to improve the specific results.” When asked what type of feedback they would like to receive about digital health metrics (example VPIT), HCPs emphasized graphical visualizations and clinically relevant parameters such as endurance, time, movement precision (e.g., smoothness), and daily-life applicability. Several HCPs requested simple graphical trends, such as line charts, for tracking changes. Patients primarily focused on viewing their own progress over time, especially regarding speed and differences between admission and post-therapy sessions.

3.3 Feedback session 2 – presentation of digital health metrics

The primary goal of this feedback session was to assess which digital health metrics patients and HCPs consider valuable, and how these metrics could be visualized in an intuitive and engaging manner. In the second feedback session, participants evaluated three sets of feedback concepts: Set 1 focused on how to generally present digital health metrics using the example of movement smoothness; set 2 focused on which metrics to present based on the condensed VPIT metrics; and set 3 addressed how to communicate worsening of assessment outcomes. Each concept in sets 1 and 2 was rated on visual appeal, informativeness, and understandability, each concept in set 3 on visual appeal and motivational aspects (Figure 5, Supplementary Table S4), while concepts in sets 1 and 3 were additionally ranked by preference ().

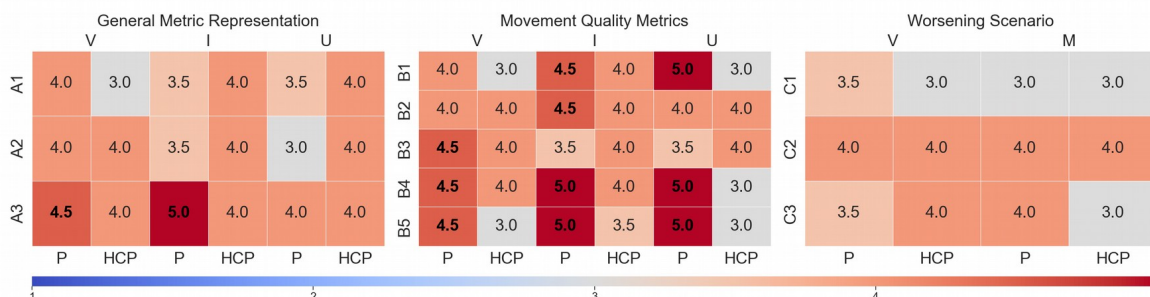


Figure 5: Median Likert-scale ratings for feedback all concepts in feedback session 2. Participants rated eleven feedback concepts (A1-A3; B1-B5; C1-C3) from three sets: Set 1 (presentation of a single digital health metric), set 2 (overview of all five grouped metrics) and set 3 (how to present worsening results). Each concept was evaluated on visual appeal (V), each concept for set 1 and 2 on informativeness (I), and understandability (U), and on motivational (M) for set 3 using a 5-point Likert scale. Ratings are shown separately for patients (P) and healthcare practitioners (HCP). Color intensity reflects the median score (red = favorable, blue = less favorable). Interquartile ranges are provided in Supplementary Table S4. Numbers in **bold** mark ratings above 4.5.

Table 3: Median concept rankings across set 1 and 3 for patients (P) and healthcare practitioners (HCP). The table summarizes rank preferences (median [IQR]) for each feedback concept shown during feedback session 2. Concepts are grouped by set and ordered as shown to reflect increasing informational complexity. Lower ranks indicate higher preference (1 = most preferred). Numbers in **bold** mark the best ranked concepts.

Set	Concept	Rank (P, n=6)	Rank (HCP, n=8)
General Metric Representation	A1 (text + illustration)	2.5 [2.0–3.0]	3.0 [3.0–3.0]
	A2 (extended text + illustration)	2.0 [1.5–2.0]	2.0 [1.5–2.0]
	A3 (progress bar)	2.0 [1.0–3.0]	1.0 [1.0–2.0]
Worsening Scenarios	C1 (text)	3.0 [2.0–3.0]	3.0 [2.0–3.0]
	C2 (progress bar)	1.5 [1.0–2.0]	2.5 [2.0–3.0]
	C3 (emoji + progress bar)	2.0 [1.0–3.0]	2.0 [1.0–2.0]

Set 1 – General metric representation: Patients assigned the highest median rank to concept B2 (extended text + illustration; 2.0 [1.5–2.0]) and B3 (progress bar; 2.0 [1.0–3.0]), while HCPs most frequently ranked concept B3 (progress bar) as their top choice (1.0 [1.0–2.0]). Among HCPs, concept B3 also received the highest Likert-scale ratings across visual appeal, informativeness, and understandability. Patients rated concepts B2 and B3 similarly on the understandability and informativeness dimensions. Qualitative feedback indicated varied preferences. One HCP described concept B2 as containing “too much information,” whereas concept B3 was described as “short, concise, and to the point.” Four additional HCPs commented that concept A3 was “simple, but understandable and very clear,” and noted that it was “very well suited for patients.” Two HCPs suggested a combination of concepts B2 and B3. One patient commented that concept B2 included “too much text,” while concept B3 was “very well understandable.” Two other patients similarly stated that concept B3 was “very clear.”

Set 2 – Digital health metrics: Participants evaluated five grouped digital health metrics: time and speed (B1), smoothness (B2), accuracy (B3), efficiency (B4), and grabbing and holding (B5). Across all evaluated dimensions, patients generally provided higher Likert-scale ratings compared to HCPs. For visual appeal, accuracy (B3) and efficiency (B4) received the highest median ratings from both groups, with scores of 4.5 among patients and 4.0 among HCPs. In contrast, time and speed (B1) was rated lower, particularly by HCPs, with median scores of 4.0 for patients and 3.0 for HCPs. Regarding informativeness, efficiency (B4) and grabbing and holding (B5) were rated most favorably by patients, each achieving median scores of 5.0. HCPs rated all metrics slightly lower, with median scores around 4.0, and a lower rating for grabbing and holding (approximately 3.5). In terms of understandability, time and speed (B1), efficiency (B4), and grabbing and holding (B5) received the highest median scores among patients (5.0), while these metrics showed the lowest median scores among HCPs (3.0). Participant comments were specific to individual visual elements, with several suggestions for refinement. For example, one HCP proposed adding a tortoise and rabbit to the speed visualization to increase clarity.

Set 3 – Worsening scenarios: Participants evaluated three concepts for communicating worsening outcomes: C1 (text only), C2 (progress bar), and C3 (emoji + progress bar). Among patients, C2 received the highest median ranking (1.5 [1.0–2.0]), while HCPs preferred C3 (2.0 [1.0–2.0]). C1 was ranked lowest by both groups (3.0 [2.0–3.0]). Likert-scale ratings showed that concepts with visual elements (C2 and C3) were generally rated higher in visual appeal, informativeness, and understandability compared to the text-only C1. Across most dimensions, patients gave higher ratings than HCPs. Participant comments added context: HCPs described C3 as a “good visualization,” appreciating the emphasis on positive aspects, though some felt the worsening was shown too strongly. Several suggested revising the emoji design, e.g., using “neutral colored, smiling emojis.” For C1, HCPs stressed the need for quantitative data: “always show values,” and “we always want to see exact values.” Patients also preferred visuals. For C3, they liked the color-coded improvements but recommended neutral smiles. One noted, “too much indicated the negative parts.” Regarding C1, a patient said, “I don’t want text, it is too negative, I prefer emojis and colors.” Participants also rated a text explanation stating that worsening results may reflect normal daily fluctuations. Understandability was rated at 4.0 [3.0–4.5] by HCPs and 4.5 [3.5–5.0] by patients; motivational value was 4.0 [3.0–4.0] and 4.5 [4.0–5.0], respectively. HCPs suggested making the message more encouraging and replacing

“measurement variability” with relatable terms like environmental or daily performance factors. Patients asked for simpler language and generally found the text helpful. General feedback about the app development was also positive. One HCP noted, “The interface is a very nice opportunity for everyone, I am glad to be part of the optimizing process,” while a patient commented, “Exciting project with a lot of informative insights into the therapy progress.”

3.4 Feedback session 3 – usability evaluation of the app prototype

The main goal of this feedback session was to evaluate the usability of the first functional app prototype. During this final structured feedback session, participants interacted with the fully functional application prototype, which featured short animations in place of static illustrations to visualize digital health metrics. Usability was evaluated using the System Usability Scale (SUS). Patients reported a mean SUS score of 93.6 ± 6.4 , while HCPs reported a mean score of 80.9 ± 8.1 , indicating high usability ratings across both groups (Figure 6). Participants were also invited to suggest improvements to the prototype. One HCP requested the inclusion of absolute values alongside percentages in performance bars: “I would like to see an absolute number with units in every bar, not only the percentage indicating improvement or decline.” Another HCP highlighted the clarity of the interface, stating: “The implementation of the user interface is really nice. It shows just the right amount of key information, it’s super easy to get a good overview.” Further suggestions included: “Consider adding brief explanations when milestones are not reached, or positive reinforcement when progress is made. I’d also like to see training recommendations as videos, such as: ‘Watch this video again to improve parameter A/B/C.’” Another comment was: “Could the system automatically assign milestones based on initial assessments? And for chronic patients, could past data from earlier years be integrated?” Patients also contributed feedback. One commented: “Compliment! There’s nothing I would want to change.” Another noted: “Can you include both the absolute values and percentage improvements in the bars?”. Participants were asked about the perceived value of structured feedback in a clinical setting. Patients were specifically questioned on whether they believed such feedback could enhance their motivation during therapy, while HCPs were asked if it could aid clinical decision-making. This question was previously posed to participants of feedback session 1, before any feedback concepts were introduced, enabling

a pre–post comparison. Median ratings from patients increased from 4.0 [4.0–5.0] to 5.0 [4.0–5.0], while HCPs' ratings rose from 4.0 [3.5–4.0] to 4.5 [4.0–5.0]. Additionally, HCPs rated the potential of structured feedback to enhance patient motivation at 4.5 [4.0–5.0]. Participants were also asked whether they would find value in performance comparisons to (i) age-matched healthy individuals or (ii) other patients with similar conditions. Patients rated comparisons to healthy individuals at 5.0 [5.0–5.0] and to other patients at 4.0 [3.0–5.0]. HCPs gave corresponding ratings of 4.0 [4.0–4.5] and 3.0 [2.0–4.0]. Lastly, the perceived benefit of celebrating therapeutic milestones within clinical settings was evaluated. Patients rated this concept at 4.0 [3.0–5.0], while HCPs provided a slightly lower median rating of 3.5 [3.0–4.5].

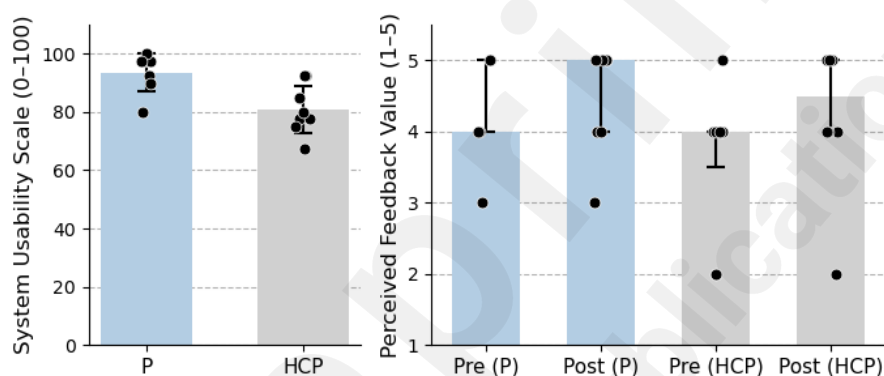


Figure 6: Evaluation of app usability and perceived value of structured feedback. The left panel shows System Usability Scale (SUS) scores for the app prototype as rated by patients (P) and healthcare practitioners (HCPs) in feedback session 3. The right panel presents pre–post ratings on the perceived value of structured assessment feedback in clinical routine, separately for patients and HCPs. Specifically, patients were asked if structured feedback about performed assessments could enhance their motivation to perform therapy on a more regular basis and HCPs were asked if such feedback could aid them in decision-making processes. Patients and HCPs rated feedback value on a 5-point Likert scale before exposure to the feedback concepts (feedback session 1) and again after app testing (feedback session 3). Bars for the SUS Score represent group means; error bars indicate standard deviations. Bars for the perceived value of structured feedback represent medians, error bars interquartile ranges. Individual responses are shown as jittered black dots.

3.5 Final implementation of the digital feedback application

The final digital feedback application incorporated input from both patients and HCPs across all three structured feedback sessions. It was implemented as a mobile iOS application with two distinct user interfaces: one for patients and one for HCPs. The primary difference lies in data management functionality, which is available only to HCPs. This includes the ability to create new patient accounts and manage user credentials, such as resetting passwords.

The application was developed using React Native (version 15.1.3)⁶², leveraging the Expo framework (version 52.0.23)⁶³ to facilitate efficient cross-platform development, deployment, and maintenance. It was optimized for tablet use to ensure seamless integration into clinical

workflows, including bedside consultations. The backend infrastructure was implemented with Django (version 5.1.5) ⁶⁴ and a structured SQLite3 (version 3.45.3) database ⁶⁵, providing robust mechanisms for secure data storage, user authentication, and real-time data synchronization.

Data from technology-based assessments (e.g., the VPIT) is automatically analyzed and then visualized in the app interface. By contrast, data from conventional assessments is currently extracted manually from the hospital information system and exported to the app. The final digital feedback application integrates all iterative refinements from the three performed structured feedback sessions (Figure 7).

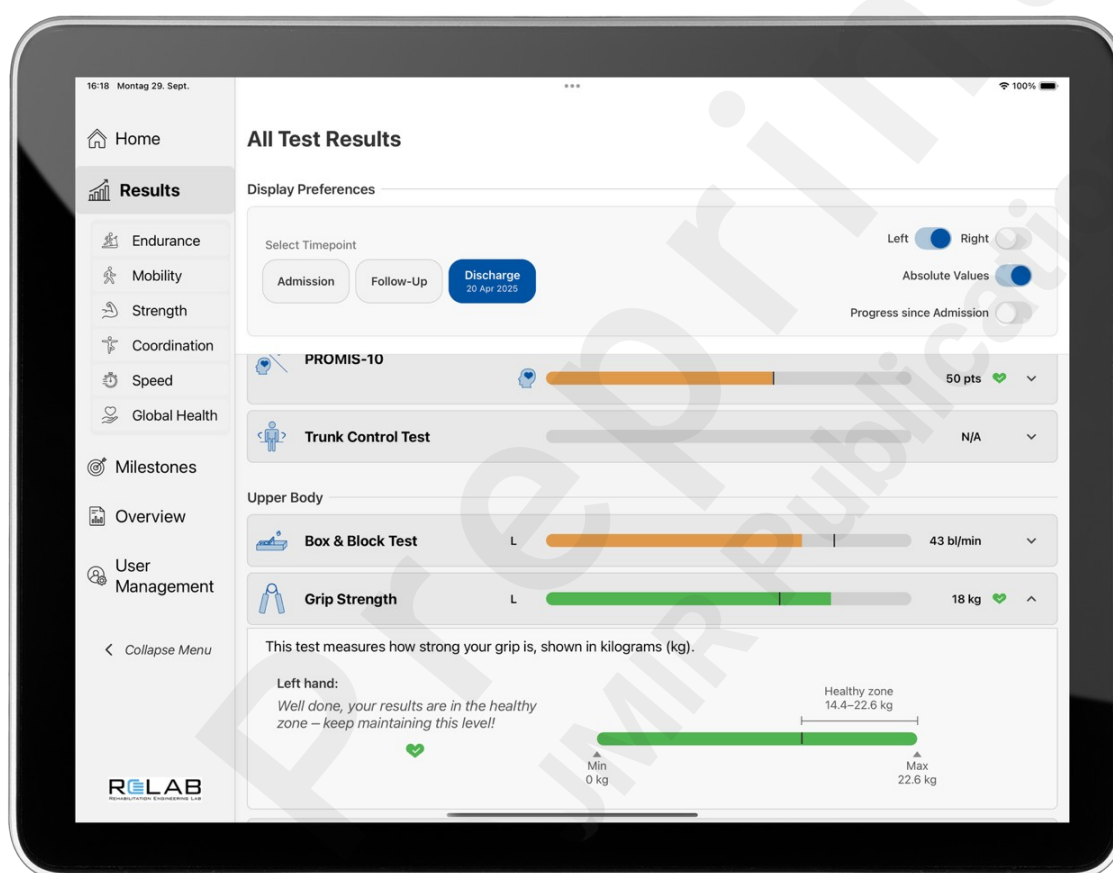


Figure 7: Final feedback application. Screenshot of the app interface as presented in feedback session 3. The app is in normative mode, i.e. shown progress bars indicate the users' current functionality compared to an assessment-specific normative population including, age-, gender-, and handedness factors. Drop-down menus provide assessment and visualization explanations. Assessments are grouped by body region and can be filtered by functional domain (left panel). The left panel can be collapsed to show icons only.

3.5.1 Implementation - results feedback session 1

The results of the first feedback session - conducted with both patients and HCPs - were subsequently reviewed in a joint discussion with HCPs. As a result of this discussion, a unified user interface for both patients and HCPs was implemented to support clinical workflows and

improve communication.

To improve usability, the navigation depth was reduced from three to two levels. The start screen now corresponds to a login screen and leads directly to an overview screen displaying all completed assessments. To enhance structure and clarity, assessments were grouped into three categories: upper body, lower body, and general assessments. A side navigation panel was added to enable filtering of assessment results by category (e.g., strength, coordination), as well as access to personal health data (e.g., BMI, weight) and the three primary therapy goals, which HCPs define together with their patient at admission.

Based on user feedback, we selected color-coded bar charts as the primary method for visualizing assessment outcomes, supplemented by emoji-based indicators where appropriate. To further support user understanding, info buttons were added to each assessment entry to provide brief explanations of what each measure represents. These buttons were later replaced with expandable drop-down menus following additional feedback in session 3.

3.5.2 Implementation - results feedback session 2

Based on insights gathered during feedback session 2, we implemented several interface enhancements to improve the understandability, usability, and interpretability of digital health metrics for patients. We introduced a standardized format for the metric representation: Each digital health metric is now displayed using a combination of a recognizable icon, a concise textual explanation, and an animated visualization illustrating key performance features (e.g., jerky vs. smooth movements). To maintain a clean interface, the explanatory text and animation are accessible on demand via an information button associated with each metric. The information button was later replaced by a dropdown menu following feedback from feedback session 3. To ensure a consistent and intuitive experience, we aligned the visualization style of the VPIT metrics with that of the conventional assessments. All assessment entries, including conventional ones, were supplemented with icons to support faster recognition and navigation. Several icons and associated descriptions were also refined to enhance clarity and ease of interpretation. To reduce cognitive load in bimanual assessments, we added a left–right toggle that allows users to filter results by limb. This feature improves legibility and reduces visual clutter in cases where assessments are conducted on both hands.

In response to concerns about how to present worsening outcomes, we adopted a neutral feedback strategy. Performance declines are now indicated solely through color changes, avoiding emotive symbols such as emojis. To prevent misinterpretation, we included an explanatory note clarifying that such fluctuations may reflect normal day-to-day variability and are not necessarily clinically meaningful.

Finally, we added an overview panel that displays assessment results in a tabular format without graphical visualization. This table includes absolute values for each metric, color-coded based on normative reference data. Directional arrows were included to highlight improvements or deteriorations over time, supporting rapid identification of clinically relevant trends.

3.5.3 Implementation - results feedback session 3

Following feedback session 3, we implemented several refinements to improve clarity, reduce visual clutter, and better align the app with the informational needs of both patients and HCPs. To reduce on-screen complexity, the previously used information buttons were replaced with expandable drop-down menus positioned next to each assessment entry. These menus provide contextual information on what each assessment measures and how to interpret the corresponding visualizations, allowing users to access explanations only when needed. To ensure a balanced presentation and prevent overemphasis on the VPIT, we grouped its individual metrics into a single assessment block with distinct subcategories. This approach preserved detail while aligning with the structure used for other assessments in the application. In response to requests from patients and HCPs for greater numerical clarity, we added value labels next to the performance bars. These are available in both absolute units and percentage format. A toggle allows users to switch between the two representations, accommodating varying levels of familiarity with specific measurement units. To facilitate longitudinal tracking, we introduced a tab-based navigation system, enabling users to browse outcomes from all previously completed assessments. By default, the most recent assessments are displayed, providing a current snapshot while preserving access to historical data.

The users also expressed interest in understanding their performance relative to normative data. To address this, we implemented two display modes: a progress mode and a normative mode. The progress mode is shown by default and displays the patient's changes in

performance over time relative to their baseline measurement (assessed at the time of entry to their clinical stay). This format emphasizes individual improvement and is potentially more motivating for patients. To prevent overinterpretation of minor changes, we implemented a color-coding scheme based on the minimal detectable change, or when not available, the minimal clinically important difference. Improvements exceeding the thresholds are shown as green bars extending to the right; changes within the threshold ranges are shown in orange; and significant deteriorations beyond the thresholds are shown in red, extending to the left. The normative mode compares the patient's performance to a healthy reference population matched by assessment-specific criteria, including age, gender, clinical condition, and hand dominance. Each outcome is mapped to a standardized 0–100 scale, where 0 indicates no function and 100 represents the median performance of a healthy reference group. The lower bound of the healthy range is defined as the median minus two standard deviations. On this scale, scores within the healthy range appear in green; values below the healthy threshold but above 50% are shown in orange; and scores below 50% are displayed in red. This system supports intuitive interpretation of how a patient's performance compares to expected norms. For the VPIT, normative values were derived from a healthy reference database ($n = 120$) matched by age and gender. Norm values for conventional assessments were drawn from previously published reference data, primarily listed in ⁹.

4 Discussion

4.1 Principal findings

This study presents a three-step, iterative, user-centered approach to designing a digital application that delivers structured feedback on assessment results to patients and HCPs in neurorehabilitation. After defining clinical and user requirements, we conducted three structured feedback sessions to evaluate: (i) the type, amount, and visual format of information presented; (ii) the integration and representation of digital health metrics from technology-based assessments; and (iii) the usability of a digital app prototype delivering this feedback. The final application presents assessment outcomes in a co-designed format tailored to both stakeholder groups. Likert-scale ratings indicated a growing perceived value across sessions. Although participant samples varied between sessions, patients' ratings increased to a median of 5.0 [4.0–5.0], while HCPs' ratings rose from 4.0 to 4.5.

4.2 Comparison with prior work

The motivational and clinical benefits of feedback are well supported in rehabilitation literature. Feedback that is timely, personalized, and performance-related has been shown to improve engagement, motor outcomes, and reduce patient anxiety^{27,28,30,31}. While most prior work has focused on real-time feedback during therapy, our application delivers post-assessment feedback. Nevertheless, we expect similar benefits, including increased motivation, better understanding of progress, and enhanced shared decision-making. Structured feedback has also been recognized as a tool for more evidence-based, value-based healthcare (VBHC)³², which our co-designed digital app supports.

To ensure these expected benefits translate effectively in practice, our main focus was to evaluate how feedback is presented to users. Across the performed structured feedback sessions, participants consistently favored simple, visually guided formats, especially progress bars and icon-supported charts, over text-heavy or abstract representations. They preferred clear visuals such as bar charts with optional, layered explanations over more complex formats like line or spider graphs, which can be harder to interpret^{58,59}. In contrast to many digital health tools that offer static or device-specific reports, our digital feedback app applies strategies like expandable content design and visual simplification⁵⁹ to preserve clinical utility without sacrificing usability. Compared to typical device-generated reports, our solution emphasizes personalization, clarity, and emotional neutrality. For example, in feedback session 1, both patients and HCPs favored the start screen featuring a progress bar (concept A3). This preference remained strong in later sessions, where other designs incorporating progress bars were again rated highest for representing general metrics. These preferences align with usability literature, especially for patients and older adults - key populations given the increasing prevalence of age-related neurological disorders. They tend to favor icon-driven navigation, high-contrast layouts, and intuitive visuals^{46,49,50,66}. However, some divergence from expectations emerged. Despite HCPs often recommending simple formats such as plain text or emojis, patients showed a clear preference for richer, more informative visuals. This contrast underscores the importance of iterative, inclusive stakeholder engagement throughout the design process. Features like toggleable views - between absolute and normalized scores, or between progress and normative benchmarks - allowed personalization without cognitive overload. These solutions addressed users' varied informational needs while preserving interpretability

and clinical relevance. Ramesh et al.⁶⁷ used a symbolic, metaphor-driven approach that omitted numerical data and involved patients only at the final evaluation stage. In contrast, our participants were engaged throughout the design process and explicitly requested access to raw scores. By combining quantitative metrics with intuitive visuals, our tool meets diverse informational needs and promotes transparency. Moreover, our larger, more diverse sample supports a broader generalizability. A key design insight was the need to balance simplicity with sufficient depth.

These design preferences not only influenced visual choices but also highlighted the importance of emotional tone and clinical context in delivering feedback. Participants were not just passive recipients of information, but active stakeholders who were concerned with how their performance data was presented, especially when it involved sensitive or potentially discouraging results. As such, our co-design process extended beyond aesthetics to encompass the psychological and clinical dimensions of feedback delivery. Both groups emphasized the sensitivity required when presenting performance declines. While emojis were appreciated for reinforcing positive progress, participants favored more neutral, color-coded indicators and supportive language when feedback indicated worsening. This reflects prior findings that uncontextualized or overly negative feedback can increase patient anxiety^{68,69} and should be carefully framed. Co-design allowed visual elements to be emotionally contextualized - emojis for improvement, neutral tones for decline - supporting engagement and comprehension. To ensure appropriate support, this digital feedback application is designed for use during patient consultations rather than in isolation.

In addition to emotional sensitivity, participants expressed a desire for actionable guidance and goal-setting tools. Patients valued reassurance during setbacks and the ability to track their own progress, while HCPs stressed the importance of clinically meaningful metrics and normative comparisons. This aligns with evidence supporting goal setting as a key motivator^{27,70} and broader recommendations for embedding digital tools within VBHC frameworks²⁶. Despite VBHC's potential, its implementation remains challenging due to fragmented infrastructure and limited outcome tracking^{3,33}.

By combining conventional and digital health metrics in a co-designed tool, our approach addresses these gaps and supports data-driven, patient-centered care. A central aim of this work was to demonstrate how digital health metrics can be meaningfully integrated into clinical routine. Digital health metrics, such as those exemplified by the VPIT, contribute objective,

fine-grained, and repeatable measurements that can enhance the sensitivity and clinical utility of outcome tracking^{11,21,51,54}. Their standardized format allows for automated longitudinal monitoring - eliminating the need for manual digitization or additional data processing - and supports meaningful comparisons across patients and time points. While the VPIT specifically targets upper limb sensorimotor function, its metric design is conceptually aligned with other technology-based assessments, such as those using wearable sensors for upper limb monitoring^{71,72}. These approaches share the goal of capturing movement quality and functional performance through quantifiable, reproducible measures. Moreover, the principles guiding the presentation and interpretation of VPIT metrics are broadly applicable to other rehabilitation domains, e.g., including lower limb function and cognitive assessments, where wealth of digital health metrics has been proposed⁷³⁻⁷⁶. As such, the use of digital health metrics in feedback tools offers significant potential to advance VBHC by improving outcome transparency, personalization, and clinical efficiency.

The final digital application, refined through iterative feedback, achieved high usability scores from both stakeholder groups (SUS: patients = 93.6 ± 6.4 , HCPs = 80.9 ± 8.1). Users highlighted ease of navigation, clear interface design, and appropriate feedback granularity. Importantly, even participants with low digital literacy reported no major interaction issues, supporting the tool's accessibility. Implementation feedback prioritized interoperability with electronic health records, inclusion of therapy recommendations, and milestone tracking. These additions, including automatic goal setting and benchmarking, enhance clinical integration and self-management.

4.3 Limitations

While promising, this study is subject to several limitations. The present findings are based on structured feedback sessions conducted with a limited number of participants within a single healthcare setting. Although the diversity of perspectives yielded valuable design insights, the generalizability of these results is constrained. Broader validation across different clinical environments and with larger, more heterogeneous populations, beyond the MS cohort included here, should be considered. Moreover, while the current evaluation was conducted using mock data, future studies should assess the application's real-world impact on clinical decision-making, patient adherence, and therapeutic outcomes. The clinical utility of the selected conventional and technology-based assessments also remains to be established through routine

use and empirical testing. Additionally, interpreting the digital health metrics, particularly their prognostic relevance and implications for therapy adaptation, requires larger datasets and further investigation. Understanding how these metrics translate into actionable clinical decisions will be critical for the tool's long-term adoption and effectiveness.

4.4 Conclusions

In summary, we demonstrate the feasibility and value of a co-designed tool for neurorehabilitation. Through structured feedback and iterative design, we developed a digital application that aligns with clinical workflows, has potential to enhance patient engagement, and advances digital health integration. While these findings are promising, especially in terms of usability and stakeholder acceptance, further research is needed to evaluate the tool's impact in real-world settings, across broader patient populations, and in fully integrated clinical infrastructures. Our results offer a foundational step toward embedding patient-centered feedback systems into clinical routine, with potential to support data-driven decision-making and VBHC.

Declarations

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to data protection reasons but are available from the corresponding author on reasonable request.

Code availability

The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Ethics approval and consent to participate

All procedures involving human participants were approved by the Clinical Trial Board of the

Kliniken Valens and conducted in accordance with the Declaration of Helsinki and the Swiss Human Research Act (HRA). The study adhered to all relevant guidelines and regulations. Informed consent was obtained from all participants through the institutional general consent process.

Competing interests

All authors declare no financial or non-financial competing interests.

Author contributions

Study design: ND, OL, RG, JB, RS. Data collection: ND, KK. Participant Recruitment: RS. Data analysis: ND. Data interpretation: ND, OL, RG, JB. App Development ND, JS, KK. Manuscript writing: ND. Manuscript review: OL, RG, JB, JS. All authors read and approved the final manuscript.

Acknowledgements

We thank all patients and healthcare practitioners who participated in this study for their valuable contributions. We also thank Franziska Riegel and Daniel Riese for their input during discussions on the study design.

This study was funded by: the Swiss Multiple Sclerosis (MS) Society, Grant Number 2021-06, the Swiss National Science Foundation as partner of the EJP RD, under grant agreement No02424 – 000, and the National Research Foundation Singapore (NRF) under its Campus for Research Excellence and Technological Enterprise (CREATE) program.

The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Abbreviations

HCP	Healthcare practitioner
HRA	Human Research Act
HST	Health Science and Technology
ICF	International Classification of Functioning, Disability and Health framework
IQR	Interquartile range
KINAR M	Kinesiological Instrument for Normal and Altered Reaching Movement
MS	Multiple Sclerosis
OT	Occupational therapist
PT	Physiotherapist
ST	Sports therapist / scientist
SUS	System Usability Scale
VBHC	Value-based healthcare
VPIT	Virtual Peg Insertion Test

Supplementary information

This paper includes four supplementary tables, namely Supplementary Table S1, S2, S3 and S4 which can be found in Multimedia Appendix 1.

References

1. Khan F, Turner-Stokes L, Ng L, Kilpatrick T, Amatya B. Multidisciplinary rehabilitation for adults with multiple sclerosis. *Cochrane Database of Systematic Reviews*. 2007;2011(12). doi:10.1002/14651858.CD006036.pub2
2. Westerink HJ, Bresser CC, Garvelink MM, et al. The use of outcome data in patient consultations from the healthcare professionals' and patients' perspectives: A mixed methods study. *Patient Educ Couns*. 2024;118:108043. doi:10.1016/j.pec.2023.108043
3. van der Nat PB. The new strategic agenda for value transformation. *Health Serv Manage Res*. 2022;35(3):189-193. doi:10.1177/09514848211011739
4. Porter ME. Defining an introducing value in health care . In: *Evidence-Based Medicine and the Changing Nature of Health Care: 2007 IOM Meeting Summary* . The national academies press; 2008:161-172.
5. Berwick DM, Nolan TW, Whittington J. The Triple Aim: Care, Health, And Cost. *Health Aff*. 2008;27(3):759-769. doi:10.1377/hlthaff.27.3.759
6. Mathiowetz V, Weber K, Kashman N, Volland G. Adult Norms for the Nine Hole Peg Test of Finger Dexterity. <http://dx.doi.org/10.1177/153944928500500102>. 1985;5(1):24-38. doi:10.1177/153944928500500102
7. Podsiadlo D, Richardson S. The Timed "Up & Go": A Test of Basic Functional Mobility for Frail Elderly Persons. *J Am Geriatr Soc*. 1991;39(2):142-148. doi:10.1111/j.1532-5415.1991.tb01616.x
8. Keith RA, Granger C V, Hamilton BB, Sherwin FS. The functional independence measure: a new tool for rehabilitation. *Adv Clin Rehabil*. 1987;1:6-18.
9. Schädler S, Kool J, Lüthi H, Marks D, Oesch P. *Assessments in Der Rehabilitation Band 1: Neurologie*. 4th ed. Hogrefe AG; 2019.
10. World Health Organization. *International Classification of Functioning, Disability and Health (ICF)*. World Health Organization; 2001. Accessed September 29, 2025. <https://www.who.int/classifications/icf/en/>
11. Kanzler CM, Schwarz A, Held JPO, Luft AR, Gassert R, Lambercy O. Technology-aided assessment of functionally relevant sensorimotor impairments in arm and hand of post-stroke individuals. *J Neuroeng Rehabil*. 2020;17(1). doi:10.1186/s12984-020-00748-5
12. Gladstone DJ, Danells CJ, Black SE. The Fugl-Meyer Assessment of Motor Recovery after Stroke: A Critical Review of Its Measurement Properties. *Neurorehabil Neural Repair*. 2002;16(3):232-240. doi:10.1177/154596802401105171
13. Hawe RL, Scott SH, Dukelow SP. Taking Proportional Out of Stroke Recovery. *Stroke*. 2019;50(1):204-211. doi:10.1161/STROKEAHA.118.023006
14. Murphy MA, Willén C, Sunnerhagen KS. Kinematic Variables Quantifying Upper-Extremity Performance After Stroke During Reaching and Drinking From a Glass. *Neurorehabil Neural Repair*. 2011;25(1):71-80. doi:10.1177/1545968310370748
15. Alt Murphy M, Murphy S, Persson HC, Bergström UB, Sunnerhagen KS. Kinematic Analysis Using 3D Motion Capture of Drinking Task in People With and Without Upper-extremity Impairments. *J Vis Exp*. 2018;(133). doi:10.3791/57228
16. Unger T, Moslehian AS, Peiffer JD, et al. Differentiable Biomechanics for Markerless Motion Capture in Upper Limb Stroke Rehabilitation: A Comparison with Optical Motion Capture. Published online November 22, 2024.
17. Unger T, de Sousa Ribeiro R, Mokni M, et al. Upper limb movement quality measures: comparing IMUs and optical motion capture in stroke patients performing a drinking task. *Front Digit Health*. 2024;6. doi:10.3389/fdgth.2024.1359776
18. Dukelow SP, Herter TM, Moore KD, et al. Quantitative Assessment of Limb Position Sense

- Following Stroke. *Neurorehabil Neural Repair*. 2010;24(2):178-187. doi:10.1177/1545968309345267
19. Graser J V., Prospero L, Liesch M, Keller U, van Hedel HJA. Test–retest reliability of upper limb robotic exoskeleton assessments in children and youths with brain lesions. *Sci Rep*. 2022;12(1):16685. doi:10.1038/s41598-022-20588-8
 20. Fluet MC, Lambercy O, Gassert R. Upper limb assessment using a Virtual Peg Insertion Test. In: *IEEE International Conference on Rehabilitation Robotics*. 2011. doi:10.1109/ICORR.2011.5975348
 21. Kanzler CM, Rinderknecht MD, Schwarz A, et al. A data-driven framework for selecting and validating digital health metrics: use-case in neurological sensorimotor impairments. *NPJ Digit Med*. 2020;3(1). doi:10.1038/s41746-020-0286-7
 22. Schwarz A, Kanzler CM, Lambercy O, Luft AR, Veerbeek JM. Systematic review on kinematic assessments of upper limb movements after stroke. *Stroke*. Lippincott Williams and Wilkins. 2019;50(3):718-727. doi:10.1161/STROKEAHA.118.023531
 23. Balasubramanian S, Melendez-Calderon A, Burdet E. A Robust and Sensitive Metric for Quantifying Movement Smoothness. *IEEE Trans Biomed Eng*. 2012;59(8):2126-2136. doi:10.1109/TBME.2011.2179545
 24. Rohrer B, Fasoli S, Krebs HI, et al. Movement Smoothness Changes during Stroke Recovery. *The Journal of Neuroscience*. 2002;22(18):8297-8304. doi:10.1523/JNEUROSCI.22-18-08297.2002
 25. Hope TMH, Friston K, Price CJ, Leff AP, Rotshtein P, Bowman H. Recovery after stroke: not so proportional after all? *Brain*. 2019;142(1):15-22. doi:10.1093/brain/awy302
 26. van Staalduinen DJ, van den Bekerom P, Groeneveld S, Kidanemariam M, Stiggelbout AM, van den Akker-van Marle ME. The implementation of value-based healthcare: a scoping review. *BMC Health Serv Res*. 2022;22(1):270. doi:10.1186/s12913-022-07489-2
 27. Oyake K, Suzuki M, Otaka Y, Momose K, Tanaka S. Motivational Strategies for Stroke Rehabilitation: A Delphi Study. *Arch Phys Med Rehabil*. 2020;101(11):1929-1936. doi:10.1016/j.apmr.2020.06.007
 28. Muraoka K, Takimoto Y, Nakazawa E, Tsuji T, Liu M. Stroke Survivors' Experiences and Needs during the Decision-making Process Considering Rehabilitation Options: A Pilot Descriptive Study in Japan. *Prog Rehabil Med*. 2022;7:20220024. doi:10.2490/prm.20220024
 29. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. *Cochrane Database of Systematic Reviews*. 2017;2018(1). doi:10.1002/14651858.CD008349.pub4
 30. Dobkin BH, Plummer-D'Amato P, Elashoff R, Lee J. International Randomized Clinical Trial, Stroke Inpatient Rehabilitation With Reinforcement of Walking Speed (SIRROWS), Improves Outcomes. *Neurorehabil Neural Repair*. 2010;24(3):235-242. doi:10.1177/1545968309357558
 31. Popović MD, Kostić MD, Rodić SZ, Konstantinović LM. Feedback-Mediated Upper Extremities Exercise: Increasing Patient Motivation in Poststroke Rehabilitation. *Biomed Res Int*. 2014;2014:1-11. doi:10.1155/2014/520374
 32. McGlinchey MP, Faulkner-Gurstein R, Sackley CM, McKevitt C. Factors guiding therapist decision making in the rehabilitation of physical function after severely disabling stroke – an ethnographic study. *Disabil Rehabil*. 2024;46(4):672-684. doi:10.1080/09638288.2023.2172463
 33. van Veghel D, Daeter EJ, Bax M, et al. Organization of outcome-based quality improvement in Dutch heart centres. *Eur Heart J Qual Care Clin Outcomes*. Published online April 26, 2019. doi:10.1093/ehjqcco/qcz021
 34. Vijverberg JRG, Daniels K, Steinmann G, et al. Mapping the extent, range and nature of research activity on value-based healthcare in the 15 years following its introduction (2006–2021): a scoping review. *BMJ Open*. 2022;12(8):e064983. doi:10.1136/bmjopen-2022-064983
 35. Zielasek J, Reinhardt I, Schmidt L, Gouzoulis-Mayfrank E. Adapting and Implementing Apps for Mental Healthcare. *Curr Psychiatry Rep*. 2022;24(9):407-417. doi:10.1007/s11920-022-01350-3
 36. King DR, Emerson MR, Tartaglia J, Nanda G, Tatro NA. Methods for Navigating the Mobile Mental

- Health App Landscape for Clinical Use. *Curr Treat Options Psychiatry*. Published online May 24, 2023;1-15. doi:10.1007/s40501-023-00288-4
37. Phillips JC, Alfano AR, Barfield LC, et al. Exploring Maternal and Infant Health App Development and Effectiveness Research: Scoping Review. *JMIR Pediatr Parent*. 2024;7:e46973. doi:10.2196/46973
38. 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. *International Journal of Behavioral Nutrition and Physical Activity*. 2016;13(1):127. doi:10.1186/s12966-016-0454-y
39. Higgins JP. Smartphone Applications for Patients' Health and Fitness. *Am J Med*. 2016;129(1):11-19. doi:10.1016/j.amjmed.2015.05.038
40. Torous J, Andersson G, Bertagnoli A, et al. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry*. 2019;18(1):97-98. doi:10.1002/wps.20592
41. International Organization for Standardization. *ISO 9241-210:2019 Ergonomics of Human-System Interaction - Part 210: Human-Centred Design for Interactive.*; 2019. Accessed September 29, 2025. <https://www.iso.org/standard/77520.html>
42. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. *J Biomed Inform*. 2004;37(1):56-76. doi:10.1016/j.jbi.2004.01.003
43. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral Functionality of Mobile Apps in Health Interventions: A Systematic Review of the Literature. *JMIR Mhealth Uhealth*. 2015;3(1):e20. doi:10.2196/mhealth.3335
44. Matthew-Maich N, Harris L, Ploeg J, et al. Designing, Implementing, and Evaluating Mobile Health Technologies for Managing Chronic Conditions in Older Adults: A Scoping Review. *JMIR Mhealth Uhealth*. 2016;4(2):e29. doi:10.2196/mhealth.5127
45. IDEO.org. *The Field Guide to Human-Centered Design*. IDEO; 2015. Accessed September 29, 2025. <https://www.designkit.org/resources/1>
46. Stonbraker S, Porras T, Schnall R. Patient preferences for visualization of longitudinal patient-reported outcomes data. *Journal of the American Medical Informatics Association*. 2020;27(2):212-224. doi:10.1093/jamia/ocz189
47. Porter ME. What Is Value in Health Care? *New England Journal of Medicine*. 2010;363(26):2477-2481. doi:10.1056/NEJMp1011024
48. Draper SW., Norman DA. *User Centered System Design : New Perspectives on Human-Computer Interaction*. CRC Press; 1986.
49. Zhang S, Song J. An empirical investigation into the preferences of the elderly for user interface design in personal electronic health record systems. *Front Digit Health*. 2024;5. doi:10.3389/fdgth.2023.1289904
50. Asfandiyar S, Hough S, Agugo V, Brixey JJ, Gong Y. Intuitive Data Visualization for Enhanced Patient Comprehension. In: 2024. doi:10.3233/SHTI240102
51. Kanzler CM, Lamers I, Feys P, Gassert R, Lambercy O. Personalized prediction of rehabilitation outcomes in multiple sclerosis: a proof-of-concept using clinical data, digital health metrics, and machine learning. *Med Biol Eng Comput*. 2022;60(1):249-261. doi:10.1007/s11517-021-02467-y
52. Kanzler CM, Sylvester R, Gassert R, Kool J, Lambercy O, Gonzenbach R. Goal-directed upper limb movement patterns and hand grip forces in multiple sclerosis. *Mult Scler J Exp Transl Clin*. 2022;8(3). doi:10.1177/20552173221116272
53. Kanzler CM, Averta G, Schwarz A, et al. A low-dimensional representation of arm movements and hand grip forces in post-stroke individuals. *Sci Rep*. 2022;12(1). doi:10.1038/s41598-022-11806-4
54. Kanzler CM, Lessard I, Gassert R, Brais B, Gagnon C, Lambercy O. Reliability and validity of digital health metrics for assessing arm and hand impairments in an ataxic disorder. *Ann Clin Transl Neurol*. 2022;9(4):432-443. doi:10.1002/acn3.51493
55. Kanzler CM, Lessard I, Gassert R, Brais B, Gagnon C, Lambercy O. Digital health metrics reveal upper limb impairment profiles in ARSACS. *J Neurol Sci*. 2023;448:120621.

- doi:10.1016/J.JNS.2023.120621
56. Likert R. A technique for the measurement of attitudes. *Archives of Psychology*. 1932;140:1-55.
 57. Roman LC, Ancker JS, Johnson SB, Senathirajah Y. Navigation in the electronic health record: A review of the safety and usability literature. *J Biomed Inform*. 2017;67:69-79. doi:10.1016/j.jbi.2017.01.005
 58. Turchioe MR, Myers A, Isaac S, et al. A Systematic Review of Patient-Facing Visualizations of Personal Health Data. *Appl Clin Inform*. 2019;10(4):751-770. doi:10.1055/s-0039-1697592
 59. Snyder C, Smith K, Holzner B, et al. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Qual Life Res*. 2019;28(2):345-356. doi:10.1007/s11136-018-2020-3
 60. Figma. Figma: collaborative interface design tool. Published online 2025. Accessed September 29, 2025. <https://www.figma.com>
 61. Brooke J. Usability evaluation in industry - SUS: a “quick and dirty” usability scale. In: Jordan PW TBWBM AL, ed. Taylor & Francis; 1996.
 62. Facebook. React Native . 2025. Accessed September 29, 2025. <https://reactnative.dev>
 63. Expo. Expo documentation. Published online 2025. Accessed September 29, 2025. <https://docs.expo.dev/>
 64. Holovaty A WS. The Django web framework. In: *Proceedings of the 2005 Python Conference*. 2005. Accessed September 29, 2025. <https://www.djangoproject.com>
 65. ISO/IEC. *ISO/IEC 9075:2016 – Information Technology – Database Languages – SQL.*; 2016. Accessed September 29, 2025. <https://www.iso.org/standard/63555.html>
 66. Dumurgier J, Tzourio C. Epidemiology of neurological diseases in older adults. *Rev Neurol (Paris)*. 2020;176(9):642-648. doi:10.1016/j.neurol.2020.01.356
 67. Ramesh SH, Ouskine A, Khorasani E, et al. A Data Visualization Tool for Patients and Healthcare Providers to Communicate during Inpatient Stroke Rehabilitation. In: *Graphics Interface*. ACM; 2024:1-14. doi:10.1145/3670947.3670978
 68. Choe EK, Duarte ME, Suh H, Pratt W, Kientz JA. Communicating Bad News: Insights for the Design of Consumer Health Technologies. *JMIR Hum Factors*. 2019;6(2):e8885. doi:10.2196/humanfactors.8885
 69. Fallowfield L, Jenkins V. Communicating sad, bad, and difficult news in medicine. *The Lancet*. 2004;363(9405):312-319. doi:10.1016/S0140-6736(03)15392-5
 70. Oyake K, Yamauchi K, Inoue S, et al. A multicenter explanatory survey of patients’ and clinicians’ perceptions of motivational factors in rehabilitation. *Communications Medicine*. 2023;3(1). doi:10.1038/s43856-023-00308-7
 71. Zadeh SM, MacDermid J, Johnson J, Birmingham TB, Shafiee E. Applications of wearable sensors in upper extremity MSK conditions: a scoping review. *J Neuroeng Rehabil*. 2023;20(1):158. doi:10.1186/s12984-023-01274-w
 72. Kim GJ, Parnandi A, Eva S, Schambra H. The use of wearable sensors to assess and treat the upper extremity after stroke: a scoping review. *Disabil Rehabil*. 2022;44(20):6119-6138. doi:10.1080/09638288.2021.1957027
 73. Peters DM, O’Brien ES, Kamrud KE, et al. Utilization of wearable technology to assess gait and mobility post-stroke: a systematic review. *J Neuroeng Rehabil*. 2021;18(1):67. doi:10.1186/s12984-021-00863-x
 74. Hulleck AA, Menoth Mohan D, Abdallah N, El Rich M, Khalaf K. Present and future of gait assessment in clinical practice: Towards the application of novel trends and technologies. *Front Med Technol*. 2022;4. doi:10.3389/fmedt.2022.901331
 75. Salchow-Hömmen C, Skrobot M, Jochner MCE, Schauer T, Kühn AA, Wenger N. Review-Emerging Portable Technologies for Gait Analysis in Neurological Disorders. *Front Hum Neurosci*. 2022;16:768575. doi:10.3389/fnhum.2022.768575
 76. Polk SE, Öhman F, Hassenstab J, et al. A scoping review of remote and unsupervised digital

cognitive assessments in preclinical Alzheimer's disease. *NPJ Digit Med.* 2025;8(1):266.
doi:10.1038/s41746-025-01583-5



Supplementary Files

Multimedia Appendixes

Untitled.

URL: <http://asset.jmir.pub/assets/2d3da4206e0b8042733168486fb02ece.pdf>